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

(21) BR 112020001453-2 A2

(22) Data do Depósito: 17/07/2018

(43) Data da Publicação Nacional: 28/07/2020

(54) Título: APARELHO E MÉTODO PARA GERAR UMA REPRESENTAÇÃO DE UMA CENA EM IMAGEM TRIDIMENSIONAL EM MOSAICO E PRODUTO DE PROGRAMA DE COMPUTADOR

(51) Int. Cl.: H04N 13/128; G06T 15/00; G06T 17/00.

(30) Prioridade Unionista: 25/07/2017 EP 17182985.6.

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

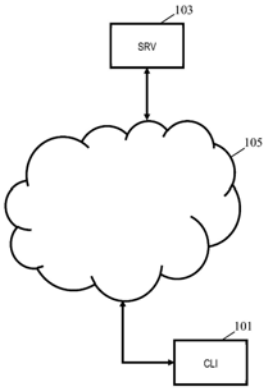
(72) Inventor(es): CHRISTIAAN VAREKAMP.

(86) Pedido PCT: PCT EP2018069387 de 17/07/2018

(87) Publicação PCT: WO 2019/020433 de 31/01/2019

(85) Data da Fase Nacional: 23/01/2020

(57) Resumo: A presente invenção refere-se a um aparelho para gerar uma representação de uma cena em imagem tridimensional, o qual compreende um receptor (301) que recebe uma representação de uma cena em imagem tridimensional em mosaico proveniente de um primeiro ponto de vista, sendo que a representação compreende uma pluralidade de blocos interconectados, sendo que cada bloco compreende um mapa de profundidade e um mapa de textura representando uma janela de visão da cena a partir do primeiro ponto de vista, e sendo que os blocos formam um padrão de mosaico. Um primeiro processador (311) determina as regiões de borda vizinhas em ao menos um primeiro bloco e num segundo bloco em resposta ao padrão de mosaico. Um segundo processador (309) modifica pelo menos um primeiro valor de profundidade de uma primeira região de borda do primeiro bloco, em resposta a pelo menos um segundo valor de profundidade em uma segunda região de borda do segundo bloco, sendo que as regiões de borda são regiões vizinhas. Os mapas de profundidade modificados podem ser usados para gerar uma malha com base na qual podem ser geradas as imagens.



APARELHO E MÉTODO PARA GERAR UMA REPRESENTAÇÃO DE UMA CENA EM IMAGEM TRIDIMENSIONAL EM MOSAICO E PRODUTO DE PROGRAMA DE COMPUTADOR

CAMPO DA INVENÇÃO

[001] A invenção se refere a um aparelho e a um método para gerar uma representação de uma cena em imagem tridimensional em mosaico e, em particular, mas não exclusivamente, para gerar uma melhor representação de uma cena em imagem tridimensional em mosaico adequada à geração de imagens com base em uma conversão para uma representação em malha.

ANTECEDENTES DA INVENÇÃO

[002] Tradicionalmente, o uso e o processamento técnico de imagens tem sido baseado em imageamento bidimensional mas, cada vez mais, a terceira dimensão está sendo explicitamente considerada no processamento de imagens.

[003] Por exemplo, foram desenvolvidas telas tridimensionais (3D) que acrescentam uma terceira dimensão à experiência de visão mediante o fornecimento, aos dois olhos de um observador, de diferentes vistas da cena sendo observada. Isso pode ser obtido fazendo com que o usuário use óculos para separar duas vistas que são mostradas. Entretanto, como isso pode ser considerado inconveniente para o usuário, em muitos cenários é preferencial o uso de telas autoestereoscópicas que usam meios na tela (como lentes lenticulares ou barreiras) para separar vistas e para enviá-las em diferentes direções onde podem individualmente chegar aos olhos do usuário. Para telas estereoscópicas são exigidas duas vistas, enquanto as telas autoestereoscópicas geralmente exigem mais vistas (como, por exemplo, nove vistas).

[004] Em muitas modalidades, pode ser desejável gerar imagens de vista para novas direções de observação. Embora sejam conhecidos vários algoritmos que geram essas novas imagens com base em informações de imagem e de profundidade, eles tendem a ser altamente dependentes da precisão da informação de profundidade fornecida (ou derivada).

[005] De fato, as informações de imagem tridimensional são frequentemente fornecidas por uma pluralidade de imagens correspondendo a diferentes direções de observação para uma cena. Especificamente, conteúdos de vídeo, como filmes ou programas de televisão, são cada vez mais gerados de modo a incluir algumas informações de 3D. Essas informações podem ser capturadas com o uso de câmeras 3D ou detectoras de profundidade dedicadas, como as câmeras que capturam duas imagens simultâneas provenientes de posições de câmera ligeiramente deslocadas.

[006] Em muitas aplicações, porém, as imagens fornecidas podem não corresponder diretamente às direções desejadas, ou mais imagens podem ser necessárias. Por exemplo, para telas autoestereoscópicas, mais de duas imagens são necessárias e, de fato, frequentemente são usadas de 9 a 26 imagens de vista.

[007] A fim de gerar imagens correspondendo a diferentes direções de observação, pode ser usado um processamento de deslocamento de ponto de vista. Isso é tipicamente executado por um algoritmo de deslocamento de vista que usa uma imagem para uma única direção de observação juntamente com informações de profundidade associadas.

[008] Um exemplo específico de uma aplicação que é baseada em processamento de imagens tridimensionais é uma

aplicação de realidade virtual. Em experiências típicas de realidade virtual, a imagem de vista do olho direito e a imagem de vista do olho esquerdo podem ser geradas continuamente, por exemplo, para um headset de realidade virtual para haver a correspondência do movimento e da mudança de orientação por parte do usuário. Essa geração de vistas dinâmicas de realidade virtual baseia-se, tipicamente, no processamento dos dados da imagem 3D que representam uma dada cena correspondente ao ambiente de realidade virtual. Por exemplo, um servidor de realidade virtual pode gerar imagens de vista para visualizações específicas com base em um modelo tridimensional ou em imagens tridimensionais, como imagens representadas por imagens de intensidade de luz e mapas de profundidade, ou mapas de textura e malhas de profundidade.

[009] Para aplicações como aplicações de realidade virtual, uma sequência de imagens pode ser gerada, por exemplo, para refletir as visões de um usuário à medida que estas mudam quando o usuário se move virtualmente ou muda de orientação/direção de observação no ambiente virtual. Em algumas aplicações, as imagens podem ser geradas para refletir alterações na orientação do observador mas sem refletir a movimentação pelo ambiente. Um vídeo que reflita tal cenário é chamado, com frequência, de vídeo omnidirecional. Em outras aplicações, uma posição de visualização em movimento pode também ser suportada para refletir o movimento virtual de um usuário no ambiente de realidade virtual. Um vídeo que reflita tal cenário é chamado, com frequência, de vídeo imersivo. A vista atual de um usuário pode ser representada por um vetor de vista que descreve parâmetros relevantes de posição e de direção para um ponto de vista.

[0010] Para um vídeo omnidirecional, um vetor de vista tipicamente descreve uma orientação de acordo com três graus de liberdade (3DoF - “3 Degrees of Freedom”), tipicamente mediante o fornecimento de valores de guinada, arfagem e rolamento (ou azimute, elevação e inclinação).

[0011] Para vídeo imersivo, um vetor tipicamente descreve tanto uma orientação quanto uma posição, de acordo com seis graus de liberdade (6DoF), tipicamente ao fornecer valores de guinada, inclinação, rotação e para três dimensões espaciais.

[0012] No entanto, um desafio particular ao tentar desenvolver e dar suporte a aplicações flexíveis de vídeo e imagem que suportem posições e/ou direções de vista variáveis é que, de preferência, as mesmas não se limitem a um subconjunto de posições e/ou direções, mas que todas as posições e/ou direções sejam idealmente suportadas. Por exemplo, para vídeo imersivo com 6DoF, um observador pode estar visualizando a cena a partir de qualquer posição e em qualquer direção. Isso requer que informações 3D estejam disponíveis para todas as partes da cena e a partir de todas as posições e para todas as direções. Este requisito é difícil ou impossível de atender em muitas aplicações práticas, como aplicações específicas nas quais os dados 3D são baseados em capturas de uma cena do mundo real.

[0013] A combinação da posição e a direção/orientação de um objeto é, no campo, geralmente chamada de um posicionamento ou uma pose. Dessa forma, um vetor de posicionamento ou pose pode compreender seis valores/componentes com cada valor/componente descrevendo uma propriedade individual da posição/localização ou da

orientação/direção do objeto correspondente. É claro que, em muitas situações, os vetores de posicionamento ou pose podem ter um menor número de componentes para representar a posição e a orientação, por exemplo se um ou mais componentes forem considerados fixos (por exemplo se todos os objetos forem considerados estarem na mesma altura e nivelados, quatro componentes podem fornecer uma representação completa da pose de um objeto).

[0014] Um dos desafios mais difíceis de se vencer é o de como representar a cena eficazmente, de modo que a representação forneça informações suficientes para possibilitar que imagens de ponto de vista adequadas sejam geradas com qualidade suficientemente alta, e ainda assim, possibilitem que a representação seja facilmente comunicada, armazenada, adaptada e/ou processada.

[0015] Uma abordagem usada é o fornecimento de uma imagem omnidirecional para uma dada posição de vista, com a imagem omnidirecional fornecendo, por exemplo, informações visuais como uma esfera ou semi-esfera centralizada na posição de vista correspondente onde a esfera tem uma imagem sobreposta que reflete a propriedade visual da cena na direção de visão correspondente. Em muitas aplicações, essa abordagem foi modificada por, em vez de representar a vista como uma textura de um objeto esférico circundando o ponto de vista, mapear a textura sobre uma estrutura tridimensional baseada em polígono, como tipicamente um cubo posicionado em torno do ponto de vista. Em tais casos, a vista omnidirecional a partir de um determinado ponto de vista é consequentemente mapeada por textura sobre um cubo posicionado em torno do ponto de vista. Ao gerar uma imagem correspondente a uma direção de

visão específica para um observador posicionado no ponto de vista, o sistema gerador de imagens pode simplesmente mapear sobre a imagem correspondente as partes adequadas do cubo para a dada janela de visão.

[0016] Entretanto, embora esse tipo de abordagem possa fornecer uma abordagem vantajosa em muitas modalidades, ela pode não ser ideal em todos os cenários, e pode ter algumas desvantagens associadas. Por exemplo, isso pode não fornecer qualidade ótima de imagem em todas as situações e, em particular, o processamento da textura mapeada sobre diferentes lados do cubo pode resultar em artefatos.

[0017] Consequentemente, seria vantajosa uma abordagem aprimorada para gerar uma representação de uma cena em imagem tridimensional. Em particular, seria vantajosa uma abordagem que possibilitasse uma operação aprimorada, maior flexibilidade, uma implementação facilitada, operação facilitada, complexidade reduzida, demanda reduzida de recursos e/ou desempenho aprimorado.

SUMÁRIO DA INVENÇÃO

[0018] Consequentemente, a presente invenção busca, de preferência, mitigar, aliviar ou eliminar uma ou mais dentre as desvantagens mencionadas acima, individualmente ou em qualquer combinação.

[0019] De acordo com um aspecto da invenção, é fornecido um aparelho para gerar uma representação de uma cena em imagem tridimensional em mosaico, sendo que o aparelho compreende: um receptor para receber uma representação de uma cena em imagem tridimensional em mosaico de um primeiro ponto de vista, sendo que a representação de imagem tridimensional em mosaico compreende uma pluralidade de blocos

interconectados, sendo que cada bloco compreende um mapa de profundidade e um mapa de textura representando uma janela de visão da cena a partir do primeiro ponto de vista, e sendo que os blocos formam um padrão em mosaico; um primeiro processador para determinar as regiões de borda vizinhas em ao menos um primeiro bloco e num segundo bloco em resposta ao padrão de mosaico, sendo que o primeiro bloco e o segundo bloco são blocos vizinhos; um segundo processador para modificar pelo menos um primeiro valor de profundidade de uma primeira região de borda do primeiro bloco, em resposta a pelo menos um segundo valor de profundidade em uma segunda região de borda do segundo bloco, de modo que uma diferença entre o primeiro valor de profundidade e o segundo valor de profundidade seja reduzida para ao menos alguns valores do primeiro valor de profundidade e do segundo valor de profundidade, sendo que a primeira região de borda e a segunda região de borda são regiões de borda vizinhas.

[0020] A invenção pode proporcionar um desempenho aprimorado em muitos cenários e pode, por exemplo, fornecer uma representação de imagem tridimensional que pode ser comunicada, armazenada e processada eficazmente. A invenção pode, particularmente em muitos cenários específicos, fornecer uma abordagem de baixa complexidade para gerar uma representação de imagem tridimensional que pode muitas vezes fornecer uma representação com taxa de dados mais baixa para uma dada qualidade de imagem, e que pode além disso suportar, por exemplo, distribuição flexível e parcial de dados, o que pode possibilitar, por exemplo, que clientes remotos recebam dados e gerem localmente as imagens representando a cena a partir de diferentes pontos de vista e direções. A abordagem

pode, em particular, possibilitar uma experiência de realidade virtual altamente eficiente onde, por exemplo, um cliente remoto de realidade virtual pode ser suprido com dados adequados de maneira flexível, como e quando for necessário.

[0021] O receptor pode receber uma representação de imagem tridimensional em mosaico de entrada e, a partir desta, gerar uma representação de imagem tridimensional em mosaico de saída, mediante a modificação de ao menos o primeiro valor de profundidade e, tipicamente, de um grande número de valores de profundidade de borda de bloco.

[0022] A abordagem pode, em muitos cenários, proporcionar uma qualidade de imagem aprimorada e pode, em particular, reduzir artefatos e imprecisões. A abordagem pode refletir o entendimento de que os erros e artefatos podem, em particular, ocorrer para janelas de visão sobrepostas a blocos diferentes, e que esses erros e artefatos podem, frequentemente, resultar de inconsistências entre valores de profundidade nas zonas de borda dos diferentes blocos, como pode resultar da codificação e decodificação com perdas que é tipicamente executada em mapas de profundidade. A abordagem pode, em muitos cenários, reduzir ou mesmo eliminar muitos desses artefatos e erros. A abordagem pode, em muitos cenários, proporcionar uma experiência de usuário mais suave, sendo que as transições entre blocos são imperceptíveis.

[0023] A representação de imagem tridimensional da cena pode, em muitas modalidades, ser considerada um modelo tridimensional de ao menos parte da cena.

[0024] O segundo processador pode, em muitas modalidades, estar disposto de modo a restringir uma diferença entre o primeiro valor de profundidade e o segundo valor de

profundidade, mediante a modificação do primeiro valor de profundidade de modo que uma diferença entre o primeiro valor de profundidade e o segundo valor de profundidade esteja restrita para todos os potenciais valores da primeira profundidade. Especificamente, o segundo processador pode, em muitas modalidades, estar disposto de modo a modificar o primeiro valor de profundidade a fim de que uma diferença entre o primeiro valor de profundidade e o segundo valor de profundidade não exceda um limite.

[0025] A modificação pelo segundo processador pode alterar o valor do primeiro valor de profundidade para apenas alguns valores do primeiro valor de profundidade e do segundo valor de profundidade. Para alguns valores do primeiro valor de profundidade e do segundo valor de profundidade, o primeiro valor de profundidade pode não ser alterado, isto é, o valor de profundidade modificado pode ser igual ao valor de profundidade original. Por exemplo, se o primeiro e o segundo valores de profundidade forem iguais, o primeiro valor de profundidade pode não ser alterado pelo segundo processador.

[0026] Em muitas modalidades, o padrão de mosaico pode ser um padrão de mosaico predeterminado/predefinido. O padrão de mosaico é não-planar. Especificamente, o padrão de mosaico pode ser tridimensional e pode ter extensões em todas as três dimensões. Os blocos e o padrão de mosaico podem formar uma estrutura tridimensional.

[0027] Cada valor de profundidade pode corresponder a um pixel de profundidade em um mapa de profundidade (tipicamente) retangular, ou especificamente quadrado, isto é, cada valor de profundidade pode ser um valor

de um pixel de profundidade em uma posição de pixel correspondente no mapa de profundidade correspondente.

[0028] O aparelho pode estar disposto de modo a modificar uma pluralidade de primeiros valores de profundidade, como especificamente todos os valores de profundidade na primeira região de borda. Especificamente, a abordagem pode ser repetida para uma pluralidade de pixels de profundidade e, possivelmente, para cada pixel de profundidade na primeira região de borda.

[0029] O primeiro e o segundo valores de profundidade podem ser selecionados como valores de profundidade correspondentes a uma mesma direção de visão a partir do primeiro ponto de vista. O primeiro e o segundo valores de profundidade podem ser selecionados de modo que uma medida de diferença para a diferença entre o vetor/direção de visão do primeiro ponto de vista em direção ao primeiro valor de profundidade e o vetor/direção de visão do primeiro ponto de vista em direção ao segundo valor de profundidade esteja abaixo de um limite. Por exemplo, pode ser necessário que um ângulo de visão entre os vetores de vista do primeiro ponto de vista em direção a posições para respectivamente o primeiro e o segundo valores de profundidade seja menor que um limite ou seja o menor para qualquer par de valores de profundidade no primeiro bloco e no segundo bloco.

[0030] Em muitas modalidades, o segundo valor de profundidade pode ser um valor de profundidade para um pixel de profundidade na segunda região mais próxima de um pixel de profundidade para o primeiro valor de profundidade.

[0031] Um bloco interconectado pode ter uma borda em comum com um e, tipicamente, com mais outros blocos

interconectados. Os blocos interconectados podem, para o padrão de mosaico, fornecer cobertura contígua para uma janela de visão, sendo que com cada bloco interconectado fornece uma sub-janela de visão que cobre uma subárea/região da totalidade da janela de visão. A totalidade da janela de visão pode tipicamente ser grande, por exemplo pode ser um quarto de esfera, uma semiesfera ou mesmo uma esfera completa.

[0032] As regiões de borda vizinhas podem ser regiões de dois blocos, onde as regiões compartilham uma borda entre os dois blocos. A primeira e a segunda regiões de borda podem, consequentemente, compartilhar uma borda entre o primeiro e o segundo blocos, e ambas podem especificamente conectar-se à/tocar na borda, ou uma ou ambas podem compreender a borda (em caso de blocos sobrepostos).

[0033] De acordo com um recurso opcional da invenção, o segundo processador está disposto de modo a definir o primeiro valor de profundidade e o segundo valor de profundidade para um mesmo valor.

[0034] Pode-se possibilitar baixa complexidade e baixa demanda de recursos, e ainda assim resultar em desempenho vantajoso com alta qualidade. Em particular, pode-se reduzir a ocorrência de artefatos, erros e degradações resultantes de inconsistências nos valores de profundidade de diferentes blocos, por exemplo devido a artefatos de codificação ou decodificação. Pode-se, em particular, possibilitar uma geração de malha 3D exata, consistente e robusta, com base em mapas de profundidade.

[0035] De acordo com um recurso opcional da invenção, o mesmo valor é um valor médio do primeiro valor de profundidade e do segundo valor de profundidade.

[0036] Pode-se possibilitar baixa complexidade e baixa demanda de recursos, e ainda assim resultar em desempenho vantajoso com alta qualidade. Adicionalmente, embora possa ser esperado que resulte em desvios de profundidade (em relação à profundidade original) que são maiores do que, por exemplo, aqueles para cálculo da média de profundidade, pode-se proporcionar suporte aprimorado à disposição assimétrica dos blocos da representação de imagem tridimensional. Por exemplo, facilita-se e reduz-se a carga computacional para uma abordagem na qual apenas um subconjunto de blocos é dinamicamente atualizado, e alguns dos blocos não sofrem alterações.

[0037] De acordo com um recurso opcional da invenção, o segundo processador está disposto de modo a definir o primeiro valor de profundidade para o segundo valor de profundidade.

[0038] Pode-se possibilitar baixa complexidade e baixa demanda de recursos, e ainda assim resultar em desempenho vantajoso com alta qualidade. Em particular, embora possa ser esperado que resulte em desvios de profundidade (em relação à profundidade original) que são maiores do que, por exemplo, aqueles para cálculo da média de profundidade, pode-se proporcionar suporte aprimorado à disposição assimétrica dos blocos da representação de imagem tridimensional. Por exemplo, facilita-se e reduz-se a carga computacional para uma abordagem na qual apenas um subconjunto de blocos é dinamicamente atualizado, e alguns dos blocos não sofrem alterações.

[0039] O aparelho pode impor condições de contorno para mapas de profundidade, como o mapa de profundidade do primeiro bloco, com base nas propriedades de profundidade dos mapas de profundidade de blocos vizinhos,

como o mapa de profundidade do segundo bloco. A abordagem pode, por exemplo, impor condições de contorno sobre um novo bloco de modo que ele seja consistente com, e se equipare a, blocos vizinhos existentes sem que estes precisem ser processados (por exemplo filtrados).

[0040] De acordo com um recurso opcional da invenção, o aparelho compreende adicionalmente um gerador de malha para gerar uma malha tridimensional a partir de mapas de profundidade dos blocos.

[0041] A abordagem pode possibilitar que seja gerada uma malha tridimensional aprimorada para a cena, com base nos mapas de profundidade dos blocos. Em particular, o risco de que ocorram inconsistências e especificamente orifícios na malha em torno das bordas de bloco pode ser substancialmente reduzido ou mesmo eliminado em muitos cenários.

[0042] A abordagem pode possibilitar que uma conversão ou tradução aprimorada de uma representação da cena baseada em imagem para uma representação espacial geométrica da cena por meio da malha tridimensional.

[0043] De acordo com um recurso opcional da invenção, o gerador de malha está disposto de modo a selecionar um subconjunto de valores de profundidade da primeira região de borda para uso na geração da malha tridimensional; e o segundo processador está disposto de modo a modificar valores de profundidade da primeira região de borda somente se pertencerem ao subconjunto de valores de profundidade.

[0044] Uma redução na complexidade e no uso de recursos computacionais pode ser obtida sem sacrificar a qualidade. A abordagem pode possibilitar uma consistência e

uma correspondência mais próximas entre a representação em bloco e a representação em malha gerada a partir da mesma.

[0045] De acordo com um recurso opcional da invenção, o segundo processador está disposto de modo a aplicar um filtro espacial a um mapa de profundidade do primeiro bloco, em seguida à modificação do primeiro valor de profundidade.

[0046] Pode-se, com isso, possibilitar uma qualidade aprimorada e, em particular, fornecer consistência melhorada entre regiões de borda modificadas e outras regiões dos blocos. A aplicação do filtro espacial pode ser limitada apenas valores de profundidade do mapa de profundidade do primeiro bloco.

[0047] De acordo com um recurso opcional da invenção, o aparelho compreende adicionalmente um comunicador de fonte remota disposto de modo a receber blocos para representação de imagem tridimensional em mosaico provenientes de uma fonte remota, e um armazenamento local para armazenar blocos para a representação de imagem tridimensional em mosaico; e o receptor está disposto de modo a recuperar ao menos o segundo valor de profundidade do armazenamento local, em resposta ao recebimento do primeiro bloco proveniente do receptor de fonte remota.

[0048] A abordagem pode possibilitar uma abordagem altamente eficiente para suporte à distribuição dinâmica de representações de uma cena, por exemplo para uma aplicação de realidade virtual. A abordagem pode, por exemplo, possibilitar o uso de uma largura de banda substancialmente reduzida, por só exigir que sejam transmitidos os blocos que tenham sido modificados, ao mesmo tempo em que possibilita que isto seja eficazmente combinado com blocos localmente

armazenados recebidos anteriormente ou, por exemplo, sendo blocos nominais predeterminados ou predefinidos que representa uma parte estática da cena.

[0049] O armazenamento local pode estar disposto de modo a armazenar blocos para a representação de imagem tridimensional em mosaico anteriormente recebida da fonte remota.

[0050] De acordo com um recurso opcional da invenção, o armazenamento local está disposto de modo a armazenar um mapa de profundidade ou uma malha para o primeiro bloco, em seguida à modificação do primeiro valor de profundidade.

[0051] Pode-se, com isso, possibilitar uma abordagem eficiente, a qual pode reduzir, por exemplo, os requisitos de comunicação para muitas aplicações. Pode-se, com isso, possibilitar uma atualização gradual de uma representação localmente armazenada de uma cena, e ao mesmo tempo fornecer consistência entre blocos novos e anteriores.

[0052] De acordo com um recurso opcional da invenção, o aparelho compreende adicionalmente uma entrada de dados pelo usuário para determinar uma direção de visão atual; e sendo que o comunicador de fonte remota está disposto de modo a transmitir à fonte remota uma indicação da direção de visão atual; e sendo que o primeiro bloco é recebido em resposta à indicação.

[0053] A abordagem pode, por exemplo, fornecer um sistema cliente/servidor muito eficiente para fornecer, por exemplo, uma experiência de realidade virtual.

[0054] De acordo com um recurso opcional da invenção, o armazenamento local está disposto de modo a

armazenar uma malha para os blocos armazenados, e o gerador de malha está disposto de modo a adicionar um vértice e ao menos uma borda a uma malha armazenada para o segundo bloco, em resposta a uma detecção de que uma medida de diferença excede um limite, sendo que a medida de diferença é indicativa de uma diferença de profundidade entre um valor de profundidade de borda do primeiro bloco e um correspondente valor de profundidade de borda da malha armazenada para o segundo bloco.

[0055] Pode-se, com isso, proporcionar um desempenho aprimorado em muitas modalidades.

[0056] De acordo com um recurso opcional da invenção, o segundo processador está disposto de modo a determinar o primeiro valor de profundidade em resposta a apenas um valor de profundidade mais próximo de cada bloco vizinho ao primeiro bloco, em uma posição do primeiro valor de profundidade.

[0057] Pode-se, com isso, possibilitar uma abordagem de baixa complexidade e, ainda assim, altamente eficiente e vantajosa, com alto desempenho na maioria dos cenários e para muitas aplicações.

[0058] De acordo com um recurso opcional da invenção, cada bloco é um polígono plano e a pluralidade de blocos interconectados forma uma estrutura tridimensional.

[0059] Pode-se, com isso, possibilitar uma abordagem de baixa complexidade e, ainda assim, altamente eficiente e vantajosa, com alto desempenho na maioria dos cenários e para muitas aplicações.

[0060] De acordo com um aspecto da invenção, é fornecido um método para gerar uma representação de uma cena em imagem tridimensional em mosaico, sendo que o método

compreende: receber uma representação de uma cena em imagem tridimensional em mosaico de um primeiro ponto de vista, sendo que a representação de imagem tridimensional em mosaico compreende uma pluralidade de blocos interconectados, sendo que cada bloco compreende um mapa de profundidade e um mapa de textura representando uma janela de visão da cena a partir do primeiro ponto de vista, e sendo que os blocos formam um padrão em mosaico; determinar as regiões de borda vizinhas em ao menos um primeiro bloco e num segundo bloco em resposta ao padrão de mosaico, sendo que o primeiro bloco e o segundo bloco são blocos vizinhos; modificar pelo menos um primeiro valor de profundidade de uma primeira região de borda do primeiro bloco, em resposta a pelo menos um segundo valor de profundidade em uma segunda região de borda do segundo bloco, de modo que uma diferença entre o primeiro valor de profundidade e o segundo valor de profundidade seja reduzida para ao menos alguns valores do primeiro valor de profundidade e do segundo valor de profundidade, sendo que a primeira região de borda e a segunda região de borda são regiões de borda vizinhas.

[0061] Esses e outros aspectos, recursos e vantagens da invenção ficarão evidentes e serão elucidados com referência às uma ou mais modalidades descritas a seguir.

BREVE DESCRIÇÃO DOS DESENHOS

[0062] As modalidades da invenção serão descritas, apenas a título de exemplo, com referência aos desenhos, nos quais:

[0063] A Figura 1 ilustra um exemplo de um sistema de realidade virtual cliente-servidor que compreende algumas modalidades da invenção;

[0064] A Figura 2 ilustra um exemplo de elementos de uma representação de uma cena em um mosaico de cubo;

[0065] A Figura 3 ilustra um exemplo de um aparelho gerador de imagens, de acordo com algumas modalidades da invenção;

[0066] A Figura 4 ilustra um exemplo de elementos de uma representação de uma cena em um mosaico de cubo;

[0067] A Figura 5 ilustra um exemplo de elementos de uma representação de uma cena em um mosaico de cubo;

[0068] A Figura 6 ilustra um exemplo de um aparelho gerador de imagens, de acordo com algumas modalidades da invenção; e

[0069] A Figura 7 ilustra um exemplo de elementos de uma representação de uma cena em um mosaico de cubo.

DESCRIÇÃO DETALHADA DAS MODALIDADES

[0070] A descrição a seguir enfoca as modalidades da invenção aplicáveis à geração de imagens para uma aplicação de realidade virtual. Entretanto, deve-se considerar que a invenção não se limita a essa aplicação, mas que ela pode ser utilizada em, por exemplo, muitas outras aplicações diferentes de processamento e geração de imagens.

[0071] A Figura 1 ilustra um exemplo de um sistema que pode incorporar exemplos e variantes do conceito da invenção.

[0072] No exemplo, um cliente sob a forma de um aparelho gerador de imagens 101 está disposto de modo a gerar imagens representando uma cena a partir de diferentes pontos de vista, isto é, podem ser geradas imagens correspondentes a diferentes janelas de visão da cena. O aparelho gerador de imagens 101 está disposto de modo a receber uma representação

de uma cena em imagem tridimensional e, então, gerar as imagens a partir dessa representação. No exemplo, a representação de imagem tridimensional é recebida de um servidor remoto 103 através de uma rede 105 que pode ser, por exemplo, a internet. No exemplo específico, o servidor remoto 103 é um servidor de realidade virtual que fornece uma representação de imagem tridimensional de um ambiente tridimensional para que o aparelho gerador de imagens 101 gere vistas correspondentes, por exemplo, ao movimento de um usuário virtual no ambiente.

[0073] Em muitas aplicações, a representação de imagem tridimensional pode ser gerada a partir de uma captura de uma cena ou de um ambiente do mundo real mediante o uso de câmeras detectoras de profundidade. Possibilita-se, com isso, que as propriedades visuais sejam capturadas juntamente com informações tridimensionais. A fim de capturar uma cena suficientemente, adota-se, quase sempre, o uso de uma pluralidade e, frequentemente, usa-se um alto número de câmeras. Em algumas aplicações, são usadas 10, 20 ou mesmo mais câmeras para fornecer uma captura satisfatória.

[0074] A representação de dados virtuais de uma cena é um fator de importância crítica no fornecimento de uma experiência de usuário vantajosa. É necessário que os dados que descrevem a cena forneçam uma representação exata tanto das propriedades visuais como das propriedades espaciais. Ao mesmo tempo, é de importância crítica reduzir a quantidade de dados necessária para representar a cena, já que, em muitas aplicações, isso tende a ser um fator limitante para a qualidade que pode ser obtida.

[0075] Além disso, a conversão das capturas das câmeras detectoras de profundidade para uma representação de

dados do ambiente/cena é quase sempre muito desafiadora e pode introduzir erros ou artefatos. Por exemplo, em algumas aplicações, os dados capturados podem ser usados para desenvolver um modelo tridimensional da cena do mundo real. As imagens de vista para um usuário ao qual é fornecida uma experiência de realidade virtual tridimensional podem, então, ser geradas mediante a avaliação do modelo a partir de um ponto de vista específico. Em outras aplicações, as imagens para janelas de visão ou pontos de vista específicos podem ser diretamente geradas a partir das imagens capturadas e das informações de profundidade, por exemplo mediante a seleção de uma ou mais dentre as imagens capturadas mais próximas e a execução de um deslocamento do ponto de vista para que corresponda ao ponto de vista desejado ou, em alguns casos, mediante o uso direto das imagens capturadas.

[0076] No presente caso, a cena é representada no domínio de imagem por um conjunto de imagens, ao invés de sob a forma de um modelo tridimensional. Entretanto, cada imagem é adicionalmente dotada de mapas de profundidade que fornecem informações de profundidade para os objetos nas imagens. Especificamente, é fornecida uma representação da cena em imagem tridimensional compreendendo um conjunto de imagens e mapas de profundidade associados.

[0077] Em particular, no sistema da Figura 1, uma representação de uma cena em imagem tridimensional em mosaico de um dado primeiro ponto de vista é usada para (ao menos parcialmente) representar a cena. Para um dado ponto de vista, é fornecida uma representação na qual um conjunto de blocos interconectados representa a cena, sendo que cada bloco compreende um mapa de textura/imagem que fornece informações

de intensidade de luz (inclusive, tipicamente, cor). Especificamente, uma dada janela de visão (grande) a partir do dado ponto de vista é dividida em uma pluralidade de subjanelas de visão menores, sendo que cada uma dessas subjanelas de visão é representada por um bloco.

[0078] Cada um dos blocos fornece, assim, uma representação de uma parte da cena, sendo que os blocos juntos fornecem uma representação de uma parte maior da cena. Tipicamente, os blocos se interconectam ao longo de linhas e junções, sendo que há uma borda distinta entre os blocos. Entretanto, em algumas modalidades, os blocos individuais podem ter uma sobreposição (geralmente muito pequena).

[0079] Dessa forma, para uma dada direção de visão a partir do primeiro ponto de vista, a cena será representada pelo bloco que representa uma subjanela de visão incluindo essa direção de visão. Na maioria das modalidades, há uma correspondência direta de um para um entre a direção de visão e um bloco que contém as informações relevantes. Entretanto, conforme mencionado, em alguns casos pode haver uma pequena sobreposição entre blocos, caso no qual pode haver potencialmente mais de um bloco associado à direção de visão.

[0080] Juntos, os blocos formam um padrão de mosaico. Assim, a grande janela de visão representada pelos blocos é dividida em blocos individuais em conformidade com o padrão de mosaico. Em muitas modalidades, o padrão de mosaico pode ser um padrão de mosaico predeterminado que é conhecido tanto pelo servidor remoto 103 como pelo aparelho gerador de imagens 101. Em outras modalidades, o padrão de mosaico pode por exemplo ser dinamicamente determinado pelo servidor remoto 103, e o padrão de mosaico pode ser comunicado ao aparelho

gerador de imagens 101. O padrão de mosaico tipicamente cobre a totalidade da janela de visão contiguamente e, assim, para cada direção de visão na janela de visão, o padrão de mosaico terá (ao menos um) bloco.

[0081] O padrão de blocos pode consequentemente formar uma janela de visão no mundo/sistema de coordenadas da cena, sendo que cada bloco corresponde a uma janela de visão (menor, ou subjanela) da cena. Em outras palavras, as janelas de visão dos blocos podem, juntas, formar uma janela de visão combinada para o primeiro ponto de vista. Os blocos podem fornecer cobertura contígua da janela de visão combinada. A janela de visão combinada pode circundar total ou parcialmente o primeiro ponto de vista.

[0082] Os blocos interconectados podem compartilhar bordas no mundo/sistema de coordenadas da cena (inclusive bordas sobrepostas). Em muitas modalidades, cada um dentre dois blocos vizinhos ou adjacentes pode representar diferentes regiões da cena, sendo que os dois blocos vizinhos ou adjacentes representam, juntos, uma única região contígua (correspondente às duas regiões individuais combinadas). A combinação de duas regiões representadas por dois blocos adjacentes ou vizinhos pode formar uma única região. Cada bloco do padrão de mosaico pode representar uma janela de visão para uma região diferente da cena, sendo que o padrão de mosaico formado pelos blocos representa uma janela de visão para uma região contígua da cena. A combinação de blocos correspondentes a diferentes janelas de visão para diferentes regiões da cena pode formar uma janela de visão combinada para uma região contígua da cena. Uma região contígua pode ser uma região correspondente a um único formato tridimensional na cena

tridimensional. As regiões da cena representadas pelos blocos do padrão de mosaico podem se combinar a um único formato/região tridimensional na cena (ao contrário de, por exemplo, as regiões combinando-se em dois ou mais formatos separados que não se tocam nem se sobrepõem).

[0083] Cada bloco pode compreender um mapa de profundidade e um mapa de textura representando a projeção da cena sobre uma área/janela de visão (no sistema de coordenadas da cena) correspondente ao bloco. Os blocos do padrão de mosaico podem refletir a projeção da cena sobre uma janela de visão correspondente à combinação/soma das janelas de visão dos blocos do padrão de mosaico. Os blocos do padrão de mosaico correspondem a janelas de visão para o mesmo ponto de vista.

[0084] Uma janela de visão para um ponto de vista pode corresponder a uma área (ou superfície 3D) na cena (sistema de coordenadas) sobre a qual uma projeção da cena a partir do ponto de vista pode ser fornecida. Um ponto de vista pode ser uma pose de vista ou, em muitas modalidades, uma posição de vista.

[0085] O termo ponto de vista pode, por exemplo, ser considerado refletir um ponto de referência para as vistas/janelas de visão, e é geralmente considerado como correspondente ao observador/câmera nominal da cena. Isso é frequentemente considerado como correspondente a uma posição, mas pode também incluir um aspecto direcional. Isso pode ser considerado, de modo mais genérico, como correspondente a uma pose. Por exemplo, um ponto de vista pode refletir uma posição de um observador e uma direção na qual ele está olhando. Uma representação em mosaico da cena para esse ponto de vista pode, por exemplo, ser uma estrutura substancialmente semiesférica

de blocos centralizados naquela direção (isto é, a representação em mosaico específica pode ser para uma posição e uma direção específicas).

[0086] De fato, em muitas modalidades práticas, uma representação em mosaico pode representar apenas uma parte de uma cena, ao invés de uma representação esférica completa. Por exemplo, para uma representação em cubo, um ou mais dos lados/blocos do cubo pode estar faltando. Assim, uma representação em mosaico para uma cena a partir de um dado ponto de vista pode não só refletir a posição onde o observador/câmera nominalmente se encontra, mas também uma direção de visão (por exemplo uma faixa de possíveis direções de visão).

[0087] O termo janela de visão pode, por exemplo, ser considerado corresponder a uma área ou janela para a cena, a partir de um dado ponto de vista. Por exemplo, uma janela de visão pode ser vista como uma superfície na qual parte da cena é projetada.

[0088] Em adição ao mapa de imagem/textura da intensidade de luz, cada bloco fornece também um mapa de profundidade que fornece informações de profundidade para os pixels do mapa de imagem/textura da intensidade de luz. O mapa de profundidade pode, tipicamente, compreender valores de profundidade que são indicativos da distância do primeiro ponto de vista ao objeto na direção de visão correspondente à posição do valor de profundidade. Um valor de profundidade pode, por exemplo, ter um valor crescente para uma distância crescente do ponto de vista ao objeto ou pode ter um valor decrescente para uma distância crescente do ponto de vista ao objeto. Em

muitas modalidades, os valores de profundidade podem ser fornecidos sob a forma de valores de disparidade.

[0089] O fornecimento de informações tanto de profundidade como de imagem (textura) pode fornecer informações adicionais que possibilitam um processamento aprimorado pelo aparelho gerador de imagens 101. Em particular, isso pode possibilitar, facilitar ou otimizar a geração de imagens de vista para outros pontos de vista diferentes do dado ponto de vista para o qual é fornecida a representação de imagem tridimensional.

[0090] Os blocos são, em muitas modalidades, polígonos planos, como retângulos (ou frequentemente quadrados), que juntos formam uma estrutura tridimensional, isto é, os blocos planos são posicionados e orientados no espaço tridimensional de modo que a estrutura resultante se estenda em três dimensões, ao invés de apenas se estender em um plano. A totalidade da janela de visão fornecida pela representação de imagem tridimensional em mosaico é, consequentemente, uma estrutura/janela de visão tridimensional.

[0091] Em muitas modalidades, a janela de visão tridimensional pode ser esférica, isto é, há um bloco para qualquer direção de visão a partir do primeiro ponto de vista. Especificamente, considerando-se uma esfera (correspondente a uma porta de visualização esférica) posicionada em torno do primeiro ponto de vista e dentro da estrutura formada pelo padrão de mosaico, uma linha que parte do primeiro ponto de vista e atravessa qualquer ponto a esfera se estenderá até uma posição em um bloco. Em algumas modalidades, a janela de visão pode não cobrir todas as possíveis direções de visão, mas pode

potencialmente cobrir apenas uma parte da esfera, como uma semiesfera ou um quarto de esfera.

[0092] Um exemplo específico de um padrão de mosaico é um padrão de mosaico em cubo como, por exemplo, é conhecido no mapeamento de cubo. Em um padrão de mosaico em cubo, seis blocos quadrados podem formar os seis lados de um cubo circundando o primeiro ponto de vista. Assim, nesse tipo de exemplo, a representação de imagem tridimensional em mosaico pode fornecer seis blocos quadrados onde cada bloco compreende um mapa de textura e um mapa de profundidade.

[0093] Os blocos podem frequentemente ser fornecidos como uma representação plana combinada. Entretanto, como os blocos formam uma estrutura tridimensional (um cubo), essa representação plana irá inerentemente separar ao menos alguns blocos vizinhos. Uma representação plana de um padrão de mosaico em cubo é ilustrado na Figura 2, onde são indicadas as representações planas dos mapas de textura e dos mapas de profundidade.

[0094] No sistema da Figura 1, o aparelho gerador de imagens 101 pode consequentemente receber uma representação de imagem tridimensional em mosaico de entrada, como uma representação de cubo, para uma cena de um dado primeiro ponto de vista. A partir dessa representação de imagem tridimensional em mosaico de entrada, o aparelho gerador de imagens 101 pode então gerar imagens correspondentes a uma direção de visão específica a partir do primeiro ponto de vista, ou de fato correspondentes a uma direção de visão específica a partir de um ponto de vista diferente. Consequentemente, no exemplo, o aparelho gerador de imagens 101 está disposto de modo a sintetizar imagens para diferentes pontos de vista e direções

a partir da representação de imagem tridimensional em mosaico. Por exemplo, o aparelho gerador de imagens 101 pode executar um aplicativo de realidade virtual que gera imagens para diferentes pontos de vista e direções, em resposta a uma entrada de dados pelo usuário como, por exemplo, a partir de um headset de realidade virtual.

[0095] A Figura 3 ilustra um exemplo do aparelho gerador de imagens 101 da Figura 1.

[0096] O aparelho gerador de imagens 101 compreende um receptor 301 que recebe a representação da cena em imagem tridimensional de entrada a partir do primeiro ponto de vista. No exemplo, o receptor 301 recebe a entrada de representação de imagem tridimensional em mosaico do servidor remoto 103 e, assim, o receptor 301 compreende uma interface de rede adequada e funcionalidade para se comunicar com o servidor remoto 103. Deve-se considerar que, em outras modalidades, a representação de imagem tridimensional em mosaico de entrada pode ser recebida de outras fontes internas ou externas, como de um armazenamento local ou leitor de disco de dados.

[0097] No exemplo, o aparelho gerador de imagens 101 compreende adicionalmente um gerador de imagens 303 que está disposto de modo a gerar imagens da representação de imagem tridimensional em mosaico recebida. O gerador de imagens 303 pode especificamente estar disposto de modo a gerar imagens para pontos de vista diferentes do primeiro ponto de vista para o qual a representação de imagem tridimensional em mosaico é fornecida.

[0098] No exemplo específico, o aparelho gerador de imagens 101 executa um aplicativo de realidade virtual e é

acoplado a uma interface de usuário 305, a qual pode controlar a posição do ponto de vista e da direção de visão para a qual são geradas as imagens de vista.

[0099] No exemplo específico, o aparelho gerador de imagens 101 é um dispositivo de realidade virtual que compreende uma interface de usuário 305 disposta de modo a receber uma entrada de dados pelo usuário, a qual pode refletir um movimento e/ou orientação desejada na cena de realidade virtual. Por exemplo, um usuário pode navegar através de uma cena de realidade virtual com o uso de um joystick e/ou teclado, ou, por exemplo, a orientação da vista pode ser determinada a partir de um headset de realidade virtual usado pelo usuário e que compreende sensores adequados. Em muitas modalidades, a interface de usuário 305 pode compreender uma funcionalidade para converter a entrada de dados pelo usuário em um ponto de vista de renderização desejado, o qual indica a posição a partir da qual a cena é observada e/ou a orientação da vista a partir de uma dada posição de vista (o ponto de vista pode ser considerado se referir a uma indicação de posição, uma indicação de orientação ou ambas, indicando assim a vista desejada para um usuário).

[00100] Será reconhecido que em algumas modalidades, pode ser gerada uma pluralidade de vistas para um dado ponto de vista de renderização ou em algumas modalidades, pode ser gerada uma pluralidade de pontos de vista a partir da mesma entrada de usuário. Por exemplo, o dispositivo de realidade virtual pode ser disposto para gerar uma imagem correspondente a uma visão pelo olho direito e uma imagem correspondente a uma visão pelo olho esquerdo, fornecendo assim um efeito 3D. Para fins de brevidade e clareza, a descrição a

seguir terá como foco a geração de uma imagem para um dado ponto de vista de renderização.

[00101] A interface de usuário 305 pode estar especificamente disposta de modo a gerar um vetor de vista de renderização, o qual é indicativo de um ponto de vista alvo/desejado de renderização. Por exemplo, pode ser gerado um vetor de vista de renderização o qual compreende três componentes/valores que definem uma orientação do observador como, por exemplo, parâmetros de arfagem, guinada e rolamento. Em muitas modalidades, o vetor de vista alvo pode compreender, adicional ou alternativamente, três valores que definem uma posição tridimensional na cena, por exemplo representada por um valor x, y, z. Deve-se considerar que em algumas modalidades, a posição e/ou a orientação podem ser representadas por menos de três valores. Por exemplo, uma posição pode ser indicada por uma distância ao longo de uma rota predeterminada através da cena, e dessa forma, a posição pode ser indicada por um valor único. De modo similar, uma inclinação e uma guinada podem ser presumidas constantes e a orientação pode ser dada apenas como um valor de rotação.

[00102] O vetor de vista de renderização gerado é fornecido ao gerador de imagens 303, que está disposto de modo a gerar uma imagem para o ponto de vista refletido pelo vetor de vista de renderização. A geração da imagem depende da representação de imagem tridimensional em mosaico.

[00103] Em algumas modalidades, a imagem gerada pode então ser fornecida a um driver de vídeo para acionar uma tela adequada como, por exemplo, uma tela de um headset de realidade virtual para o olho direito ou esquerdo. O headset de realidade virtual pode compreender adicionalmente sensores

que fornecem uma entrada de dados pelo usuário à unidade de interface de usuário 305.

[00104] Dessa forma, o dispositivo de realidade virtual pode gerar imagens que representam vistas a partir de pontos de vista específicos em uma cena virtual representada pela representação de imagem tridimensional em mosaico. O ponto de vista é controlado por uma entrada de usuário que permite que o usuário tipicamente se mova e olhe ao redor na cena virtual.

[00105] Consequentemente, as imagens são renderizadas pelo gerador de imagens 303, com base na representação de imagem tridimensional em mosaico de entrada recebida. Entretanto, o aparelho gerador de imagens 101 não usa diretamente a representação de imagem tridimensional em mosaico de entrada, mas a modifica antes de realizar a síntese de imagem. Assim, a renderização é executada com base em uma representação de imagem tridimensional em mosaico modificada, gerada mediante a modificação de ao menos um valor de profundidade da representação de imagem tridimensional em mosaico de entrada.

[00106] No exemplo específico, o gerador de imagens 303 não usa diretamente as informações do mapa de profundidade mas, em vez disso, executa uma conversão das mesmas em um formato de malha. Embora a geração de uma malha tridimensional a partir do mapa de profundidade, com o subsequente processamento/síntese de imagem tendo por base a malha, seja um recurso opcional, é possível fornecer uma operação e/ou implementação vantajosa em muitas modalidades.

[00107] Na verdade, embora os mapas de profundidade frequentemente formem um formato de captura e

compressão tridimensional natural, a maioria dos hardwares e softwares de 3D usados para, por exemplo, reprodução de realidade virtual, usa uma representação tridimensional baseada em malha que consiste em bordas e coordenadas de vértice tridimensional. Converter um mapa de profundidade em uma malha é, portanto, uma operação importante para muitos algoritmos offline e online (em tempo real).

[00108] No exemplo da Figura 3, o aparelho gerador de imagens 101 compreende, consequentemente, um gerador de malha 307 que está disposto de modo a gerar uma malha tridimensional com base nos mapas de profundidade dos blocos. Assim, o gerador de imagens 303 é alimentado com os mapas de textura de bloco provenientes do receptor 301, e a malha é gerada a partir do gerador de malha 307. O versado na técnica estará ciente de muitos algoritmos diferentes para sintetização de imagens a partir de mapas de textura e malhas, e deve-se considerar que qualquer destes pode ser usado sem depreciar a invenção.

[00109] Em muitas aplicações gráficas, a representação de uma cena se dá por meio de uma combinação de um mapa de textura e uma malha tridimensional e, portanto, hardware dedicado é frequentemente usado para gerar imagens a partir dessas representações. Uma abordagem particularmente eficaz em muitos cenários consiste em representar objetos de imagem ou, de fato, a cena como um todo, por meio de uma malha poligonal onde polígonos de um conjunto são conectados por suas bordas ou cantos (vértices) em comum, os quais são dados por meio de posições tridimensionais. A malha poligonal tridimensional combinada, consequentemente fornece um modelo eficaz de objetos tridimensionais, incluindo, possivelmente,

uma descrição tridimensional de uma imagem completa. A malha poligonal é frequentemente uma malha triangular formada por triângulos tendo cantos comuns dados em espaço tridimensional.

[00110] Na prática, os mapas de profundidade são, consequentemente, com frequência convertidos em uma malha tridimensional. A malha tridimensional é um formato gráfico bem conhecido, que é suportado pela maioria dos drivers de hardware (por exemplo, baseados em OpenGL).

[00111] Como um exemplo, uma câmera estéreo pode registrar uma imagem de uma cena de um dado ponto de vista. Para cada pixel, uma estimativa de disparidade pode ser feita para estimar a distância até o objeto representado pelo pixel. Isso pode ser executado para cada pixel, fornecendo, assim, uma posição tridimensional de x, y, z para cada pixel. Essas posições podem, então, ser usadas como vértices de uma malha triangular com dois triângulos sendo formados para cada grupo de 2 x 2 pixels. Como isso pode resultar em um grande número de triângulos, o processo pode incluir combinar alguns triângulos iniciais em triângulos maiores (ou, em alguns cenários, mais genericamente em polígonos maiores). Isso reduzirá o número de triângulos, mas também diminuirá a resolução espacial da malha. Consequentemente, a granularidade da malha é tipicamente dependente das variações de profundidade e polígonos maiores aparecem predominantemente em áreas mais planas.

[00112] Cada vértice é adicionalmente associado a um valor de intensidade de luz do mapa de textura. O mapa de textura fornece essencialmente a intensidade de luz/cor na cena para o objeto na posição de pixel para o vértice. Tipicamente, uma imagem de intensidade de luz/mapa de textura

é fornecido juntamente com a malha, sendo que cada vértice contém dados que representam a posição x, y, z do vértice e dados u,v que identificam uma posição no mapa de textura, isto é, a mesma aponta para a intensidade de luz na posição x, y, z conforme capturada no mapa de textura.

[00113] Nessas representações, a malha poligonal é usada para fornecer informações da geometria tridimensional dos objetos, enquanto a textura é geralmente fornecida como uma estrutura de dados separada. Especificamente, a textura é frequentemente fornecida como um mapa bidimensional separado que, pelo algoritmo de processamento, pode ser sobreposto à geometria tridimensional.

[00114] O uso de malhas triangulares é particularmente adequado para processamento e manipulação por algoritmos de computação gráfica, e muitas soluções eficientes de software e hardware foram desenvolvidas e estão disponíveis no mercado. Em muitos dos sistemas de processamento, uma eficiência computacional substancial é obtida pelo algoritmo que processa os vértices individuais como um conjunto em uma pluralidade de polígonos, em vez de processar cada um separadamente. Por exemplo, para uma malha triangular típica, cada vértice é frequentemente comum a vários triângulos (em geral de 3 a 8). O processamento de um único vértice pode, consequentemente, ser aplicável a um número relativamente alto de triângulos, reduzindo assim substancialmente o número de pontos em uma imagem ou outro objeto que está sendo processado.

[00115] No aparelho gerador de imagens 101 da Figura 3, o gerador de malha 307 gera malhas com base nos mapas de profundidade de bloco recebidos e, consequentemente, executa uma transformação de uma representação baseada em

imagem para uma representação baseada em geometria (a malha pode ser considerada uma estrutura tridimensional geométrica).

[00116] No entanto, ao invés de usar, diretamente, os mapas de profundidade de bloco recebidos, o aparelho gerador de imagens 101 está disposto de modo a processá-los e modificar alguns dos valores de profundidade nos mapas de profundidade.

[00117] Especificamente, o aparelho gerador de imagens 101 compreende um primeiro processador 311 que é acoplado ao receptor 301 e que está disposto de modo a determinar regiões de borda vizinhas em ao menos um primeiro bloco e num segundo bloco em resposta ao padrão de mosaico, sendo que o primeiro bloco e segundo bloco são blocos vizinhos. O primeiro processador 311 pode, com base no padrão de mosaico, identificar uma região de borda ou margem em um bloco e uma região de borda ou margem correspondente em outro bloco, com as duas regiões de borda sendo vizinhas e, especificamente, onde as duas regiões de borda ou se tocam/conectam ou se sobrepõem. Especificamente, para um pixel de borda de um primeiro bloco, ao menos um pixel adjacente/vizinho estará em outro bloco. O pixel de borda e o pixel adjacente/vizinho pertencem, consequentemente, a regiões de borda diferentes de dois blocos/mapas de profundidade diferentes. Como o padrão de mosaico é conhecido pelo aparelho gerador de imagens 101 (por exemplo, por ser predeterminado ou por ser comunicado a partir do servidor remoto 103), o primeiro processador 311 é capaz de determinar quais blocos são conectados/vizinhos, e também quais pixels em outro bloco/mapa de profundidade são adjacentes e vizinhos a um dado pixel de borda/valor de profundidade em um bloco/mapa de profundidade atual.

[00118] Em muitas modalidades, o primeiro processador 311 pode estar disposto de modo a determinar a primeira região de borda do primeiro bloco como todos os pixels de borda/valores de profundidade que estão na borda do primeiro mapa de profundidade/bloco, e para os quais o pixel adjacente/mais próximo (em ao menos uma geração) está no segundo mapa de profundidade/bloco. Em algumas modalidades, a segunda região de borda (isto é a região de borda no segundo bloco) pode ser determinada como o conjunto de pixels adjacentes no segundo bloco.

[00119] O primeiro processador 311 é acoplado a um segundo processador 309, o qual é acoplado ao receptor 301 e recebe deste o mapa de profundidade. Ele é adicionalmente acoplado ao primeiro processador 311 e recebe deste as informações da primeira e da segunda regiões de borda.

[00120] O segundo processador 309 está disposto de modo a processar ao menos alguns dos pixels da primeira região de borda, isto é, está disposto de modo a modificar os valores de profundidade da região de borda identificada no primeiro bloco. A modificação dos valores de profundidade é tal que eles dependem dos valores de profundidade na segunda região de borda, isto é, os valores de profundidade modificados na região de borda do primeiro bloco dependem dos valores de profundidade na região de borda vizinha do segundo bloco. Assim, ao menos um primeiro valor de profundidade da primeira região de borda do primeiro bloco é modificado com base em ao menos um segundo valor de profundidade na segunda região de borda do segundo bloco (onde a primeira região de borda e a segunda região de borda são regiões de borda vizinhas). O primeiro valor de profundidade modificado pode ser determinado como uma função

do segundo valor de profundidade, isto é, ele é uma função de ao menos um valor de profundidade em um outro bloco da representação em mosaico. Deve-se considerar que, para alguns valores do primeiro e do segundo valores de profundidade, o valor de profundidade modificado pode ser igual ao valor de profundidade original.

[00121] Em muitas modalidades, o processo é repetido para ao menos todos os pixels de borda de ao menos um bloco. Adicionalmente, em muitas modalidades o valor de profundidade modificado pode tipicamente depender de apenas um valor de profundidade em um bloco vizinho (e possivelmente do valor de profundidade original, isto é, do valor de entrada para o pixel sendo modificado). Em particular, em muitas modalidades o valor de profundidade para um dado pixel de borda pode ser determinado como uma função apenas do valor de profundidade mais próximo em um outro bloco e, possivelmente, o próprio valor de profundidade original.

[00122] Assim, considerando que é de fato possível, em muitas modalidades, considerar uma pluralidade de valores de profundidade em um bloco vizinho (adjacente) (por exemplo, calculando-se a média de vários valores de profundidade próximos no bloco vizinho), em muitas modalidades o segundo processador 309 considerará apenas um valor de profundidade em cada bloco. Adicionalmente, para a maioria dos valores de profundidade sendo modificados, apenas um bloco vizinho é considerado e, portanto, apenas um valor de profundidade de um outro bloco é considerado. Entretanto, para alguns valores de profundidade, como por exemplo um valor de profundidade de um canto para uma representação em mosaico de cubo, o segundo processador 309 pode considerar um valor de

profundidade em cada um dos blocos vizinhos. Ele pode, por exemplo, minimizar a soma das distâncias quadráticas entre o valor de profundidade modificado e os correspondentes valores de profundidade de canto nos dois blocos. Como outro exemplo, ele pode definir todos os três valores de profundidade para o mesmo valor, como por exemplo para o valor de profundidade médio.

[00123] O valor de profundidade que é considerado em um outro bloco, quando apenas um valor de profundidade naquele bloco é considerado, pode especificamente ser o valor de profundidade mais próximo. Assim, é selecionado o valor de profundidade que estiver mais próximo do valor de profundidade sendo modificado. O valor de profundidade mais próximo pode ser o valor de profundidade para o qual a métrica de distâncias de uma cena, entre as posições nas janelas de visão para os valores de profundidade, seja a menor.

[00124] O segundo processador 309 modifica assim um ou mais dentre os mapas de profundidade recebidos na representação de imagem tridimensional em mosaico de entrada, gerando assim uma representação de imagem tridimensional em mosaico modificada. O conjunto de mapas de profundidade resultante é, então, fornecido ao gerador de malha 307, o qual procede à geração da malha com base nesses mapas de profundidade modificados.

[00125] A abordagem pode fornecer um desempenho substancialmente aprimorado em muitas modalidades, e pode possibilitar consistência aprimorada entre diferentes blocos em muitos cenários. Em particular, a abordagem pode proporcionar uma geração substancialmente aprimorada da malha em muitas modalidades onde, em particular, uma malha mais

consistente pode ser gerada com um risco substancialmente reduzido de, por exemplo, presença de orifícios na malha de profundidade gerada.

[00126] A seguir, vários aspectos da abordagem descrita serão descritos com referência ao exemplo específico de uma representação em mosaico de cubo, mas deve-se considerar que este é meramente um exemplo.

[00127] As realizações do inventor podem ser ilustradas ao se considerar a representação de cubo em mosaico da Figura 2. Tipicamente, cada lado de uma representação de cubo corresponde a uma projeção de perspectiva diferente, e é frequentemente gerada separadamente dos outros lados do cubo. Por exemplo, uma representação de cubo em mosaico pode ser gerada com o uso de seis câmeras detectoras de profundidade com deslocamentos de 90° de rotação. Para se fazer a representação de profundidade corresponder aos lados do cubo, é melhor que seja primeiro convertida em uma representação de distância/faixa, de modo que o mapa de profundidade codifique por bloco a distância entre o centro do cubo e o ponto do objeto. Para limitar a profundidade de bits, essa distância/faixa em metros pode ser codificada como uma constante dividida pela distância, já que uma relação inversa resulta em uma codificação mais compacta. Após essa etapa de transformação e quantificação, o mapa de profundidade pode ser adicionalmente comprimido com perda espacial por bloco. Isso resultará, geralmente, em uma variação de profundidade entre diferentes lados do cubo.

[00128] Como uma alternativa, os diferentes blocos podem ser combinados em uma única estrutura plana, como aquela da Figura 2, para codificação. Os algoritmos de compressão de

imagens e vídeo introduzem erros de quantificação e de codificação, e percebeu-se que isso pode introduzir inconsistências específicas de borda nas representações de bloco onde a codificação não sobrepõe os diferentes blocos.

[00129] Estes efeitos são, em particular, prevalentes onde um único padrão plano de blocos é codificado. De fato, nesses casos, a conversão de três dimensões para duas dimensões resulta em lacunas e rupturas na estrutura. Por exemplo, para uma representação de cubo, a topologia tridimensional é rompida nos lados do cubo, conforme ilustrado na Figura 2. Como resultado, os artefatos de codificação terão um efeito significativo e ficarão frequentemente visíveis. Especificamente, já que os pontos tridimensionais sobre uma borda de um lado de um cubo não serão consistentes (em termos de posição tridimensional) com os pontos tridimensionais sobre uma borda de um outro lado do cubo, a malha resultante também pode acabar sendo inconsistente e potencialmente pode ter lacunas.

[00130] A Figura 4 ilustra onde os potenciais erros de malha têm probabilidade de ocorrer, quando é usado um formato em mosaico de cubo. O exemplo mostra uma representação de profundidade em mosaico de cubo, codificada sob a forma de seis blocos separados. As setas indicam quais bordas de blocos terão a possibilidade de causar erros de malha devido à codificação com perdas de mapas de profundidade.

[00131] Na abordagem do aparelho gerador de imagens 101 da Figura 3, esses efeitos podem ser mitigados ou contornados por valores de profundidade em uma dada borda de bloco sendo restritos de modo que sejam muito similares ou, tipicamente, iguais aos valores de profundidade na borda de um

ou mais outros blocos. Isso pode assegurar que a malha subsequentemente construída seja consistente e não sofra erros de compressão de profundidade.

[00132] A Figura 5 ilustra um exemplo de malha gerada mediante a formação de triângulos entre todos os pixels/valores de profundidade. No exemplo, as localizações de pixel do mapa de profundidade que se tornarão vértices da malha são indicadas por círculos. Tipicamente, é usada uma grade regular de localizações de pixel (por exemplo cada segundo ou quarto de pixel). Um subconjunto de todas essas localizações de pixel é mostrado na Figura 5. No exemplo, o bloco $D_{\text{para cima}}$ tem a borda de bloco A, e o bloco $D_{\text{para a frente}}$ tem a borda de bloco B, e pixels em cada borda são indexados como $D_{A,i}$ e $D_{B,j}$, respectivamente. No exemplo, as bordas A e B são bordas provenientes de dois diferentes blocos, e a compressão de imagens com perdas dos mapas de profundidade pode tornar inconsistentes essas amostras, isto é, os correspondentes valores de profundidade nos diferentes blocos podem acabar tendo valores diferentes, os quais resultariam em erros na malha, já que os triângulos gerados para o bloco $D_{\text{para cima}}$ não se alinharão àqueles gerados para $D_{\text{para a frente}}$. Entretanto, mediante a definição dos valores de profundidade dos locais de borda correspondentes para o mesmo valor de profundidade, antes de gerar a malha, uma malha consistente é produzida.

[00133] Deve-se considerar que, em modalidades diferentes, funções diferentes podem ser usadas para determinar os valores de profundidade. Entretanto, em muitas modalidades os valores de profundidade podem ser determinados de modo que os valores de profundidade de borda vizinha em dois blocos vizinhos sejam iguais, ou ao menos de modo que a

diferença entre eles seja reduzida. Especificamente, se o primeiro valor de profundidade e o segundo valor de profundidade forem valores de borda, os dois serão definidos para o mesmo valor. Em algumas modalidades, o primeiro valor de profundidade pode ser definido para o valor do segundo valor mas, em muitas modalidades, um valor de profundidade é determinado como uma função tanto do primeiro como do segundo valores de profundidade (e possivelmente de outros parâmetros ou valores de profundidade) e, subsequentemente, tanto o primeiro como o segundo valores de profundidade são definidos para esse valor, isto é, ambos os valores são modificados.

[00134] Em algumas modalidades, os dois valores de profundidade podem não ser definidos para o mesmo valor, mas podem por exemplo ser definidos para valores que estão mais próximos uns dos outros, isto é, de modo que a diferença entre eles seja reduzida. Como um exemplo, o primeiro valor de profundidade pode ser definido como igual ao valor de profundidade médio para o primeiro e o segundo valores de profundidade, desde que a alteração no primeiro valor de profundidade não o altere em mais do que uma dada quantidade. Dessa forma, a alteração no valor de profundidade é restrita a um nível máximo. Em alguns cenários isso pode, por exemplo, resultar em valores de profundidade modificados que ainda resultam em uma lacuna na malha gerada, mas com essa lacuna sendo potencialmente reduzida de maneira substancial. Entretanto, ao mesmo tempo, é assegurado que os valores de profundidade nas regiões de borda não sejam demasiadamente modificados em relação a valores de profundidade no primeiro bloco fora da região de borda. Isso pode fornecer um mapa de profundidade mais consistente dentro do bloco. Deve-se

considerar que a exata função usada pode depender das preferências e dos requisitos da modalidade individual e, especificamente, pode depender do equilíbrio desejado entre consistência de profundidade dentro de um bloco e consistência de profundidade entre blocos.

[00135] Ao se definir os valores de profundidade dos pixels de borda para que sejam iguais pode-se fornecer um alto nível de consistência entre triângulos da malha gerados a partir dos dois mapas de profundidade diferentes e, consequentemente, fornecerá uma malha consistente.

[00136] Em algumas modalidades, o valor aplicado tanto ao primeiro como ao segundo valor de profundidade pode ser determinado como um valor médio dos dois valores de profundidade. Por exemplo, para o exemplo da Figura 5, os valores de profundidade na borda dos blocos pode ser determinado como:

$$D'_{A,i} \leftarrow \frac{D_{A,i} + D_{B,j}}{2} \; e \; D'_{B,i} \leftarrow \frac{D_{A,i} + D_{B,j}}{2}$$

[00137] O novo valor de profundidade é, assim, definido no valor de profundidade médio, o qual é o mesmo para as localizações de pixel de borda correspondentes. Quando uma malha é agora construída com o uso das coordenadas de pixel e o valor de profundidade, os pixels de borda provenientes das bordas A e B vão acabar na mesmo localização de 3D. No exemplo, a operação de restrição de profundidade calcula a média de duas amostras, mas deve-se considerar que podem ser usadas mais amostras. Nos cantos, o cálculo da média pode tipicamente ser acima de três valores de profundidade (no caso de um mapa em cubo) ou quatro valores de profundidade (para outros formatos em mosaico onde quatro blocos se encontram). Em outras palavras, o valor médio pode ser determinado mediante o cálculo

da média dos valores de profundidade dos pixels mais próximos em cada bloco vizinho ao pixel atual.

[00138] Como uma alternativa, $D'_{A,i}$ pode ser definido para a média de um número de localizações de amostra em torno da posição j. Além disso, as localizações de pixel um pouco mais distantes da borda poderiam ser modificadas.

[00139] No exemplo, o segundo processador 309 está consequentemente disposto de modo a determinar o primeiro valor de profundidade em resposta a apenas o valor de profundidade de um pixel mais próximo de cada bloco vizinho ao primeiro bloco, em uma posição do primeiro valor de profundidade.

[00140] Em muitas modalidades, o segundo processador 309 pode adicionalmente estar disposto de modo a aplicar um filtro espacial ao mapa de profundidade do primeiro bloco, em seguida à modificação do primeiro valor de profundidade.

[00141] Assim, após o segundo processador 309 ter modificado valores de profundidade nas regiões de borda de um dado bloco, ele pode prosseguir para aplicar um filtro espacial ao mapa de profundidade modificado. O filtro espacial pode especificamente ser aplicado à totalidade do mapa de profundidade, incluindo tanto as regiões de borda como as regiões internas. Em outras modalidades, porém, o filtro espacial pode apenas ser aplicado às regiões de borda e às partes das regiões internas que estão adjacentes às regiões de borda.

[00142] A modificação dos valores de profundidade nas uma ou mais regiões de borda pode possibilitar consistência aprimorada entre blocos. Entretanto, há um risco de que isso possa reduzir a consistência interna dentro do mapa de

profundidade e a filtragem espacial passa-baixa pode ser usada para assegurar ou aumentar essa consistência.

[00143] Por exemplo, para reduzir o risco de que a modificação introduza uma nova inconsistência entre os pixels de borda e os pixels no interior do bloco mas não na borda, pode ser desejável filtrar espacialmente os valores de profundidade das fileiras e colunas que ficam (imediatamente) dentro da borda, de modo que se tornem espacialmente consistentes com os novos valores de profundidade na borda.

[00144] Em muitas modalidades, essa filtragem espacial pode não ser aplicada aos valores de profundidade modificados. Por exemplo, os pixels de borda podem ser modificados com base nos valores de profundidade em outros blocos e, subsequentemente, um filtro espacial pode ser aplicado aos pixels que estiverem próximos aos pixels de borda, mas não incluindo estes. Assim, os pixels de borda manterão seus valores mas, devido à filtragem dos valores de profundidade internos que estão próximos aos pixels de borda, estes serão modificados para serem mais consistentes com os novos valores para os pixels de borda (o kernel de filtragem para esses pixels tipicamente incluirá os pixels de borda).

[00145] Em muitas modalidades, o gerador de malha 307 pode estar disposto de modo a usar apenas um subconjunto dos pixels para a geração da malha, isto é, apenas um subconjunto dos pixels pode ser selecionado para corresponder a vértices dos polígonos (tipicamente triângulos) da malha. Por exemplo, um vértice pode ser gerado para apenas cada segundo ou cada quarto pixel do mapa de profundidade. Esse tipo de abordagem pode frequentemente ser aplicada para gerar

uma malha de menor complexidade, a qual fornece uma taxa de dados e uma exigência de recursos computacionais mais baixas.

[00146] Em muitas dessas modalidades, o gerador de malha 307 pode estar disposto de modo a indicar ao segundo processador 309 que apenas um subconjunto de valores de profundidade é usado para a geração da malha. Em resposta, o segundo processador 309 pode prosseguir para modificar apenas os valores de profundidade do mapa de profundidade/região de borda que são usados para a subsequente geração da malha. Em muitas modalidades, isso pode reduzir de modo substancial a complexidade e o uso de recursos computacionais do segundo processador 309.

[00147] Essa abordagem pode ser particularmente adequada para situações nas quais são usadas malhas irregulares. Especificamente, em modalidades nas quais são usadas malhas regulares, a topologia é fixa e independente das malhas dinâmicas que podem ser geradas e, especificamente, é independente de malhas geradas para outros blocos. Portanto, o segundo processador 309 pode ter informações predeterminadas sobre quais valores de profundidade modificar.

[00148] Entretanto, caso sejam usadas malhas irregulares, o subconjunto de valores de profundidade que são de fato convertidos em vértices da malha irá variar, e nesse tipo de situação, a seleção do subconjunto de valores de profundidade a modificar pode ser baseada nos vértices da malha irregular.

[00149] Por exemplo, o gerador de malha 307 pode, antes de qualquer modificação dos valores de profundidade, gerar uma primeira malha irregular (interna) com base apenas nos valores de profundidade do primeiro bloco, isto é, pode

gerar uma primeira malha inicial para o primeiro bloco. De modo similar, o gerador de malha 307 pode, antes de qualquer modificação, gerar uma segunda malha irregular (interna) com base apenas nos valores de profundidade do segundo bloco, isto é, pode gerar uma segunda malha inicial para o segundo bloco.

[00150] A fim de combinar essas duas malhas irregulares, o gerador de malha 307 pode identificar todos os vértices que ficam ao longo da borda entre o primeiro e o segundo bloco. Cada valor de profundidade que corresponda a um vértice da primeira ou da segunda malha irregular pode, então, ser incluído no subconjunto que é indicado ao segundo processador 309. O segundo processador 309 pode então executar a modificação/restrição de profundidade, e pode especificamente definir os valores de profundidade para que sejam iguais nos dois blocos. Os valores de profundidade resultantes podem, então, ser realimentados ao gerador de malha 307, o qual pode prosseguir para modificar os valores de profundidade dos vértices existentes nas malhas, bem como adicionar à outra malha qualquer vértices (e correspondentes bordas) que estejam presentes apenas em uma malha. Como resultado, as duas malhas irregulares resultantes corresponderão exatamente ao longo das linhas da borda entre os dois blocos e, assim, podem ser facilmente integradas em uma única malha.

[00151] No exemplo anterior, o servidor remoto 103 gerou e transmitiu a representação de imagem tridimensional em mosaico como uma única estrutura combinada. Entretanto, em muitas modalidades, o servidor remoto 103 pode estar disposto de modo a transmitir apenas um subconjunto de blocos e, especificamente, apenas um bloco de cada vez. Nessas

modalidades, o aparelho gerador de imagens 101 pode estar disposto de modo a receber um bloco ou um subconjunto de blocos, e combinar estes com blocos provenientes de outras fontes, como especificamente com outros blocos que estejam localmente armazenados. Esses blocos localmente armazenados podem ser blocos que foram anteriormente recebidos provenientes do servidor remoto 103, ou que poderiam potencialmente, por exemplo, ser blocos nominais predeterminados representando uma cena nominal predeterminada que é, então dinamicamente atualizada e adaptada mediante o recebimento de blocos de substituição provenientes do servidor remoto 103.

[00152] Nesse tipo de modalidade, o aparelho gerador de imagens 101 pode estar disposto de modo a usar uma abordagem correspondente para restringir os valores de profundidade do bloco recém-recebido.

[00153] Como um exemplo específico, o aparelho gerador de imagens 101 pode, conforme ilustrado na Figura 6, compreender um comunicador de fonte remota 601 e um armazenamento local 603, os quais são acoplados ao receptor 301. Nesse exemplo, o receptor 301 pode, portanto, recuperar blocos da representação de imagem tridimensional em mosaico, seja do armazenamento local 603, seja do comunicador de fonte remota 601.

[00154] No exemplo, o armazenamento local 603 compreende um conjunto completo de blocos correspondentes a uma representação de imagem tridimensional em mosaico, e o receptor 301 pode estar disposto de modo a recuperar os blocos adequados quando for necessário. O conjunto armazenado de

blocos pode ser formado especificamente por blocos que foram anteriormente recebidos do servidor remoto 103.

[00155] O comunicador de fonte remota 601 pode estar disposto de modo a receber um único bloco (ou potencialmente um subconjunto de blocos) do servidor remoto 103. O novo bloco pode ser uma substituição para um bloco existente atualmente armazenado no armazenamento local 603.

[00156] O novo bloco pode, por exemplo, ser recebido em resposta a uma mensagem transmitida ao servidor remoto 103, proveniente do aparelho gerador de imagens 101. De fato, no exemplo específico, a interface de usuário 305 é acoplada ao comunicador de fonte remota 601 e fornece o ponto de vista (direção e/ou posição de visão) atualmente determinado ao comunicador de fonte remota 601. Em resposta, o comunicador de fonte remota 601 transmite uma indicação do ponto de vista para o servidor remoto 103, o qual determina se é necessário uma atualização adequada de qualquer dos blocos armazenados. Por exemplo, se o ponto de vista atual corresponde a uma direção que não foi atualizada durante um intervalo de tempo significativo, o servidor remoto 103 pode decidir transmitir um novo bloco de substituição. Em outras modalidades, o próprio servidor remoto 103 pode determinar que um bloco de substituição deveria ser transmitido, por exemplo em resposta à detecção de que uma alteração dinâmica substancial ocorreu na cena.

[00157] No exemplo, quando o comunicador de fonte remota 601 recebe um novo bloco (frequentemente um bloco de substituição), ele pode prosseguir para extrair os blocos vizinhos armazenados e, então, prosseguir para executar a operação anteriormente descrita para o bloco recebido, por

exemplo, para um ou mais pixels de borda/valores de profundidade do bloco recebido, ele pode modificar os valores de profundidade de modo que correspondam entre os novos blocos e os blocos vizinhos armazenados.

[00158] Entretanto, em muitas modalidades, os valores de profundidade são modificados apenas para o novo bloco. Assim, ao invés de, por exemplo, determinar o valor de profundidade médio e atribuí-lo aos pixels tanto do novo mapa recebido como do mapa de profundidade vizinho armazenado, o segundo processador 309 pode definir o valor de profundidade do novo mapa de profundidade para que seja igual ao valor de profundidade vizinho do mapa de profundidade armazenado. Assim, o mapa de profundidade do novo bloco é modificado para ser consistente com os mapas de profundidade localmente armazenados, mas sem modificar estes últimos. Consequentemente, qualquer processamento já executado nos mapas de profundidade armazenados não precisa ser refeito, por exemplo, não há necessidade de filtrar novamente esses mapas de profundidade. Ao invés disso, o mapa de profundidade do novo bloco é modificado e pode, depois disso, ser usado com os mapas de profundidade existentes.

[00159] Em seguida a essa modificação, o novo mapa de profundidade recebido pode ser armazenado no armazenamento local, atualizando assim a representação de imagem tridimensional localmente armazenada.

[00160] A abordagem pode proporcionar um sistema muito eficiente e pode, em particular, reduzir a largura de banda de comunicação, já que apenas blocos individuais podem ser transmitidos como e quando for apropriado. Adicionalmente, pode ser reduzido o risco de que esse novo bloco resulte em

qualidade de imagem degradada e erros ou artefatos. Além disso, o recurso computacional pode ser mantido baixo, já que não há necessidade de processamento ou modificação dos mapas de profundidade vizinhos armazenados.

[00161] De fato, em algumas modalidades podem ter sido geradas malhas para a cena, e os valores de profundidade dos outros blocos podem não estar armazenadas sob a forma de mapas de profundidade explícitos, mas propriamente sob a forma de malhas. Nesse tipo de situação, uma malha para um dado bloco é tipicamente gerada de modo a ter um número de vértices ao longo da borda ou margem do bloco. A profundidade desses vértices pode agora ser usada como os valores de profundidade da região de borda do bloco vizinho e, portanto, os valores de profundidade do bloco recém-recebido podem ser modificados para que sejam consistentes com os valores de profundidade de borda da malha existente para o correspondente bloco vizinho.

[00162] O conjunto de valores de profundidade que são modificados no novo mapa de profundidade podem ser adicionalmente selecionados em resposta à malha e, especificamente para cada valor de profundidade/vértice da malha adjacente ao novo bloco, o valor de profundidade correspondente (mais próximo) no novo mapa de profundidade pode ser selecionado e definido para o mesmo valor do valor de profundidade do vértice. Esses pixels/valores de profundidade podem então ser usados como vértices para polígonos (tipicamente triângulos) da malha, com outros vértices sendo selecionados dentro do novo bloco, de acordo com um critério de seleção adequado.

[00163] Dessa maneira, é gerada uma malha para o novo bloco, a qual é consistente com as malhas já geradas dos

blocos armazenados. Consequentemente, essas malhas vizinhas não precisam ser modificadas, e podem ser usadas diretamente, reduzindo assim substancialmente a carga de processamento.

[00164] Entretanto, embora esse tipo de abordagem possa proporcionar um desempenho bem alto em muitas modalidades, em algumas situações ela pode resultar em desvios de profundidade. Em particular, em muitas modalidades os vértices de borda do bloco vizinho podem estar a alguma distância uns dos outros, por exemplo devido a vários triângulos menores da malha terem sido combinados em uma malha maior, em consequência da variação de profundidade ser pequena nessa área. Entretanto, se um novo bloco que tem uma variação de profundidade relativamente grande ao longo desta linha for agora recebido, uma linha reta de um único triângulo da malha não é capaz de refletir essa variação de profundidade. Em algumas modalidades, o segundo processador 309 pode estar disposto de modo a detectar que uma diferença de profundidade entre um valor de profundidade da borda do novo bloco e um valor de profundidade correspondente de um polígono de malha existente de um bloco vizinho armazenado excede um limiar e, em resposta a essa detecção, pode adicionar um novo vértice de borda. Esse novo vértice de borda formará então a base dos polígonos na malha que é gerada para o novo bloco, mas também será um vértice para o bloco vizinho. Consequentemente, um novo vértice é adicionado à malha anteriormente gerada, e o correspondente polígono da malha existente é dividido em dois polígonos. Assim, uma ou mais bordas são adicionadas à malha existente, entre o novo vértice e um ou mais vértices existentes no interior da malha existente para o bloco vizinho (o número de bordas necessário para dividir o polígono

anteriormente existente em polígonos menores com base no novo vértice dependerá da topologia específica).

[00165] Como um exemplo específico, quando é recebido um novo bloco de profundidade, em muitas modalidades, é desejável combiná-lo com a malha já existente. Essa situação é ilustrada na Figura 7. Nesse exemplo, os mapas de profundidade dos blocos vizinhos já foram convertidos em uma malha e agora deseja-se produzir uma malha para o bloco recém-chegado. Isso pode ser feito mediante a análise de valores de profundidade na borda dos blocos. Se o novo mapa de profundidade tiver um valor de profundidade de borda (por exemplo sendo determinado como um candidato a um vértice devido às variações de profundidade interna no novo bloco) com uma grande diferença para a borda de triângulo já existente naquela posição, então um novo vértice é criado e um ou mais triângulos são divididos no bloco vizinho, criando assim uma malha mais detalhada nessa área.

[00166] Entretanto, se não for esse o caso, então a topologia da malha dos blocos vizinhos pode ser usada diretamente, e uma malha é formada no interior do novo bloco, com o uso dos vértices de borda existentes da malha vizinha.

[00167] A abordagem mais complexa é particularmente adequada para exemplos que usam uma malha irregular. No exemplo da Figura 7, a conversão do mapa de profundidade em malha para o bloco armazenado foi baseada em uma análise da variação do mapa de profundidade e resultou em uma malha irregular. Ao se criar a malha para o novo bloco, alguns vértices não são modificados (círculos fechados). Entretanto, quando o novo mapa de profundidade indica que o modelo de malha precisa ser modificado, então um novo vértice

é posicionado. Essa decisão pode basear-se na comparação da posição de vértices candidatos gerados a partir do novo mapa de profundidade na borda com os blocos vizinhos com a malha existente. Se a diferença de profundidade ou de posição for muito grande (indicando que a diferença de profundidade entre o novo mapa de profundidade e a malha existente é muito grande), então um novo vértice é introduzido, e os triângulos de blocos vizinhos são atualizados.

[00168] Deve-se considerar que, embora a descrição acima tenha focalizado na geração de uma imagem com base na representação de imagem tridimensional em mosaico e, especificamente, mediante a geração de uma malha a partir dos mapas de profundidade, nenhuma destas é característica essencial. De fato, gerar uma representação de imagem tridimensional em mosaico aprimorada pode ser vantajoso em muitas aplicações diferentes, e não se limita ao uso subsequente específico dessa representação de imagem tridimensional em mosaico aprimorada, conforme aqui descrito.

[00169] Os exemplos acima se concentraram em modalidades nas quais um primeiro valor de profundidade de borda de um primeiro mapa de profundidade de um primeiro bloco é modificado com base no valor de profundidade de borda mais próximo em um segundo mapa de profundidade de um segundo bloco. Entretanto, deve-se considerar que, em muitas modalidades o primeiro valor de profundidade de borda pode, adicional ou alternativamente, ser modificado com base em um valor de profundidade do segundo mapa de profundidade, que não é o mais próximo do primeiro valor de profundidade de borda. Em algumas modalidades, o primeiro valor de profundidade de borda pode ser modificado com base em uma pluralidade de valores de

profundidade em uma região de borda do segundo mapa de profundidade, incluindo valores de profundidade que não são necessariamente valores de profundidade de borda, e potencialmente até excluindo o valor de profundidade da borda mais próxima. Por exemplo, o primeiro valor de profundidade de borda pode ser modificado com base na pluralidade de valores de profundidade no segundo mapa, que consiste em todos os valores de profundidade que tenham uma distância até o primeiro valor de profundidade da borda abaixo de um dado limite.

[00170] De modo similar, deve-se considerar que o primeiro valor de profundidade, isto é o valor de profundidade sendo modificado, não precisa ser um valor de profundidade de borda.

[00171] Deve-se considerar que, para maior 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 ou organização lógica ou física estrita.

[00172] A invenção pode ser implementada em qualquer forma adequada, incluindo hardware, software,

firmware ou qualquer combinação dos mesmos. A invenção pode, opcionalmente, ser 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 física, funcional e logicamente implementados 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.

[00173] Embora a presente invenção tenha sido descrita em conexão com algumas modalidades, não se pretende limitá-la à forma específica aqui apresentada. Ao invés 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 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.

[00174] 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 na limitação a tal categoria, mas, ao invés 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 exemplos esclarecedores e não devem ser interpretados como limitadores do escopo das reivindicações de forma alguma.

REIVINDICAÇÕES

1. APARELHO PARA GERAR UMA REPRESENTAÇÃO DE UMA CENA EM IMAGEM TRIDIMENSIONAL EM MOSAICO, caracterizado por compreender:

um receptor (301) para receber uma representação de uma cena em imagem tridimensional em mosaico de um primeiro ponto de vista, sendo que a representação de imagem tridimensional em mosaico compreende uma pluralidade de blocos interconectados, sendo que cada bloco compreende um mapa de profundidade e um mapa de textura representando uma janela de visão da cena a partir do primeiro ponto de vista, e sendo que os blocos formam um padrão em mosaico;

um primeiro processador (311) para determinar as regiões de borda vizinhas em ao menos um primeiro bloco e num segundo bloco em resposta ao padrão de mosaico, sendo que o primeiro bloco e o segundo bloco são blocos vizinhos;

um segundo processador (309) para modificar pelo menos um primeiro valor de profundidade de uma primeira região de borda do primeiro bloco, em resposta a pelo menos um segundo valor de profundidade em uma segunda região de borda do segundo bloco, de modo que uma diferença entre o primeiro valor de profundidade e o segundo valor de profundidade seja reduzida para ao menos alguns valores do primeiro valor de profundidade e do segundo valor de profundidade, sendo que a primeira região de borda e a segunda região de borda são regiões de borda vizinhas.

2. APARELHO, de acordo com a reivindicação 1, caracterizado pelo segundo processador (309) estar disposto de modo a definir o primeiro valor de profundidade e o segundo valor de profundidade para um mesmo valor.

3. APARELHO, de acordo com a reivindicação 2, caracterizado pelo mesmo valor ser um valor médio do primeiro valor de profundidade e do segundo valor de profundidade.

4. APARELHO, de acordo com qualquer uma das reivindicações 1 ou 2, caracterizado pelo segundo processador (309) estar disposto de modo a definir o primeiro valor de profundidade para o segundo valor de profundidade.

5. APARELHO, de acordo com qualquer uma das reivindicações 1 a 4, caracterizado por compreender adicionalmente um gerador de malha (307) para gerar uma malha tridimensional a partir de mapas de profundidade dos blocos.

6. APARELHO, de acordo com a reivindicação 5, caracterizado pelo gerador de malha (307) estar disposto de modo a selecionar um subconjunto de valores de profundidade da primeira região de borda para uso na geração da malha tridimensional; e por o segundo processador estar disposto de modo a modificar valores de profundidade da primeira região de borda somente se pertencerem ao subconjunto de valores de profundidade.

7. APARELHO, de acordo com qualquer uma das reivindicações 1 a 6, caracterizado pelo segundo processador (309) estar disposto de modo a aplicar um filtro espacial a um mapa de profundidade do primeiro bloco, em seguida à modificação do primeiro valor de profundidade.

8. APARELHO, de acordo com qualquer uma das reivindicações 1 a 7, caracterizado por compreender adicionalmente um comunicador de fonte remota (601) disposto de modo a receber blocos para representação de imagem tridimensional em mosaico provenientes de uma fonte remota (103), e um armazenamento local (603) para armazenar blocos

para a representação de imagem tridimensional em mosaico; e por o receptor (301) estar disposto de modo a recuperar ao menos o segundo valor de profundidade do armazenamento local, em resposta ao recebimento do primeiro bloco proveniente do receptor de fonte remota (103).

9. APARELHO, de acordo com a reivindicação 8, caracterizado pelo armazenamento local (603) estar disposto de modo a armazenar um mapa de profundidade ou uma malha para o primeiro bloco, em seguida à modificação do primeiro valor de profundidade.

10. APARELHO, de acordo com qualquer uma das reivindicações 8 ou 9, caracterizado por compreender adicionalmente uma entrada de dados pelo usuário (305) para determinar uma direção de visão atual; e por o comunicador de fonte remota (601) estar disposto de modo a transmitir à fonte remota uma indicação da direção de visão atual; e por o primeiro bloco ser recebido em resposta à indicação.

11. APARELHO, de acordo com qualquer uma das reivindicações 8 a 10, como dependentes da reivindicação 5, caracterizado pelo armazenamento local (603) estar disposto de modo a armazenar uma malha para os blocos armazenados, e o gerador de malha (307) estar disposto de modo a adicionar um vértice e ao menos uma borda a uma malha armazenada para o segundo bloco, em resposta a uma detecção de que uma medida de diferença excede um limite, sendo que a medida de diferença é indicativa de uma diferença de profundidade entre um valor de profundidade de borda do primeiro bloco e um correspondente valor de profundidade de borda da malha armazenada para o segundo bloco.

12. APARELHO, de acordo com qualquer uma das reivindicações 1 a 11, caracterizado pelo segundo processador (309) estar disposto de modo a determinar o primeiro valor de profundidade em resposta apenas ao valor de profundidade de um pixel mais próximo de cada bloco vizinho ao primeiro bloco, em uma posição do primeiro valor de profundidade.

13. APARELHO, de acordo com qualquer uma das reivindicações 1 a 12, caracterizado por cada bloco ser um polígono plano e por a pluralidade de blocos interconectados formar uma estrutura tridimensional.

14. MÉTODO PARA GERAR UMA REPRESENTAÇÃO DE UMA CENA EM IMAGEM TRIDIMENSIONAL caracterizado por compreender:

receber uma representação de uma cena em imagem tridimensional em mosaico de um primeiro ponto de vista, sendo que a representação de imagem tridimensional em mosaico compreende uma pluralidade de blocos interconectados, sendo que cada bloco compreende um mapa de profundidade e um mapa de textura representando uma janela de visão da cena a partir do primeiro ponto de vista, e sendo que os blocos formam um padrão em mosaico;

determinar as regiões de borda vizinhas em ao menos um primeiro bloco e num segundo bloco em resposta ao padrão de mosaico, sendo que o primeiro bloco e o segundo bloco são blocos vizinhos;

modificar pelo menos um primeiro valor de profundidade de uma primeira região de borda do primeiro bloco, em resposta a pelo menos um segundo valor de profundidade em uma segunda região de borda do segundo bloco, de modo que uma diferença entre o primeiro valor de profundidade e o segundo valor de profundidade seja reduzida para ao menos alguns

valores do primeiro valor de profundidade e do segundo valor de profundidade, sendo que a primeira região de borda e a segunda região de borda são regiões de borda vizinhas.

15. 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 14, quando o dito programa for executado em um computador.

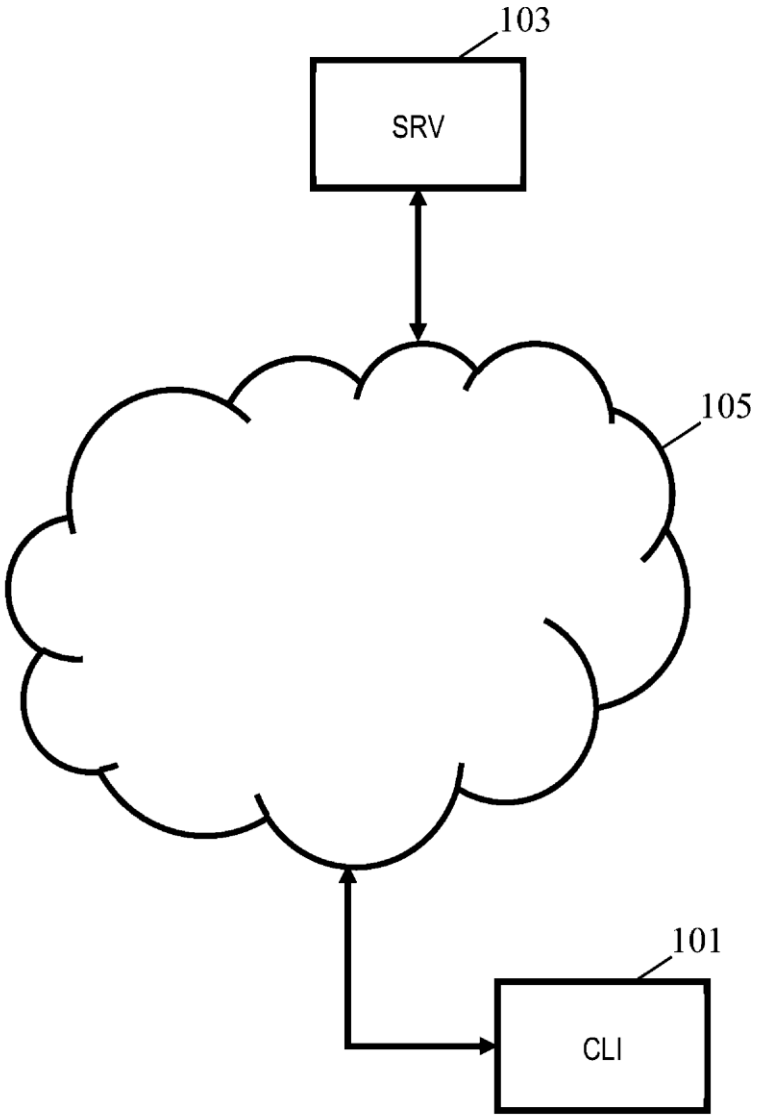


FIG. 1

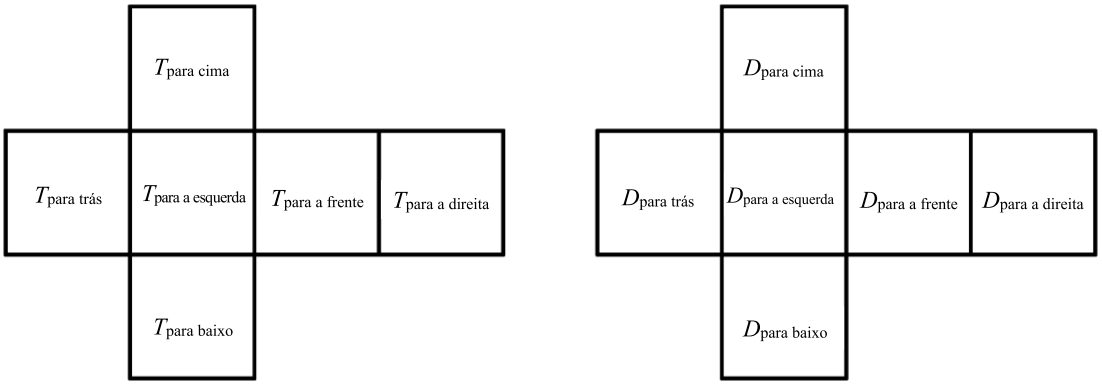

$T_{\text{para cima}}$
$T_{\text{para trás}}$
$T_{\text{para a esquerda}}$
$T_{\text{para a frente}}$
$T_{\text{para a direita}}$
$T_{\text{para baixo}}$
TexM
$D_{\text{para cima}}$
$D_{\text{para trás}}$
$D_{\text{para a esquerda}}$
$D_{\text{para a frente}}$
$D_{\text{para a direita}}$
$D_{\text{para baixo}}$
Prof

FIG. 2

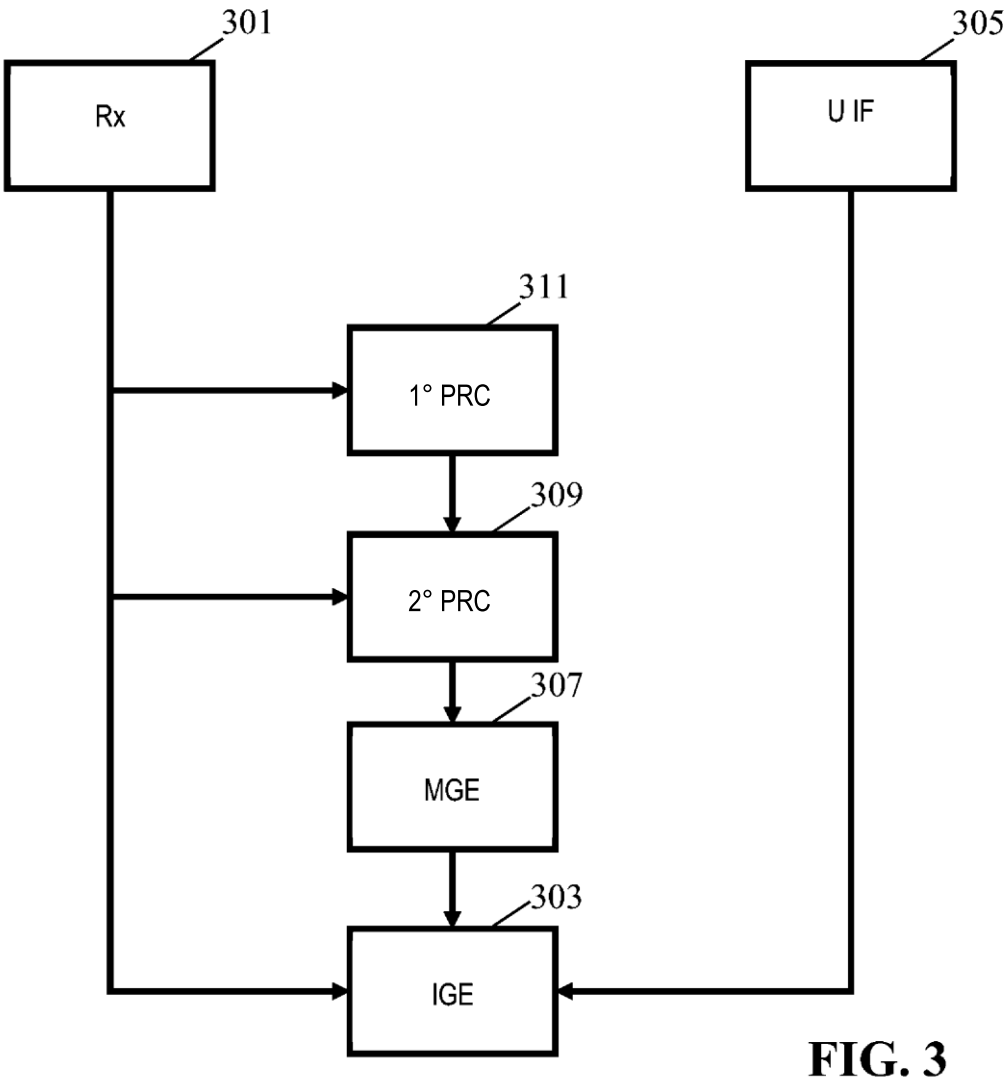


FIG. 3

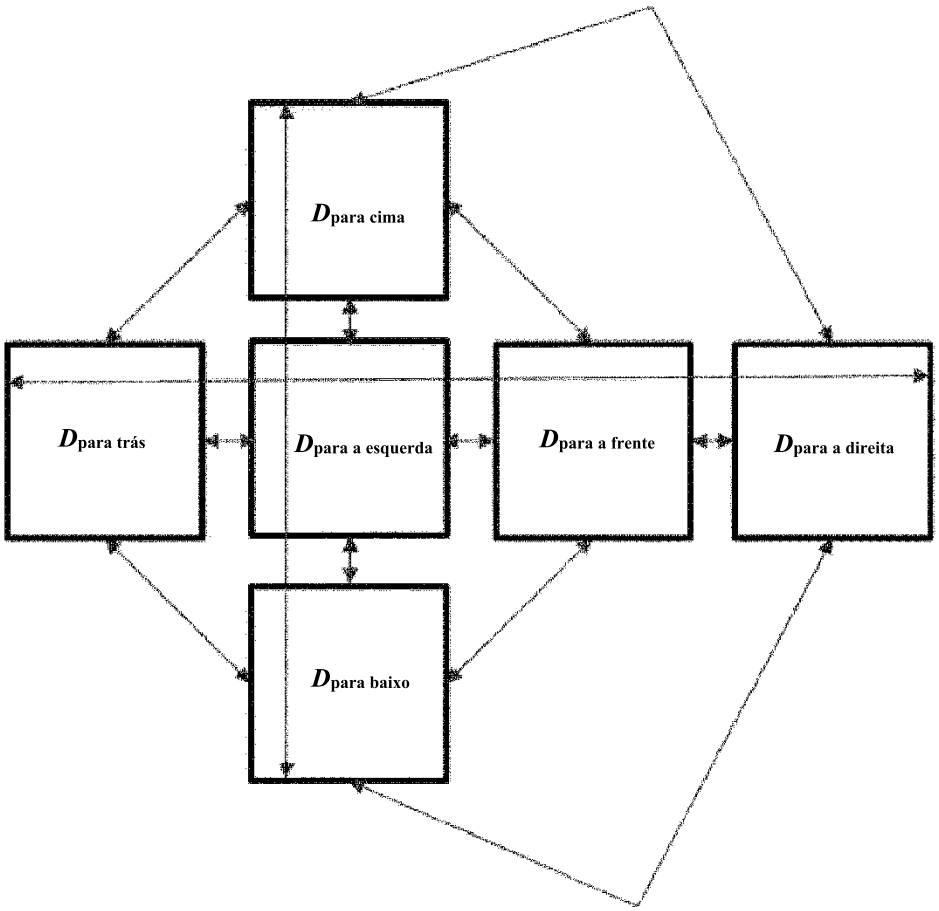

D para cima
D para trás
D para a esquerda
D para a frente
D para a direita
D para baixo

FIG. 4

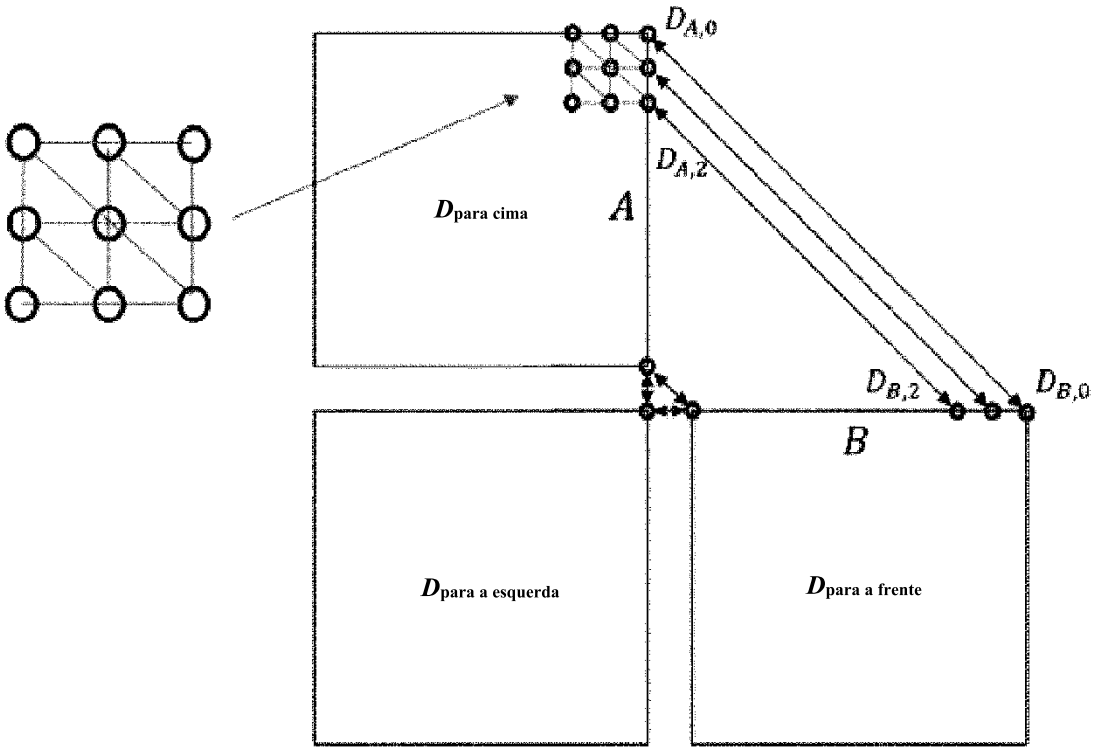

$D_{A,0}$
$D_{A,2}$
$D_{\text{para cima}}$
A
$D_{B,2}$
$D_{B,0}$
B
$D_{\text{para a esquerda}}$
$D_{\text{para a frente}}$

FIG. 5

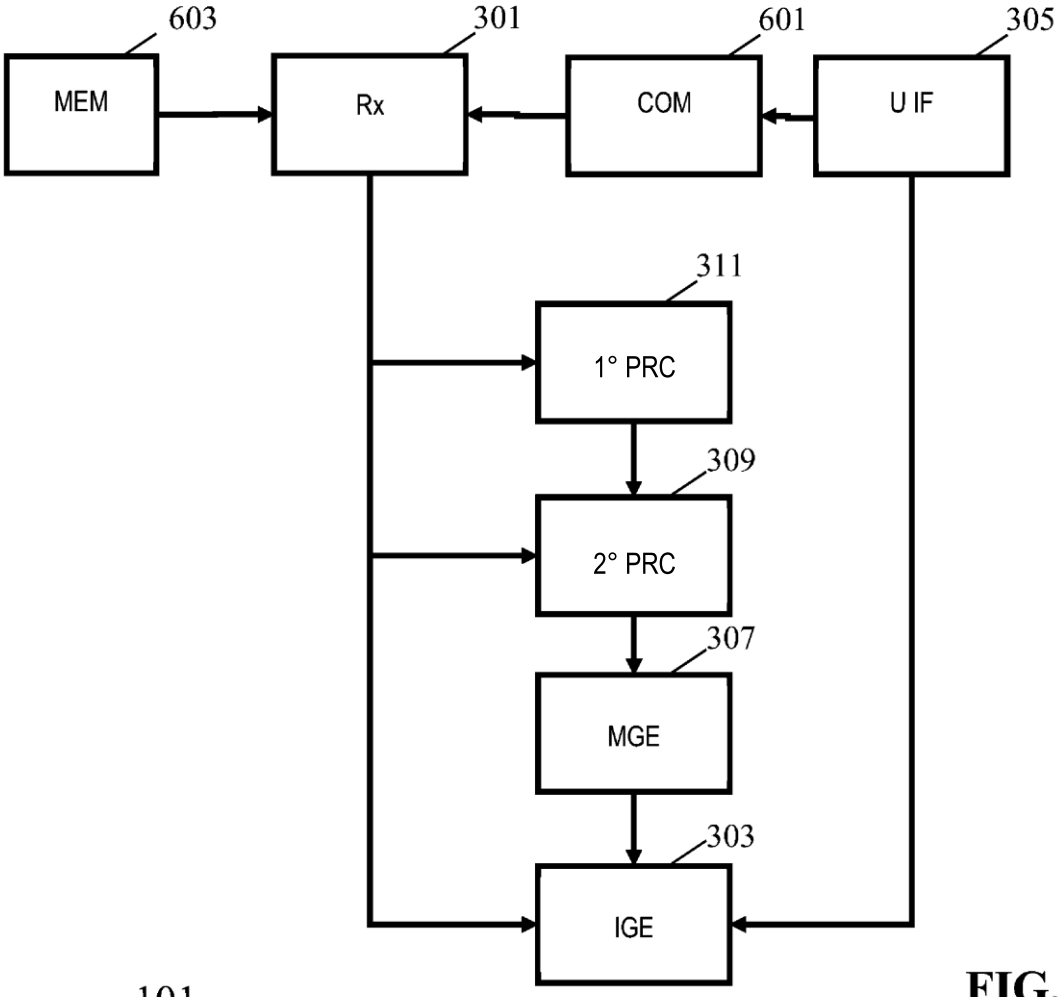

603
301
601
305
MEM
Rx
COM
U IF
311
1° PRC
309
2° PRC
307
MGE
303
IGE
101

FIG. 6

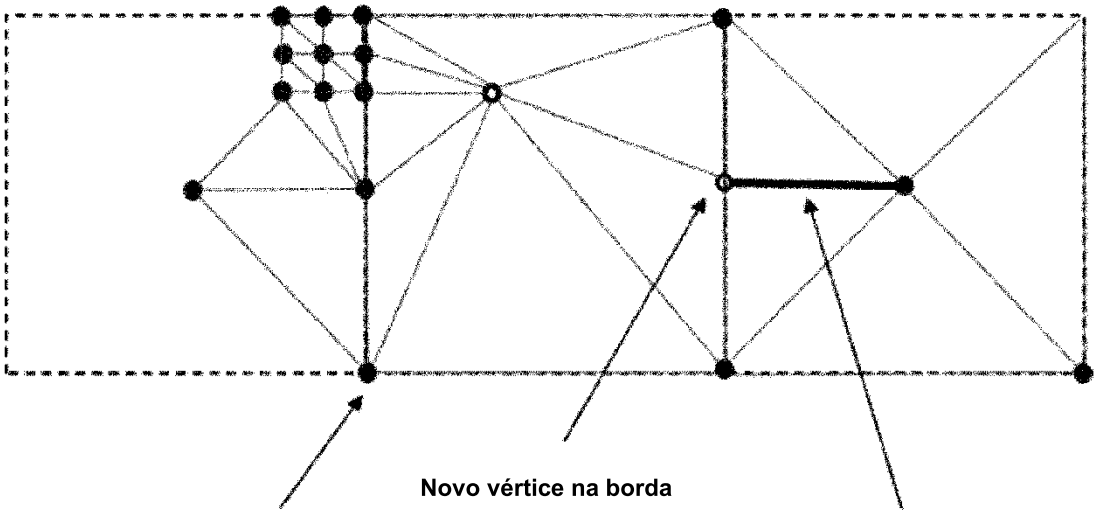

Novo vértice na borda
Vértice fixo da malha vizinha
Divisão do triângulo no bloco vizinho

FIG. 7

RESUMO

APARELHO E MÉTODO PARA GERAR UMA REPRESENTAÇÃO DE UMA CENA EM IMAGEM TRIDIMENSIONAL EM MOSAICO E PRODUTO DE PROGRAMA DE COMPUTADOR

A presente invenção refere-se a um aparelho para gerar uma representação de uma cena em imagem tridimensional, o qual compreende um receptor (301) que recebe uma representação de uma cena em imagem tridimensional em mosaico proveniente de um primeiro ponto de vista, sendo que a representação compreende uma pluralidade de blocos interconectados, sendo que cada bloco compreende um mapa de profundidade e um mapa de textura representando uma janela de visão da cena a partir do primeiro ponto de vista, e sendo que os blocos formam um padrão de mosaico. Um primeiro processador (311) determina as regiões de borda vizinhas em ao menos um primeiro bloco e num segundo bloco em resposta ao padrão de mosaico. Um segundo processador (309) modifica pelo menos um primeiro valor de profundidade de uma primeira região de borda do primeiro bloco, em resposta a pelo menos um segundo valor de profundidade em uma segunda região de borda do segundo bloco, sendo que as regiões de borda são regiões vizinhas. Os mapas de profundidade modificados podem ser usados para gerar uma malha com base na qual podem ser geradas as imagens.