



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

(11) PI 1011212-0 B1



(22) Data do Depósito: 29/05/2010

(45) Data de Concessão: 05/05/2020

(54) Título: MÉTODO PARA USAR UM ATALHO DO GESTO EM UM SISTEMA QUE UTILIZA GESTOS DO USUÁRIO COMO ENTRADA PARA UMA APLICAÇÃO, MEIO DE ARMAZENAMENTO LEGÍVEL POR COMPUTADOR E SISTEMA PARA USAR UM ATALHO DO GESTO EM UM SISTEMA PARA USAR GESTOS DO USUÁRIO COMO ENTRADA PARA UMA APLICAÇÃO

(51) Int.Cl.: G06F 3/01; G06F 3/03.

(30) Prioridade Unionista: 29/05/2009 US 12/474.781.

(73) Titular(es): MICROSOFT TECHNOLOGY LICENSING, LLC.

(72) Inventor(es): STEPHEN LATTA; KEVIN GEISNER; JOHN CLAVIN; KUDO TSUNODA; KATHRYN STONE PEREZ; RELJA MARKOVIC; GREGORY N. SNOOK; ALEX KIPMAN.

(86) Pedido PCT: PCT US2010036774 de 29/05/2010

(87) Publicação PCT: WO 2010/138952 de 02/12/2010

(85) Data do Início da Fase Nacional: 17/11/2011

(57) Resumo: MÉTODO PARA USAR UM ATALHO DO GESTO EM UM SISTEMA QUE UTILIZA GESTOS DO USUÁRIO COMO ENTRADA PARA UMA APLICAÇÃO, MEIO DE ARMAZENAMENTO LEGÍVEL POR COMPUTADOR E SISTEMA PARA USAR UM ATALHO DO GESTO EM UM SISTEMA PARA USAR GESTOS DO USUÁRIO COMO ENTRADA PARA UMA APLICAÇÃO. A presente invenção refere-se a sistemas, métodos e meios legíveis por computador para atalhos para gesto. O movimento de um usuário ou a posição do corpo é capturado por um dispositivo de captura de um sistema, e é utilizado como entrada para controlar o sistema. Para um gesto reconhecido pelo sistema, pode existir uma versão completa do gesto e um atalho do gesto. Onde o sistema reconhece que a versão completa do gesto ou o atalho do gesto foi executado, ele envia uma indicação que o gesto reconhecido pelo sistema foi observado para uma aplicação correspondente. Onde o atalho compreende um subconjunto da versão completa do gesto, e tanto o atalho como a versão completa do gesto são reconhecidos à medida que o usuário executa a versão completa do gesto, o sistema reconhece que somente uma única execução do gesto ocorreu, e indica para a aplicação como tal.

Relatório Descritivo da Patente de Invenção para "MÉTODO PARA USAR UM ATALHO DO GESTO EM UM SISTEMA QUE UTILIZA GESTOS DO USUÁRIO COMO ENTRADA PARA UMA APLICAÇÃO, MEIO DE ARMAZENAMENTO LEGÍVEL POR COMPUTADOR E SISTEMA PARA USAR UM ATALHO DO GESTO EM UM SISTEMA PARA USAR GESTOS DO USUÁRIO COMO ENTRADA PARA UMA APLICAÇÃO"

ANTECEDENTES

[0001] Várias aplicações de computação, tais como jogos de comutador, aplicações multimídia, aplicações de escritório e assim por diante, utilizam controles para permitir aos usuários manipularem os personagens do jogo ou outros aspectos de uma aplicação. Tipicamente, tais controles são informados utilizando, por exemplo, controladores, controles remotos, teclados, mouses e assim por diante. Infelizmente, tais controles podem ser difíceis de aprender, assim criando uma barreira entre um usuário e tais jogos e aplicações. Adicionalmente, tais controles podem ser diferentes das ações reais do jogo ou das ações de outra aplicação para as quais os controles são utilizados. Por exemplo, um controle de jogo que causa que um personagem do jogo move um bastão de beisebol pode não corresponder a um movimento real de mover o bastão de beisebol.

SUMÁRIO DA INVENÇÃO

[0002] Neste documento são revelados sistemas e métodos para atalhos para gesto.

[0003] Em uma concretização, um dispositivo de computação recebe uma série de dados de imagem a partir de uma câmera. Esta câmera pode compreender uma câmera colorida (tal como uma vermelho – verde - azul ou RGB), uma câmera com sensor de profundidade, e uma câmera tridimensional (3D). Estes dados podem compreender imagens com percepção de profundidade e imagem

coloridas separadas, uma imagem combinada que incorpora informações de profundidade e de cor, ou uma imagem dividida, onde os objetos são identificados, tais como pessoas que têm seu esqueleto mapeado. Estes dados capturam movimentos ou posturas feitas por pelo menos um usuário. Estes movimentos ou posturas podem ser reconhecidos pelo dispositivo de computação como um tipo de entrada – informação sobre gesto. Para um dado gesto (por exemplo, navegar para cima), pode existir uma versão completa do gesto que o usuário pode fazer e um atalho do gesto que o usuário pode fazer, o atalho do gesto geralmente requerendo menos tempo, movimento, ou dificuldade de movimento para o usuário.

[0004] Onde o dispositivo de computação reconhece que o atalho do gesto ou a versão completa do gesto foi executado pelo usuário, ele envia uma indicação disto para uma aplicação que utiliza os gestos como entrada.

[0005] Em uma concretização, onde o atalho compreende um subconjunto da versão completa do gesto, e tanto o atalho como a versão completa do gesto são reconhecidos à medida que o usuário executa a versão completa do gesto, o dispositivo de computação reconhece que somente uma única execução do gesto ocorreu, e indica para a aplicação como tal.

[0006] O precedente é um sumário e assim contém, por necessidade, simplificações, generalizações e omissões dos detalhes. Pode ser apreciado que o sumário é somente ilustrativo e não é pretendido para de qualquer modo ser limitação.

BREVE DESCRIÇÃO DOS DESENHOS

[0007] Os sistemas, métodos e meios legíveis por computador para atalhos para gesto de acordo com este relatório descritivo são adicionalmente descritos com referência aos desenhos acompanhantes, nos quais:

[0008] As FIGs. 1A e 1B ilustram uma concretização ilustrativa de um sistema de reconhecimento, análise e rastreamento de alvo com um usuário jogando um jogo.

[0009] A FIG. 2 ilustra uma concretização ilustrativa de um dispositivo de captura que pode ser utilizado em um sistema de reconhecimento, análise e rastreamento de alvo.

[00010] A FIG. 3A ilustra uma concretização ilustrativa de um ambiente de computação que pode ser utilizado para interpretar um ou mais gestos em um sistema de reconhecimento, análise e rastreamento de alvo.

[00011] A FIG. 3B ilustra outra concretização ilustrativa de um ambiente de computação que pode ser utilizado para interpretar um ou mais gestos em um sistema de reconhecimento, análise e rastreamento de alvo.

[00012] A FIG. 4A ilustra um mapeamento de esqueleto de um usuário que foi gerado a partir do sistema de reconhecimento, análise e rastreamento de alvo da FIG. 2.

[00013] A FIG. 4B ilustra detalhes adicionais da arquitetura de reconhecimento de gesto apresentada na FIG. 2.

[00014] As FIGs. 5A e 5B ilustram como os filtros de gesto podem ser empilhados para criarem filtros de gesto mais complexos.

[00015] As FIGs. 6A, 6B, 6C, 6D e 6E ilustram um gesto exemplo que um usuário pode fazer para sinalizar um gesto de “fair catch” (“sinal de mão feito sobre a cabeça, sinalizando que não tentará o retorno, terminando ali a jogada e determinando onde se iniciará a próxima campanha ofensiva de sua equipe”) em um jogo de vídeo de futebol americano.

[00016] As FIGs. 7A, 7B, 7C, 7D e 7E ilustram o gesto de “fair catch” ilustrativo das FIGs. 6A até 6E à medida que cada quadro de dados de imagem foi analisado para produzir um mapa do esqueleto do usuário.

[00017] A FIG. 8A ilustra um usuário fazendo um gesto completo de correr.

[00018] A FIG. 8B ilustra um usuário fazendo um gesto de atalho de correr, o gesto de atalho compreendendo um subconjunto do movimento do gesto de correr completo da FIG. 8A.

[00019] A FIG. 8C ilustra um usuário fazendo um segundo tipo de gesto de atalho de correr, o segundo tipo de gesto de atalho de correr compreendendo movimento separado do gesto completo de correr da FIG. 8A.

[00020] A FIG. 9 representa procedimentos operacionais ilustrativos para os atalhos de gesto.

DESCRÍÇÃO DETALHADA DAS CONCRETIZAÇÕES ILUSTRATIVAS

[00021] Como será descrito neste documento, um usuário pode controlar uma aplicação executando em um ambiente de computação tal como um console de jogo, um computador, ou coisa parecida, por executar um ou mais gestos. De acordo com uma concretização, os gestos podem ser recebidos, por exemplo, por um dispositivo de captura. Por exemplo, o dispositivo de captura pode capturar uma imagem com percepção de profundidade de uma cena. Em uma concretização, o dispositivo de captura pode determinar se um ou mais alvos ou objetos na cena correspondem a um alvo humano tal como o usuário. Para determinar se um alvo ou objeto na cena corresponde a um alvo humano, cada um dos alvos pode ser preenchido por inundação e comparado com um padrão de um modelo de corpo humano. Cada alvo ou objeto que se associa com o modelo do corpo humano pode então ser lido oticamente para gerar um modelo de esqueleto associado com o mesmo. O modelo de esqueleto pode então ser proporcionado para o ambiente de computação de modo que o ambiente de computação possa rastrear o modelo de esqueleto, produzir uma representação on-line de uma pessoa (“avatar”) associada com o

modelo de esqueleto, e pode determinar quais controles executar para executar em uma aplicação executando no ambiente de computador baseado, por exemplo, nos gestos do usuário que foram reconhecidos a partir do modelo de esqueleto. Um mecanismo de reconhecimento de gesto, cuja arquitetura é descrita em maiores detalhes abaixo, é utilizado para determinar quando um gesto particular foi feito pelo usuário.

[00022] As FIGs. 1A e 1B ilustram uma concretização ilustrativa de uma configuração de um sistema de reconhecimento, análise e rastreamento de alvo 10 com um usuário 18 jogando um jogo de luta de boxe. Em uma concretização ilustrativa, o sistema de reconhecimento, análise e rastreamento de alvo 10 pode ser utilizado para reconhecer, analisar, e/ou rastrear um alvo humano tal como o usuário 18.

[00023] Como apresentado na FIG. 1A, o sistema de reconhecimento, análise e rastreamento 10 pode incluir um ambiente de computação 12. O ambiente de computação 12 pode ser um computador, um sistema de jogo ou console, ou coisa parecida. De acordo com uma concretização ilustrativa, o ambiente de computação 12 pode incluir componentes de hardware e/ou componentes de software de modo que o ambiente de computação 12 pode ser utilizado para executar aplicações tais como aplicações de jogos, aplicações que não são de jogos, e assim por diante.

[00024] Como apresentado na FIG. 1A, o sistema de reconhecimento, análise e rastreamento de alvo 10 pode adicionalmente incluir um dispositivo de captura 20. O dispositivo de captura 20 pode ser, por exemplo, uma câmera que pode ser utilizada para visualmente monitorar um ou mais usuários, tal como o usuário 18, de modo que os gestos executados pelo um ou mais usuários possam ser capturados, analisados e rastreados para executar um ou mais controles ou ações dentro de uma aplicação, como será descrito abaixo

em maiores detalhes.

[00025] De acordo com uma concretização, o sistema de reconhecimento, análise e rastreamento de alvo 10 pode ser conectado com um dispositivo audiovisual 16 tal como uma televisão, um monitor, uma televisão de alta definição (HDTV), e assim por diante, que possa proporcionar jogo ou aplicações visuais e/ou áudio para um usuário tal como o usuário 18. Por exemplo, o ambiente de computação 12 pode incluir um adaptador de vídeo tal como uma placa gráfica e/ou um adaptador de áudio tal como uma placa de som que pode proporcionar sinais audiovisuais associados com a aplicação de jogo, com a aplicação que não é de jogo, e assim por diante. O dispositivo audiovisual 16 pode receber os sinais audiovisuais a partir do ambiente de computação 12 e então pode emitir os elementos visuais e/ou de áudio do jogo ou da aplicação associados com os sinais áudio visuais para o usuário 18. De acordo com uma concretização, o dispositivo audiovisual 16 pode ser conectado com o ambiente de computação 12 via, por exemplo, um cabo S-Video, um cabo coaxial, um cabo HDMI, um cabo DVI, um cabo VGA ou coisa parecida.

[00026] Como apresentado nas FIGs. 1A e 1B, o sistema de reconhecimento, análise e rastreamento de alvo 10 pode ser utilizado para conhecer, analisar e/ou rastrear um alvo humano tal como o usuário 18. Por exemplo, o usuário 18 pode ser rastreado utilizando o dispositivo de captura 20 de modo que os movimentos do usuário 18 possam ser interpretados como controles que podem ser utilizados para afetar a aplicação sendo executada pelo ambiente de computador 12. Assim, de acordo com uma concretização, o usuário 18 pode mover seu corpo para controlar a aplicação.

[00027] Como apresentado nas FIGs. 1A e 1B, em uma concretização ilustrativa, a aplicação executando no ambiente de computação 12 pode ser um jogo de boxe que o usuário 18 está

jogando. Por exemplo, o ambiente de computação 12 pode utilizar o dispositivo audiovisual 16 para proporcionar uma representação visual de um oponente de luta de boxe 22 para o usuário 18. O ambiente de computação 12 também pode utilizar um dispositivo audiovisual 16 para proporcionar uma representação visual de um avatar do usuário 24 que o usuário 18 pode controlar com seus movimentos. Por exemplo, como apresentado na FIG. 1B, o usuário 18 pode dar um soco no espaço físico para causar que um avatar do jogador 24 dê um soco no espaço do jogo. Assim, de acordo com uma concretização ilustrativa, o ambiente de computador 12 e o dispositivo de captura 20 do sistema de reconhecimento, análise e rastreamento de alvo 10 podem ser utilizados para reconhecer e analisar o soco do usuário 18 no espaço físico de modo que o soco possa ser interpretado como um controle de jogo do avatar do jogador 24 no espaço do jogo.

[00028] Outros movimentos pelo usuário 18 também podem ser interpretados como outros controles ou ações, tais como controles para balançar, avançar dando “voltas”, arrastar os pés, bloquear, esmurrar, ou desferir socos com diferentes potências. Adicionalmente, alguns movimentos podem ser interpretados como controles que podem corresponder a ações diferentes de controlar o avatar do jogador 24. Por exemplo, o jogador pode utilizar movimentos para terminar, pausar, ou salvar um jogo, selecionar um nível, ver os maiores placares, se comunicar com um amigo, etc.

[00029] Nas concretizações ilustrativas, o alvo humano tal como o usuário 18 pode possuir um objeto. Em tais concretizações, o usuário de um jogo eletrônico pode estar segurando o objeto de modo que os movimentos do jogador e do objeto podem ser utilizados para ajustar e/ou controlar os parâmetros do jogo. Por exemplo, o movimento de um jogador segurando uma raquete pode ser rastreado e utilizado para controlar uma raquete na tela em um jogo eletrônico de esporte. Em

outra concretização ilustrativa, o movimento de um jogador segurando um objeto pode ser rastreado e utilizado para controlar uma arma na tela em um jogo eletrônico de combate.

[00030] De acordo com outras concretizações ilustrativas, o sistema de reconhecimento, análise e rastreamento de alvo 10 pode ser adicionalmente utilizado para interpretar os movimentos do alvo como controles do sistema operacional e/ou da aplicação que estão fora do âmbito do jogo. Por exemplo, virtualmente, qualquer aspecto controlável de um sistema operacional e/ou aplicação pode ser controlado pelos movimentos do alvo, tal como o usuário 18.

[00031] A FIG. 2 ilustra uma concretização ilustrativa do dispositivo de captura 20 que pode ser utilizado no sistema de reconhecimento, análise e rastreamento de alvo 10. De acordo com uma concretização ilustrativa, o dispositivo de captura 20 pode ser configurado para capturar vídeo com a informação de profundidade incluindo uma imagem com percepção de profundidade que pode incluir valores de profundidade via qualquer técnica adequada incluindo, por exemplo, tempo de trânsito, luz estruturada, imagem estéreo, ou coisa parecida. De acordo com uma concretização, o dispositivo de captura 20 pode organizar a informação de profundidade calculada em “camadas Z”, ou em camadas que podem ser perpendiculares a um eixo geométrico Z se estendendo a partir da câmara com sensor de profundidade ao longo de sua linha de visão.

[00032] Como apresentado na FIG. 2, o dispositivo de captura 20 pode incluir um componente de câmera de imagem 22. De acordo com uma concretização ilustrativa, o componente de câmera de imagem 22 pode ser uma câmera com sensor de profundidade que pode capturar a imagem com percepção de profundidade de uma cena. A imagem com percepção de profundidade pode incluir uma área de pixel bidimensional (2-D) da cena capturada onde cada pixel na área de pixel 2-D pode

representar um comprimento, por exemplo, em centímetros, milímetros, ou coisa parecida, de um objeto na cena capturada a partir da câmera.

[00033] Como apresentado na FIG. 2, de acordo com uma concretização ilustrativa, o componente de câmera de imagem 22 pode incluir um componente de luz IR 24, uma câmera tridimensional (3-D), e uma câmera RGB 28 que podem ser utilizados para capturar a imagem com percepção de profundidade de uma cena. Por exemplo, na análise de tempo de trânsito, o componente de luz IR 24 do dispositivo de captura 20 pode emitir uma luz infravermelha sobre a cena e então pode utilizar sensores (não apresentado) para detectar a luz retrodispersa a partir da superfície de um ou mais alvos e objetos na cena utilizando, por exemplo, a câmera 3-D 26 e/ou a câmera RGB 28. Em algumas concretizações, luz infravermelha pulsada pode ser utilizada de modo que o tempo entre o pulso de luz de saída e um pulso de luz entrante correspondente possa ser medido e utilizada para determinar uma distância física a partir do dispositivo de captura 20 até uma localização particular nos alvos ou nos objetos na cena. Adicionalmente, em outras concretizações ilustrativas, a fase da onda de luz de saída pode ser comparada com a fase da onda de luz entrante para determinar um deslocamento de fase. O deslocamento de fase pode então ser utilizado para determinar uma distância física a partir do dispositivo de captura até uma localização particular nos alvos ou objetos.

[00034] De acordo com outra concretização ilustrativa, a análise de tempo de trânsito pode ser utilizada para indiretamente determinar uma distância física a partir do dispositivo de captura 20 até uma localização particular nos alvos ou objetos pela análise da intensidade do feixe de luz refletido através do tempo via várias técnicas, incluindo, por exemplo, a representação de imagem de pulso de luz fechado.

[00035] Em outra concretização ilustrativa, o dispositivo de captura 20 pode utilizar uma luz estruturada para capturar a informação de

profundidade. Em tal análise, a luz padronizada (isto é, exibida como uma padrão conhecido tal como um padrão de grade ou um padrão de listra) pode ser projetada sobre a cena, por exemplo, via o componente de luz IR 24. Quando da chegada na superfície de um ou mais alvos ou objetos na cena, o padrão pode se tornar deformado em resposta. Tal deformação do padrão pode ser capturado, por exemplo, pela câmera 3-D 26 e/ou pela câmera RGB 28 e então podem ser analisada para determinar uma distância física a partir do dispositivo de captura até uma localização particular nos alvos ou objetos.

[00036] De acordo com outra concretização, o dispositivo de captura 20 pode incluir duas ou mais câmeras fisicamente separadas que podem ver uma cena a partir de diferentes ângulos, para obterem dados visuais estéreo que podem ser resolvidos para gerar informação de profundidade.

[00037] O dispositivo de captura 20 pode adicionalmente incluir um microfone 30. O microfone 30 pode incluir um transdutor ou sensor que pode receber e converter som em um sinal elétrico. De acordo com uma concretização, o microfone 30 pode ser utilizado para reduzir realimentação entre o dispositivo de captura 20 e o ambiente de computação 12 no sistema de reconhecimento, análise e rastreamento de alvo 10. Adicionalmente, o microfone 30 pode ser utilizado para receber sinais de áudio que também podem ser proporcionados pelo usuário para controlar as aplicações, tais como aplicações de jogo, aplicações que não são de jogo, ou coisa parecida, que podem ser executadas pelo ambiente de computação 12.

[00038] Na concretização ilustrativa, o dispositivo de captura 20 adicionalmente pode incluir um processador 32 que pode estar em comunicação operativa com o componente de câmera de imagem 22. O processador 32 pode incluir um processador padronizado, um processador especializado, um microprocessador, e assim por diante,

que pode executar instruções que podem incluir instruções para receber a imagem com percepção de profundidade, determinar se um alvo adequado pode estar incluído na imagem com percepção de profundidade, converter o alvo adequado para uma representação ou modelo de esqueleto do alvo, ou qualquer outra instrução adequada.

[00039] O dispositivo de captura 20 adicionalmente pode incluir um componente de memória 34 que pode armazenar as instruções que podem ser executadas pelo processador 32, imagens ou quadros de imagens capturados pela câmera 3-D ou câmera RGB, ou qualquer outra informação adequada, imagens ou coisa parecida. De acordo com uma concretização ilustrativa, o componente de memória 34 pode incluir memória de acesso aleatório (RAM), memória somente para leitura (ROM), cache, memória Flash, um disco rígido, ou qualquer componente de armazenamento adequado. Como apresentado na FIG. 2, em uma concretização, o componente de memória 34 pode ser um componente separado em comunicação com o componente de captura de imagem 22 e com o processador 32. De acordo com outra concretização, o componente de memória 34 pode ser integrado no processador 32 e/ou no componente de captura de imagem 22.

[00040] Como apresentado na FIG. 2, o dispositivo de captura de imagem 20 pode estar em comunicação com o ambiente de computação 12 via uma ligação de comunicação 36. A ligação de comunicação 36 pode ser uma conexão com fios incluindo, por exemplo, uma conexão USB, uma conexão Firewire, uma conexão a cabo Ethernet, e assim por diante, e/ou uma conexão sem uso de fios tal como uma conexão sem uso de fios 802.11b, g, a, ou n. De acordo com uma concretização, o ambiente de computação 12 pode proporcionar um relógio para o dispositivo de captura 20 que pode ser utilizado para determinar quando capturar, por exemplo, uma cena via a ligação de comunicação 36.

[00041] Adicionalmente, o dispositivo de captura 20 pode

proporcionar a informação de profundidade e imagens capturadas, por exemplo, pela câmera 3-D e/ou pela câmera RGB, e um modelo de esqueleto que pode ser gerado pelo dispositivo de captura 20 para o ambiente de computação 12 via a ligação de comunicação 36. O ambiente de computação 12 então pode utilizar o modelo de esqueleto, a informação de profundidade, e as imagens capturadas, por exemplo, para reconhecer gestos do usuário e em resposta, controlar uma aplicação tal como um jogo ou processador de texto. Por exemplo, como apresentado na FIG. 2, o ambiente de computação 12 pode incluir um mecanismo de reconhecimento de gestos 190. O mecanismo de reconhecimento de gestos 190 pode incluir uma coleção de filtros de gestos, cada um compreendendo informação com respeito a um gesto que pode ser executado pelo modelo de esqueleto (á medida que o usuário se move). Os dados capturados pelas câmeras 26, 28 e pelo dispositivo 20 na forma do modelo de esqueleto e dos movimentos associados com o mesmo podem ser comparados com os filtros de gesto no mecanismo de reconhecimento de gesto 190 para identificar quando um usuário (como representado pelo modelo de esqueleto) executou um ou mais gestos. Estes gestos podem ser associados com vários controles de uma aplicação. Assim, o ambiente de computação 12 pode utilizar o mecanismo de reconhecimento de gesto 190 para interpretar os movimentos do modelo de esqueleto e para controlar uma aplicação baseado nos movimentos.

[00042] A Flg. 3A ilustra uma concretização ilustrativa de um ambiente de computação que pode ser utilizado para implementar o ambiente de computação 12 das Figuras 1A até 2. O ambiente de computação pode ser um console multimídia 100, tal como um console de jogo. Como apresentado na FIG. 3A, o console multimídia 100 possui uma unidade central de processamento (CPU) 101 possuindo uma cache de nível 1 102 e uma cache de nível 2 104, e uma Rom

(Memória Somente para Leitura) flash 106. A cache de nível 1 102 e a cache de nível 2 104 temporariamente armazenam dados e por consequência, reduzem o número de ciclos de acesso à memória, desse modo melhorando a velocidade de processamento e a produtividade. A CPU 101 pode ser proporcionada possuindo mais do que um núcleo e assim, cachês de nível 1 e de nível 2 adicionais 102 e 104. A ROM flash 106 pode armazenar código executável que é carregado durante a fase inicial de um processo de inicialização quando o console multimídia 100 é ligado.

[00043] Uma unidade de processamento gráfico (GPU) 108 e um codificador de vídeo / codec (codificador / decodificados) de vídeo 114 formam um canal de processamento de vídeo para processamento gráfico de alta velocidade e de alta resolução. Os dados são transportados a partir da unidade de processamento gráfico 108 para o codificador de vídeo / codec de vídeo 114 via um barramento. O canal de processamento de vídeo emite dados para uma porta A/V (áudio / vídeo) 140 para transmissão para uma televisão ou outro dispositivo de exibição. Um controlador de memória 110 está conectado com a GPU 108 para facilitar o acesso do processador aos vários tipos de memória 112, tal como, mas não limitada a uma RAM (Memória de Acesso Aleatório).

[00044] O console multimídia 100 inclui um controlador de E/S 120, um controlador de gerenciamento de sistema 122, uma unidade de processamento de áudio 123, um controlador de interface de rede 124, um primeiro controlador hospedeiro USB 126, um segundo controlador USB 128 e uma submontagem de E/S do painel frontal 130 que de preferência são implementados em um módulo 118. Os controladores USB 126 e 128 servem como hospedeiros para controladores de periféricos 141(1) até 142(2), um adaptador sem uso de fios 148, e uma dispositivo de memória externo 146 (por exemplo, memória flash,

unidade de CD / DVD ROM, mídia removível, etc.). A interface de rede 124 e/ou o adaptador sem uso de fios 148 proporcionam acesso a uma rede (por exemplo, Internet, rede doméstica, etc.) e podem ser qualquer um dentre uma ampla variedade de vários componentes adaptadores com fio ou sem fio, incluindo placa Ethernet, um modem, um módulo Bluetooth, um modem a cabo e assim por diante.

[00045] A memória do sistema 143 é proporcionada para armazenar dados da aplicação que são carregados durante o processo de inicialização. Uma unidade de mídia 144 é proporcionada e pode compreende uma unidade de DVD / CD, disco rígido, ou outra unidade de mídia removível, etc. A unidade de mídia 144 pode ser interna ou externa ao console multimídia 100. Os dados da aplicação podem ser acessados via a unidade de mídia 144 para execução, reprodução, etc., pelo console multimídia 100. A unidade de mídia 144 é conectada com o controlador de E/S 120 via um barramento, tal como um barramento serial ATA ou outra conexão de alta velocidade (por exemplo, IEEE 1394).

[00046] O controlador de gerenciamento do sistema 122 proporciona uma variedade de funções de serviço relacionadas com garantir a disponibilidade do console multimídia 100. A unidade de processamento de áudio 123 e um codec de áudio 132 formam um canal de processamento de áudio correspondente com processamento com alta fidelidade e estéreo. Os dados de áudio são transportados entre a unidade de processamento de áudio 123 e o codec de áudio 132 via uma ligação de comunicação. O canal de processamento de áudio emite dados para a porta de A/V 140 para reprodução por um reproduutor de áudio externo ou por um dispositivo possuindo capacidades de áudio.

[00047] A submontagem de E/S do painel frontal 130 suporta a funcionalidade do botão de ligar 150 e do botão de ejectar 152, bem como quaisquer LEDs (diodos de emissão de luz) ou outros indicadores

expostos na superfície externa do console multimídia 100. Um módulo de fonte de energia do sistema 136 proporciona energia para os componentes do console multimídia 100. Uma ventoinha 138 resfria o conjunto de circuitos dentro do console multimídia 100.

[00048] A CPU 101, a GPU 108, o controlador de memória 110, e vários outros componentes dentro do console multimídia 100 são interconectados via um ou mais barramentos, incluindo barramentos seriais e paralelos, um barramento de memória, um barramento de periférico, e um barramento do processador ou local utilizando qualquer uma dentre várias arquiteturas. A título de ilustração, tais arquiteturas podem incluir um barramento de Interconexão de Componente Periférico (PCI), barramento PCI-Express, etc.

[00049] Quando o console multimídia 100 é ligado, os dados da aplicação podem ser carregados a partir da memória do sistema 143 para a memória 112 e/ou para as caches 102, 104 e executados na CPU 101. A aplicação pode apresentar uma interface gráfica com o usuário que proporciona uma experiência consistente com o usuário quando navegando para tipos de mídia diferentes disponíveis no console multimídia 100. Em operação, as aplicações e/ou outras mídias contidas dentro da unidade de mídia 144 podem ser ativadas ou reproduzidas a partir da unidade de mídia 144 para proporcionar funcionalidades adicionais para o console multimídia 100.

[00050] O console multimídia 100 pode ser operado como um sistema independente por simplesmente conectar o sistema com uma televisão ou outro dispositivo de exibição. Neste modo independente, o console multimídia 100 permite a um ou mais usuários interagirem com o sistema, assistir filmes, ou escutar música. Entretanto, com a integração de conectividade de banda larga feita disponível através da interface de rede 124 ou do adaptador sem uso de fios 148, o console multimídia 100 adicionalmente pode ser operado como um participante

em uma comunidade maior de rede.

[00051] Quando o console multimídia 100 é ligado, uma quantidade estabelecida de recursos de hardware é reservada para uso pelo sistema pelo sistema operacional do console multimídia. Estes recursos podem incluir uma reserva de memória (por exemplo, 16 MB), CPU e ciclos de GPU (por exemplo, 5%), largura de banda para ligação em rede (por exemplo, 8 kbs), etc. Devido a estes recursos serem reservados na hora da inicialização do sistema, os recursos reservados não existem a partir da visão da aplicação.

[00052] Em particular, a reserva de memória é grande o suficiente para conter o núcleo de ativação, as aplicações simultâneas do sistema e os controladores. A reserva da CPU de preferência é constante de modo que se a utilização reservada da CPU não for utilizada pelas aplicações do sistema, uma encadeamento inativo irá consumir quaisquer ciclos não utilizados.

[00053] Com respeito à reserva da GPU, mensagens leves geradas pelas aplicações do sistema (por exemplo, exibição imediata) são exibidas pela utilização de uma interrupção da GPU para programar código para produzir exibição imediata em uma sobreposição. A quantidade de memória requerida para uma sobreposição depende do tamanho da área da sobreposição e a sobreposição de preferência é proporcional à resolução da tela. Onde uma interface com o usuário completa é utilizada pela aplicação simultânea do sistema, é preferível utilizar uma resolução independente da resolução da aplicação. Um escalador pode ser utilizado para estabelecer esta resolução, o qual não precisa alterar a frequência e causar que uma resincronização de TV seja eliminada.

[00054] Após o console multimídia 100 inicializar e os recursos do sistema serem reservados, as aplicações simultâneas do sistema executam para proporcionar as funcionalidades do sistema. As

funcionalidades do sistema são encapsuladas em um conjunto de aplicações do sistema que executam dentro dos recursos reservados do sistema descritos acima. O núcleo do sistema operacional identifica encadeamentos que são encadeamentos de aplicação do sistema versus encadeamentos de aplicação de jogo. As aplicações do sistema de preferência são programadas para funcionarem na CPU 101 em tempos e intervalos predeterminados de modo a proporcionarem uma visão consistente do recurso do sistema para a aplicação. A programação é para minimizar a interrupção da cache para a aplicação de jogo funcionando no console.

[00055] Quando uma aplicação do sistema simultânea requer áudio, o processamento de áudio é programado de forma assíncrona à aplicação de jogo devido à sensibilidade de tempo. Um gerenciador de aplicação de console multimídia (descrito abaixo) controla o nível de áudio da aplicação de jogo (por exemplo, mudo, atenuado) quando as aplicações do sistema estão ativas.

[00056] Os dispositivos de entrada (por exemplo, os controladores 142(1) e 142(2)) são compartilhados pelas aplicações de jogo e pelas aplicações do sistema. Os dispositivos de entrada não são recursos reservados, mas são para serem trocados entre as aplicações do sistema e a aplicação de jogo de modo que cada um terá um foco do dispositivo. O gerenciador de aplicação de preferência controla a troca do fluxo de entrada, sem conhecimento do conhecimento da aplicação de jogo e um controlador mantém a informação de condição com respeito às trocas de foco. As câmeras 26, 28 e o dispositivo de captura 20 podem definir dispositivos de entrada adicionais para o console 100.

[00057] A FIG. 3B ilustra outra concretização ilustrativa de um ambiente de computação 220 que pode ser o ambiente de computação 12 apresentado nas FIGs. 1A até 2 utilizado para interpretar um ou mais gestos em um sistema de reconhecimento, análise e rastreamento de

alvo. O ambiente do sistema de computação 220 é somente um exemplo de um ambiente de computação adequado e não é pretendido para sugerir qualquer limitação quanto ao e escopo de uso ou quanto à funcionalidade do assunto presentemente revelado. O ambiente de computação 220 também não deve ser interpretado como possuindo qualquer dependência ou requerimentos se relacionando com qualquer componente ou combinação de componentes ilustrado no ambiente operacional ilustrativo 220. Em algumas concretizações, os vários elementos de computação representados podem incluir o conjunto de circuitos configurado para instanciar aspectos específicos da presente revelação. Por exemplo, o termo conjunto de circuitos utilizado na revelação pode incluir componentes de hardware especializados configurados para executarem função (funções) por firmware ou chaves. Em outras concretizações ilustrativas, o termo conjunto de circuitos pode incluir uma unidade de processamento de propósito geral, memória, etc., configurado por instruções de software que incorporam lógica operável para executar função (funções). Nas concretizações ilustrativas onde o conjunto de circuitos inclui uma combinação de hardware e software, uma pessoa que implementa pode gravar código fonte incorporando lógica e o código fonte pode ser compilado em código legível por máquina que pode ser processado pela unidade de processamento de propósito geral. Desde que se pode apreciar que o estado da técnica evoluiu até um ponto onde existe pouca diferença entre hardware, software, ou uma combinação de hardware / software, a seleção de hardware em relação a software para efetuar funções específicas é uma escolha de projeto deixada para uma pessoa que implementa. Mais especificamente, os versados na técnica podem apreciar que um processo de software pode ser transformado em uma estrutura de hardware equivalente, e uma estrutura de hardware pode ser ela própria transformada em um processo de software equivalente.

Assim, a seleção de uma implementação de hardware em relação a uma implementação de software é uma escolha de projeto e deixada para a pessoa que implementa.

[00058] Na FIG. 3B, o ambiente de computação 220 compreende um computador 241, o qual tipicamente inclui uma variedade de meios legíveis por computador. Os meios legíveis por computador podem ser qualquer meio disponível que possa ser acessado pelo computador 241, e inclui tanto meio volátil como não volátil, meio removível e não removível. A memória do sistema 222 inclui um meio de armazenamento do computador na forma de memória volátil e/ou não volátil, tal como memória somente para leitura (ROM) 223 e a memória de acesso aleatório (RAM) 260. Um sistema básico de entrada / saída (BIOS) 224, contendo as rotinas básicas que ajudam a transferir informações entre os elementos dentro do computador 241, tal como durante a inicialização, tipicamente é armazenado na ROM 223. A RAM 260 tipicamente contém dados e/ou módulos de programa que são imediatamente acessíveis e/ou atualmente sendo operados pela unidade de processamento 259. A título de exemplo e não de limitação, a FIG. 3B ilustra o sistema operacional 225, os programas de aplicação 226, outros módulos de programa 227 e os dados de programa 228.

[00059] O computador 241 também pode incluir outros meios de armazenamento removíveis / não removíveis, voláteis / não voláteis. Somente a título de exemplo, a FIG. 3B ilustra uma unidade de disco rígido 238 que lê ou grava junto a um meio magnético não removível, não volátil, uma unidade de disco magnético 239 que lê ou grava junto a um disco magnético removível, não volátil 254, e uma unidade de disco ótico 240 que lê ou grava junto a um disco ótico removível, não volátil 253, tal como um CD ROM ou outro meio ótico. Outros meios de armazenamento do computador removíveis / não removíveis, voláteis / não voláteis que podem ser utilizados no ambiente operacional

ilustrativo incluem, mas não estão limitados aos cassetes de fita magnética, aos cartões de memória flash, aos discos versáteis digitais, às fitas de vídeo digitais, à RAM de estado sólido, à ROM de estado sólido, e assim por diante. A unidade de disco rígido 238 tipicamente está conectada com o barramento do sistema 221 através de uma interface de memória não removível, tal como a interface 234, e a unidade de disco magnético 239 e a unidade de disco ótico 240 tipicamente estão conectadas com o barramento do sistema 221 por uma interface de memória removível, tal como a interface 235.

[00060] As unidades e seus meios de armazenamento do computador associados discutidos acima e ilustrados na FIG. 3B proporcionam armazenamento de instruções legíveis por computador, estruturas de dados, módulos de programa e outros dados para o computador 241. Na FIG. 3B, por exemplo, a unidade de disco rígido 238 é ilustrada como armazenando o sistema operacional 258, os programas de aplicação 257, outros módulos de programa 256 e os dados de programa 255. Observe que estes componentes podem ser os mesmos ou diferentes do sistema operacional 225, dos programas de aplicação 226, dos outros módulos de programa 227 e dos dados de programa 228. O sistema operacional 258, os programas de aplicação 257, os outros módulos de programa 256 e os dados de programa 255 recebem números diferentes aqui para ilustrar que, no mínimo, eles são cópias diferentes. Um usuário pode entrar com comandos e informações no computador 241 através dos dispositivos de entrada tais como um teclado 251 e um dispositivo de apontamento 252, normalmente referido como mouse, trackball ou teclado sensível ao toque. Outros dispositivos de entrada (não apresentados) podem incluir um microfone, um controle de jogo, console de jogo, antena de satélite, scanner, ou coisa parecida. Estes e outros dispositivos de entrada frequentemente são conectados com a unidade de processamento 259

através de uma interface de entrada com o usuário 236 que está acoplada com o barramento do sistema, mas podem ser conectados por outra interface e estruturas de barramento, tais como uma porta paralela, porta de jogo ou um barramento serial universal (USB). As câmeras 26, 28 e o dispositivo de captura 20 podem definir dispositivos de entrada adicionais para o console 100. Um monitor 242 ou outro tipo de dispositivo de exibição também está conectado com o barramento do sistema 221 via uma interface, tal como uma interface de vídeo 232. Em adição ao monitor, os computadores também podem incluir outros dispositivos periféricos de saída tais como alto-falantes 244 e a impressora 243, os quais podem ser conectados através de uma interface de periférico de saída 233.

[00061] O computador 241 pode operar em um ambiente em rede utilizando conexões lógicas com um ou mais computadores remotos, tais como o computador remoto 246. O computador remoto 246 pode ser um computador pessoal, um servidor, um roteador, um PC de rede, um dispositivo par ou outro nó comum de rede, e tipicamente inclui vários ou todos os elementos descritos acima com relação ao computador 241, apesar de somente um dispositivo de armazenamento em memória 247 ter sido ilustrado na FIG. 3B. As conexões lógicas representadas na FIG. 3B incluem uma rede de área local (LAN) 245 e uma rede de longa distância (WAN) 249, mas também podem incluir outras redes. Tais ambientes em rede são normais em escritórios, em redes de computadores de grandes empresas, intranets e na Internet.

[00062] Quando utilizado em um ambiente em rede LAN, o computador 241 é conectado com a LAN 245 através de uma interface de rede ou adaptador 237. Quando utilizado em um ambiente de rede WAN, o computador 241 tipicamente inclui um modem 250 ou outro dispositivo para estabelecer comunicações através da WAN 249, tal como a Internet. O modem 250, o qual pode ser interno ou externo,

pode ser conectado com o barramento do sistema 221 via a interface de entrada do usuário 236, ou via outro mecanismo apropriado. Em um ambiente em rede, os módulos de programa representados em relação ao computador 241, ou partes dos mesmos, podem ser armazenados no dispositivo remoto de armazenamento em memória. A título de exemplo, e não de limitação, a FIG. 3B ilustra os programas de aplicação remotos 248 como residindo no dispositivo de memória 247. Será apreciado que as conexões de rede apresentadas são ilustrativas e outros dispositivos para estabelecer uma ligação de comunicação entre os computadores podem ser utilizados.

[00063] A FIG. 4A representa um mapeamento de esqueleto de um usuário ilustrativo que pode ser gerado a partir do dispositivo de captura 20. Nesta concretização, várias articulações e ossos são identificados: cada mão 302, cada antebraço 304, cada cotovelo 306, cada bíceps 308, cada ombro 310, cada quadril 312, cada coxa 314, cada joelho 316, cada perna dianteira 318, cada pé 320, a cabeça 322, o tronco 324, a parte de cima 326 e a parte de baixo 328 da espinha, e a cintura 330. Quando mais pontos são rastreados, aspectos adicionais podem ser identificados, tal como os ossos ou articulações dos dedos da mão ou do pé, ou aspectos individuais da face, tais como nariz e olhos.

[00064] Através do movimento de seu corpo, um usuário pode criar gestos. Um gesto compreende um movimento ou pose por um usuário que pode ser capturado como dados de imagem e dividido em relação ao significado. Um gesto pode ser dinâmico, compreendendo um movimento, tal como imitando arremessar uma bola. Um gesto pode ser uma posição estática, tal como mantendo os braços cruzados 304 em frente de seu tronco 324. Um gesto também pode incorporar acessórios, tais como uma espada de imitação. Um gesto pode compreender mais do que uma parte do corpo, tal como bater palmas 302, ou movimento discreto, tal como fazer beicinho.

[00065] Os gestos podem ser utilizados para entrada em um contexto geral de computação. Por exemplo, vários movimentos das mãos 302 ou de outras partes do corpo podem corresponde às tarefas amplas do sistema comuns tal como navegar para cima e para baixo em uma lista hierárquica, abrir um arquivo, fechar um arquivo e salvar um arquivo. Os gestos também podem ser utilizados em um contexto específico de jogo de vídeo, dependendo do jogo. Por exemplo, com um jogo de dirigir, vários movimentos das mãos 302 e dos pés 320 podem corresponder a dirigir um veículo em uma direção, passar as marchas, acelerar e frear.

[00066] Um usuário pode gerar um gesto que corresponde a andar ou correr, por andar ou correr no próprio local. O usuário alternativamente pode levantar e abaixar cada perna 312 até 320 para imitar andar sem se mover. O sistema pode analisar este gesto por analisar cada quadril 312 e cada coxa 314. Um passo pode ser reconhecido quando um ângulo quadril – coxa (como medido em relação a uma linha vertical, onde uma perna na vertical possui um ângulo quadril – coxa de 0^0 , e uma perna estendida horizontalmente para frente possui um ângulo quadril – coxa de 90^0) excede certo limite em relação à outra coxa. Um andar ou corrida pode ser reconhecido após algum número de passos consecutivos por pernas alternadas. O tempo entre os dois passos mais recentes pode ser pensado como um período. Após algum número de períodos onde o ângulo limite não é alcançado, o sistema pode determinar que o gesto de andar ou de correr parou.

[00067] Dado um gesto de “andar ou correr”, uma aplicação pode estabelecer valores para parâmetros associados com este gesto. Estes parâmetros podem, incluir o ângulo limite acima, o número de passos requeridos para iniciar um gesto de andar ou de correr, um número de períodos onde não ocorre passos para terminar o gesto, e um período limite que determinar se o gesto é uma caminhada ou uma corrida. Um

período rápido pode corresponde a uma corrida, à medida que o usuário estará movendo suas pernas rapidamente, e um período mais lento pode corresponder a uma caminhada.

[00068] Um gesto pode ser associado com um conjunto de parâmetros preestabelecidos a princípio, os quais a aplicação pode substituir com seus próprios parâmetros. Neste cenário, uma aplicação não é forçada a proporcionar parâmetros, mas ao invés disso, pode utilizar um conjunto de parâmetros preestabelecidos que permitem que o gesto seja reconhecido na ausência de parâmetros definidos pela aplicação.

[00069] Existe uma variedade de saída que podem ser associadas com o gesto. Pode existir uma linha de base “sim ou não” quanto a se um gesto está ocorrendo. Também pode existir um nível de confiança, o qual corresponde à probabilidade de que o movimento do usuário rastreado corresponde ao gesto. Isto poderia ser uma escala linear que varia através de números de ponto flutuante entre 0 e 1, inclusive. Onde uma aplicação recebendo esta informação de gesto não pode aceitar positivos falsos como entrada, ela pode utilizar somente estes gestos reconhecidos que possuem um alto nível de confiança, tal como pelo menos 0,95. Onde uma aplicação deve reconhecer cada instancia do gesto, mesmo ao custo de positivos falsos, ela pode utilizar gestos que possuem pelo menos uma nível de confiança muito menor, tal como estes meramente maiores do que 0,2. O gesto pode possuir uma saída para o tempo entre os dois passos mais recentes, e onde somente um primeiro passo foi registrado, este pode ser estabelecido para um valor reservado, tal como -1 (desde que o tempo entre dois passos deve ser positivo). O gesto também pode possuir uma saída para o ângulo da coxa mais elevado alcançado durante o passo mais recente.

[00070] Outro gesto ilustrativo é um “salto de sustentação no calcanhar”. Neste, um usuário pode criar um gesto por elevar seus

calcanhares para fora do solo, mas mantendo os dedos do pé plantados. Alternativamente, o usuário pode saltar, onde seus pés 320 deixam totalmente o solo. O sistema pode analisar o esqueleto em relação a este gesto por analisar a relação de ângulo dos ombros 310, dos quadris 312 e dos joelhos 316 para ver se eles estão em uma posição de alinhamento igual a ficar em pé reto. Então, estes pontos e os pontos superior 326 e inferior 328 da espinha podem ser monitorados em relação a qualquer aceleração para cima. Uma combinação suficiente de aceleração pode ativar um gesto de salto.

[00071] Dado este gesto de “salto sustentado no calcanhar”, uma aplicação pode estabelecer valores para os parâmetros associados com este gesto. Os parâmetros podem incluir o limite de aceleração acima, o qual determinar o quanto rápido algumas combinação dos ombros 310, dos quadris 312 e dos joelhos 316 do usuário deve se mover para cima para ativar o gesto, bem como um ângulo máximo de alinhamento entre os ombros 310, os quadris 312 e os joelhos 316 nos qual um salto ainda pode ser ativado.

[00072] As saídas podem compreender um nível de confiança, bem como o ângulo do corpo do usuário na hora do salto.

[00073] Estabelecer parâmetros para um gesto baseado nos particulares da aplicação que irá receber o gesto é importante ao precisamente identificar os gestos. Apropriadamente identificar os gestos e a intenção de um usuário ajuda muito em criar uma experiência positiva para o usuário. Onde um sistema de reconhecimento de gesto é muito sensível, e mesmo um ligeiro movimento para frente da mão 302 é interpretado como um arremesso, o usuário pode ser tornar frustrado devido aos gestos serem reconhecidos onde ele não tem intenção de fazer um gesto, e assim, ele perde o controle em relação ao sistema. Onde um sistema de reconhecimento de gesto não é sensível o suficiente, o sistema pode não reconhecer tentativas conscientes pelo

usuário de fazer um gesto de arremesso, frustrando o mesmo de uma maneira similar. Em qualquer extremidade do espectro de sensibilidade, o usuário se torna frustrado devido ao mesmo não poder proporcionar apropriadamente entrada para o sistema.

[00074] Outro parâmetro para um gesto pode ser uma distância movida. Onde os gestos de um usuário controlam as ações de um avatar em um ambiente virtual, este avatar pode estar o comprimento do braço a partir de uma bola. Se o usuário deseja interagir com a bola e apanhar a mesma, isto pode requerer que o usuário estenda seu braço 302 – 310 toda a extensão enquanto fazendo o gesto de apanhar. Nesta situação, um gesto de apanhar similar onde o usuário somente parcialmente estende seu braço 302 – 310 pode não obter o resultado de interatividade com a bola.

[00075] Um gesto ou uma parte do mesmo pode ter como um parâmetro um volume de espaço no qual ele deve ocorrer. Este volume de espaço tipicamente pode ser expresso em relação ao corpo onde um gesto compreende o movimento do corpo. Por exemplo, um gesto de arremesso de futebol americano para um usuário destro pode ser reconhecido somente no volume de espaço não menor do que o ombro direito 310a, e no mesmo lado da cabeça 322 como o braço que arremessa 301a – 310a. Pode não ser necessário definir todos os limites de um volume, tal como dentro deste gesto de arremesso, onde um limite externo longe do corpo é indefinido à esquerda, e o volume excede indefinidamente, ou a borda da cena que está sendo monitorada.

[00076] A FIG. 4B proporciona detalhes adicionais de uma concretização ilustrativa do mecanismo de reconhecimento de gesto 190 da FIG. 2. Como apresentado, o mecanismo de reconhecimento de gesto 190 pode compreender pelo menos um filtro 418 para determinar um gesto ou gestos. Um filtro 418 compreende informação

definindo um gesto 426 (daqui para frente referido como “gesto”) junto com os parâmetros 428, ou metadados, para este gesto. Por exemplo, um arremesso, o qual compreende o movimento de uma das mãos a partir de trás da traseira do corpo até passar pela frente do corpo, pode ser implementado como um gesto 426 compreendendo informação representando o movimento de uma das mãos do usuário a partir de atrás do corpo passando a frente do corpo, à medida que este movimento seria capturado pela câmera com sensor de profundidade. Então, os parâmetros 428 podem ser estabelecidos para este gesto 426. Onde o gesto 426 é um arremesso, um parâmetro 428 pode ser uma velocidade limite que a mão tem que alcançar, uma distância que a mão deve percorrer (absoluto, ou relativo ao tamanho do usuário como um todo), e uma classificação de confiança pelo mecanismo de reconhecimento que o gesto ocorreu. Estes parâmetros 428 para o gesto 426 podem variar entre as aplicações, entre os contextos de uma única aplicação, ou dentro de um contexto de uma aplicação com o passar do tempo.

[00077] Os filtros podem ser modulares ou intercambiáveis. Em uma concretização, um filtro possui uma série de entradas, cada uma destas entradas possuindo um tipo, e uma série de saídas, cada uma destas saídas possuindo um tipo. Nesta situação, um filtro pode ser substituído por um segundo filtro que possui o mesmo número e tipos de entradas e saídas como o primeiro filtro sem alterar qualquer outro aspecto da arquitetura do mecanismo de reconhecimento. Por exemplo, pode existir um primeiro filtro para o ato de dirigir que pega como entrada dados do esqueleto e emite uma confiança de que o gesto associado com o filtro está ocorrendo e um ângulo de direção. Onde é desejado substituir este primeiro filtro do ato de dirigir por um segundo filtro do ato de dirigir – talvez devido ao segundo filtro do ato de dirigir seja mais eficiente e exija menos recursos de processamento – pode-se fazer isto

simplesmente por substituir o primeiro filtro pelo segundo filtro, contanto que o segundo filtro possua estas mesmas entradas e saídas – uma entrada do tipo dados do esqueleto e duas saídas do tipo confiança e do tipo ângulo.

[00078] Um filtro não precisa possuir um parâmetro. Por exemplo, um filtro de “altura do usuário” que retorna a altura do usuário pode não permitir quaisquer parâmetros que possam se ajustados. Um filtro de “altura do usuário” alternativo pode possuir parâmetros que podem ser ajustados – tal como se considera o tamanho do pé do usuário, o penteado, acessório para a cabeça e a postura ao determinar a altura do usuário.

[00079] Entradas para um filtro podem compreender entradas tais como dados de articulação sobre uma posição da articulação do usuário, ângulos iguais formados pelos ossos que se encontram na articulação, dados de cor RGB a partir da cena, e a taxa de alteração de uma característica do usuário. As saídas a partir de um filtro podem compreender saídas tais como a confiança de que um dado gesto está sendo feito, a velocidade na qual o movimento do gesto é feito, e um tempo no qual um movimento de gesto é feito..

[00080] Um contexto pode ser um contexto cultural, e ele pode ser um contexto ambiental. Um contexto cultural se refere á cultura de um usuário utilizando um sistema. Diferentes culturas podem utilizar gestos similares para comunicarem significados marcadamente diferentes. Por exemplo, um usuário Americano que desejar falar para outro usuário “olhar” ou “utilizar seus olhos” pode colocar seu dedo indicador na sua cabeça próximo do lado distante de seus olhos. Entretanto, para um usuário Italiano, este gesto pode ser interpretado como uma referência à máfia.

[00081] De forma similar, podem existir diferentes contextos entre diferentes ambientes de uma única aplicação. Pegue um jogo de

atirador em primeira pessoa que envolve operar um veículo motorizado. Enquanto o usuário está em pé, apontando com os dedos em direção ao solo e estendendo o dedo apontado para frente e para longe do corpo pode representar um gesto de perfuração. Enquanto o usuário está no contexto de dirigir, este mesmo movimento pode representar um gesto de “troca de marcha”. Também pode existir um ou mais ambientes de menu, onde o usuário pode salvar seu jogo, selecionar dentre o equipamento do personagem ou executar ações similares que não comprehende, jogar o jogo diretamente. Neste ambiente, este mesmo gesto pode possuir um terceiro significado, tal como selecionar alguma coisa ou avançar para outra tela.

[00082] O mecanismo de reconhecimento de gesto 190 pode possuir um mecanismo de reconhecimento base 416 que proporciona funcionalidade para um filtro de gesto 418. Em uma concretização, a funcionalidade que o mecanismo de reconhecimento 416 implementa inclui um arquivo de entrada através do tempo que rastreia gestos reconhecidos e outras entradas, uma implementação de Modelo de Hidden Markov (onde o sistema modelado é assumido como sendo um processo Markov – um sistema onde um estado presente encapsular qualquer informação de estado passado necessária para determinar um estado futuro, de modo que nenhuma outra informação de estado passado deve ser mantida para este propósito – com parâmetros desconhecidos, e parâmetros ocultos são determinados a partir dos dados observáveis), bem como outra funcionalidade requerida para resolver instâncias particulares de reconhecimento de gesto.

[00083] Os filtros 418 são carregados e implementados na parte de cima do mecanismo de reconhecimento base 416 e podem utilizar serviços proporcionados pelo mecanismo 416 para todos os filtros 418. Em uma concretização, o mecanismo de reconhecimento base 416 processa os dados recebidos para determinar se eles atendem aos

requerimentos de qualquer filtro 418. Desde que estes serviços proporcionados, tais como analisar a entrada, são proporcionados uma vez pelo mecanismo de reconhecimento base 416 ao invés do que por cada filtro 418, tal serviço precisa ser somente processado uma vez em um período de tempo, oposto a uma vez por filtro 418 para este período, de modo que o processamento requerido para determinar gestos é reduzido.

[00084] Uma aplicação pode utilizar os filtros 418 proporcionados pelo mecanismo de reconhecimento 190, ou ela pode proporcionar seu próprio filtro 418, o qual se acopla com o mecanismo de reconhecimento base 416. Em uma concretização, todos os filtros 418 possuem uma interface comum para permitir esta característica de acoplamento. Adicionalmente, todos os filtros 418 podem utilizar parâmetros 428, de modo que uma única ferramenta de gesto como descrita abaixo pode ser utilizada para depurar e ajustar todo o sistema de filtros 418.

[00085] Estes parâmetros 428 podem ser ajustados para uma aplicação ou para um contexto de uma aplicação por uma ferramenta de gesto 420. Em uma concretização, a ferramenta de gesto 420 comprehende vários controles deslizantes 422, cada controle deslizante 422 correspondendo a um parâmetro 428, bem como uma representação pictórica de um corpo 424. À medida que um parâmetro 428 é ajustado com um controle deslizante correspondente 422, o corpo 424 pode demonstrar ambas as ações que seriam reconhecidas como o gesto com estes parâmetros 428 e ações que não seriam reconhecidas como o gesto com estes parâmetros 428, identificadas como tais. Esta visualização dos parâmetros 428 de gestos proporciona um dispositivo eficaz tanto para depurar como para o ajuste fino de um gesto.

[00086] A FIG. 5 representa gestos ou filtros 418 mais complexos criados a partir de gestos ou filtros 418 empilhados. Os gestos pode se

acumular em cima um do outro. Ou seja, mais do que um gesto pode ser expresso por um usuário em um único momento. Por exemplo, ao invés de não permitir qualquer entrada sem ser um arremesso quando um gesto de arremesso é feito, ou exigir que um usuário permaneça sem se mover salvo para os componentes do gesto (por exemplo, ficar em pé parado enquanto fazendo um gesto de arremesso que envolve somente um braço). Onde os gestos se acumulam, um usuário pode fazer um gesto de salto e um gesto de arremesso simultaneamente, e ambos os gestos serão reconhecidos pelo mecanismo de gesto.

[00087] A FIG. 5A representa um filtro de gesto simples 418 de acordo com o paradigma de empilhamento. O filtro IFilter 502 é um filtro básico 418 que pode ser utilizado em cada filtro de gesto. O IFilter 502 pega os dados de posição do usuário 504 e emite um nível de confiança 506 que um gesto ocorreu. Ele também alimenta estes dados de posição 504 para um filtro de Volante 508 que pega os mesmos como uma entrada e emite um ângulo para o qual o usuário está dirigindo (por exemplo, 40 graus para a direita do ângulo de direção atual do usuário) 510.

[00088] A FIG. 5B representa um gesto mais complexo que empilha os filtros 418 sobre o filtro de gesto da FIG. 5A. Em adição ao IFilter 502 e ao SteeringWheel 508, existe um filtro ITracking 512 que recebe dados de posição 504 a partir do IFilter 502 e emite a quantidade de progresso que o usuário fez através de um gesto 514. O ITracking 512 também alimenta os dados de posição 504 para GreaseLightning 516 e EBRAKE 519, os quais são filtros 418 com respeito a outros gestos que podem ser feitos ao operar um veículo, tal como utilizar o freio de emergência.

[00089] A FIG. 6 representa um gesto ilustrativo que um usuário 602 pode fazer para sinalizar um “fair catch” em um jogo de vídeo de futebol americano. Estas figuras representam o usuário neste ponto no tempo,

com a FIG. 6A sendo o primeiro ponto no tempo, e a FIG. 6E sendo o último ponto no tempo. Cada uma destas figuras pode corresponder a um instantâneo ou quadro de dados de imagem capturados por uma câmera com sensor de profundidade 402, apesar de não necessariamente quadros consecutivos de dados de imagem, à medida que a câmera com sensor de profundidade 402 pode estar apta a capturar quadros mais rapidamente do que o usuário pode cobrir a distância. Por exemplo, este gesto pode ocorrer através de um período de 3 segundos, e onde uma câmera com sensor de profundidade captura dados em 40 quadros por segundo, ela iria capturar 60 quadros de dados de imagem enquanto o usuário 602 fazia este gesto de “fair catch”.

[00090] Na FIG. 6A, o usuário 602 começa com seus braços 604 para baixo em seus lados. Ele então eleva os mesmos acima de seus ombros como representado na FIG. 6B e então adicionalmente para cima, até o nível aproximado de sua cabeça, como representado na FIG. 6C. De lá, ele abaixa seus braços 604 até o nível do ombro, como representado na FIG. 6D, e então novamente os eleva, até o nível aproximado de sua cabeça, como representado na FIG. 6E. Onde um sistema captura estas posições pelo usuário 602 sem qualquer posição intermediária que possa sinalizar que o gesto está cancelado, ou outro gesto está sendo feito, ele pode ter o filtro de gesto de “fair catch” emitido com um alto nível de confiança de que o usuário 602 fez o gesto de “fair catch”.

[00091] A FIG. 7 representa o exemplo do gesto de “fair catch” da FIG. 5 à medida que cada quadro dos dados de imagem foram analisados para produzirem um mapa do esqueleto do usuário. O sistema, tendo produzido o mapa do esqueleto a partir da imagem com percepção de profundidade do usuário, pode agora determinar como o corpo deste usuário se move com o passar do tempo, e a partir disto,

analisar o gesto.

[00092] Na FIG. 7A, os ombros do usuário 310, estão acima de seus cotovelos 306, os quais, por sua vez, estão acima de suas mãos 302. Os ombros 310, os cotovelos 306 e as mãos 302 estão então em um nível uniforme na FIG. 7B. Então, o sistema detecta na FIG. 7C que as mãos 302 estão acima dos cotovelos, os quais estão acima dos ombros 310. na FIG. 7D, o usuário retornou para a posição da FIG. 7B, onde os ombros 310, os cotovelos 306 e as mãos 302 estão em um nível uniforme. Na posição final do gesto, apresentada na FIG. 7E, o usuário retorna para a posição da FIG. 7C, onde as mãos 302 estão acima dos cotovelos, os quais estão acima dos ombros 310.

[00093] Enquanto o dispositivo de captura 20 captura uma série de imagens paradas, de modo que em qualquer imagem o usuário parece estar estacionário, o usuário está se movendo no curso da execução deste gesto (oposto a um gesto estacionário, como discutido supra). O sistema está apto a pegar esta série de poses em cada imagem parada, e partir destas determinar o nível de confiança do gesto de movimento que o usuário está fazendo.

[00094] Ao executar o gesto, um usuário é improvável de estar apto a criar um ângulo como formado por seu ombro direito 310a, cotovelo direito 306a e mão direita 302a, por exemplo, de entre 140° e 145° . Assim, a aplicação utilizando o filtro 418 para o gesto de “fair catch” 426 pode ajustar os parâmetros associados 428 para melhor servirem aos elementos específicos da aplicação. Por exemplo, as posições nas FIGs. 7C e 7E podem ser reconhecidas a qualquer tempo que o usuário tenha suas mãos 302 acima de seus ombros 310, sem considerar a posição do cotovelo 306. Um conjunto de parâmetros que são mais rigorosos pode requerer que as mãos 302 estejam acima da cabeça 310 e que os cotovelos 306 estejam ambos acima dos ombros 310 e entre a cabeça 3422 e as mãos 302. Adicionalmente, os parâmetros 428 para

um gesto de “fair catch” 426 podem requerer que o usuário se mova a partir da posição da FIG. 7A através da posição da FIG. 7E dentro de um período de tempo especificado, tal como 1,5 segundos, e se o usuário levar mais do que 1,5 segundos para se mover através destas posições, ela não será reconhecida como a “fair catch” 418, e um nível de confiança muito baixo pode ser emitido.

[00095] As FIGs. 8A até 8C ilustram um usuário fazendo o mesmo gesto de correr reconhecido pelo sistema através de movimentos e poses capturados de forma diferente.

[00096] A FIG. 8A ilustra um usuário fazendo um gesto completo de corrida. O usuário 18 é capturado pelo dispositivo de captura 20. O usuário 18 cria o gesto completo de corrida por correr no local – alternadamente levantando cada um de seus joelhos até aproximadamente a altura da cintura e então abaixando a perna até o solo. Esta versão do gesto completo de correr é um gesto periódico no qual o usuário 18 repete os movimentos que compreendem o gesto durante a duração que ele deseja que o gesto continue.

[00097] A FIG. 8B ilustra um usuário fazendo um gesto de atalho de correr, o gesto de atalho compreendendo um subconjunto do movimento do gesto completo de correr da FIG. 8A. Para fazer esta versão do gesto de atalho de correr, o usuário 18 levanta uma de suas pernas de modo que seu joelho fique aproximadamente no nível do quadril, e mantém esta pose. Este atalho de gesto compreende um subconjunto do movimento do gesto completo da FIG. 8A – onde o usuário na FIG. 8A repetidamente levanta e abaixa seus joelhos, aqui, o usuário levanta seu joelho uma vez e mantém esta pose. Enquanto o gesto completo da FIG. 8A envolve o movimento periódico, nesta concretização, o gesto de atalho envolve uma série de movimentos não repetidos, ou uma série de movimentos onde a série como um todo não é repetida. Em uma concretização, o usuário 18 abaixa seu joelho até

uma pose em pé quando ele deseja terminar o gesto. Em uma concretização, este ato de abaixar o joelho também pode compreender um subconjunto do gesto completo. Em uma concretização, o ambiente de computação 12 determina que este movimento é para terminar o atalho do gesto ao invés do que produzir o gesto completo quando o usuário 18 mantém seu joelho aproximadamente no nível do quadril por mais do que uma duração de tempo especificada.

[00098] A FIG. 8C ilustra um usuário fazendo um segundo tipo de gesto de atalho de correr, o segundo tipo de gesto de atalho de correr compreendendo o movimento separado a partir do gesto completo de correr da FIG. 8A. Aqui, o usuário 18 dá um passo para frente e mantém esta pose com um pé na frente do outro, ambos os pés no solo, durante o tempo que ele deseja produzir o gesto de correr. Esta posição não é encontrada no gesto completo de correr da FIG. 8A. O usuário 18 pode terminar o gesto pô andar para trás para uma posição em pé. Este gesto é similar a este da FIG. 8B pelo fato de que ambos envolvem o movimento para iniciar o gesto, então mantendo uma pose para manter o gesto, e o movimento para terminar o gesto.

[00099] A FIG. 9 ilustra procedimentos operacionais ilustrativos para atalhos de gesto. Como discutido acima, uma entrada de gesto para um dispositivo de computação pode ser reconhecida pelo dispositivo de computação como resultado de vários modos executados por um usuário. Em uma concretização, estes vários modos nos quais um gesto pode ser executado por um usuário compreendem uma versão completa do gesto e um atalho do gesto.

[000100] Os atalhos para gesto podem ser utilizados em vários dos contextos da aplicação. Por exemplo, os atalhos de gesto de correr podem ser utilizados em aplicações que envolvem correr, tal como um jogo de atletismo. Os atalhos de entrada de texto podem ser utilizados em um contexto de entrada de texto de uma aplicação. Por exemplo, o

usuário pode utilizar gestos da linguagem de sinais para entrar texto. Uma versão completa de um gesto de palavra pode compreender sinalizar cada letra da palavra, tal como H-E-A-R-T. Um atalho para o gesto da palavra “heart” pode compreender um único gesto para heart, tal como a partir do usuário colocando a mão em uma representação de um coração. Tal linguagem de sinais pode compreender a American Sign Language (ASL).

[000101] Um atalho para gesto pode envolver diferentes partes do corpo diferente da versão completa correspondente de um gesto. Por exemplo, onde um usuário não utiliza suas pernas, e a versão completa de um gesto de correr envolve correr no local, o atalho do gesto pode envolver imitar um movimento de corrida com as mãos do usuário.

[000102] A operação opcional 902 representa receber entrada do usuário correspondendo a definir o atalho do gesto. Por exemplo, o usuário, por ser solicitado pelo dispositivo de computação, ou através da indicação para o dispositivo de computação de seu desejo de fazer isto, pode fazer um movimento ou pose que é capturado por um dispositivo de captura e armazenado como um modo de executar o gesto.

[000103] Em uma concretização, onde o usuário definiu um atalho para gesto através de seu movimento ou pose, ele pode então refinar o atalho para o gesto no sistema. Por exemplo, onde os gestos são reconhecidos utilizando filtros e parâmetros correspondentes, ele pode ajustar os parâmetros para seu atalho para o gesto de modos como discutidos acima.

[000104] Em uma concretização, o atalho do gesto corresponde a uma versão completa de um segundo gesto. Um atalho do gesto pode corresponde a vários gestos completos, e onde o usuário define um atalho, ele pode indicar que o atalho é para corresponder ao vários gestos completos. Por exemplo, em um contexto de imprimir um arquivo

de texto, o usuário pode definir um atalho do gesto que corresponde ao gesto completo para selecionar a orientação do papel a ser vertical, o gesto completo para selecionar quatro cópias para impressão, e o gesto completo para selecionar uma impressora específica a partir de qual imprimir.

[000105] Em uma concretização, onde os gestos são reconhecidos através de filtros de gesto e parâmetros, o atalho de um gesto e a versão completa de um gesto podem utilizar o mesmo filtro de gesto, mas um valor diferente para um ou mais parâmetros. Por exemplo, a versão completa de um gesto de “arremessar bola” pode requerer que o usuário move sua mão a partir de atrás de seu tronco até aproximadamente o comprimento do braço na frente de seu tronco. O atalho pode reduzir a distância requerida que a mão deve percorrer de modo que a mão não deve ser estendida tão longe para trás nem tão longe para frente. Isto pode ser efetuado pela alteração de um valor ou valores de parâmetros, tal como um para “distância mínima da mão”.

[000106] A operação 904 representa receber dados capturados por um dispositivo de captura, os dados correspondendo a um gesto executado pelo usuário. O dispositivo de captura pode capturar uma cena que contém todo o usuário, tal como a partir do piso até o teto e até a parede em cada lateral de um ambiente na distância na qual o usuário está localizado. O dispositivo de captura também pode capturar uma cena que contém somente parte do usuário, tal como o usuário a partir do abdômen para cima, à medida que ele ou ela senta em uma escrivaninha. O dispositivo de captura também pode capturar um objeto controlado pelo usuário, tal como uma câmera portátil que o usuário segura em sua mão.

[000107] A operação opcional 906 representa determinar a partir da aplicação para processar os dados com o atalho do gesto. Uma aplicação pode limitar os atalhos que são utilizados como entrada de

alguma maneira. Por exemplo, em um jogo de atletismo, correr pode ser considerado integrante para o processo e a aplicação pode desabilitar ou habilitar um atalho para um gesto de correr, requerendo que um usuário faça um gesto completo de correr quando ele deseja correr. Em contraste, em um jogo de atirador em primeira pessoa, correr pode ser considerado auxiliar para o uso do jogo, de modo que um atalho para um gesto de correr pode ser permitido. Neste jogo de atirar em primeira pessoa, a mecânica para descarregar uma arma de fogo pode ser considerada integrante para o processo, e a aplicação pode proibir ou permitir atalhos para um gesto de mirar ou de disparar.

[000108] Em uma concretização, um usuário pode executar tanto atalhos para gestos como versões completas dos gestos que são reconhecidos, da mesma maneira que um usuário pode simultaneamente executar múltiplos gestos, como discutido anteriormente. Utilizando o exemplo do atirador em primeira pessoa, o usuário pode simultaneamente executar o atalho para o gesto de correr, e a versão completa do gesto de mirar.

[000109] Em uma concretização, esta determinação para processar os dados com o atalho dos gestos se origina a partir do usuário. Por exemplo, quais atalhos processar pode corresponder ao nível de dificuldade da aplicação que o usuário seleciona. Onde o usuário escolhe o nível de dificuldade mais baixo, todos os atalhos podem ser processados. À medida que o usuário aumenta o nível de dificuldade, o número de atalhos permitidos para o processo pode diminuir, até o nível de dificuldade mais elevado, onde nenhum atalho é processado.

[000110] Esta determinação pode alterar com o passar do tempo para ser adaptável à habilidade do usuário. Por exemplo, um parâmetro preestabelecido de atalhos permitido pode ser implementado no início de uma sessão da aplicação, e os atalhos permitidos podem ser aumentados ou diminuídos durante a sessão à medida que o usuário

apresenta sua habilidade em executar gestos, ou carece de tal habilidade. Adicionalmente, à medida que o usuário se cansa durante o curso de uma sessão, ou aumenta em competência, os atalhos permitidos podem ser aumentados ou diminuídos para corresponderem ao seu estado atual de habilidade.

[000111] A operação 908 representa processar os dados para determinar uma saída correspondendo a se o usuário executou um atalho de um gesto, o atalho do gesto correspondendo a uma versão completa do gesto. Em uma concretização, esta saída pode compreender um nível de confiança de que o gesto ocorreu. Em uma concretização, isto pode compreender uma indicação quanto a se uma versão completa de um gesto ou um atalho do gesto foi observado.

[000112] A operação 910 representa enviar a saída correspondendo ao atalho do gesto para a aplicação. Onde as presentes operações são executadas pela aplicação, a saída pode ser enviada para um componente da aplicação que pega a entrada do usuário processada e mapeia a mesma para ações na aplicação.

[000113] A operação opcional 912 representa processar os dados para determinar uma saída correspondendo a se o usuário executou a versão completa do gesto, e enviar a saída correspondente para a versão completa do gesto para a aplicação.

[000114] Em uma concretização, o atalho do gesto compreende o movimento do usuário que compreende um subconjunto de movimentos do usuário que compreendem a versão completa do gesto.

[000115] Em uma concretização, a saída correspondendo ao atalho do gesto corresponde a uma alta probabilidade de que o usuário ativou o gesto, a saída correspondendo à versão completa do gesto corresponde a um alta probabilidade de que o usuário ativou o gesto, e a aplicação somente reconhece um gesto do usuário. Onde o atalho do gesto compreende um subconjunto da versão completa do gesto,

quando o usuário executa a versão completa do gesto, ele também irá executar o atalho do gesto. Assim, para uma entrada de gesto pretendida, dois gestos podem ser reconhecidos. Em uma concretização, onde tanto o atalho do gesto como a versão completa do gesto são reconhecidos dentro de um período de tempo prescrito (um que pode ser específico do gesto e/ou do usuário), somente um é utilizado como entrada e o outro é desconsiderado.

[000116] Em uma concretização, a saída correspondendo ao atalho do gesto corresponde a um alta probabilidade de que o usuário ativou o gesto, a saída correspondendo à versão completa do gesto corresponde a uma alta probabilidade de que o usuário ativou o gesto, e a aplicação utiliza uma saída correspondendo à versão completa do gesto para adicionar detalhes para o gesto. Onde o atalho do gesto é reconhecido e processado (tal como uma animação ou resultado correspondente é exibido em um dispositivo de exibição), como uma entrada pela aplicação, e então a versão completa do gesto é reconhecida enquanto o atalho do gesto está sendo processado, a saída a partir da versão completa do gesto pode ser utilizada no processamento.

[000117] Por exemplo, uma versão completa de um gesto de “salto” pode compreender o usuário saltando. Um atalho deste gesto de “salto” pode compreender os movimentos iniciais da versão completa do gesto – agachar e levantar – e a saída do nível de confiança de que o gesto foi executado. Como resultado de observar o atalho do gesto de salto, a aplicação pode processar isto por exibir o avatar do usuário como saltando. Enquanto isto ocorre, se o usuário completar a versão completa do gesto de saltar por continuar a subir e deixar o solo com ambos os pés, a aplicação pode utilizar a altura que o usuário fisicamente salta para exibir o avatar como saltando uma altura correspondente para adicionar detalhes para o atalho de salto correntemente processado. Se o usuário executou somente o atalho

do gesto de saltar, a aplicação pode ter o avatar saltando uma altura preestabelecida.

[000118] Em uma concretização, onde o usuário executa um atalho de um gesto, ele pode corresponde a uma realização a menos na aplicação do que se o usuário executasse a versão completa do gesto. Por exemplo, em um jogo de skate que classifica a performance de um usuário com pontos no placar, onde o usuário executa uma dada pírueta utilizando o atalho de um gesto de pírueta ao andar de skate, o usuário pode receber menos pontos do se ele tivesse executado a dada pírueta utilizando a versão completa do gesto de pírueta ao andar de skate.

Conclusão

[000119] Enquanto a presente revelação foi descrita em conexão com os aspectos preferidos, como ilustrados nas várias figuras, é entendido que outros aspectos similares podem ser utilizados ou modificações e adições podem ser feitas para os aspectos descritos para executar a mesma função da presente revelação sem desvio dos mesmos. Portanto, a presente revelação não deve ser limitada a qualquer aspecto único, mas ao invés disso, construída em amplitude e escopo de acordo com as reivindicações anexas. Por exemplo, os vários procedimentos descritos neste documento podem ser implementados com hardware ou software, ou uma combinação dos mesmos. Assim, os métodos e aparelho das concretizações reveladas, ou certos aspectos ou partes dos mesmos, podem assumir a forma de código de programa (isto é, instruções) incorporado em meio tangível, tal como discos flexíveis, CD-ROMs, unidades de disco rígido, ou em qualquer outro meio de armazenamento legível por máquina. Quando o código de programa é carregado e executado por uma máquina, tal como um computador, a máquina se torna um aparelho configurado para praticar as concretizações reveladas. Em adição ás implementações específicas explicitamente expostas neste documento, outros aspectos e

implementados serão aparentes a partir da consideração do relatório descritivo revelado neste documento. É pretendido que o relatório descritivo e as implementações ilustradas sejam considerados somente como exemplos.

REIVINDICAÇÕES

1. Método para usar um atalho do gesto em um sistema que utiliza gestos do usuário (426) como entrada para uma aplicação, **caracterizado pelo fato de que** compreende as etapas de:

receber (904) dados capturados por um dispositivo de captura (20), os dados correspondentes a um movimento ou pose do usuário;

determinar, com base nos dados, uma primeira saída identificando que o movimento ou pose do usuário corresponde ao usuário executando um atalho de um gesto completo, e uma segunda saída correspondente a um gesto completo, em que desempenho do atalho do gesto completo e desempenho do gesto completo identifique uma entrada para a aplicação, o atalho do gesto completo compreendendo um subconjunto de um movimento ou pose do usuário compreendendo o gesto completo;

enviar a primeira saída para a aplicação;

enviar a segunda saída para a aplicação;

reconhecer a primeira saída e a segunda saída como identificando a mesma entrada de gesto pretendida para a aplicação; e

com base, pelo menos, no reconhecimento do atalho do gesto completo e do gesto completo dentro de um período de tempo predeterminado como identificando a mesma entrada de gesto pretendida para a aplicação, a aplicação usando apenas o atalho do gesto completo ou o gesto completo como entrada de gesto pretendida para a aplicação.

2. Método, de acordo com a reivindicação 1, **caracterizado pelo fato de que** ainda compreende:

receber (902) entrada do usuário correspondente à definição de um movimento ou pose do usuário usado para executar o atalho do gesto completo (426).

3. Método, de acordo com a reivindicação 1, **caracterizado pelo fato de que** o atalho do gesto completo (426) corresponde a um filtro de gesto (418) e pelo menos um parâmetro, o gesto completo (426) corresponde ao filtro de gesto (418) e pelo menos um parâmetro (428), o parâmetro tendo um valor diferente para o atalho do gesto completo do que para o gesto completo e em que determinar, com base nos dados, a primeira saída identificando se o movimento ou pose do usuário corresponde ao usuário que executa o atalho do gesto completo ainda comprehende:

processar os dados com o filtro de gesto e o parâmetro.

4. Método, de acordo com a reivindicação 1, **caracterizado pelo fato de que** determinar (906), com base nos dados, a primeira saída identificando se o movimento ou pose do usuário corresponde ao usuário (18, 602) que executa o atalho de um gesto completo ainda comprehende:

determinar a partir da aplicação para processar os dados para determinar se o atalho do gesto completo (426) foi executado antes de processar os dados para determinar a primeira saída.

5. Método, de acordo com a reivindicação 1, **caracterizado pelo fato de que** determinar, com base nos dados, a primeira saída identificando se o movimento ou pose do usuário corresponde ao usuário (18, 602) que executa o atalho do gesto completo (426) ainda comprehende:

determinar, com base nos dados, a primeira saída que identifica se o movimento ou pose do usuário corresponde ao usuário que executa o atalho de um gesto com base, pelo menos, no recebimento de uma indicação a partir de um usuário de modo que os atalhos de gestos devem ser reconhecidos.

6. Meio de armazenamento legível por computador (223, 238, 239, 240, 253, 254, 260) **caracterizado pelo fato de que** possui o

método de qualquer uma das reivindicações 1 a 5.

7. Sistema (12, 100) para usar um atalho do gesto em um sistema que utiliza gestos do usuário (426) como entrada para uma aplicação, **caracterizado pelo fato de que** compreende:

um processador (101, 259); e

uma memória acoplada comunicativamente ao processador quando o sistema estiver operacional, as instruções executáveis do processador contendo a memória que, com base pelo menos no que está sendo executado no processador, fazem com que o sistema pelo menos:

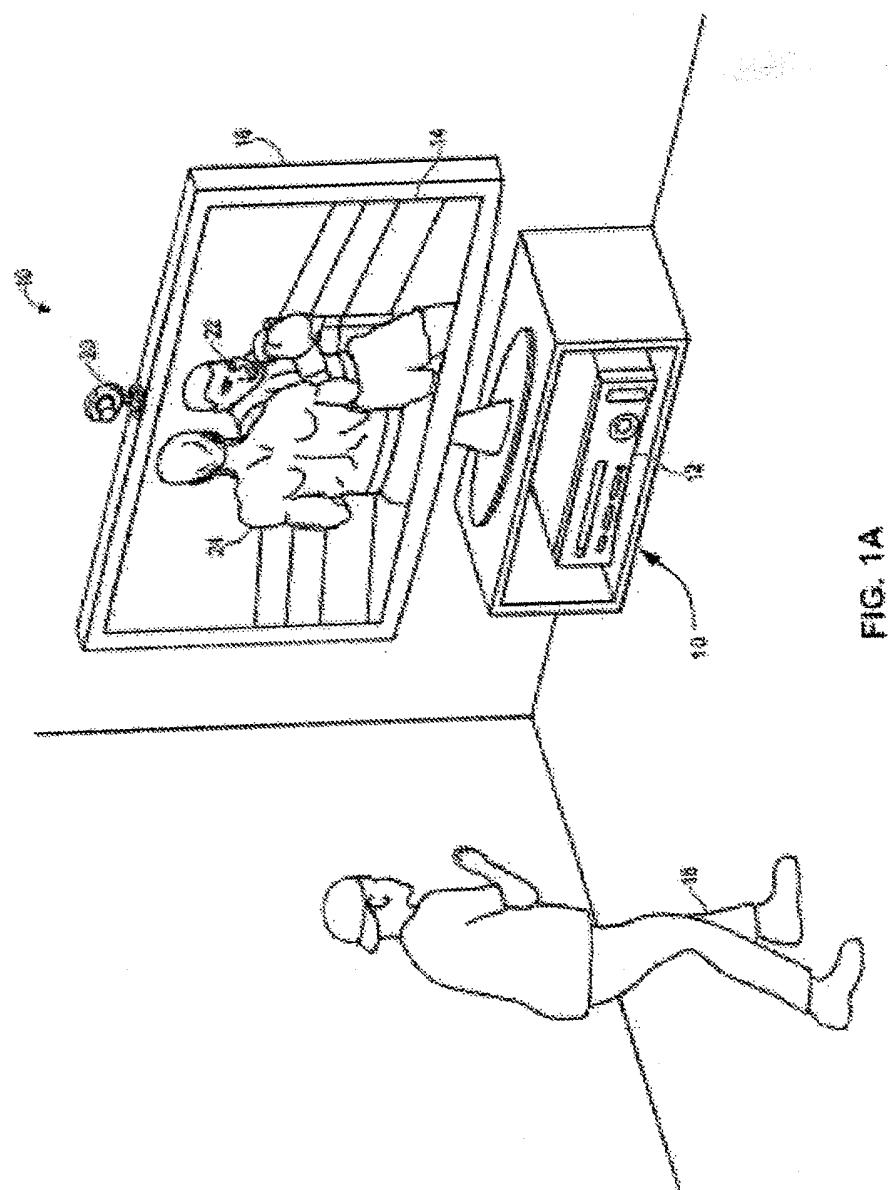
receber (904) dados capturados por um dispositivo de captura (20), os dados correspondentes a um movimento ou pose do usuário;

determinar, com base nos dados, uma primeira saída identificando que o movimento ou pose do usuário corresponde ao usuário (18, 602) executando um atalho de um gesto completo (426) e uma segunda saída correspondente a um gesto completo, em que um desempenho do atalho do gesto completo (426) e desempenho do gesto completo identificam uma entrada para a aplicação, o atalho do gesto completo compreendendo um subconjunto de um movimento ou pose do usuário compreendendo o gesto completo;

enviar (910) a primeira saída para a aplicação; enviar (912) a segunda saída para a aplicação;

reconhecer a primeira saída e a segunda saída como identificando a mesma entrada de gesto pretendida para a aplicação; e

com base, pelo menos, no reconhecimento do atalho do gesto completo e do gesto completo dentro de um período de tempo predeterminado como identificando a mesma entrada de gesto pretendida para a aplicação, usar apenas o atalho do gesto completo ou o gesto completo como a entrada de gesto pretendida para a aplicação.



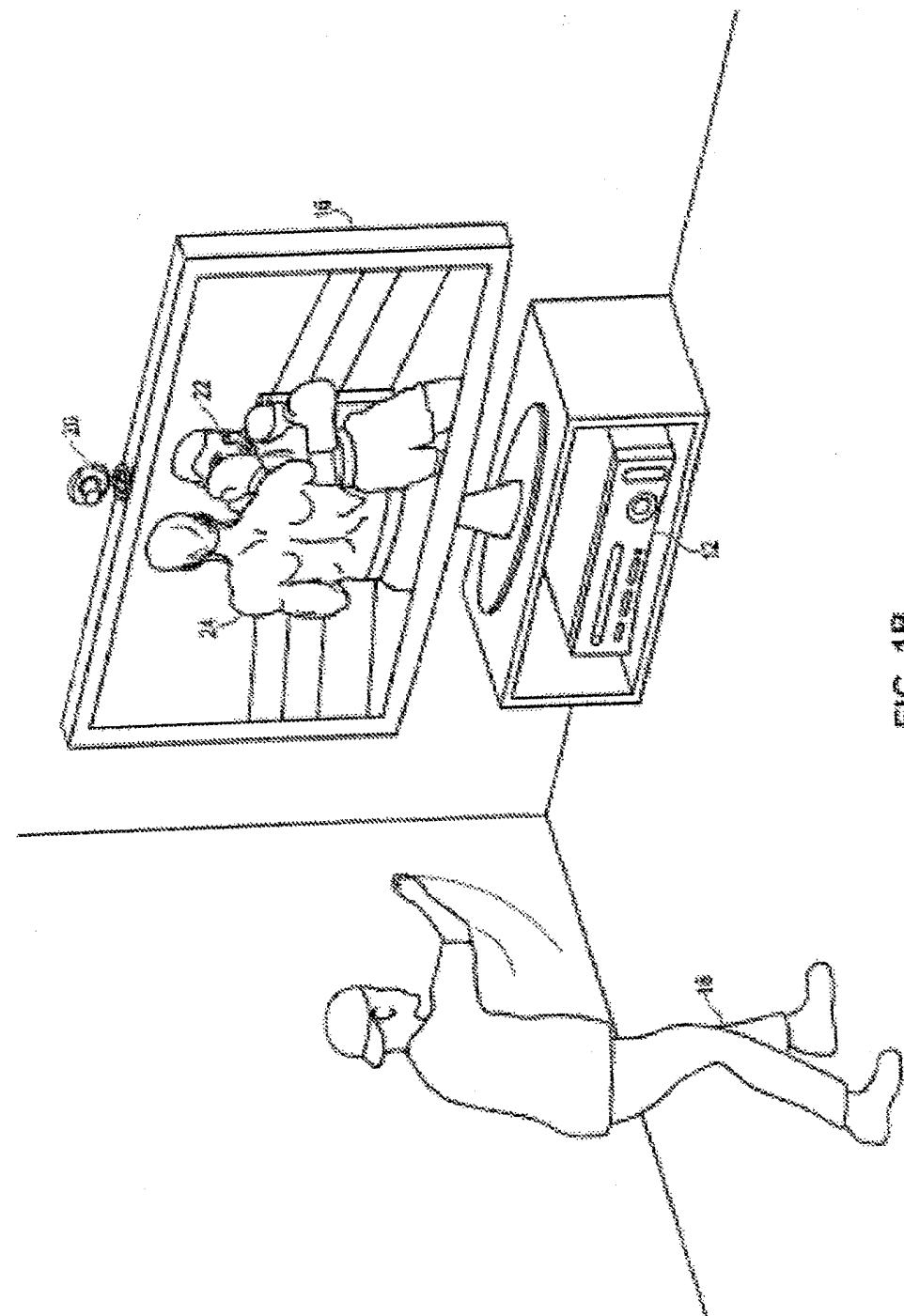


FIG. 1B

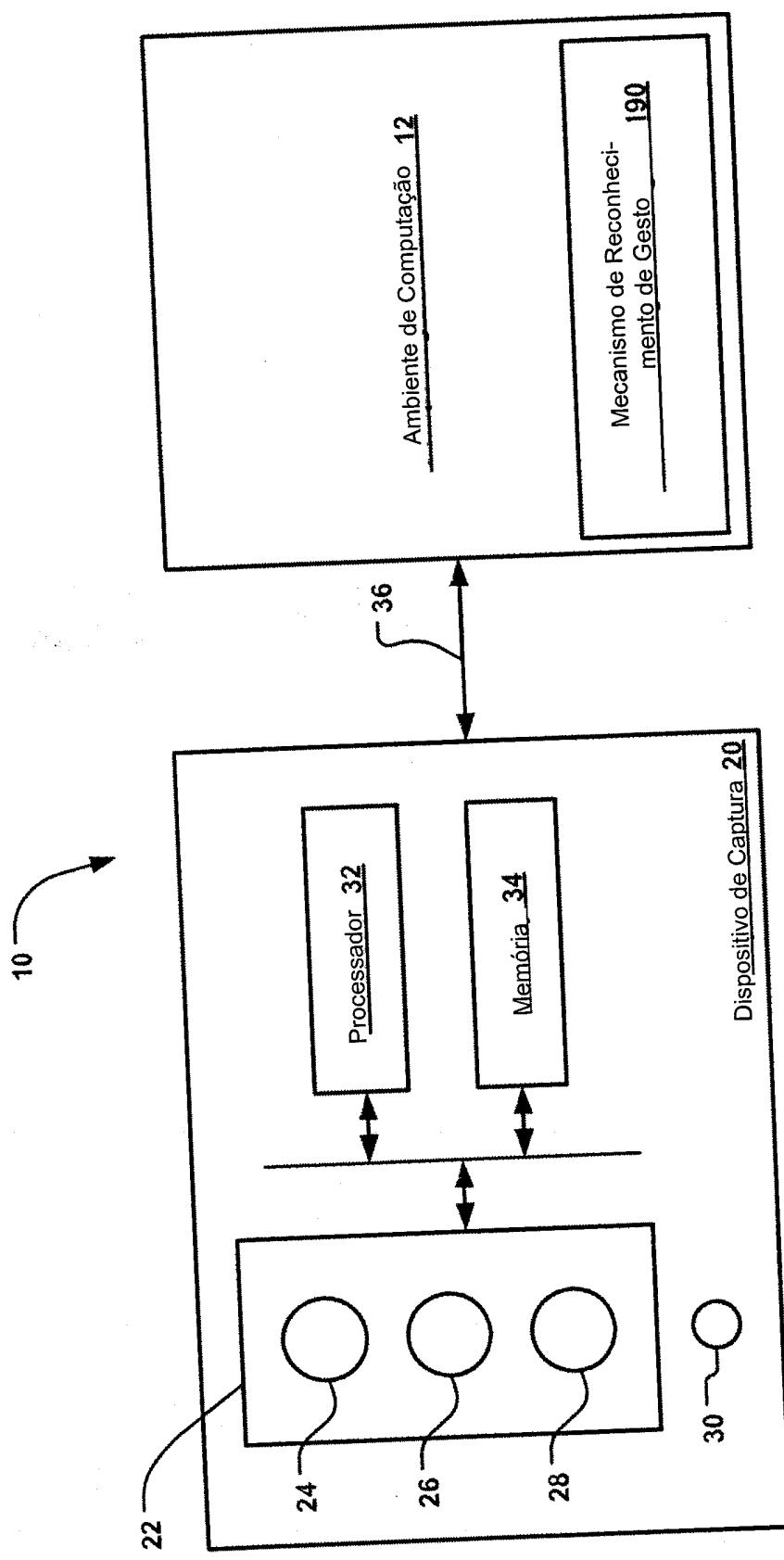


FIG. 2

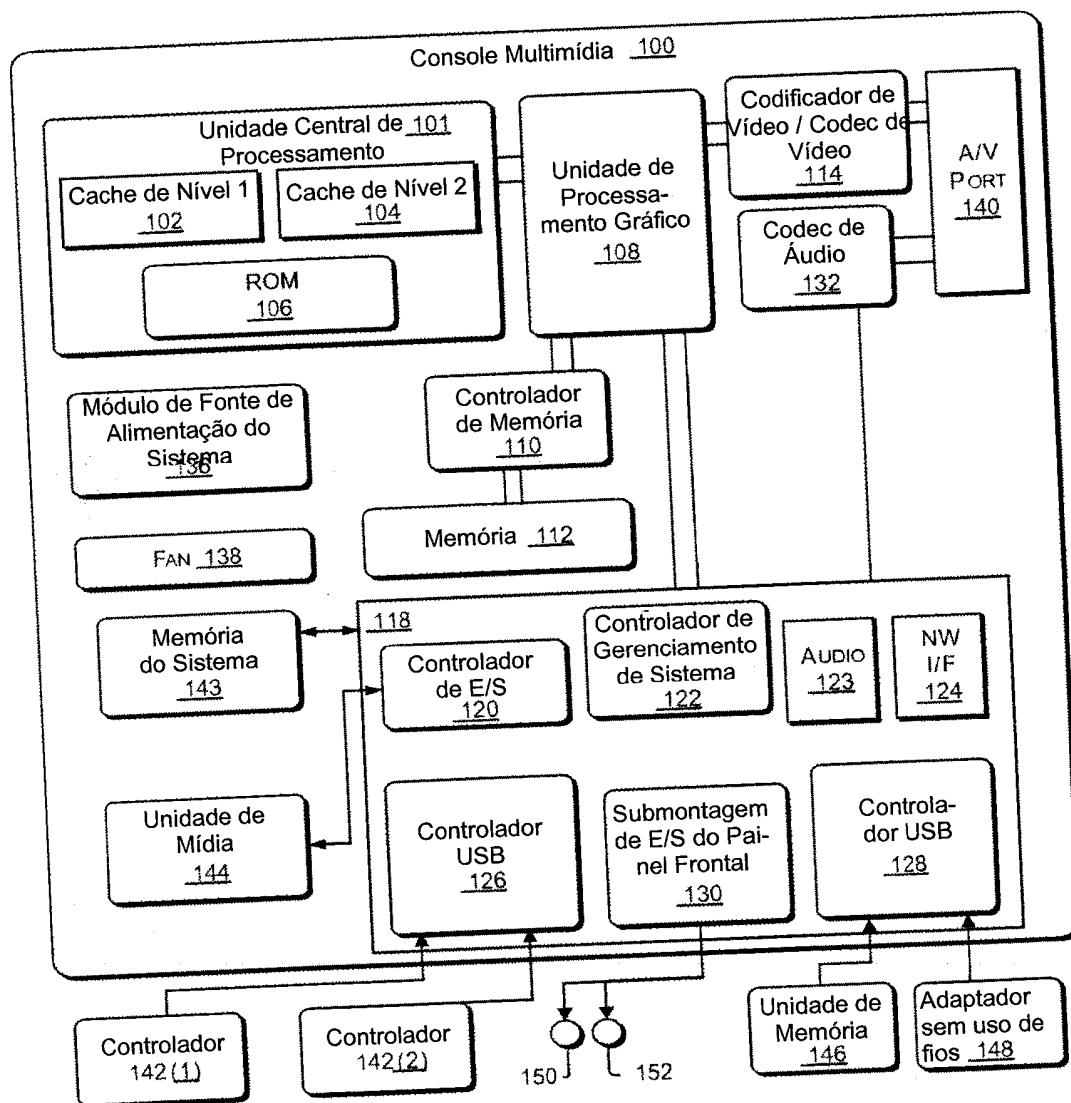


FIG. 3A

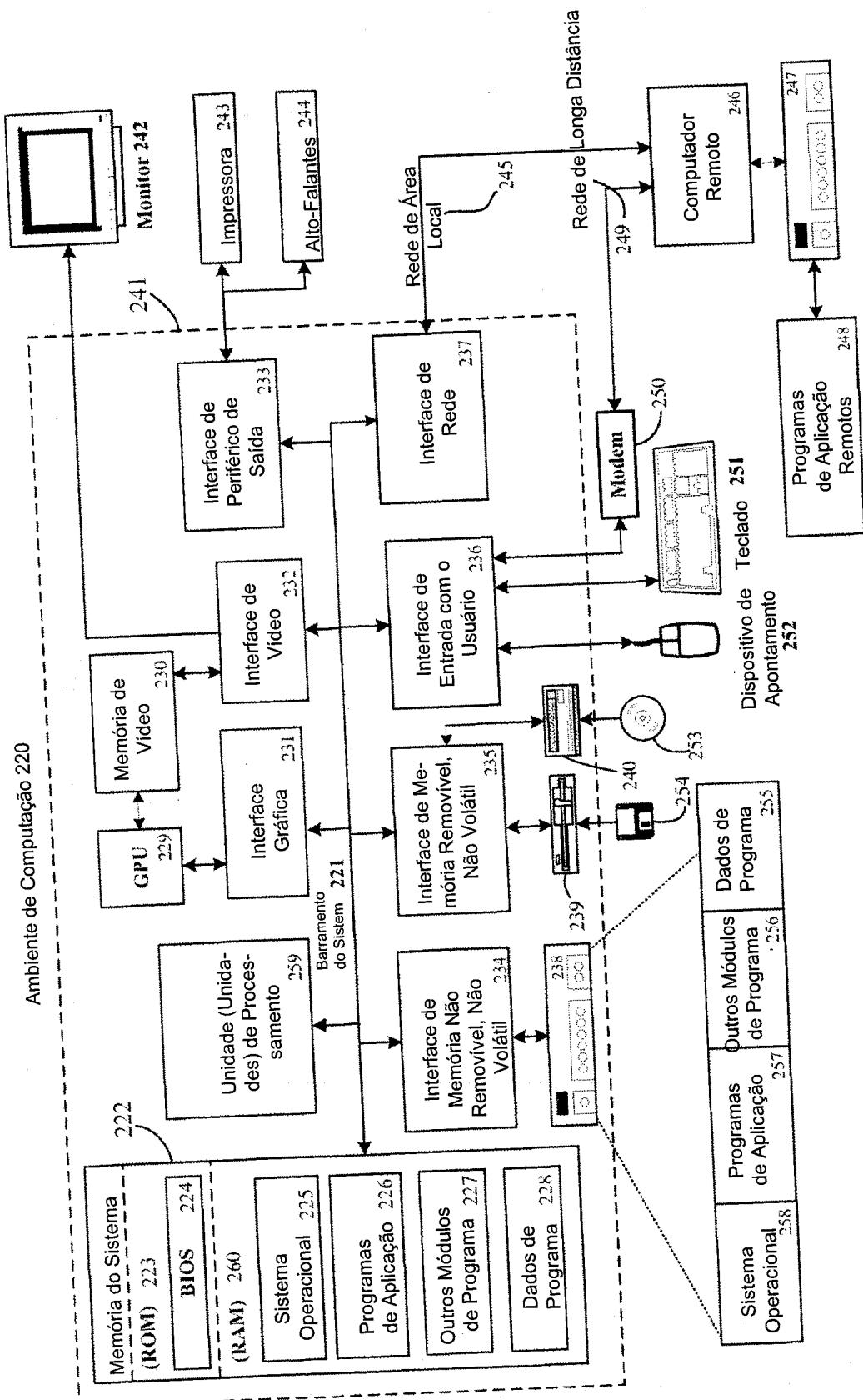


FIG. 3B

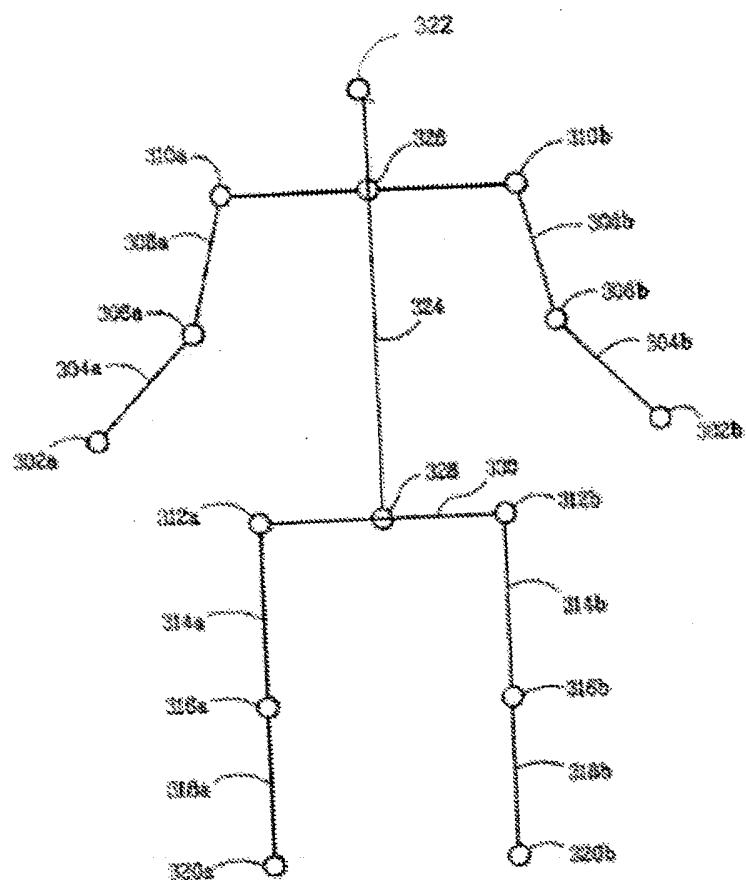


FIG. 4A

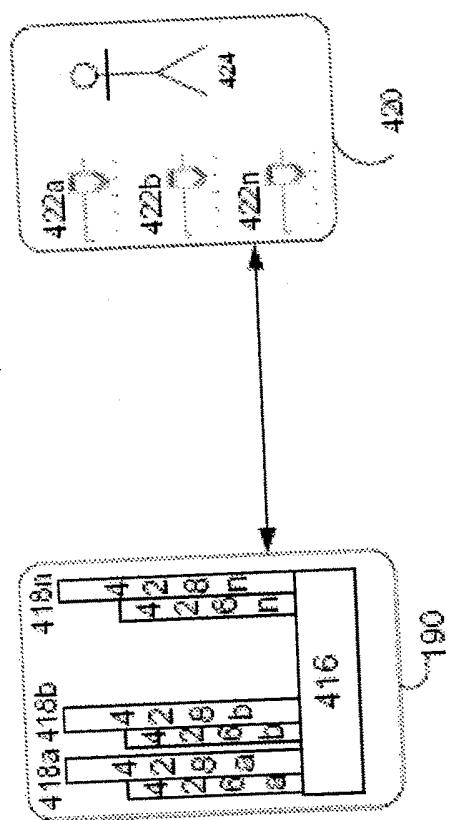


FIG. 4B

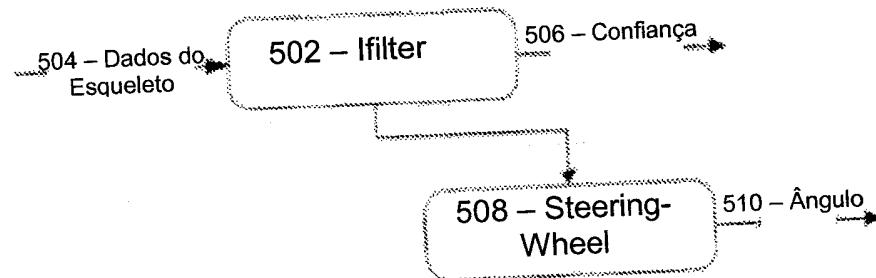


Fig. 5A

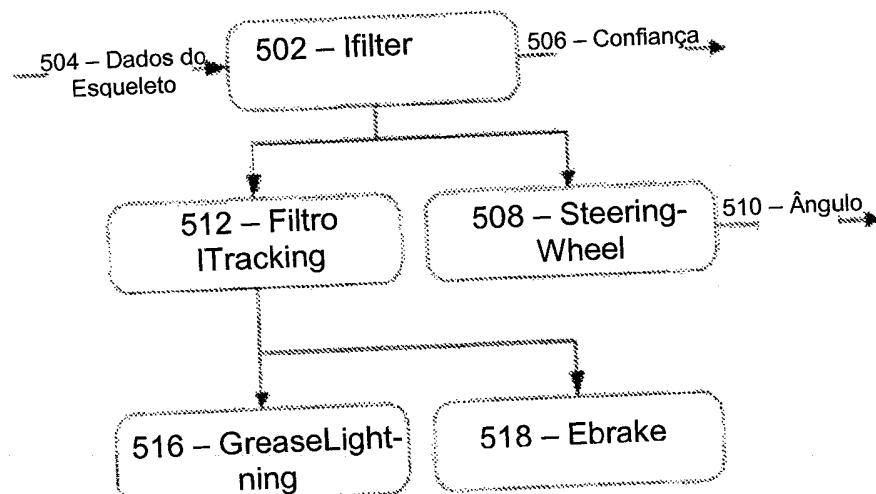


Fig. 5B

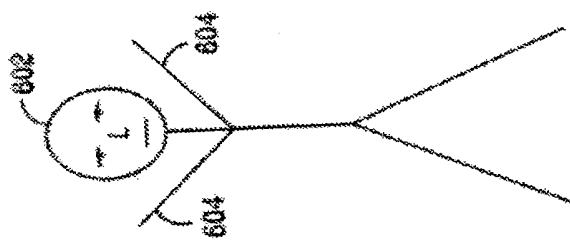


FIG. 6E

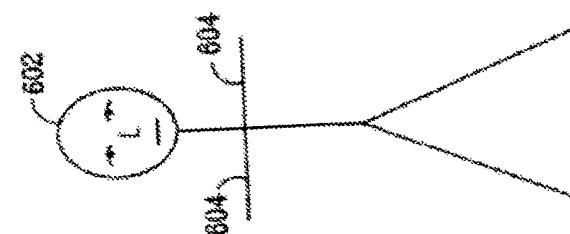


FIG. 6D

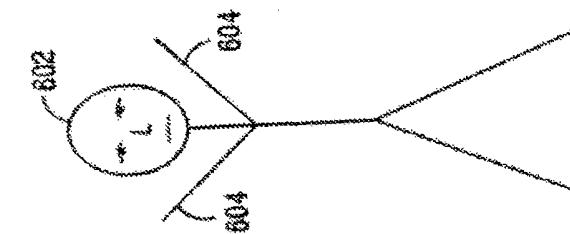


FIG. 6C

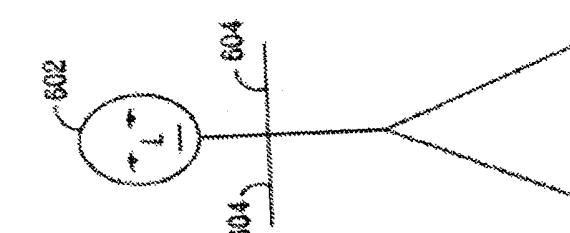


FIG. 6B

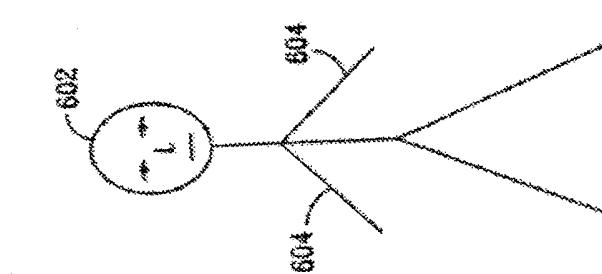


FIG. 6A

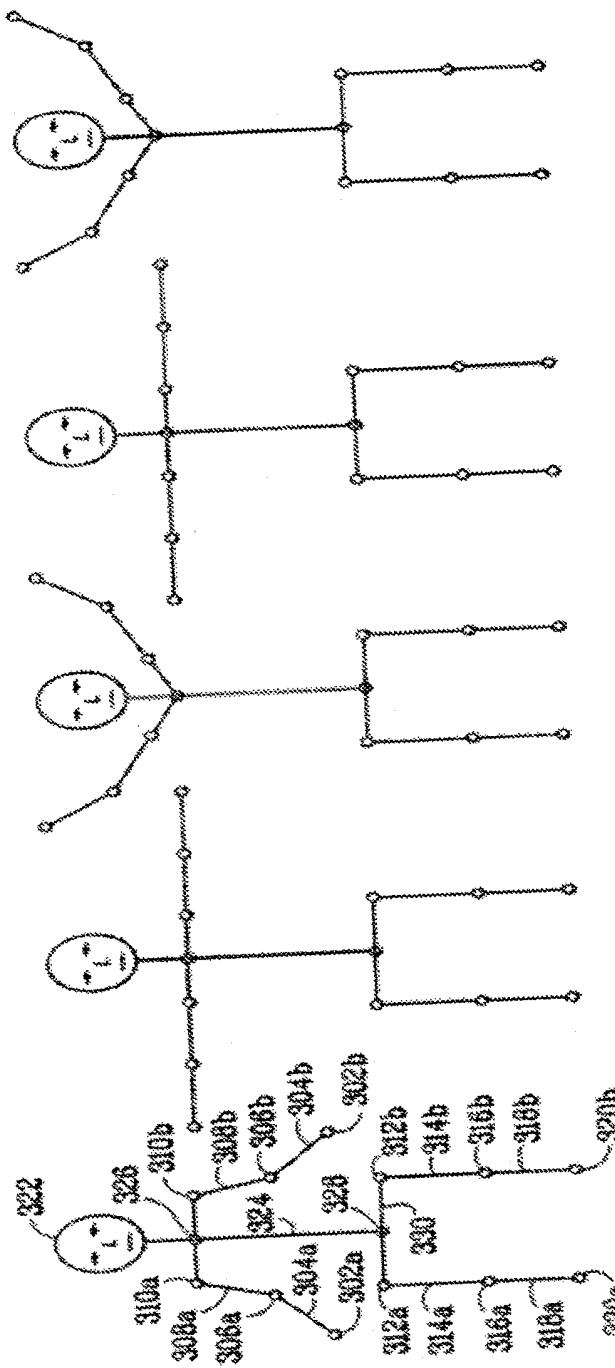


FIG. 7A

FIG. 7B

FIG. 7C

FIG. 7D

FIG. 7E

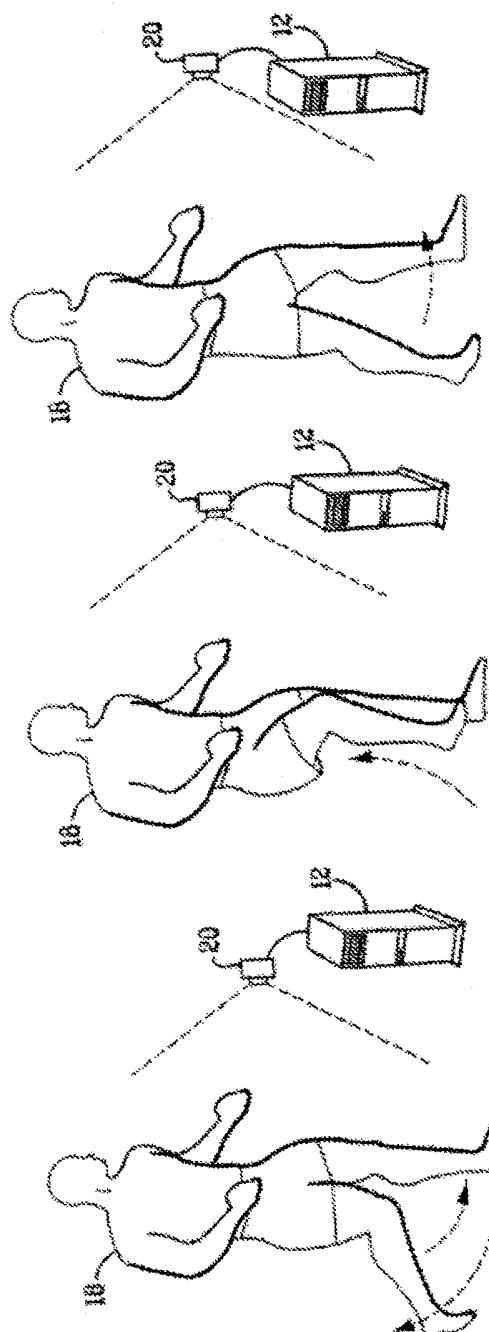


FIG. 8C

FIG. 8B

FIG. 8A

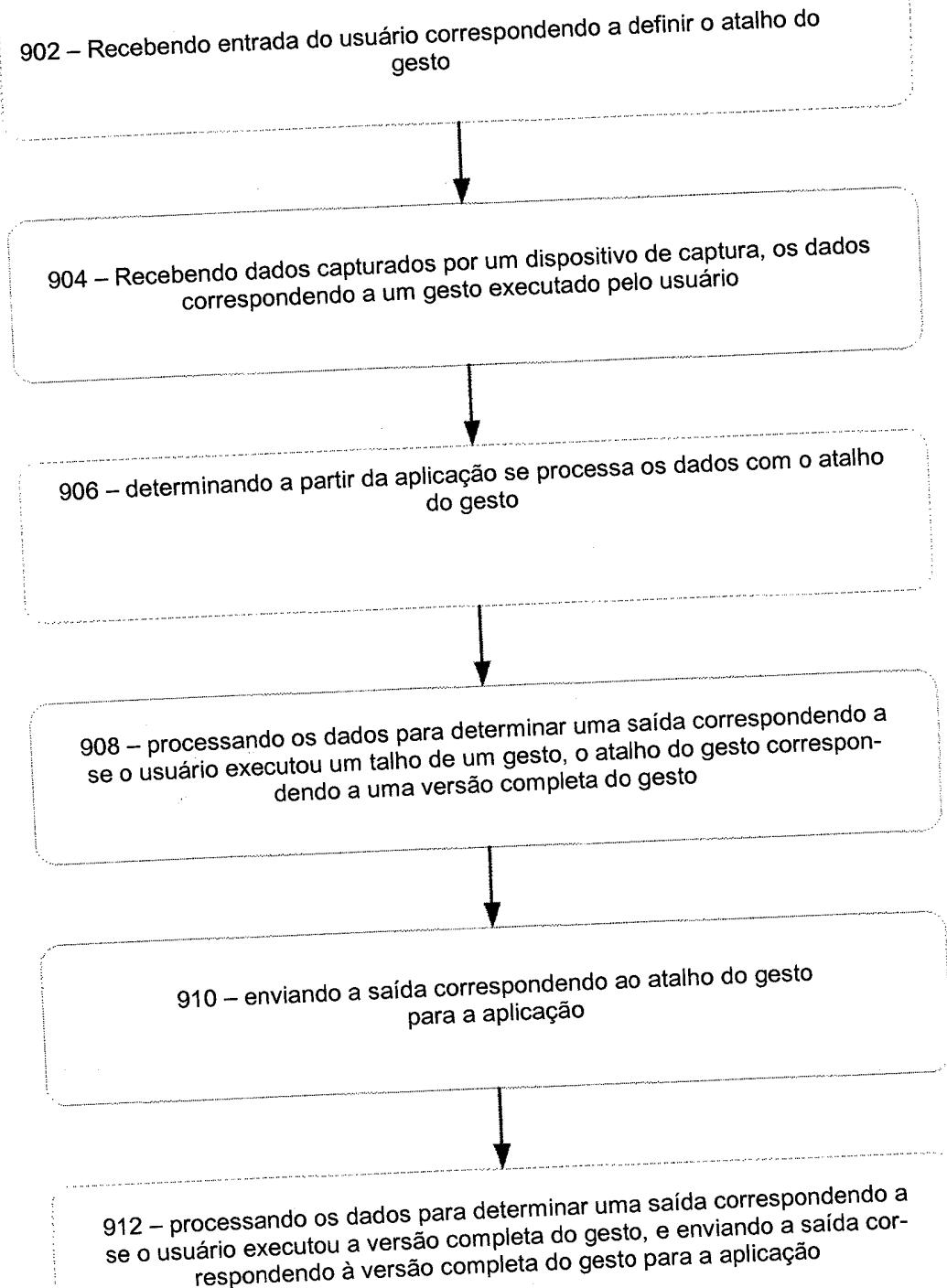


Fig. 9