



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

(11) BR 112012005477-5 B1



(22) Data do Depósito: 10/09/2010

(45) Data de Concessão: 29/06/2021

(54) Título: INSERÇÕES VIRTUAIS EM VÍDEIO 3D

(51) Int.Cl.: H04N 13/00.

(30) Prioridade Unionista: 11/09/2009 US 61/241,687.

(73) Titular(es): DISNEY ENTERPRISES, INC..

(72) Inventor(es): SHELDON KATZ; GREGORY HOUSE; HOWARD KENNEDY.

(86) Pedido PCT: PCT US2010048427 de 10/09/2010

(87) Publicação PCT: WO 2011/031968 de 17/03/2011

(85) Data do Início da Fase Nacional: 12/03/2012

(57) Resumo: INSERÇÕES VIRTUAIS EM VÍDEIO 3D. As modalidades referem-se às inserções em vídeo 3D. Os modelos de câmera virtual possibilitam que as inserções sejam reconciliadas em relação aos canais esquerdo e direito do vídeo 3D para maximizar a precisão 3D e o realismo das inserções. As câmaras são formadas como composto, e podem ser derivadas de outros modelos. Os modelos de câmara podem ser baseados em uma análise visual do vídeo 3D e podem ser baseados nos dados da câmera 3D incluindo a convergência ("toe-in") e o espaçamento ocular. Os dados da câmera podem ser derivados da informação coletada usando a instrumentação conectada a uma sistema de câmera 3D, derivados com base na análise visual do vídeo 3D, ou derivados usando uma combinação da informação coletada usando a instrumentação e a análise visual do vídeo 3D. As inserções podem ser ajustadas no espaço 3D e/ou transmitidos separadamente para um sítio remoto. As inserções podem ser ajustadas no espaço 3D com base em um tipo de inserção, de composição da cena de vídeo 3D, e/ou do retorno do usuário, incluindo o ajustamento interativo das inserções 3D e os ajustamentos em vista da sensibilidade do usuário ao cansaço visual. (...).

Relatório Descritivo da Patente de Invenção para "INSERÇÕES VIRTUAIS EM VÍDEO 3D"

CAMPO DA INVENÇÃO

As modalidades referem-se a inserções virtuais em vídeo 3D.

ANTECEDENTES DA INVENÇÃO

Os métodos para transmitir conteúdo de vídeo aos observadores podem utilizar a estereoscopia para projetar o conteúdo do programa em um campo 3D. Os sistemas 3D capazes podem transmitir canais separados para imagens do olho esquerdo e do olho direito, fornecendo visões em paralaxe das cenas. Embora os métodos para adicionar inserções virtuais em vídeo 2D convencional sejam conhecidos, tais métodos 2D podem não ser adequados para fornecer experiências de visualização ideais aos observadores de vídeo 3D. Consequentemente, há uma necessidade de fornecer inserções virtuais mais realistas em vídeo 3D que parecem, para os observadores, como se fossem parte da produção original.

BREVE DESCRIÇÃO DOS DESENHOS

Os desenhos em anexo são incluídos para fornecer um entendimento adicional, são incorporados e constituem uma parte desta especificação e ilustram as modalidades que, juntas com a descrição, servem para explicar os princípios da invenção. Nos desenhos:

A FIG. 1 é uma ilustração esquemática de uma modalidade para gerar inserções e aprimoramentos no vídeo 2D.

A FIG. 2 é uma ilustração esquemática de uma modalidade para gerar inserções e aprimoramentos no vídeo 3D.

A FIG. 3A ilustra uma primeira vista de um método de oclusão 3D exemplificado de acordo com uma modalidade.

A FIG. 3B ilustra uma segunda vista de um método de oclusão 3D exemplificado de acordo com a FIG. 3A.

A FIG. 4A ilustra uma primeira vista de um método de oclusão 3D exemplificado de acordo com uma modalidade.

A FIG. 4B ilustra uma segunda vista do método de oclusão 3D exemplificado de acordo com a FIG. 4A.

A FIG. 5 é uma ilustração esquemática de uma modalidade para gerar inserções e aprimoramentos no vídeo 3D.

A FIG. 6 ilustra um canal de produção e distribuição de vídeo 3D exemplificado de acordo com uma modalidade.

A FIG. 7 é um diagrama de bloco de um sistema de computador exemplificado no qual as modalidades podem ser implementadas.

As presentes modalidades serão agora descritas com relação aos desenhos em anexo. Nos desenhos, números de referência similares podem
5 indicar elementos funcionalmente similares ou idênticos.

DESCRIÇÃO DETALHADA

Enquanto a presente invenção é descrita aqui com relação às modalidades ilustrativas para aplicações particulares, dever-se-ia entender que a invenção não está limitada a essas. Os versados na técnica com acesso aos
10 ensinamentos fornecidos aqui reconhecerão modificações, aplicações e modalidades adicionais dentro do escopo da invenção e dos campos adicionais nos quais a invenção seria de utilidade significativa.

As modalidades incluem a inserção de aprimoramentos, tal como logomarcas de propaganda, caixas de escore e a primeira linha de descida em
15 jogos de futebol, em conteúdo de vídeo 3D. As modalidades referem-se à mídia 3D, incluindo, mas não limitadas a: vídeo, televisão (transmissão aberta, por cabo, por satélite ou por fibra óptica), cinemas, Internet, dispositivos móveis (telefone celular ou outro dispositivo sem fio), e outras plataformas de transmissão contínua de mídia de vídeo 3D.

20 As inserções e aprimoramentos no vídeo 2D podem ser integrados com o vídeo tal que eles possam parecer mais realisticamente como sendo parte do vídeo original. As inserções podem ser implementadas, por exemplo, como descrito na Patente US. No. 5.264.933 de Rosser e Outros, depositada em 28 de janeiro de 1992 e intitulada "Television Displays Having Selected Inserted Indicia",
25 cujos conteúdos são incorporados integralmente no presente por referência. Um sistema de inserção virtual para vídeo 2D pode usar qualquer número de técnicas de busca para reconhecer uma cena e construir um modelo de câmera virtual para a cena para adicionar as inserções virtuais. Um modelo de câmera pode incluir uma posição da câmera e outros parâmetros, que possibilitam que uma câmera seja localizada em
30 relação a uma cena. Uma vez que a cena é reconhecida, modelos de quadro subsequentes podem ser calculados por qualquer número de métodos para acompanhar as cenas de vídeo. O processamento de oclusão possibilita que os objetos em primeiro plano dentro do vídeo bloqueiem as inserções adicionadas ao plano de fundo da cena. Os cálculos de oclusão podem ser executados com base na
35 cor da cena, por exemplo, em sistemas de "chroma key". A inserção precisa ser renderizada usando um renderizador de gráficos disponível comercialmente, por

exemplo, antes de ser misturada com o vídeo de programa. As modalidades citadas aqui podem ser utilizadas em combinação com os sistemas de movimento, onde a informação de movimento é extraída de uma cena e usada para corresponder um movimento de uma inserção ao movimento da cena.

5 O modelo de câmera pode conter múltiplos parâmetros que se referem às medidas físicas de uma câmera montada sobre um tripé, tais como, rotação panorâmica, inclinação, rotação, distância da imagem, posição x, posição y e posição z. Outros parâmetros, tal como parâmetros para a distorção radial da lente, por exemplo, podem ser utilizados. Os parâmetros de dados da câmera podem ser
10 derivados de dados coletados usando instrumentação conectada a um sistema de câmera 3D, podem ser derivados com base na análise visual do vídeo 3D, ou podem ser derivados usando uma combinação dos dados coletados usando a instrumentação conectada ao sistema de câmera 3D e a análise visual do vídeo 3D. O próprio modelo de câmera pode conter parte de toda a informação necessária para descrever o
15 campo de visão de um ou ambos os canais de vídeo 3D. Por exemplo, ele pode conter um único parâmetro tal como zoom ou a distância da imagem associada ou com o canal direito ou com o canal esquerdo. Os parâmetros de único canal alternativos incluem, mas não estão limitados a foco, rotação, distorção da lente, etc. Os parâmetros de dados da câmera determinados para uma visão de canal podem ser
20 derivados independente de outra visão de canal. Também, os modelos de câmera podem ser limitados a um ou mais parâmetros associados com ambos os canais de vídeo 3D (posição da câmera, rotação panorâmica, inclinação). Os parâmetros de dados da câmera podem ser determinados usando similaridades entre as visões de canais individuais. Ademais, os modelos de câmera podem ser limitados aos
25 parâmetros que descrevem a relação dos canais esquerdo e direito (espaçamento ocular, ângulo de convergência, etc.). Os parâmetros de dados da câmera podem ser determinados usando as diferenças entre as visões de canais individuais. Dever-se-ia entender que os parâmetros de dados da câmera podem ser representados com uma ampla faixa de unidades ou dispositivos de medição. Um modelo de câmera composto
30 para vídeo 3D pode ser compreendido de modelos de câmera para os canais individuais do vídeo 3D.

Em uma modalidade desta invenção, um modelo de câmera pode ser representado como uma ou duas matrizes 3x3. Em outra modalidade, o modelo de câmera pode ser gerado usando outras matrizes dimensionais. Os
35 elementos ou parâmetros da representação de matriz podem ser considerados os parâmetros de dados da câmera. Os parâmetros de matriz podem incluir os

parâmetros externos da câmera, tal como, as coordenadas de posição da câmera, e os parâmetros internos, tal como, os fatores de escala horizontal e vertical do sensor. Outros métodos, tal como, por exemplo, os métodos baseados em homografia, podem ser utilizados e a presente invenção não conta com um dispositivo particular para calcular um modelo de câmera. Por exemplo, o modelo de câmera pode fornecer simplesmente uma relação homográfica entre as visões de canal da câmera atuais e alguma referência física, tal como o plano que contém a quadra de basquetebol. Em outro exemplo, o modelo de câmera pode incluir um mapeamento homográfico entre a visão de entrada e uma imagem de referência da cena, onde a imagem de referência é usada para definir a localização de uma inserção gráfica. Os elementos ou parâmetros do mapeamento homográfico podem ser considerados os parâmetros de dados da câmera. Em modalidades adicionais desta invenção, os modelos de câmera podem ser uma representação da localização de um objeto, um grupo de objetos, ou uma parte da cena nos canais de vídeo 3D. Como com todos os modelos de câmera, a localização do objeto no campo de visão pode ser atualizada ao longo do tempo.

As modalidades baseadas na adição de inserções virtuais ao vídeo 2D podem ser incorporadas para gerar inserções separadas para os canais de vídeo do olho esquerdo e do olho direito, como precisa ser usado em sistemas de vídeo 3D. Tais modalidades podem abordar os erros de inserção que podem ocorrer em cada modelo de câmera do canal esquerdo e direito. Estes erros de modelo podem ocorrer devido aos dados de pixel com ruído na busca ou acompanhamento dos canais de vídeo, por exemplo. Os acompanhadores que usam blocos modelo podem ter um componente aleatório, isto é, os blocos podem ser selecionados aleatoriamente, e podem não fornecer um comportamento consistente entre os canais em outro exemplo. Quando se adiciona inserções virtuais ao vídeo 2D, por exemplo, um erro do modelo de busca pode fazer com que um anúncio virtual, localizado em uma quadra de basquetebol, seja mal localizado por volta de 0,5 metros na quadra. Os observadores do vídeo 2D podem até não achar isto desagradável, especialmente se a posição da logomarca estiver relativamente longe das principais características próximas, tais como, as linhas de intersecção da quadra.

No vídeo 3D, entretanto, erros de má localização similares para as imagens do olho esquerdo e do olho direito podem ser considerados desagradáveis devido às imagens do olho esquerdo e do olho direito estarem mal localizadas entre si, especialmente se os erros não tendem a se acompanhar entre si. Consequentemente, relacionar os modelos de câmera do olho esquerdo e do olho direito a uma referência ou manter uma diferença relativa entre os modelos do canal direito e esquerdo, como

descrito aqui, pode aprimorar as experiências do observador que assiste às inserções virtuais no vídeo 3D.

A FIG. 1 é uma ilustração esquemática de uma modalidade para gerar inserções e aprimoramentos em vídeo 2D, de modo que eles possam parecer realisticamente aos observadores como sendo parte do vídeo original. Uma fonte de vídeo, tal como a entrada de vídeo/alimentação de vídeo de programa, é a entrada para o subsistema como a entrada de vídeo 101. A entrada de vídeo 101 pode ser modificada para incluir inserções e aprimoramentos, e emitido como a saída de vídeo 121. Um controlador principal 103 representa um módulo de hardware e/ou software que pode controlar e coordenar os blocos de subsistema 103-113. O bloco de busca 105 representa um módulo de hardware e/ou software que pode analisar a entrada de vídeo 101 para calcular os modelos de câmera e computar a geometria de cena para as cenas de vídeo de programa. O bloco de acompanhamento 107 representa um módulo de hardware e/ou software que pode acompanhar objetos dentro do vídeo de programa para reduzir as exigências de processamento para o bloco de busca 105 e possibilitar o acompanhamento mais suave de inserções e aprimoramentos associados com um plano de fundo da entrada de vídeo 101. O bloco de oclusão 109 representa um módulo de hardware e/ou software que pode determinar quando os objetos em primeiro plano deveriam bloquear inserções e aprimoramentos e gerar uma chave de oclusão para possibilitar que o misturador 113 exiba inserções e aprimoramentos bloqueados. O bloco de renderização 111 representa um módulo de hardware e/ou software que pode receber os modelos de câmera, as localizações de inserção, as chaves de oclusão e outra informação para renderizar inserções e aprimoramentos para misturar com a saída de vídeo 121. Qualquer tipo de gráfico real ou virtual que seja combinado ou, de outra forma, misturado com o vídeo 3D pode ser considerado a inserção de um aprimoramento no vídeo 3D. Esta modalidade exemplificada é somente para ilustração e as modalidades podem ser implementadas com várias outras arquiteturas que compreendem hardware, software ou combinações de hardware e software para um ou mais blocos. Os múltiplos blocos de renderização 111 e os misturadores 113 podem, por exemplo, servir como processadores de retaguarda para fornecer múltiplas versões de inserções e aprimoramentos aos diferentes observadores. Em algumas modalidades, o bloco de Busca 105 e o bloco de Acompanhamento 107 podem ser combinados, por exemplo. Em outras modalidades, o bloco de Busca 105, o bloco de Acompanhamento 107 e o bloco de Oclusão 109 podem ser combinados.

As abordagens básicas para gerar os modelos de câmera

podem incluir sensores físicos conectados ao sistema de câmera, a análise de processamento de imagem e visão de computador dos canais de vídeo, ou uma combinação das medições do sensor físico e o processamento de análise de vídeo. O processamento de visão pode ser empregado pelo bloco de Busca 105, pelo bloco de Acompanhamento 107 ou pelo bloco de Oclusão 109 na FIG. 1. Para o bloco de Busca 105, a análise visual pode ser usada para derivar as localizações de imagem ou de tela das características visuais na cena. Os modelos de câmera podem ser gerados por quadros particulares através da associação das localizações de característica de imagem e suas posições 3D em cena correspondentes. Tais métodos são descritos no Pedido de Patente US. No. 12/659.628, cujos conteúdos são incorporados integralmente no presente por referência. Para o bloco de Acompanhamento 107, a análise visual pode ser usada para acompanhar a localização das características ou pontos de interesse entre os quadros de uma sequência de imagens. Uma operação exemplificada é descrita no Pedido de Patente US. No. 6.741.725 por Astle, cujos conteúdos são incorporados integralmente no presente por referência. Para o bloco de Oclusão 109, a análise visual pode ser usada para distinguir os pixels em primeiro plano de uma imagem a partir dos pixels de plano de fundo. Um método baseado na cor é descrito por Jeffers e outros na Patente US. No. 7.015.978, cujos conteúdos são incorporados integralmente no presente por referência.

A FIG. 2 ilustra uma representação esquemática de uma modalidade exemplificada para gerar inserções e aprimoramentos em vídeo 3D. O controlador 290 pode empregar os métodos usados para adicionar inserções virtuais em vídeo 2D. O controlador 290 representa um módulo de hardware e/ou software que pode fazer interface com as unidades de processamento de vídeo para ambos os canais direito e esquerdo. O vídeo de programa para o canal de vídeo do olho esquerdo é inserido, como a entrada de vídeo esquerdo 201, aos subsistemas correspondentes de busca 205, de acompanhamento 207, de oclusão 209 e misturador 213. O vídeo de programa para o canal de vídeo do olho direito é inserido, como a entrada de vídeo direita em 251, aos subsistemas correspondentes de busca 255, de acompanhamento 257, de oclusão 259 e misturador 263. A entrada de vídeo esquerda/direita 201, 251 pode ser modificada para incluir inserções e aprimoramentos, e emitida como saída de vídeo esquerda 221 e saída de vídeo direita 271, respectivamente.

O controlador 290 pode controlar e coordenar os vários blocos de subsistema. Os blocos de Busca 205, 255 representam os módulos de hardware e/ou software que podem analisar as entradas de vídeo esquerda/direita 201, 251, e

calcular os modelos de câmera para as cenas de vídeo de programa. Os blocos de Acompanhamento 207, 257 representam os módulos de hardware e/ou software que podem acompanhar objetos dentro do vídeo para reduzir as exigências de processamento para os blocos de busca 205, 255, e possibilitar o acompanhamento
5 mais suave de inserções e aprimoramentos com relação aos planos de fundo das entradas de vídeo direita e esquerda 201, 251. Os blocos de Oclusão 209, 259 representam os módulos de hardware e/ou software que podem determinar quando os objetos em primeiro plano deveriam bloquear inserções e aprimoramentos, de modo a gerar uma chave de oclusão que possibilita que os misturadores 213, 263 exibam
10 inserções e aprimoramentos bloqueados. Os blocos de renderização 211, 261 representam os módulos de hardware e/ou software que podem receber a câmera e outros modelos, as localizações de inserção, as chaves de oclusão e outra informação para renderizar as inserções e aprimoramentos para misturar a saída de vídeo esquerda/direita 221, 271.

15 O controlador 290 pode incluir um gerenciador de modelo 292 que monitora os blocos de busca 205, 255 e os blocos de acompanhamento 207, 257 para determinar a informação de modelo de câmera atual para os canais de vídeos esquerdo e direito. O gerenciador de modelo 292 pode relacionar cada modelo de câmera para os canais de vídeo esquerdo e direito para reconciliar os modelos de
20 câmera esquerda e direita. Por exemplo, o gerenciador de modelo 292 pode calcular um modelo de câmera médio/referência que tem uma posição de câmera, em coordenadas 3D, situada a meio caminho entre os modelos de câmera de canal esquerdo e de canal direito. Em alguns casos, pode ser preferível o uso dos modelos de câmera de canal esquerdo ou de canal direito como uma referência comum. Usar
25 um modelo de câmera de referência comum ou média associado com ambos os canais de vídeo esquerdo e direito pode reduzir os efeitos de incompatibilidade de modelo de câmera entre os canais esquerdo e direito. Os modelos de câmera do canal esquerdo e direito, por exemplo, podem ser deslocados por quantidades fixas ou distâncias ao modelo de câmera de referência comum. Como um exemplo, os
30 modelos de câmera do canal esquerdo e direito podem ser feitos para ter distâncias espaciais fixas em coordenadas 3D a partir das coordenadas 3D do modelo de câmera de referência comum. A distância entre os modelos de câmera esquerdo e direito pode, por exemplo, corresponder à distância entre as lentes da câmera esquerda e direita para um sistema de câmera 3D conhecido. A distância entre as lentes da
35 câmera, ou a distância ou espaçamento ocular, pode variar durante as sequências de vídeo, mas uma distância média pode ser adequada para algumas aplicações. Para

outras aplicações, pode ser desejável um espaçamento ocular de modelo mais preciso com fórmulas ou aproximações conhecidas, por exemplo. Em outro exemplo, os deslocamentos do modelo de câmera de referência comum podem ser calculados usando métodos para calcular a paralaxe entre as imagens estereoscópicas. Isso pode ser executado através da análise visual dos canais esquerdo e direito, ou através da análise visual dos canais esquerdo e direito, ou análise visual dos canais esquerdo e direito em conjunto. A análise de paralaxe estereoscópica pode ser usada para determinar ou derivar a relação entre os canais de vídeo 3D. Os parâmetros de dados de câmera podem ser derivados com base na análise de paralaxe ou na análise estereoscópica dos canais de vídeo 3D. A reconciliação entre os canais pode ser usada para um subconjunto de parâmetros também. Por exemplo, os dados de zoom ou de aumento podem ser reconciliados com base em um valor de zoom médio antes dos modelos de câmera esquerda e direita serem reconciliados. Neste exemplo, os dados de zoom poderiam ser filtrados quanto ao ruído antes de aplicá-los para calcular os modelos de câmera. Alternativamente, um ajustamento por mínimos quadrados pode ser empregado para encontrar uma melhor correspondência com os parâmetros de entrada.

Limitar os parâmetros físicos para validar as faixas esperadas é outro mecanismo que pode ser empregado no processo de reconciliação. Isso pode se aplicar a pontos individuais no tempo, bem como ao longo de um período de tempo. Por exemplo, a taxa de mudança de um parâmetro particular, tal como o zoom, pode ser limitada ou suavizada. Isso pode ser alcançado em parte através do processamento de imagem dos canais de vídeo 3D ou através do processamento de sinal de medições de sensor físico. A reconciliação pode utilizar técnicas de filtragem conhecidas, métodos estatísticos, métodos de limiarização, ou outras abordagens. A reconciliação pode ser aplicada aos parâmetros de dados da câmera individuais ou a um agrupamento de parâmetros de dados da câmera. Um agrupamento de parâmetros de dados da câmera, tal como um modelo da câmera composta, pode ser consistente, ou de outra forma, reconciliado com um ou mais parâmetros de dados da câmera. Em algumas modalidades, um modelo de câmera composto e um ou mais parâmetros dos dados da câmera são consistentes, ou de outra forma, reconciliados com as estimativas iniciais para um ou mais parâmetros dos dados da câmera individuais. A reconciliação pode envolver a produção de um modelo de câmera ou de parâmetros dos dados da câmera consistentes com outros parâmetros dos dados da câmera. Em uma modalidade, um ou mais parâmetros de dados da câmera ou modelos de câmera que são reconciliados com um primeiro parâmetro de dados da

câmera podem ser gerados simultaneamente ao primeiro parâmetro de dados da câmera. Em outra modalidade, um ou mais parâmetros de dados da câmera ou modelos de câmera que são reconciliados com um primeiro parâmetro de dados da câmera são gerados de forma sequencial após o primeiro parâmetro de dados da
 5 câmera. Em modalidades alternativas, um ou mais parâmetros de dados da câmera ou modelos de câmera que são reconciliados com um primeiro e um segundo parâmetro de dados da câmera, são gerados ou simultânea ou sequencialmente após a geração do primeiro e do segundo parâmetro de dados da câmera. A reconciliação pode ser baseada nos canais do vídeo 3D, na análise visual dos canais 3D, nos parâmetros da
 10 câmera derivados dos canais de vídeo 3D, nas medições do sensor ou nos parâmetros da câmera do sistema de câmera 3D, ou quaisquer combinações dos acima. A reconciliação não está limitada a um método particular ou a um grupo de métodos.

Para algumas aplicações 3D, pode ser desejável utilizar
 15 modelos de busca mais frequentes em comparação às aplicações 2D, para minimizar os deslocamentos que podem ocorrer com o acompanhamento da cena. Pode ser também desejável minimizar os deslocamentos relativos entre os canais esquerdo e direito entre si. A precisão de busca para as aplicações de inserção 3D é desejável em vista dos potenciais erros associados com os objetos 3D que convergem para
 20 localizações incorretas dentro das cenas 3D, por exemplo. Tais erros podem fazer as inserções imprecisas parecerem notadamente não naturais em 3D, em contraste dos erros de localização em imagens 2D. Por exemplo, um erro de convergência para uma primeira linha de descida 3D em um jogo de futebol na televisão pode fazer com que a linha de campo apareça para cima ou para baixo do campo de jogo. As inserções de
 25 canal esquerdo e direito da primeira linha de descida precisam corresponder em comprimento, bem como ter posições corretas, ou as extremidades da linha podem não parecer naturais em 3D. Tipos adicionais de erros, por exemplo, erros imprevistos, podem fazer com que os objetos se movam de repente no espaço 3D. As incompatibilidades de tamanho podem causar erros de cor ou outros problemas de
 30 aparência. Em tais casos, o gerenciador de modelo 292 pode aprimorar o desempenho levando em conta as diferenças de modelo de câmera entre os canais esquerdo e direito.

Os blocos de acompanhamento 207, 257 podem utilizar os métodos 2D para acompanhar as cenas, tal como os métodos de modelo de textura.
 35 (Ver, por exemplo, Patente US. No. 6.741.725 por Astle, intitulada "Motion Tracking Using Image-Texture Templates"). Usando a análise visual dos canais de vídeo 3D, os

modelos de textura ou os blocos de acompanhamento podem ser selecionados dentro das cenas para modelos de geração e acompanhamento de cena. Os métodos de acompanhamento podem usar a informação 2D em cenas, usando os modelos de textura 2D dos canais esquerdo e direito. Outros métodos de acompanhamento podem

5 usar a textura de cena 2D, mas utilizam a informação de posição 3D para os blocos de acompanhamento. Tais métodos podem ser chamados métodos de acompanhamento 3D, embora eles utilizem os modelos de textura 2D. Em outros casos, a informação 3D derivada de visões estereoscópicas dos canais esquerdo e direito pode ser usada. Os blocos de acompanhamento 3D baseados em voxels, ou pixels 3D, podem ser usados

10 para o acompanhamento de cenas 3D. Tais métodos podem ser estendidos para outras técnicas tal como o fluxo óptico. Para muitas aplicações, o processamento 2D pode ser adequado, entretanto, e minimiza a complexidade e o custo. Em algumas modalidades, um objeto ou grupos de objetos ou uma parte da cena pode ser acompanhada nos canais individuais do vídeo 3D, ou acompanhada juntamente

15 através de ambos os canais simultaneamente.

Usando voxels, alguns erros do método de acompanhamento 2D podem ser evitados. Por exemplo, o acompanhamento do modelo 2D pode falhar quando muitos blocos de modelo, em relação aos blocos do plano de fundo, estão em objetos em primeiro plano em movimento nas cenas. Quando tais objetos em primeiro

20 plano se movem em relação a um plano de fundo, os modelos de câmera incorretos podem ser calculados. Isto pode ocorrer, por exemplo, durante um jogo de basquetebol transmitido por televisão, quando a câmera amplia a visão dos jogadores, e o acompanhamento usa os blocos sobre os jogadores. Usar voxels com coordenadas 3D conhecidas possibilita a seleção de voxels do plano de fundo com

25 base na posição 3D para o acompanhamento. No exemplo acima, os voxels podem ser selecionados sobre um plano do piso da quadra ou sobre o plano da arquibancada com observadores, por exemplo. Similar à busca, o acompanhamento pode se beneficiar através do gerenciador de modelo 292, levando em consideração as diferenças de modelo entre os canais. Os ganhos de desempenho podem também ser

30 realizados limitando-se o bloco 2D ou as buscas de voxel dentro das restrições definidas pelas relações de canal esquerdo e direito. A busca por blocos ou voxels sobre regiões similares possibilita o uso de mais elementos de acompanhamento, dando uma melhor precisão e desempenho de acompanhamento. A análise acima pode ser alcançada através da análise visual dos canais 3D, das medições de sensor

35 do sistema de câmera 3D, ou uma combinação de análise visual e medições do sensor. O uso de voxels pode ser parte de uma reconciliação da análise visual ou

medições do sensor associado com os canais de vídeo 3D.

Os blocos de oclusão 109, 259 podem executar o processamento de oclusão. O processamento de oclusão pode, por exemplo, ser feito usando métodos tal como o "chroma key". Para vídeo 3D, o processamento de oclusão poderia usar a informação 3D da cena. Os pixels dentro de uma cena podem, por exemplo, ser relacionados em ambos os canais esquerdo e direito usando métodos tal como a correspondência de modelos. A informação de posição 3D para os pixels de canal esquerdo e direito correspondentes pode então ser calculada usando, por exemplo, as técnicas de geometria epipolar. Uma vez que a informação de posição 3D para os pixels é determinada, um subsistema de oclusão pode determinar se ou não esses pixels deveriam ser bloqueados pelos objetos em primeiro plano. Como um exemplo, se um bloco de pixels em primeiro plano foi determinado como estando localizado mais próximo à câmera 3D do que o bloco de pixels no plano de fundo na cena, o bloco de pixels em primeiro plano poderia bloquear o bloco de pixels do plano de fundo.

A FIG. 3A ilustra uma primeira visão de um método de oclusão 3D exemplificado de acordo com uma modalidade. A atadura de cor verde sobre uma faixa de pulso do jogador é mostrada com a atadura 306 em um canal do olho esquerdo 302 e em um canal do olho direito 304. Usando um método de "chroma key" em cada um dos canais do olho esquerdo/direito 302, 304, essa atadura 306 pode ser difícil de distinguir da cor verde do campo de jogo no plano de fundo 310, e pode aumentar a probabilidade de sangramento. Entretanto, as modalidades podem usar um método de paralaxe para distinguir a atadura 306 do plano de fundo 310, mesmo quando cores similares, tal como a cor da atadura 306 e do plano de fundo 310, estão envolvidas. Os métodos de paralaxe podem também ser usados em conjunto com os métodos de "chroma key".

A FIG. 3B ilustra uma segunda visão do método de oclusão 3D exemplificado da FIG. 3A. Usando a paralaxe, a posição 3D da atadura verde do pulso do jogador 306 pode ser determinada como estando mais próxima à câmera/observador 312 do que a cor verde similar do campo de jogo 310. Uma inserção virtual, tal como uma primeira linha de descida amarela localizada atrás do jogador no campo de jogo 310, pode então ser bloqueada pela atadura 306 com base na determinação da paralaxe. Inserções e aprimoramentos que utilizam a paralaxe podem evitar o "sangramento" não natural da primeira linha de descida inserida sobre a atadura 306.

A FIG. 4A ilustra uma primeira visão de um método de oclusão

3D exemplificado de acordo com uma modalidade, onde a informação espacial é usada por múltiplos jogadores 406, 408. Usando uma busca, por exemplo, os jogadores 406, 408 são encontrados nos canais esquerdo e direito 402, 404. As posições dos jogadores 406, 408, incluindo uma distância a partir da câmera/observador 412, podem ser determinadas usando a paralaxe. O campo de jogo 410 pode aparecer no plano de fundo e pode ser bloqueado por outros jogadores e por inserções virtuais. Como ilustrado na FIG. 4B, a inserção 414 pode ser bloqueada pelo jogador 406, que está mais próximo da câmera/observador 412, mas não pode ser bloqueada pelo jogador 408 que está mais afastado da câmera/observador 412. A inserção virtual 414 pode parecer que está entre os jogadores 406 e 408 dentro da cena 3D, enquanto não há sangramento no plano de fundo do campo de jogo 410 atrás dos jogadores 406, 408. Este método pode ser estendido para qualquer pluralidade de jogadores ou objetos dentro das cenas 3D.

A análise estereovisual entre as visões esquerda e direita do vídeo 3D pode permitir a geração de um mapa ou máscara de profundidade, onde os pixels ou regiões da cena de vídeo podem ser representados por medições de profundidade. Vários métodos para gerar mapas de profundidade a partir de uma visão estereoscópica podem ser usados. Quando o mapa de profundidade segue precisamente o desenho dos objetos dentro da cena, ele pode ser usado para gerar uma máscara de oclusão para remover as seções dos gráficos inseridos. As seções removidas são impedidas de bloquear os objetos em primeiro plano, possibilitando que os mesmos apareçam na frente do gráfico inserido. Isso pode, por exemplo, ser derivado do mapa de profundidade fazendo uma comparação pixel por pixel entre uma distância efetiva a partir da câmera de um pixel do gráfico inserido e a distância a partir da câmera do ponto na cena associado com o pixel. O gráfico inserido pode, por exemplo, ser removido quando o pixel do objeto ou da cena estiver mais próximo da câmera do que a posição virtual do gráfico. O gráfico inserido pode, por exemplo, ser colocado no vídeo quando o pixel do objeto ou da cena estiver mais afastado da câmera do que a posição virtual do gráfico.

Os gráficos padrão que sobrepõem o vídeo nas transmissões esportivas 2D podem apresentar mais desafios nas produções de vídeo 3D. Os gráficos podem incluir camadas gráficas de escore fixo, às vezes chamadas barras de escore ou "Fox boxes", que podem exibir continuamente a hora do jogo atual, escore, e informação de jogo relevante. Os gráficos podem também incluir gráficos instantâneos temporários, às vezes chamados gráficos terço inferiores, que fornecem, informação no plano de fundo sobre os jogadores nos jogos. Uma abordagem para

inserir tais gráficos em vídeo 3D pode ser fazer com que os gráficos apareçam em localizações ou distâncias fixas em relação à câmera. Entretanto, isso pode não ser agradável para os observadores e, em alguns casos, pode causar cansaço visual, algumas vezes associado à visualização 3D. Se os gráficos parecerem agradáveis aos observadores, pode depender das profundidades dos objetos e do plano de fundo na cena de vídeo 3D em um dado momento ou durante todo um período de tempo. O posicionamento dos objetos e do plano de fundo em maiores distâncias de visualização pode ser mais fácil para que os observadores foquem, e então reduzindo o cansaço visual do observador. Adicionalmente, os gráficos localizados relativamente próximos à câmera e/ou afastados dos elementos de cena, por exemplo, distantes em frente à tela/plano de exibição e/ou próximos aos observadores, podem distrair os mesmos da cena e/ou podem parecer aos observadores irrelevantes à cena. Integrar os gráficos dentro da cena, entretanto, pode minimizar tais problemas.

As modalidades podem usar modelos de câmera para guiar a localização de gráficos virtuais, incluindo os gráficos de ajuste, em cenas de vídeo 3D de modo que podem parecer mais agradáveis aos observadores. Os modelos de câmera, junto com os modelos físicos, podem permitir que a faixa de profundidade de objetos em cenas seja determinada. Em uma modalidade, os gráficos de ajuste podem ser colocados na localização do plano/tela de exibição, parecendo estar localizados na mesma distância do observador que o plano/tela de exibição. Em outras modalidades, os gráficos de ajuste podem ser colocados em distâncias relativas comparáveis aos objetos em visualização, ou ligeiramente na frente ou atrás dos objetos em visualização. As localizações dos gráficos podem diferir com base na composição de uma cena. Por exemplo, a localização de um gráfico em uma câmera alta com uma ampla tomada de vigilância de um jogo de futebol pode diferir da localização de um gráfico para um nível de campo, tomada ampliada de um grupo de jogadores no campo. Em outra modalidade, os gráficos de ajuste podem ser localizados em uma profundidade além dos objetos ou superfície de jogo nas cenas, aparecendo em uma distância relativamente grande da câmera. Os modelos de câmera e os métodos de busca podem ser usados para determinar as localizações de tela que provavelmente seriam desbloqueadas pelos jogadores ou árbitros, ou os algoritmos podem encontrar áreas desbloqueadas diretamente. Em modalidades adicionais, a profundidade dos gráficos pode ser fixa para uma dada câmera com base na cobertura operacional esperada, tal como na vista de vigilância versus a cobertura do jogador isolado. Isso pode ser confirmado de uma forma sistemática usando sinais de contagem ou ouvindo a chamada de áudio do jogo pelo diretor da produção, por exemplo. Em outra

modalidade, os modelos de câmera podem ser usados para avaliar a adequabilidade da cena para as sobreposições de gráficos levando em consideração a distância 3D a partir da câmera. Os gráficos podem ser seletivamente habilitados usando vários critérios de modo a serem agradáveis para os observadores.

5 As modalidades podem ser estendidas para outra informação que pode ser inserida nas cenas de vídeo 3D. As legendas ou texto "closed caption" podem ser inseridos e integrados dentro das cenas 3D. As inserções podem ser posicionadas para minimizar o cansaço visual ou por outras razões estéticas ou funcionais. O texto de diálogo, por exemplo, pode ser localizado próximo aos alto-
10 falantes nas cenas. Os metadados dentro de fluxos de vídeo podem, por exemplo, possibilitar a localização automática de texto "closed caption" dentro das cenas. A colocação da inserção virtual pode ser controlada pelos observadores, e pode ser implementada, por exemplo, como descrito na Publicação do Pedido de Patente US. No. 2010/0050082 para Katz e outros, depositada em 13 de agosto de 2009 e
15 intitulada "Interactive Video Insertions, and Applications Thereof", cujos conteúdos são incorporados integralmente no presente por referência. Os observadores propensos ou sensíveis a cansaço visual quando assistindo vídeo 3D podem escolher, por exemplo, preferir as inserções em maiores distâncias de visualização.

As modalidades podem utilizar gráficos virtuais integrados nas
20 cenas para apresentar vários tipos de dados no vídeo 3D, de modo que os dados pareçam agradáveis aos observadores. A informação sobre o estado do jogo pode ser apresentada como um gráfico alfanumérico integrado no campo de jogo, por exemplo. Em uma modalidade, tais dados podem ser apresentados em uma localização de campo fixa, tal como próxima dos arremessadores em um jogo de baseball, ou como
25 parte do círculo central ou próximo da linha de fundo em um jogo de futebol. Em outra modalidade, um gráfico de informação pode ser anexado a outros gráficos virtuais, tal como os gráficos de distância e/ou de descida, que estão associados com a localização da primeira linha de descida ou linha de gráficos de "scrimmage". Os gráficos de informação podem ser apresentados em localizações alternativas da
30 produção de televisão. Isso pode incluir a parede dos fundos ou as arquibancadas de um jogo de baseball, ou uma marca suspensa a partir do andar superior da estrutura de um estádio em uma produção de jogo de futebol.

Localizar inserções virtuais em distâncias de visualização maiores pode reduzir o cansaço visual e pode reduzir as exigências de foco do olho
35 após os períodos de foco relativamente próximos. Para alguns observadores que são sensíveis a vídeo 3D e que podem desenvolver dores de cabeça, o foco em maiores

distâncias pode reduzir os sintomas desagradáveis. Ter a capacidade de controlar as distâncias de inserção virtual possibilita que as produções de vídeo reduzam o cansaço visual e outros sintomas associados com o vídeo 3D.

Os parâmetros para o sistema de vídeo 3D incluem o
5 espaçamento ocular e ângulo de convergência. O espaçamento ocular é a distância entre as lentes, e o ângulo de convergência é o ângulo de visualização relativo entre as lentes. Os parâmetros podem ser controlados manualmente por um operador. Isso pode ser feito por um indivíduo designado para suportar um ou mais operadores de câmera 3D. Os motores podem mover as câmeras para ajustar os parâmetros. Os
10 parâmetros podem ser determinados com base na distância do objeto e outra informação de cena. Os operadores podem determinar os parâmetros contando com a experiência com cenas similares, usando as linhas guia conhecidas, usando telas com imagem ao vivo, ou usando outras técnicas. As câmeras ou os controladores de câmera podem calcular esses parâmetros com base nas tabelas de consulta ou nos
15 parâmetros tais como, distância de visualização, ângulo de visualização, geometria da cena, etc. para determinar as configurações da câmera. O espaçamento ocular e o ângulo de convergência podem ser computados diretamente combinando visualmente e registrando os pontos de característica no plano de fundo entre os dois canais estéreo. O espaçamento ocular e o ângulo de convergência da câmera conhecida
20 podem ser incorporados nos cálculos para os modelos de inserção 3D. Esses parâmetros podem ser embutidos dentro do vídeo como metadados ou podem ser enviados, por exemplo, via um canal de dados diretamente a um sistema de inserção virtual.

Em uma modalidade usando dados de câmera associados com
25 as configurações/parâmetros de câmeras 3D, um controlador pode embutir os dados de câmera 3D no intervalo de apagamento vertical da gravação de vídeo produzida pela câmera. Os dados da câmera 3D podem incluir o espaçamento ocular, o ângulo de convergência, zoom, foco, extensor, e outros parâmetros ou sinais de câmera 3D. Dados adicionais de uma montagem de câmera 3D, tais como, os dados de rotação e
30 inclinação similares aos dados dos sistemas usados para vídeo 2D, também podem ser incluídos nos dados da câmera 3D. Tais dados de câmera 3D embutidos podem ser direcionados com o vídeo para localizações remotas, tal como estúdios de transmissão, para inserções virtuais 3D. As regiões de inserção no vídeo, ou os métodos de inserção, podem ser selecionados para assegurar a integridade dos dados
35 no destino final para o sistema de inserções virtuais. Em outro exemplo, os dados de câmera podem ser codificados dentro de canais de áudio não usados, dentro da

região de apagamento horizontal ou dentro da região dos dados auxiliares horizontais (HANC) do vídeo.

Outros tipos de metadados, exceto os dados de câmera, podem ser inseridos no vídeo para habilitar as inserções 3D virtuais em estágios à jusante em uma produção de vídeo e/ou no canal de distribuição. Em um exemplo, 4 pontos podem definir uma área alvo para inserções virtuais para cada um dos canais esquerdo e direito. Esses 8 pontos definem um plano retangular 3D que pode ser usado para as inserções em um estágio posterior. Outro número de pontos ou representação alternativa, tal como as bordas ou curvas ou curvas paramétricas, pode ser usado para designar a localização de aprimoramentos inseridos. Em outro exemplo, um marcador de localização alvo 3D pode ser inserido no vídeo e então substituído a jusante. O marcador de localização alvo pode representar as coordenadas de imagem para inserir um aprimoramento em um quadro particular ou quadro chave. O marcador de localização alvo para quadros intermediários pode ser interpolado ou, de outra forma, determinado a partir do marcador de localização alvo em quadros chave dentro de proximidade temporal ao quadro intermediário. Alternativamente, o marcador de localização alvo pode representar as coordenadas da imagem de um objeto físico na cena, tal como um campo de futebol, usado para determinar as coordenadas de imagem de um aprimoramento ou inserção. As coordenadas podem ser embutidas ou, de outra forma, codificadas no vídeo 3D de tal forma que elas não afetam a parte do vídeo usado para a visualização ativa. Isso pode incluir os canais de áudio não usados dentro da região de apagamento horizontal ou dentro da região de dados auxiliares horizontais (HANC) do vídeo. O marcador de localização pode ser processado de modo que as regiões de oclusão não sejam incluídas no marcador de localização. Em outros exemplos, as localizações de inserção podem ser codificadas em metadados e nas mascaras chave de oclusão que são codificadas separadamente. Vários métodos similares podem ser inventados por aqueles versados na técnica.

Em alguns casos, pode ser preferível processar o vídeo a montante, por exemplo, próximo à localização do evento, ao contrário de à jusante, por exemplo, em um estúdio. Como um exemplo, o vídeo disponível em um evento pode ser descomprimido, enquanto o vídeo em um estúdio pode ser comprimido para facilitar a transmissão via satélite. Pode ser que o processamento de oclusão forneça melhores resultados usando vídeo descomprimido, por exemplo. Em outro exemplo, pode ser desejável manter um equipamento de menor porte em um estúdio transferindo ao menos algum processamento para estágios à montante do estúdio.

Para algumas aplicações, pode ser desejável usar uma combinação de dados de câmera e processamento de visão para calcular os modelos de câmera/inserção, ou usar somente o processamento de visão para calcular os modelos e derivar os parâmetros de cena/sistema de câmera. Em um exemplo, os dados de rotação e de inclinação podem ser usados para fornecer uma localização de busca aproximada para um sistema baseado em visão que poderia refinar a busca usando os métodos baseados na visão. Em outro exemplo, uma localização de busca aproximada poderia ser derivada por processamento de visão, espaçamento ocular e ângulo de convergência utilizado a partir da câmera 3D. Várias combinações podem ser vantajosas para as diferentes aplicações e disponibilidade de sinal. O processamento baseado em visão pode ser alcançado através da análise visual dos canais de vídeo do vídeo 3D. Os canais podem ser processados de forma independente ou em combinação.

Os modelos de câmera para canais individuais ou os próprios canais podem ser usados individualmente ou em conjunto para calcular o espaçamento ocular ou os parâmetros do ângulo de convergência para os sistemas de câmera 3D. Os parâmetros calculados podem ser usados para controlar as câmeras 3D ou para aumentar o controle manual do operador de câmera. Essas configurações podem ser refinadas quando mudanças de vídeo em resposta às novas configurações e os novos parâmetros/modelos de câmera são calculadas. Os limites podem ser colocados em parâmetros tal como o zoom para evitar o enquadramento que pode ser desagradável aos observadores. O cálculo do espaçamento ocular e do ângulo de convergência pode automaticamente possibilitar a configuração mais rápida das câmeras 3D e fornecer resultados e configurações mais consistentes. O cálculo de parâmetro automático pode economizar custos de produção minimizando trabalho. As produções 3D podem favorecer visualizações relativamente próximas de cenas para dar aos observadores uma sensação mais forte da estrutura 3D. Visualizações relativamente longas, nas quais os objetos estão em distâncias de visualização relativamente grandes, podem parecer mais como vídeo 2D para os observadores, e pode-se considerar que as produções 3D não são necessárias para tais visualizações. As tomadas de ação de enquadramento em distâncias mais curtas, tal como durante transmissões televisivas de futebol, por exemplo, podem fornecer desafios para configurar rapidamente ou ler as câmeras 3D. Para tais casos, as modalidades podem ser usadas para ajustar automaticamente os parâmetros da câmera 3D para a variação das tomadas de cena. Usando esses métodos, um longo passe de futebol pode ser coberto com variações mais estreitas de tomadas de um longo passe, ao

contrário de uma tomada mais consistentemente afastada, por exemplo.

Em outra modalidade para inserções virtuais, pode ser desejável localizar as inserções em um plano particular, tal como, por exemplo, uma primeira linha de descida 3D em uma transmissão de futebol 3D. As inserções dos canais esquerdo e direito precisam ser posicionadas corretamente dentro dos quadros de modo que a inserção final convergirá para a posição 3D correta dentro da cena. Os erros de posição podem fazer com que a primeira linha de descida apareça ou sobre ou sob o plano do campo, por exemplo. Para evitar tais problemas, as posições de inserção de canal esquerdo e direito podem ser monitoradas e ajustadas para garantir que a inserção 3D final converge em um plano particular na cena. Em alguns casos pode ser aceitável ou preferível errar em direção a um lado de um plano, tal como acima de um campo de jogo ao contrário de abaixo do campo, por exemplo.

A FIG. 5 é uma ilustração esquemática de uma modalidade para gerar as inserções e aprimoramentos no vídeo 3D, demonstrando que o vídeo 3D pode ser manipulado usando uma abordagem integrada. A entrada, entrada de vídeo 3D 501, pode compreender os canais esquerdo e direito separados tal como alimentações independentes de Interface Digital Serial de Alta Definição (HD-SDI), ou pode compreender uma alimentação composta tendo os canais esquerdo e direito entrelaçados. O entrelaçamento pode incluir comprimir de forma anamórfica os canais esquerdo e direito em um único fluxo HD-SDI, ou usar um esquema alternativo para combinar as alimentações. A entrada de vídeo 3D 501 pode ser modificada para incluir as inserções e/ou aprimoramentos, e emitido como a saída de vídeo 521. As inserções e/ou aprimoramentos podem parecer aos observadores ser realisticamente parte do vídeo original. Um controlador principal (interface de controle principal integrado 503) pode controlar e coordenar os blocos de subsistema 503-513. Outros esquemas de combinar ou codificar os canais individuais em um fluxo composto são possíveis e podem ser baseados nos métodos de compressão de vídeo.

O bloco de busca integrado 505 pode analisar a entrada de vídeo 3D 501 e calcular os modelos de câmera e analisar a geometria de cena para as cenas de vídeo de programa. Os modelos de câmera e a análise podem ser derivados de um único canal e extrapolados para o segundo canal, derivado de um único canal e refinado pelo processamento do segundo canal, computados a partir de ambos os canais com um modelo que corresponde, de forma otimizada, a ambas as visualizações, ou qualquer combinação/permutação das acima. Para o bloco de busca integrado 505, a análise visual pode ser usada para derivar a localização da imagem de características visuais nos canais esquerdo e direito. Um modelo de câmera

composto pode ser gerado para um quadro particular através da associação de localizações de imagem das características de cena nos canais de vídeo 3D e através da posição 3D correspondente de características de cena. Um modelo de câmera composto pode ser gerado reconciliando as localizações de características derivadas para os canais individuais. Isso pode ser alcançado, por exemplo, computando o ajustamento de erro por mínimos quadrados para mapeamento entre as localizações de imagem de características e localizações de cena 3D.

O bloco de acompanhamento integrado 507 pode atualizar os modelos baseados em uma única visualização e extrapolar a segunda visualização. O bloco de acompanhamento integrado 507 pode atualizar os modelos com base em uma única visualização refinada pela segunda visualização, atualizar os modelos diretamente para combinar, de forma otimizada, ambas as visualizações, ou atualizar os modelos com base em qualquer combinação/permutação dos acima. A análise visual pode ser usada pelo bloco de acompanhamento integrado 507 para acompanhar a localização de características ou pontos de interesse entre os quadros de uma sequência de imagens. Isso pode ser executado em conjunto com as medições do sensor físico. O bloco de acompanhamento integrado 507 pode gerar um modelo composto de uma forma similar ao gerenciador de modelo 292 do controlador 290 (FIG. 2), ou pode acompanhar um objeto ou parte de um objeto tal como o capacete de um jogador de hóquei. Em uma modalidade, o objeto pode ser acompanhado e a cena pode ser acompanhada de forma independente, de tal forma que a posição do objeto em relação à cena pode ser derivada para um efeito gráfico tal como a colocação de um trilho atrás do objeto, por exemplo, um jogador. Ademais, o bloco de acompanhamento integrado 507 pode usar os dados extraídos pela instrumentação da câmera principal (rotação, inclinação, espaçamento ocular, ângulo de convergência, etc.) e pela informação de lente extraída (zoom, foco, duplicador, distância focal, convergência do eixo óptico, etc.), comunicando ou produzindo conexões elétricas para a câmera e/ou as lentes. Em uma modalidade, o processo de calibração de pré-tomada é usado para derivar a relação entre as visualizações esquerda e direita. Em outra modalidade, os modelos de canal esquerdo e direito são derivados com base no conhecimento da câmera física, por exemplo, com base no conhecimento dos sensores de câmera do canal esquerdo e direito, parâmetros da lente, etc. No tempo de operação, o modelo de câmera pode ser computado para um canal e o modelo de calibração pode ser usado para derivar o modelo de câmera para outro canal.

O bloco de oclusão integrado 509 pode determinar quando os

objetos em primeiro plano deveriam bloquear as inserções e aprimoramentos no vídeo 3D gerando uma chave de máscara associada com cada canal. Para a metodologia "chroma key", um ou ambos os canais podem ser usados para determinar a cor ideal da região de inserção, e isso pode ser usado para gerar as chaves independentes para ambos os canais. A correspondência entre os pixels do plano de fundo em ambos os canais pode ser usada para suavizar o ruído da tomada ou outros objetos de oclusão em chaves de máscara para um canal de câmera particular. Alternativamente, as máscaras de oclusão para ambos os canais podem ser diretamente computadas a partir de mapas de profundidade estereoscópicos. As máscaras para ambos os canais podem ser processadas para garantir que os mesmos pixels correspondentes para cada canal sejam selecionados para mascaramento. Ter alguns pixels mascarados em um canal e não mascarados no outro pode resultar em objetos tal como os erros de cor de inserção ou outros objetos gerados pela combinação imprópria de canal esquerdo e direito. A análise visual pode ser empregada pelo bloco de oclusão integrado 509 para gerar uma chave de máscara para cada canal.

O bloco de renderização integrado 511 pode executar a renderização estereoscópica das inserções com base em um modelo composto determinado a partir dos modelos de canal esquerdo e direito. Os mecanismos de renderização gráficos podem ser usados para gerar canais esquerdo e direito simultâneos para as inserções virtuais integradas no vídeo 3D. A mistura chave de máscaras de oclusão com as chaves gráficas pode ser implementada pelo bloco de renderização integrado 511 e também possivelmente a mistura final de vídeo 3D com os canais de preenchimento 3D. Também, a mistura pode ser implementada usando um bloco misturador integrado independente 513, que pode ser compreendido de dois misturadores de vídeo de transmissão. Em algumas modalidades, a mistura pode ser implementada por um único misturador de transmissão, se os canais esquerdo e direito são entrelaçados em um formato de vídeo padrão tal como HD-SDI.

Em uma modalidade, o bloco de renderização integrado 511 pode renderizar os elementos visuais de acordo com os modelos de câmera determinados pelo bloco de busca integrado 505 e pelo bloco de acompanhamento integrado 507. Em um exemplo, o elemento visual pode ser um objeto tridimensional, e o bloco de renderização integrado 511 pode renderizar o objeto tridimensional para aparecer dentro do vídeo. Neste exemplo, o bloco de renderização integrado 511 pode renderizar as figuras tridimensionais dinâmicas/animadas usando técnicas de modelagem tridimensionais, incluindo, por exemplo, o carregamento de textura, a modelagem e renderização de câmera virtual para uma porta de visualização.

Alternativamente, o objeto 3D renderizado pode ser estático, tal como uma representação 3D da primeira linha de descida. Técnicas de renderização tridimensionais podem ser usadas, tal como aquelas em aplicações de jogos. Em outros exemplos, o elemento visual inserido na entrada de vídeo 3D 501 pode ser uma
 5 imagem, vídeo, gráfico, texto, ou propaganda (tal como uma logomarca de propaganda). Os elementos visuais podem ser gerados usando as fontes de caractere, permitindo que as inserções sejam derivadas de fontes de dados, tal como, canais de transmissão de dados de jogo ou estatísticas de posição do jogador durante os eventos esportivos. Os elementos virtuais que são combinados ou, de outra forma,
 10 misturados com o vídeo 3D podem ser considerados como sendo a inserção de um aprimoramento no vídeo 3D.

Os elementos visuais inseridos na entrada de vídeo 3D 501 podem acompanhar as cenas do plano de fundo, tal como uma inserção da primeira linha de descida virtual travada para um plano de fundo da superfície de jogo de
 15 futebol. Um elemento visual pode acompanhar um ponto de interesse, tal como anéis posicionados nos pés dos jogadores durante um evento esportivo. A parte de um gráfico pode acompanhar um ponto de interesse dentro do vídeo, tal como uma seta que aponta para um jogador em um evento esportivo onde somente a ponta da seta acompanha o ponto de localização de interesse. A inserção de gráfico 3D pode ser
 20 relativa tanto às cenas de plano de fundo e quanto aos pontos de interesse do primeiro plano, por exemplo, quando os gráficos marcam o trilha (caminho) de um jogador em movimento em uma transmissão de vídeo. Neste caso, os pontos da trilha – posição dos pés através do tempo – são primeiro inicializados com base no acompanhamento do ponto de interesse e então atualizados para compensar o movimento da câmera.

Em uma modalidade, os gráficos do sistema 3D de desenhos para comentaristas de televisão ("telestrator"), que descrevem através de diagrama os movimentos dos jogadores, por exemplo, são sobrepostos em 3D na superfície de jogo. Em outra modalidade, os gráficos do sistema de desenhos para comentaristas de televisão podem ser representados como um ajuste em uma distância configurada a
 30 partir da câmera. Isso pode funcionar bem para algumas aplicações, mas pode ser limitado em outras que poderiam beneficiar-se da informação espacial da cena. Como um exemplo, um operador do sistema de desenhos para comentaristas de televisão pode colocar círculos em torno dos jogadores com base na distância da câmera ou com base no plano da tela. Para alguns jogadores, tais círculos podem parecer
 35 envolver os jogadores, mas para outros jogadores, os círculos podem parecer flutuar acima deles. A colocação de tais círculos com base na posição 3D dentro da cena, tal

como, por exemplo, próximo dos pés dos jogadores, pode fornecer relações de perspectiva aprimoradas entre os jogadores e os círculos. Problemas similares podem se aplicar a outros gráficos, incluindo as setas que apontam para os jogadores. Por exemplo, as setas colocadas em uma distância configurada atrás de um plano de tela

5 podem não parecer próxima ou obviamente ligadas aos jogadores. Em uma modalidade específica, os gráficos do sistema 3D de desenhos para comentaristas de televisão no vídeo 3D podem ser posicionados e/ou gerados em parte com base nos comandos de usuário capturados usando uma interface manual (tela sensível ao toque, mouse, dispositivo de jogos, tablet, etc.). Os gráficos do sistema 3D de

10 desenhos para comentaristas de televisão podem ser feitos para acompanhar as cenas 3D, tal como setas que seguem os jogadores, por exemplo. O acompanhamento de cena do sistema 3D de desenhos para comentaristas de televisão é possível usando os métodos descritos aqui incluindo o uso do espaçamento ocular da câmera e dos dados do ângulo de convergência, por exemplo.

15 Em um exemplo ilustrativo, os círculos poderiam ser inseridos em torno das cinturas dos jogadores. Usando os métodos 2D, a colocação dos círculos nas cenas próximo aos pés dos jogadores poderia resultar em círculos que não estão associados com os jogadores. A colocação precisa dos círculos no espaço 3D poderia corrigir isso. Os gráficos do sistema 3D de desenhos para comentaristas de televisão, que são

20 combinados ou, de outra forma, misturados com o vídeo 3D podem ser considerados como sendo a inserção de um aprimoramento no vídeo 3D.

A FIG. 6 ilustra um canal de produção e distribuição de vídeo 3D exemplificado de acordo com uma modalidade. As inserções/aprimoramentos virtuais no vídeo 3D usando a câmera e outra informação de dados podem ser

25 fornecidos em diferentes estágios do canal de vídeo 3D, como ilustrado na FIG. 6. A competição em um local de esportes 602 pode ser coberta por múltiplas câmeras de vídeo 3D 604, e as alimentações de vídeo 3D enviadas a uma produção em sítio 606. O sistema de inserção virtual 3D pode modificar uma alimentação da câmera de vídeo 3D dedicada no sistema de inserção 3D 608 a montante de um dos comutadores de

30 canal esquerdo e direito da produção em sítio 606, por exemplo. O sistema pode modificar a alimentação do programa de vídeo 3D em sítio no sistema de inserção 3D 610 à jusante da produção em sítio 606.

A informação de dados da câmera pode ser extraída pela instrumentação da câmera 3D ou pelo sistema de câmera 3D, que pode incluir as

35 lentes, o controlador e a cabeça do tripé. A informação de dados da câmera pode ser fornecida a sistema de inserção 3D através de uma conexão de dados ou através da

codificação de informação no formato de vídeo. A informação do modelo de câmara pode ser extraída diretamente pela análise do vídeo, ou por uma combinação da análise de vídeo e dos sensores de câmara. A alimentação de vídeo 3D pode ser transmitida pela transmissão de vídeo 612 a uma localização remota tal como uma produção de estúdio de transmissão 614, onde as inserções virtuais podem ser integradas no vídeo 3D usando um sistema de inserção 3D 616. Os modelos de câmara ou os parâmetros de dados podem ser transmitidos de um estágio de produção em sítio para uma localização remota, onde os dados são recebidos e utilizados para integrar um aprimoramento no vídeo 3D. O vídeo 3D com inserções virtuais pode ser distribuído através de uma distribuição de vídeo 3D 620, onde ele pode ser fornecido às plataformas incluindo a televisão 622, a internet 624 ou os dispositivos móveis 626.

Em uma modalidade, as inserções virtuais são integradas no vídeo 3D em uma localização remota a partir de uma produção em sítio, usando a análise de vídeo de um ou ambos os canais de transmissão de vídeo. A localização pode incluir, mas não está limitada a um estúdio de transmissão, uma estação transmissora a cabo regional, uma estação transmissora a cabo local, um nó de cabo, um dispositivo de conexão à Internet via TV, um sistema de computador e um dispositivo móvel. Em outra modalidade, a análise de vídeo pode acontecer em sítio ou em uma localização remota, tal como, mas não limitada a um estúdio ou a uma estação transmissora a cabo regional. A informação pode ser propagada à jusante na cadeia de distribuição para onde a inserção é integrada (estação transmissora a cabo regional, estação transmissora a cabo local, nó de cabo, dispositivo de conexão à Internet via TV). Em ainda outra modalidade, a informação do sensor de câmara pode ser derivada para a(s) câmara(s) de vídeo 3D e enviada para uma localização remota a partir do local a ser usado por um sistema de inserção virtual para integrar os gráficos no vídeo 3D.

A FIG. 7 é um diagrama esquemático de um sistema de computador 700 exemplificado usado para implementar as modalidades para inserções em vídeo 3D. Vários aspectos das várias modalidades podem ser implementados por software, suporte lógico inalterado, hardware, ou uma combinação desses. A FIG. 7 ilustra um sistema de computador 700 exemplificado no qual uma modalidade, ou partes dessa, pode ser implementada como um código legível por computador. Várias modalidades são descritas em termos deste sistema de computador 700 exemplificado. Após a leitura desta descrição, estará claro para um versado na técnica relevante como implementar as modalidades usando outros

sistemas de computador e/ou arquiteturas de computador.

O sistema de computador 700 inclui um ou mais processadores, tal como o processador 704. O processador 704 pode ser um processador de propósito geral ou propósito específico. O processador 704 é
5 conectado a uma infraestrutura de comunicação 706 (por exemplo, um barramento ou uma rede).

O sistema de computador 700 também inclui uma memória principal 708, preferencialmente uma memória de acesso aleatório (RAM), e pode também incluir uma memória secundária 710. A memória 710 pode incluir, por
10 exemplo, uma unidade de disco rígido 712 e/ou uma unidade de armazenamento removível 714. A unidade de armazenamento removível 714 pode compreender uma unidade de disquete, uma unidade de fita magnética, uma unidade de disco óptico, uma unidade de memória rápida, ou similares. A unidade de armazenamento removível 714 lê a partir e/ou grava em uma unidade de armazenamento removível
15 718 de uma forma bem conhecida. A unidade de armazenamento removível 718 pode compreender um disquete, uma fita magnética, um disco óptico, etc., que é lido e/ou gravado pela unidade de armazenamento removível 714. Como estará claro para os versados na técnica relevante, a unidade de armazenamento removível 718 inclui um meio de armazenamento legível por computador tangível que armazena software e/ou
20 dados de computador.

Em implementações alternativas, a memória secundária 710 pode incluir outros dispositivos similares que permitem que os programas de computador ou outras instruções sejam carregadas no sistema de computador 700. Tais dispositivos incluem, por exemplo, uma unidade de armazenamento removível
25 722 e uma interface 720. Exemplos de tais dispositivos podem incluir um cartucho de programa e uma interface de cartucho (tal como a encontrada nos dispositivos de vídeo game), um chip de memória removível (tal como um EPROM, ou PROM) e tomada associada, e outras unidades de armazenamento removível 722 e interfaces 720 que permitem que o software e os dados sejam transferidos da unidade de
30 armazenamento removível 722 para o sistema de computador 700.

O sistema de computador 700 pode também incluir uma interface de comunicação 724. A interface de comunicação 724 permite que o software e os dados sejam transferidos entre o sistema de computador 700 e os dispositivos externos. A interface de comunicação 724 pode incluir um modem, uma interface de
35 rede (por exemplo, uma placa Ethernet), uma porta de comunicações, uma fenda e uma placa PCMCIA, ou similares. O software e os dados transferidos via a interface de

comunicação 724 são fornecidos à própria interface de comunicação 724 via o caminho de comunicação 726. O caminho de comunicação 726 pode ser implementado usando fio ou cabo, fibras ópticas, linha telefônica, ligação de telefone celular, ligação RF ou outros canais de comunicação.

5 Neste documento, os termos “meio de programa de computador” e “meio utilizado por computador” são usados para se referir geralmente a meios tal como a unidade de armazenamento removível 718, a unidade de armazenamento removível 722 e um disco rígido instalado na unidade de disco rígido 712. O meio de programa de computador e o meio utilizado por computador podem
10 ser referir a memórias, tais como, a memória principal 708 e a memória secundária 710, que podem ser semicondutores de memória (por exemplo, DRAMs, etc.) Esses produtos de programa de computador são dispositivos que fornecem software ao sistema de computador 700.

Os programas de computador (também chamados de lógica de
15 controle de computador) são armazenados na memória principal 708 e/ou na memória secundária 710. Os programas de computador podem também ser recebidos via a interface de comunicações 724. Tais programas de computador, quando executados, possibilitam que o sistema de computador 700 implemente as modalidades discutidas aqui, tal como o sistema descrito acima. Em particular, os programas de computador,
20 quando executados, possibilitam que o processador 704 implemente os processos de modalidades. Consequentemente, tais programas de computador representam os controladores do sistema de computador 700. Quando as modalidades são implementadas usando software, o mesmo pode ser armazenado em um produto de programa de computador e carregado no sistema de computador 700 usando a
25 unidade de armazenamento removível 714, a interface 720, a unidade de disco rígido 712 ou a interface de comunicações 724.

São descritos acima os sistemas, aparelhos e métodos para inserções em vídeo 3D e suas aplicações. Estará claro que a seção Descrição
30 Detalhada, e não o Resumo, é destinada a ser usada para interpretar as reivindicações. O resumo pode apresentar uma ou mais, mas não todas as modalidades exemplificadas da presente invenção como observado pelos inventores, e assim, não são destinadas a limitar a presente invenção e as reivindicações em anexo de qualquer forma.

As modalidades foram descritas acima com o auxílio de blocos
35 de construção funcionais que ilustram a implementação de funções específicas e relações destas. Os limites desses blocos de construção funcionais foram definidos

arbitrariamente aqui para conveniência da descrição. Os limites alternativos podem ser definidos contanto que as funções e relações específicas destas são executadas de forma apropriada.

5 A descrição anterior das modalidades específicas revelará completamente a natureza geral da invenção que outros podem, aplicando o conhecimento da técnica, modificar e/ou adaptar prontamente para as várias aplicações de tais modalidades específicas, sem experimentação indevida, sem abandonar o conceito geral da presente invenção. Então, pretende-se que tais adaptações e modificações estejam dentro do ensinamento e da faixa de equivalentes das modalidades descritas, com base no ensinamento e orientação apresentados aqui. Entende-se que a fraseologia ou a terminologia apresentada aqui é usada com o propósito de descrição e não de limitação, de modo que a terminologia ou fraseologia da presente especificação seja interpretada pelos versados na técnica face aos ensinamentos e orientações.

15 A extensão e o escopo da presente invenção não deveriam ser limitados por qualquer uma das modalidades exemplificadas descritas acima, mas deveriam ser definidos somente de acordo com as reivindicações e seus equivalentes.

REIVINDICAÇÕES

1. Método **caracterizado** por compreender:

- determinar um primeiro parâmetro de dados da câmera de um primeiro modelo de câmera associado com um primeiro canal de um vídeo 3D, em que o primeiro modelo de câmera descreve o campo de visão do primeiro canal;

- determinar um segundo parâmetro de dados da câmera de um segundo modelo de câmera associado com um segundo canal do vídeo 3D, em que o segundo modelo de câmera descreve o campo de visão do segundo canal;

- em que a determinação do primeiro parâmetro de dados da câmera e do segundo parâmetro de dados da câmera é baseada em uma análise de pesquisa de pelo menos o primeiro canal, a análise de pesquisa baseada em voxels correspondendo a pelo menos o primeiro canal;

- gerar um modelo de câmera composto reconciliando o primeiro parâmetro de dados da câmera do primeiro modelo e o segundo parâmetro de dados da câmera do segundo modelo de câmera; e

- inserir um aprimoramento no vídeo 3D com base no modelo de câmera composto.

2. Método, de acordo com a reivindicação 1, **caracterizado** por compreender adicionalmente reconciliar o primeiro parâmetro da câmera e o segundo parâmetro da câmera.

3. Método, de acordo com a reivindicação 1, **caracterizado** por determinar o primeiro o parâmetro de dados da câmera e o segundo parâmetro de dados da câmera é ainda com base na análise visual de ao menos um dentre o primeiro e o segundo canal de vídeo 3D.

4. Método, de acordo com a reivindicação 1, **caracterizado** pelos parâmetros de dados da câmera incluírem o espaçamento ocular e o ângulo de convergência.

5. Método, de acordo com a reivindicação 1,

caracterizado por adicionalmente compreender calibrar automaticamente um sistema de câmera 3D associado com o vídeo 3D, com base nos parâmetros de dados da câmera.

6. Método, de acordo com a reivindicação 1, **caracterizado** por o primeiro canal de vídeo 3D e o segundo canal de vídeo 3D são obtidos a partir de um sistema de câmera 3D associado com o vídeo 3D.

7. Método, de acordo com a reivindicação 6, **caracterizado** por compreender restringir a análise de busca pesquisa com base em uma relação entre o primeiro e segundo canais de vídeo 3D.

8. Método, de acordo com a reivindicação 1, **caracterizado** por adicionalmente compreender atualizar o primeiro e o segundo parâmetro de dados da câmera com base em ao menos a análise de acompanhamento do primeiro canal.

9. Método, de acordo com a reivindicação 1, **caracterizado** por adicionalmente que compreender bloquear o aprimoramento baseado no modelo de câmera composto.

10. Método, de acordo com a reivindicação 1, **caracterizado** por adicionalmente compreender posicionar de forma interativa o aprimoramento em uma localização 3D de acordo com a entrada recebida.

11. Método, de acordo com a reivindicação 1, **caracterizado** pelo aprimoramento ser um elemento visual tridimensional renderizado.

12. Método, de acordo com a reivindicação 1, **caracterizado** por adicionalmente compreender posicionar automaticamente o aprimoramento em uma localização 3D de acordo com uma composição de cena do vídeo 3D e o tipo do aprimoramento.

13. Método **caracterizado** por compreender:

- receber um primeiro parâmetro de dados da câmera associado com um primeiro modelo de câmera, o primeiro modelo de câmera está associado a um primeiro canal de um vídeo 3D e descreve o campo de

visão do primeiro canal;

- receber um segundo parâmetro de dados de câmera associado a um segundo modelo de câmera, o segundo modelo de câmera é associado a um segundo canal do vídeo 3D e descreve o campo de visão do segundo canal;

- em que a determinação do primeiro parâmetro de dados da câmera e do segundo parâmetro de dados da câmera é baseada em uma análise de pesquisa de pelo menos o primeiro canal, a análise de pesquisa baseada em voxels correspondendo a pelo menos o primeiro canal;

- gerar um modelo de câmera composto com base em pelo menos o primeiro parâmetro de dados da câmera e o segundo parâmetro de dados da câmera; e

- inserir um aprimoramento no vídeo 3D com base no modelo de câmera composta, em que a inserção é realizada remotamente a partir de um estágio de tubulação de produção no local.

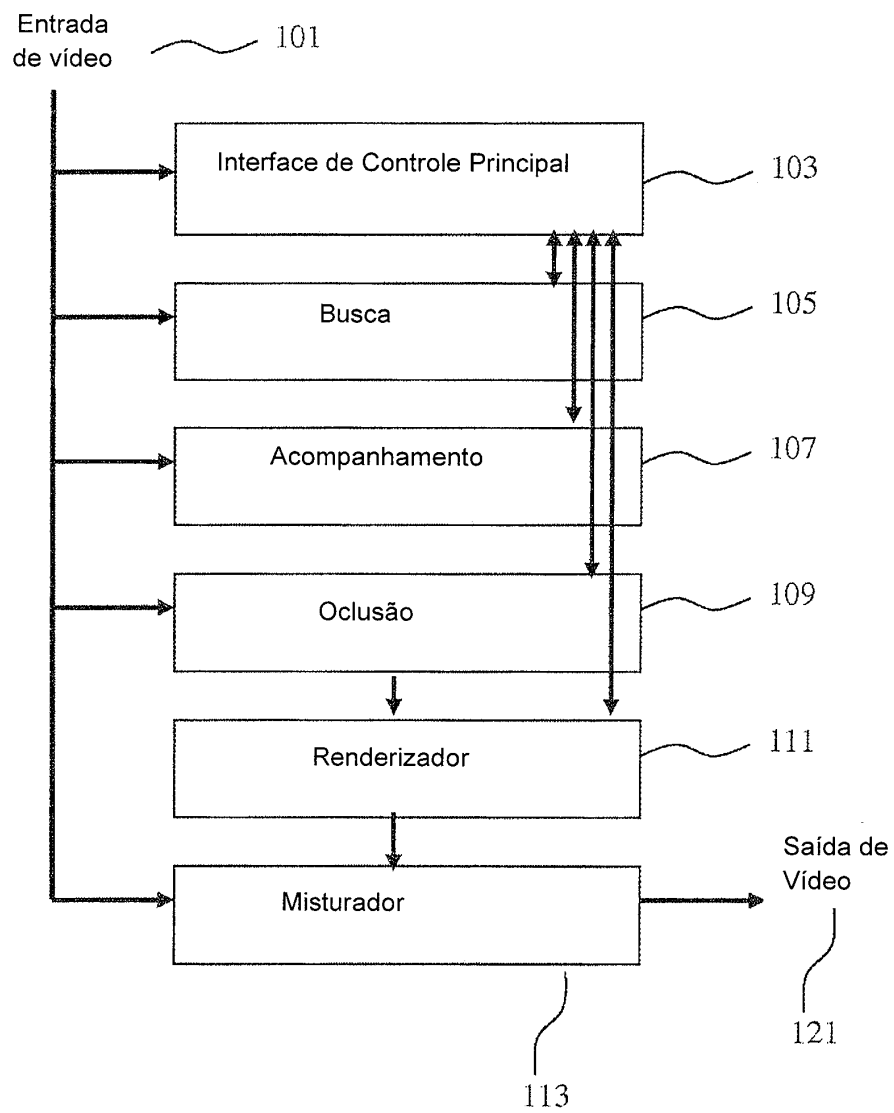


FIG. 1

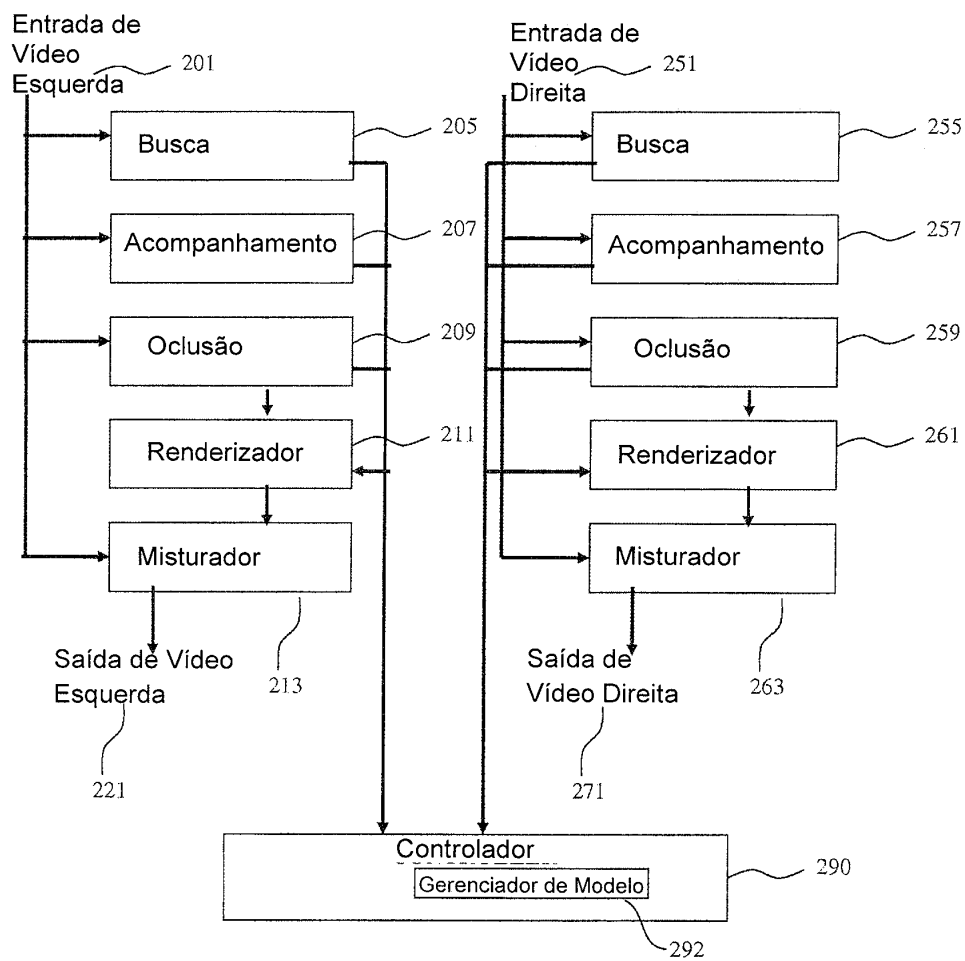


FIG. 2

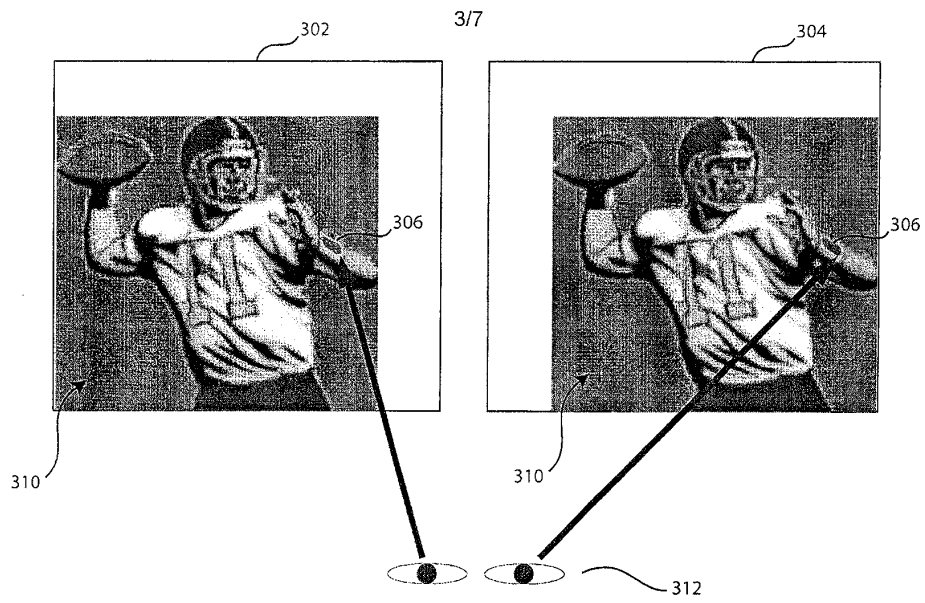


FIG. 3A

Campo de jogo verde atrás da faixa de pulso	310
---	-----

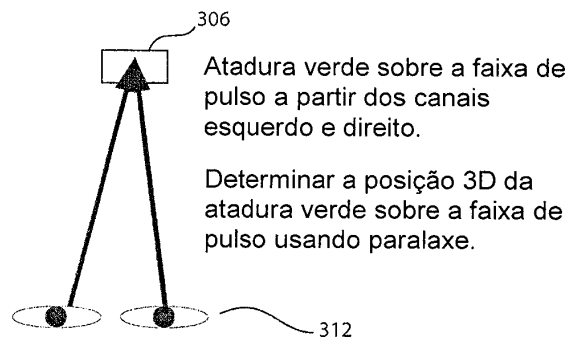


FIG. 3B

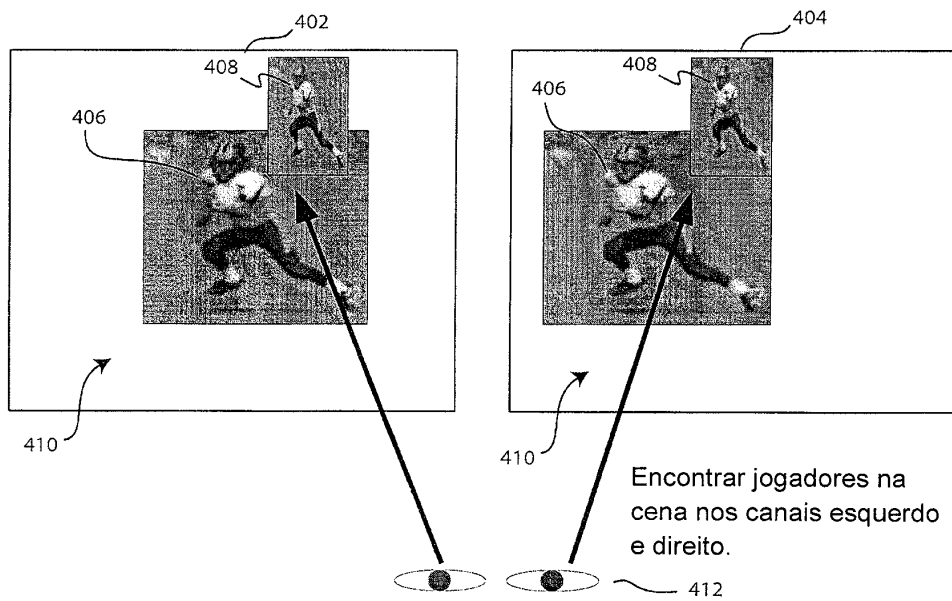


FIG. 4A

Campo de jogo verde atrás dos jogadores 410

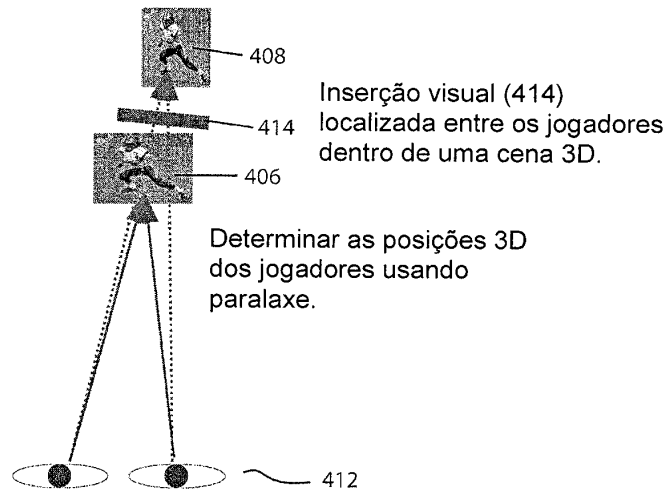
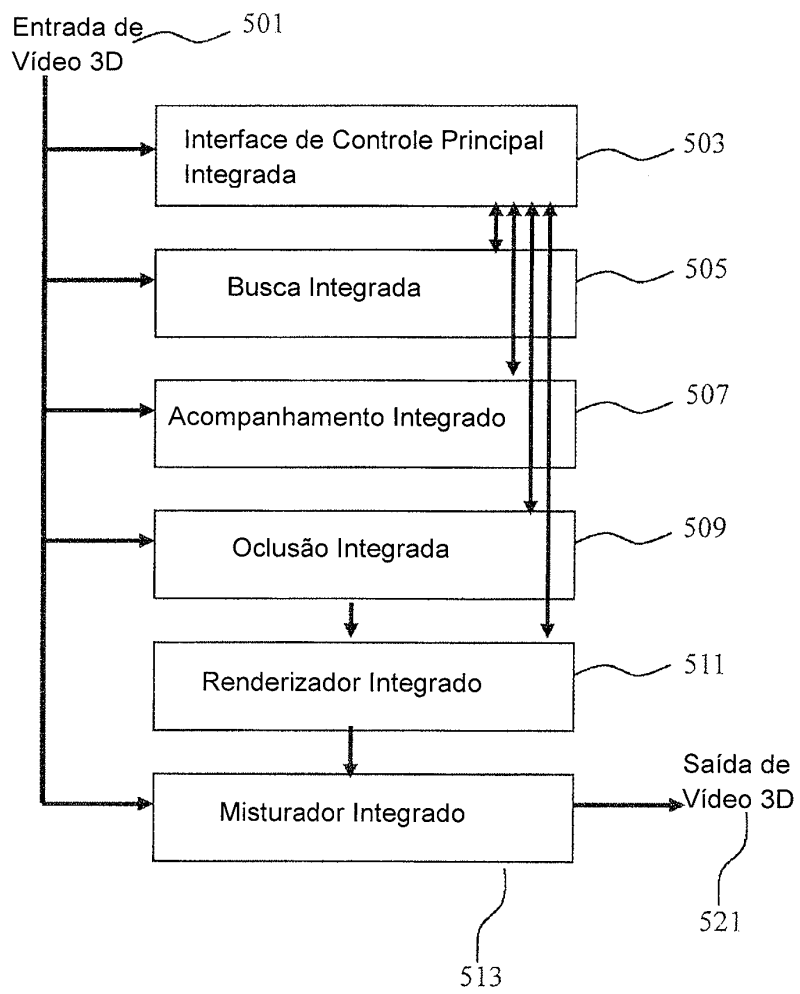


FIG. 4B

**FIG. 5**

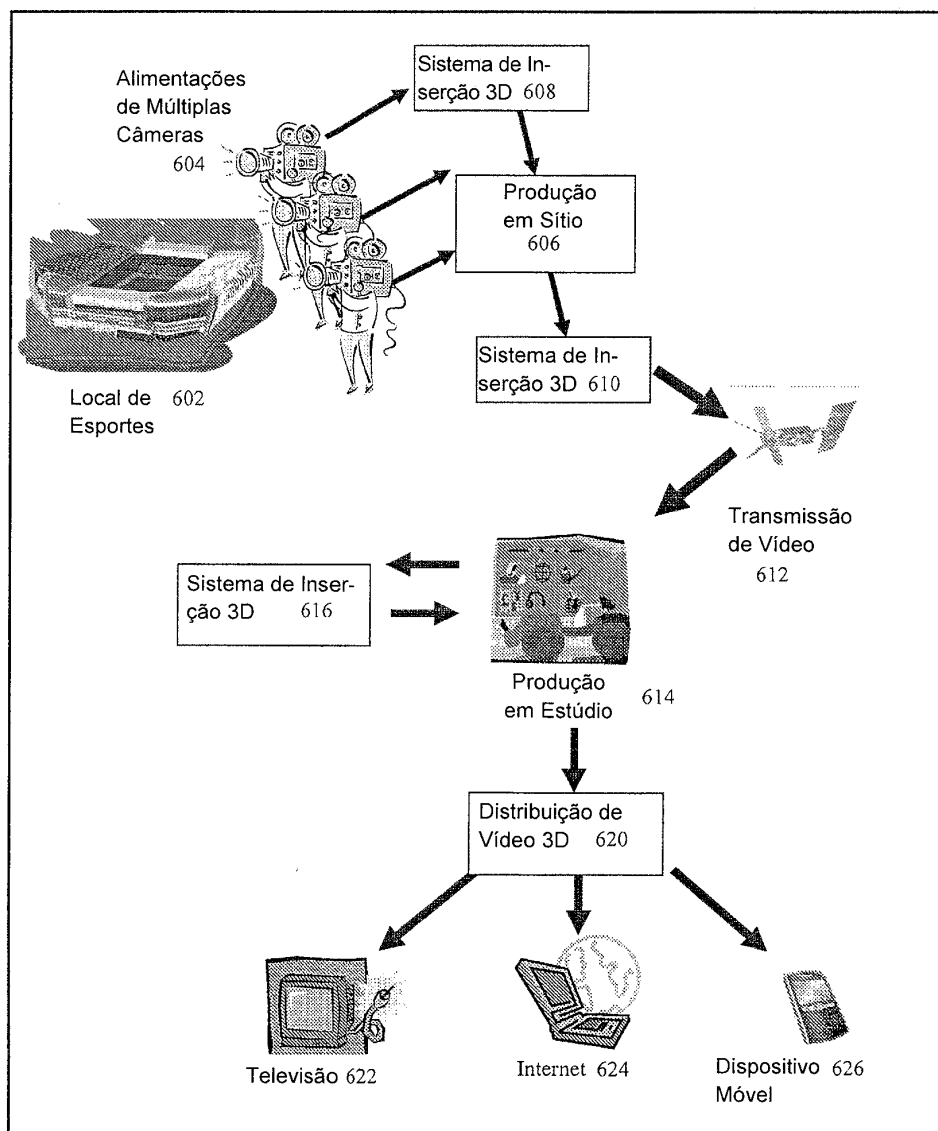


FIG. 6

