

República Federativa do Brasil
Ministério do Desenvolvimento, Indústria
e do Comércio Exterior
Instituto Nacional da Propriedade Industrial.

(21) **PI0616547-8 A2**

(22) Data de Depósito: 16/08/2006
(43) Data da Publicação: 21/06/2011
(RPI 2111)



* B R P I O 6 1 6 5 4 7 A 2 *

(51) *Int.Cl.:*
H04N 13/00 2006.01
G06K 9/00 2006.01
G03H 1/00 2006.01

(54) Título: **MÉTODO E SISTEMA DE RECONHECIMENTO E RASTREAMENTO, EM TEMPO REAL, DA POSIÇÃO OCULAR DE MÚLTIPLOS USUÁRIOS**

(30) Prioridade Unionista: 17/08/2005 DE 102005040598.3

(73) Titular(es): Seereal Technologies GmbH

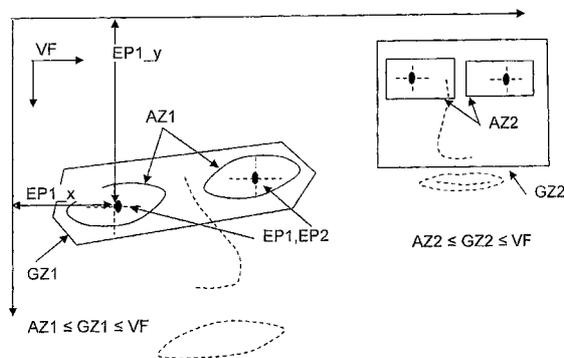
(72) Inventor(es): Alexander Schwerdtner

(74) Procurador(es): Símbolo Marcas e Patentes Ltda.

(86) Pedido Internacional: PCT DE2006001437 de 16/08/2006

(87) Publicação Internacional: WO 2007/019842 de 22/02/2007

(57) Resumo: MÉTODO E SISTEMA DE RECONHECIMENTO E RASTREAMENTO, EM TEMPO REAL, DA POSIÇÃO OCULAR DE MÚLTIPLOS USUÁRIOS, que se refere a um método e a um arranjo de circuito para a detecção e rastreamento isenta de contato de posições de olhos ou pupilas de múltiplos observadores em tempo real. Os dados de entrada consistem de uma seqüência de quadros digitais de vídeo. O método mencionado compreende os seguintes passos: a ação conjunta de um recurso de localização de faces para a detecção de faces, seguido por um recurso de localização de olhos para a detecção de regiões de olhos, e um recurso de rastreamento de olhos para detectar e rastrear pontos de referência de olhos. O objetivo da presente invenção é determinar a posição dos olhos, dentro de uma rotina hierarquicamente organizada, que tem como objetivo reduzir gradualmente a região de procura, iniciando com o quadro completo de vídeo (VF) para a região da face alvo (GZ) e subsequentemente para a região alvo dos olhos (AZ). Além disso, um recurso ou um grupo de recursos é executado em cada um dos casos em uma unidade dedicada de computação, enquanto processos separados estão sendo executados em paralelo. A invenção é usada para rastrear as posições dos olhos de usuários de displays autoestereoscópicos e em holografia de vídeo.



“MÉTODO E SISTEMA DE RECONHECIMENTO E RASTREAMENTO, EM TEMPO REAL, DA POSIÇÃO OCULAR DE MÚLTIPLOS USUÁRIOS”

Campo Técnico

A presente invenção se refere a um método e a um arranjo de circuito para a detecção e rastreamento isenta de contato de posições de olhos ou pupilas de múltiplos observadores em tempo real.

Os dados de entrada consistem de material de imagem na forma de uma seqüência de quadros digitais de vídeo obtidos por um ou por múltiplos sensores de imagem.

Pontos de referência dos olhos de múltiplos observadores podem ser determinados sem a necessidade de quaisquer meios auxiliares tais como óculos, suportes para cabeça ou focos de luz.

Ao contrário das aplicações estacionárias, por exemplo, o monitoramento de motoristas ou pilotos, nas quais o alcance dos movimentos e particularmente a escala de profundidade são muito limitados e portanto quase estacionários, esta invenção serve para detectar a posição dos olhos em uma ampla região alvo, para detectar rápidos movimentos do observador, e para determinar a coordenada de profundidade em uma escala relativamente ampla, por exemplo entre 0,5 e 3,5 m.

A eficiente e exata detecção de olhos em tempo real é realizada através de uma interface-homem-máquina principal. Um campo principal de aplicação da invenção é um dispositivo para detectar e rastrear as posições dos olhos de usuários de *displays* autoestereoscópicos. *Displays* deste tipo proporcionam uma impressão de imagem estereoscópica aos observadores sem a necessidade de quaisquer meios auxiliares, tais como óculos com vidros polarizados. Aplicações adicionais desta invenção compreendem, por exemplo, a holografia de vídeo e implementações na área de detecção de pessoas, faces ou direções de visualização.

Os *displays* autoestereoscópicos, nos quais a apresentação é rastreada pelos assim denominados dispositivos de rastreamento, proporcionam excelente mobilidade a múltiplos observadores em ampla região de visibilidade.

A detecção e o rastreamento livre de erros dos olhos, das posições dos olhos ou das pupilas é também uma importante interface-homem-máquina nestes campos de representação de imagem.

Um dispositivo de rastreamento que funciona de modo confiável e isento de erros normalmente não é percebido pelo usuário. Em muitas aplicações, no entanto, os erros dos sistemas de rastreamento provocam efeitos colaterais indesejados, que, por exemplo, no campo das aplicações 3D, provocam reconstruções com falhas e interferências cruzadas. Um dispositivo de rastreamento deve apresentar alta precisão, confiabilidade e exatidão. O sistema também deve ser adequadamente eficiente e preciso para que ele possa executar o rastreamento correto de todos os principais movimentos e assim permitir que o observador se mova tão livremente quanto possível em todas as três dimensões.

Fundamentos anteriores

Diversos tipos de sistemas de rastreamento isentos de contato encontram-se comercialmente disponíveis. Modelos simples geralmente utilizam um aplicativo básico de *software* para sistemas operacionais padronizados e empregam *hardware* e interfaces de *software* padronizadas.

No documento WO 03/079 902 A1, "Detecção e rastreamento de olhos em tempo real sob condições variadas de luz", Zhiwei Zhu Qiang Ji descreve um método de detecção de olhos em tempo real e isenta de contato, que compreende principalmente uma etapa de detecção de posição de olhos e uma etapa de rastreamento. A etapa de detecção de posição inclui uma combinação do método de iluminação ativa com o reconhecimento de um padrão. Depois que os olhos de um observador foram detectados pela primeira vez, o rastreamento é executado, e a última etapa compreende a combinação e a síntese de diversos algoritmos e técnicas. Apesar da combinação e da síntese de diversos meios, continua existindo o problema dos movimentos principais e abruptos da cabeça em todas as três dimensões não poderem ser rastreados em tempo real, e do processamento em tempo real ser impedido em função do atraso entre o fornecimento dos dados de posição e a aquisição da imagem. Isto se aplica particularmente à detecção da posição dos olhos na dimensão da profundidade em circunstâncias ambientais desfavoráveis.

Em um veículo, por exemplo, o rosto do motorista está sempre situado dentro de uma distância previsível do painel de instrumentos. Além disso, existem somente pequenas variações dos movimentos no sentido vertical e horizontal. Particularmente, o alcance real do movimento na dimensão da profundidade é muito pequeno, de tal maneira que a posição da profundidade pode ser geralmente extrapolada com suficiente exatidão, mesmo se somente uma câmera for utilizada.

O objeto da presente invenção é o de proporcionar um grande alcance de movimentos em todas as três dimensões de um espaço da visualização, com tempos reduzidos de computação. Contrastando com os fundamentos anteriores acima mencionados, é necessário detectar os olhos em todas as três dimensões, isto é, incluindo a dimensão da profundidade. A escala da profundidade deve compreender preferencialmente um grande alcance, de 0,5 a pelo menos 3,5 metros. Para determinar a profundidade, uma pluralidade de câmeras arranjadas independentemente é necessária para a tomada de imagens da região alvo de diversas perspectivas. Além disso, a detecção dos olhos a uma distância de até vários metros requer que as câmeras tenham alta definição, o que resulta na geração de uma grande quantidade de dados por câmera e por quadro de vídeo.

O problema do processamento de grandes quantidades de dados em tempo real se torna mais grave quando existem diversos observadores a serem detectados. Particularmente, etapas e processos de computação muito intensivos são necessários para permitir a detecção de observadores difíceis de serem distinguidos em função de efeitos de iluminação, reflexos ou lentes de óculos. Experiências mostraram, que a detecção de uma terceira ou quarta pessoa, que está parcialmente oculta ou que está posicionada levemente de lado, pode frequentemente ser obtida apenas com esforços computacionais intensivos e consumidores de tempo. No entanto, o esforço computacional requerido para o observador que é momentaneamente menos facilmente detectável e que é apenas detectável com grande esforço, não deve adversamente afetar o rastreamento em tempo real dos outros observadores.

Problemas com a detecção da posição de olhos decorrem do fato dos quadros de entrada de vídeo não poderem ser processados permanentemente no modo de tempo real. Um tempo máximo aceitável de computação por pessoa e por quadro pode ser excedido se lentes de óculos ou fones de ouvidos cobrirem os olhos, ou se um observador se afastar das câmeras abruptamente, mas apenas por um breve momento.

Conhecendo as desvantagens das artes anteriores, é objeto da presente invenção providenciar um método que permita detectar as posições dos olhos de múltiplos observadores em tempo real, mesmo que os observadores movam suas cabeças significativamente, abruptamente e em todas as três dimensões. O método deve detectar a posição dos olhos em uma ampla região alvo, deve compensar movimentos abruptos dos observadores e deve determinar a coordenada de profundidade em uma ampla escala. Além disso, durante a minimização da quantidade de erros, o tempo de resposta entre a aquisição de imagem, isto é, a leitura de um quadro de vídeo e o fornecimento de um resultado, isto é, o fornecimento da posição dos olhos, deve ser sustentavelmente reduzido. Além do mais, o método deve permitir resultados isentos de erros no modo de tempo real também com a utilização de câmeras de alta resolução.

20 Descrição da Invenção

O método é utilizado para a detecção e o rastreamento em tempo real de pontos de referência dos olhos de múltiplos observadores. Os dados de entrada compreendem dados de imagem na forma de uma seqüência de quadros digitais de vídeo que são adquiridos por um ou por vários sensores de imagem, por exemplo câmeras. Os pontos de referência dos olhos são as posições das pupilas e/ou dos cantos dos olhos.

O método compreende a ação conjunta de um recurso de localização de faces para a detecção de faces, seguido por um recurso de localização de olhos hierarquicamente subordinado para a detecção de regiões de olhos, e um recurso de rastreamento de olhos para detectar e rastrear pontos de referência de olhos. O recurso de rastreamento de olhos está hierarquicamente subordinado ao recurso de localização de olhos.

A presente invenção é baseada na idéia de que a localização da posição dos olhos é realizada dentro de uma rotina hierarquicamente organizada, que tem como objetivo reduzir gradualmente a região de procura iniciando com uma imagem de vídeo completa. O comportamento em tempo real é obtido graças à gradual redução e entrelaçamento hierárquicos da região de procura, iniciando com o quadro completo de vídeo do recurso de localização de faces para a região reduzida de alvo da face do recurso de localização de olhos ou do recurso de rastreamento de olhos. Além disso, um recurso ou um grupo de recursos é executado em cada um dos casos em uma unidade dedicada de computação, enquanto processos separados estão sendo executados em paralelo.

O recurso de localização de faces procura, na região de um quadro de vídeo completo, pela posição da cabeça ou da face de cada observador. O recurso portanto determina, a partir dos dados do quadro de vídeo completo, que representa a respectiva região alvo da face, uma quantidade muito menor de dados para cada face, e fornece esta região limitada para o recurso de localização de olhos.

O recurso de localização de olhos está hierarquicamente subordinado ao recurso de localização de faces. A partir dos dados fornecidos pela região alvo da face, o recurso de localização de olhos necessita processar apenas uma quantidade fortemente reduzida de dados. Nestes dados, o recurso determina a posição do olho ou dos olhos do observador e define novamente uma quantidade de dados significativamente menor do que a região alvo da face como sendo a região alvo dos olhos. Apenas esta região limitada de procura é fornecida a um recurso próximo e hierarquicamente subordinado de localização de olhos.

O recurso de localização de olhos então determina, em alta velocidade, neste conjunto de dados fortemente reduzidos da região de procura de olhos, os pontos de referência dos olhos procurados. Com o ajuste hierárquico das regiões de procura e com a redução do volume de dados, o recurso de rastreamento de olhos trabalha rapidamente e com alta eficiência.

De acordo com a presente invenção, para a redução do tempo total de atraso do processo, os recursos de localização de faces e os

recursos de localização de olhos / recursos de rastreamento de olhos devem ser executados independentemente um do outro e em processos separados e paralelos.

5 A paralelização através da atribuição de um recurso ou de um grupo de recursos a uma unidade dedicada de computação, pode ser implementada em um número de modalidades preferidas de execução.

Em uma modalidade de execução particularmente preferida da presente invenção, um recurso de localização de faces é executado para cada câmera de uma unidade dedicada de computação. Desta forma, para
10 cada observador que foi detectado por um recurso de localização de faces, é atribuída uma unidade dedicada de computação para realizar um recurso de localização de olhos, e, subseqüentemente um recurso de rastreamento de olhos. Se um recurso de localização de faces detectar uma nova face, um recurso de localização de olhos e de rastreamento de olhos recebe instruções ou é
15 inicializado imediatamente, e estes recursos são executados em uma unidade de computação dedicada e especialmente atribuída. Um rastreamento imediato de detecção de faces é também executado para faces que foram brevemente perdidas, mas novamente detectadas.

Um dos benefícios principais da presente invenção reside
20 no fato de que o recurso de localização de faces não é bloqueado ou atrasado de maneira nenhuma, por que os recursos subordinados são agora executados em unidades dedicadas de computação. O recurso de localização de faces continua a procurar por faces nos dados do corrente quadro de vídeo, enquanto mantém todos os outros recursos de computação. Resultados de procura intermediários e
25 parciais que foram determinados, são transmitidos a um recurso de controle para processamento / distribuição adicional, ou resultados parciais fornecidos pelos recursos de rastreamento de olhos / localização de olhos são recebidos pelo recurso de controle para que seja possível extrapolar a região alvo de face em um elo positivo de controle.

30 A execução imediata dos recursos reduz substancialmente o tempo de resposta do método, e forma a primeira base para um comportamento em tempo real.

O comportamento em tempo real é obtido graças à gradual redução e entrelaçamento hierárquicos da região de procura, iniciando com o quadro completo de vídeo do recurso de localização de faces para a região reduzida de alvo da face do recurso de localização de olhos ou do recurso de rastreamento de olhos.

Finalmente, de acordo com a presente invenção, o comportamento em tempo real é adicionalmente suportado e assegurado pela execução de um recurso ou de um grupo de recursos em paralelo, com processos rodando em separado em unidades de computação dedicadas. Opções adicionais são possíveis de serem realizadas com relação às paralelidades dos recursos. Como mencionado acima, um recurso de localização de faces e um recurso de localização de olhos / de rastreamento de olhos podem ser executados em unidades dedicadas de computação. Além disso, um recurso de localização de faces / de localização de olhos e um recurso de rastreamento de olhos podem ser executados em unidades dedicadas de computação. Também parece ser possível executar o recurso de localização de olhos em uma unidade dedicada de computação, entretanto, ele consiste de um recurso que requer um tempo relativamente curto de computação, de tal forma que ele pode ser preferencialmente atribuído a uma das unidades de computação utilizadas pelos recursos de localização de faces ou recursos de rastreamento de olhos que requerem computação mais intensiva.

Ambos os processos dos recursos e a troca de dados entre os recursos são preferencialmente controlados e monitorados por um recurso de controle. Particularmente, este recurso controla a atribuição de faces detectadas ou regiões alvo de face aos recursos de localização de olhos / rastreamento de olhos nas unidades dedicadas de computação. A troca de dados envolve principalmente a reinicialização dos recursos através da atribuição de regiões de procura, a troca de resultados parciais e finais de recursos, e a transmissão dos pontos resultantes de referência de olhos a uma interface externa.

Por exemplo, o recurso de controle atualiza e reinicializa os recursos de localização de olhos e de rastreamento de olhos, correspondentes

com uma face que já foi rastreada. O recurso de controle seleciona, verifica e avalia a veracidade das regiões alvo de faces e das regiões alvo de olhos encontradas. Parâmetros correspondentes de avaliação são determinados pelos recursos no decorrer do processo, e são também utilizados pelo recurso de
5 controle para realizar um recurso de processo de controle otimizado e para realizar a atribuição de unidades de computação disponíveis.

O método de acordo com a presente invenção permite a detecção da posição dos olhos de múltiplos observadores em tempo real, mesmo que os observadores movam suas cabeças significativamente, abruptamente e
10 em todas as três dimensões. Adicionalmente foi verificado, que os resultados do método podem alcançar resultados no modo de tempo real também com a alta quantidade de dados de câmeras de alta resolução.

Breve Descrição das Figuras

As figuras seguintes ilustram modalidades preferidas de
15 execução do método de acordo com a presente invenção, sendo utilizadas em conjunto com um dispositivo de rastreamento para um *display* autoestereoscópico.

A Figura 1 ilustra uma representação esquemática dos recursos entrelaçados, regiões reduzidas de procura do localizador de faces, do
20 localizador de olhos e do rastreador de olhos.

A Figura 2 ilustra um fluxograma da paralelização dos recursos estruturados hierarquicamente do método de acordo com a presente invenção.

A Figura 3 ilustra uma representação esquemática do
25 arranjo do circuito e um fluxograma da paralelização dos recursos estruturados hierarquicamente do método de acordo com a presente invenção.

Modalidades Preferidas de Execução da Invenção

A Figura 1 ilustra os recursos entrelaçados e as regiões
30 reduzidas de procura dos recursos deste método. O material de imagem, representando seqüências de quadros digitais de vídeo VF de múltiplos sensores de imagem, por exemplo, de uma câmera estéreo infravermelha, é capturado

como sendo os dados de entrada. A Figura 1 mostra esquematicamente uma porção do quadro total de vídeo VF, definido pelo sistema de coordenadas.

Um primeiro recurso de localização de faces analisa os dados do quadro total de vídeo VF, e detecta as faces dos observadores dentro do quadro total de vídeo. Na Figura 1 são mostrados os dados das duas faces. A primeira face (esquerda) está aparentemente situada próxima da câmera, enquanto que a segunda face (direita) apresenta uma maior distância da câmera.

O recurso de localização de faces determina, a partir dos dados do quadro total de vídeo VF, para cada face detectada uma região reduzida de dados que corresponde com a região da face alvo GZ. Os índices são relacionados com a primeira face, mostrada no lado esquerdo da figura. A região determinada da face alvo GZ então forma a região de procura reduzida para o recurso subsequente de localização de olhos. O recurso de localização de olhos determina a posição dos olhos nesta região de procura e reduz, como resultado, adicionalmente a quantidade de dados da região da face alvo GZ para obter uma quantidade de dados ainda menor e que corresponde com a região alvo dos olhos AZ.

Os dados da região alvo dos olhos AZ com as posições dos olhos são os dados de entrada para o recurso subsequente de rastreamento de olhos ET, que então detecta, na região alvo dos olhos AZ do quadro de vídeo corrente e de acordo com a seqüência de movimentos já determinada na região alvo dos olhos AZ, rastreada nos quadros de vídeo seguintes, pontos de referência de olhos a serem fornecidos como resultados.

A informação dos pontos de referência dos quadros de vídeo anteriores, de acordo com os movimentos do observador, é utilizada para atualizar e rastrear a região alvo dos olhos AZ, e para extrapolar as regiões dos quadros de vídeo correntes e dos quadros de vídeo subsequentes. Se o observador se mover na dimensão da profundidade, o conteúdo da imagem pode requerer redimensionamento adicional.

Conforme ilustrado na figura, a região alvo dos olhos pode compreender diversas porções não contíguas. Como ilustrado adicionalmente na figura, estas regiões alvo são irregulares, mas de formato preferencialmente

convexo, dependendo da posição da cabeça do observador e da sua direção de visualização. Em uma modalidade simples de execução, estas regiões são representadas por uma lista de superfícies geométricas parametrizadas, tais como elipses, círculos ou retângulos.

5 A Figura 2 é baseada na última modalidade preferida de execução, e mostra um fluxograma da paralelização dos recursos. A figura descreve a estrutura hierárquica do recurso de localização de faces FF, do recurso de localização de olhos EF, do recurso de rastreamento de olhos ET e da atribuição às unidades de computação dedicadas R1 a R2.

10 Três unidades de computação R1 a R3 encontram-se disponíveis nesta modalidade preferida de execução. Uma primeira unidade de computação R1 é dedicada ao recurso de localização de faces FF. Este recurso detecta a face de um primeiro observador nos dados de um quadro de vídeo e determina a região da face alvo GZ. A seguir, uma unidade de computação
15 dedicada é atribuída imediatamente à região da face alvo para a execução de um recurso de localização de olhos e, subseqüentemente, um recurso de rastreamento de olhos.

A figura ilustra o fluxo dos dados nas regiões alvo reduzidas, isto é, da região da face alvo GZ e da região do olho alvo AZ até os
20 recursos subseqüentes, respectivamente. Um recurso de rastreamento de olhos ET fornece os dados dos pontos de referência dos olhos a um recurso de controle de mais alto nível (não mostrado) ou a uma interface externa. Ao mesmo tempo, a informação dos pontos de referência detectados em quadros de vídeos anteriores é utilizada para rastrear a região do olho alvo AZ, e para extrapolar a mesma aos
25 quadros de vídeo subseqüentes caso o observador se mova. Os dados da região corrente do olho alvo e das regiões de quadros de vídeo anteriores são, portanto, ambos utilizados pelo recurso de rastreamento de olhos ET, como ilustrado na figura.

O segundo observador é detectado e rastreado da mesma
30 maneira. Se existirem mais observadores do que unidades de computação, um recurso de localização de olhos / de rastreamento de olhos é preferencialmente executado para cada observador (ou, em outras palavras, para cada região de

face alvo), de tal forma que processos múltiplos independentes e separados sejam executados onde naturalmente processos múltiplos são executados em uma unidade de computação comum.

A Figura 3 ilustra o arranjo de circuitos, o fluxograma da
5 paralelização dos recursos hierarquicamente estruturados e a paralelização do método, com a ajuda dos dados de imagem de múltiplas câmeras em diferentes posições. Para a detecção e o rastreamento dos olhos, cada câmera é baseada em um método de acordo com as modalidades preferidas de execução acima mencionadas. Cada câmara é portanto atribuída ao método de paralelização de
10 recursos conforme ilustrado na Figura 1 e Figura 2.

O sistema do lado esquerdo detecta, baseado nos dados de imagem do lado esquerdo VFL (quadro de vídeo esquerdo), a região da face alvo GZ1-L do primeiro observador com a ajuda de um recurso de localização de faces FF executado em uma primeira unidade de computação R1. O recurso de
15 localização de olhos EF e o recurso de rastreamento de olhos ET correspondentes são executados na unidade de computação R2. Com relação ao arranjo dos circuitos, estas unidades de computação são tipicamente configuradas na forma de CPUs ou DSPs.

Um segundo grupo de recursos da unidade de
20 computação R3 é atribuído a um segundo observador. Os outros recursos e unidades de computação mostrados na figura, que são sinalizados como VFR (quadro de vídeo direito) e identificados pelo indexador 'R', são relacionados à imagem do lado direito e aos recursos correspondentes ou elementos do arranjo de circuitos.

25 Uma unidade de controle implementada, que não está sendo mostrada na figura, executa a função de controlar os processos individuais e de organizar a troca de dados durante o processo. A troca de dados procede particularmente entre as unidades de computação que estão relacionadas a um observador. Por exemplo, informações já disponíveis da imagem da esquerda são
30 utilizadas para determinar e para extrapolar a posição na imagem direita, cujo conteúdo não difere substancialmente da imagem esquerda com uma tolerância aceitável. Uma transformação de resultados parciais é possível de ser realizada,

baseada na posição x-y dos *pixels* do olho da imagem esquerda, na distância do observador como determinado nos cálculos anteriores de profundidade e nos parâmetros da câmera. Por exemplo, os dados de uma região do olho alvo AZ1-L encontrados na meia imagem esquerda, são definidos como parâmetros de entrada para a meia imagem direita AZ1-R, e transformados se necessário. Conseqüentemente, é possível utilizar outros algoritmos e outros parâmetros de controle do que aqueles utilizados para o processo do lado esquerdo.

A informação necessária para estes cálculos compreende, principalmente, a resolução e o distanciamento de *pixels* das câmeras, a distância focal da lente do objeto, a distância entre a imagem da lente do objeto e a câmera, e a distância e a orientação das câmeras.

O arranjo dos circuitos compreende, principalmente, comunicações, módulos lógicos programáveis, processadores, ROMs e RAMs. As unidades de computação são preferencialmente otimizadas e configuradas apenas para o propósito pretendido, particularmente para os recursos acima mencionados. Em uma modalidade preferida de execução adicional, o arranjo de circuitos contém unidades de computação dedicadas adicionais, para a execução de processos auxiliares, tais como redimensionamentos, correções gama, etc.

Legenda das Figuras

- Na Figura 2:
- 2.1: quadro total de vídeo (VF)
 - 2.2: recurso de localização de faces (FF)
 - 2.3: recurso de localização de olhos (EF)
 - 2.4: recurso de rastreamento de olhos (ET)
 - 2.5: recurso de localização de olhos (EF)
 - 2.6: recurso de rastreamento de olhos (ET)
 - 2.7: pontos de referência de posição (EP1, ..., EPn)
- Na Figura 3:
- 3.1: quadro de vídeo esquerdo (VFL)
 - 3.2: quadro de vídeo direito (VFR)
 - 3.3: recurso de localização de faces (FF)
 - 3.4: recurso de localização de faces (FF)

- 5
- 3.5: recurso de localização de olhos (EF)
 - 3.6: recurso de rastreamento de olhos (ET)
 - 3.7: recurso de localização de olhos (EF)
 - 3.8: recurso de rastreamento de olhos (ET)
 - 3.9: recurso de localização de olhos (EF)
 - 3.10: recurso de rastreamento de olhos (ET)
 - 3.11: pontos de referência de posição (EP1, ..., EPn) do lado esquerdo
- 10
- 3.12: pontos de referência de posição (EP1(x,y,z), EP2(x,y,z), ..., EPn(x,y,z))
 - 3.13: pontos de referência de posição (EP1, ..., EPn) do lado direito

REIVINDICAÇÕES

1. MÉTODO E SISTEMA DE RECONHECIMENTO E RASTREAMENTO, EM TEMPO REAL, DA POSIÇÃO OCULAR DE MÚLTIPLOS USUÁRIOS para a detecção e o rastreamento dos pontos de referência dos olhos (EP1, ..., EPn) de múltiplos observadores, no qual os pontos de referência dos olhos são as posições das pupilas e/ou os cantos dos olhos, em dados de imagem de quadros de vídeo (VF), que são adquiridos por pelo menos um sensor de imagem, caracterizado pelo fato de compreender: recurso de localização de faces (FF) para detectar posições de faces em quadros de vídeo (FF), o mencionado recurso extraíndo uma sub-região muito menor do quadro de vídeo completo como sendo uma região da face alvo (GZ) de cada face, para transmitir informações da mencionada região da face alvo (GZ) a pelo menos um subseqüente; recurso de localização de olhos (EF) para detectar posições de olhos, o mencionado recurso extraíndo uma sub-região muito menor de cada região da face alvo (GZ) como sendo uma região do olho alvo (AZ), para transmitir informações da mencionada região do olho alvo (AZ) a pelo menos um subseqüente; recurso de rastreamento de olhos (ET) para rastrear posições de olhos, o mencionado recurso detectando pontos de referência de posição dos olhos (EP1, ..., EPn) na região do olho alvo (AZ) do quadro de vídeo corrente e dos quadros de vídeo subseqüentes, e gerando informações para um dispositivo de rastreamento; no qual cada um dos recursos ou um dos recursos e uma combinação dos dois outros recursos são executados em paralelo um em relação ao outro e em uma unidade dedicada de computação.

2. MÉTODO de acordo com a Reivindicação 1, caracterizado pelo fato de que as coordenadas dos pontos de referência de posição (EP1, ..., EPn) são posições horizontais e verticais.

3. MÉTODO de acordo com a Reivindicação 2, caracterizado pelo fato de que as coordenadas de profundidade dos olhos são determinadas baseadas em pontos de referência de posição (EP1, ..., EPn) bidimensionais de múltiplos sensores de imagem.

4. MÉTODO de acordo com a Reivindicação 3, caracterizado pelo fato de que o recurso de localização de faces e uma combinação do recurso de localização de olhos e do recurso de rastreamento de olhos, ou uma combinação do recurso de localização de faces e do recurso de

localização de olhos e do recurso de rastreamento de olhos, são executados em respectivas unidades de computação dedicadas.

5 **5. MÉTODO** de acordo com a Reivindicação 3, caracterizado pelo fato de que para cada face detectada um recurso de localização de olhos e um recurso de rastreamento de olhos são inicializados e executados em pelo menos uma unidade de computação dedicada.

10 **6. MÉTODO** de acordo com a Reivindicação 3, caracterizado pelo fato de que uma ordem de avaliação é atribuída aos observadores e/ou às posições resultantes de olhos alvo, a mencionada ordem sendo utilizada para a atribuição de unidades de computação.

7. MÉTODO de acordo com a Reivindicação 3, caracterizado pelo fato de que um ou diversos recursos de localização de faces são executados permanentemente e inicializam os recursos de localização de olhos / rastreamento de olhos respectivamente atribuídos.

15 **8. MÉTODO** de acordo com a Reivindicação 1, caracterizado pelo fato de que um recurso de controle que determina se uma região alvo de face encontrada pelo recurso de localização de faces resulta de uma face já rastreada ou de uma nova face detectada, atribui o recurso de rastreamento de olhos ou uma combinação do recurso de localização de olhos com o recurso de rastreamento de olhos às unidades de computação disponíveis, as inicializa e sincroniza a execução de todos os recursos.

20 **9. MÉTODO** de acordo com uma das reivindicações anteriores, caracterizado pelo fato de que os dados de imagem são adquiridos por múltiplos sensores de imagem, no qual um recurso de localização de olhos, um recurso de rastreamento de olhos ou uma combinação do recurso de localização de olhos com o recurso de rastreamento de olhos é executado para cada observador e para cada sensor de imagem, em unidades de computação dedicadas, e no qual resultados parciais e/ou finais são trocados e processados entre os recursos respectivamente relacionados a um observador.

30 **10. MÉTODO** de acordo com a Reivindicação 9, caracterizado pelo fato de que, conhecendo as posições dos sensores de imagem e dos observadores, bem como, parâmetros dos sensores de imagem, os resultados parciais e/ou finais dos recursos de um sensor de imagem são

transformados para os recursos de outro sensor de imagem e adaptados para a sua perspectiva.

5 **11. ARRANJO DE CIRCUITO** para a detecção e rastreamento de pontos de referência de posição (EP1, ..., EPn) de múltiplos observadores, no qual os pontos de referência de posição são as posições das pupilas e/ou dos cantos dos olhos, em dados de imagem de quadros de vídeo (VF) com múltiplas unidades de computação em comunicação (R1, ..., Rn), com um recurso de localização de faces (FF), um recurso localizador de olhos (EF) e um recurso de rastreamento de olhos (ET), caracterizado pelo fato de
10 compreender: recurso de localização de faces (FF) que serve para detectar as posições de faces em quadros de vídeo (FF) e extrair uma sub-região muito menor do quadro de vídeo completo como sendo uma região da face alvo (GZ) de cada face, e para transmitir informações da mencionada região da face alvo (GZ) a pelo menos um subseqüente; recurso de localização de olhos (EF) para
15 detectar posições de olhos, o mencionado recurso extraíndo uma sub-região muito menor de cada região da face alvo (GZ) como sendo uma região do olho alvo (AZ), para transmitir informações da mencionada região do olho alvo (AZ) a pelo menos um subseqüente; recurso de rastreamento de olhos (ET) para rastrear posições de olhos, o mencionado recurso detectando os pontos de referência de
20 posição dos olhos (EP1, ..., EPn) na região do olho alvo (AZ) do quadro de vídeo corrente e dos quadros de vídeo subseqüentes, e gerando informações para um dispositivo de rastreamento; no qual cada um dos recursos ou um dos recursos e uma combinação dos dois outros recursos são executados em paralelo um em relação ao outro e em uma unidade dedicada de computação.

25 **12. ARRANJO DE CIRCUITO** de acordo com a Reivindicação 11, caracterizado pelo fato de contar com um meio separado de computação para a escalagem ou correção dos dados de imagem com relação à escalagem, correção gama, controle de brilho, e recursos semelhantes.

30 **13. ARRANJO DE CIRCUITO** de acordo com a Reivindicação 11, caracterizado pelo fato de contar com um meio de computação que calcula as coordenadas de profundidade dos pontos de referência de posição (EP1, ..., EPn) determinadas a partir de pelo menos dois sensores de imagem.

14. ARRANJO DE CIRCUITO de acordo com a Reivindicação 11, caracterizado pelo fato de que pelo menos uma das unidades de computação é implementada como sendo uma CPU, DSP ou semelhante.

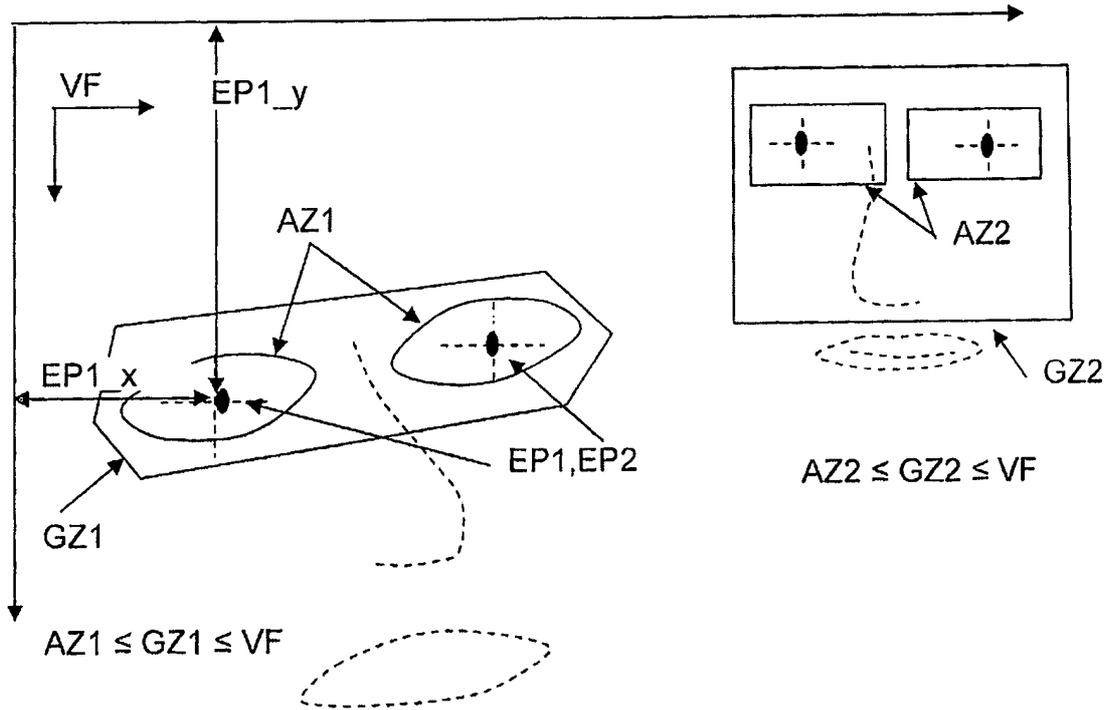


FIGURA 1

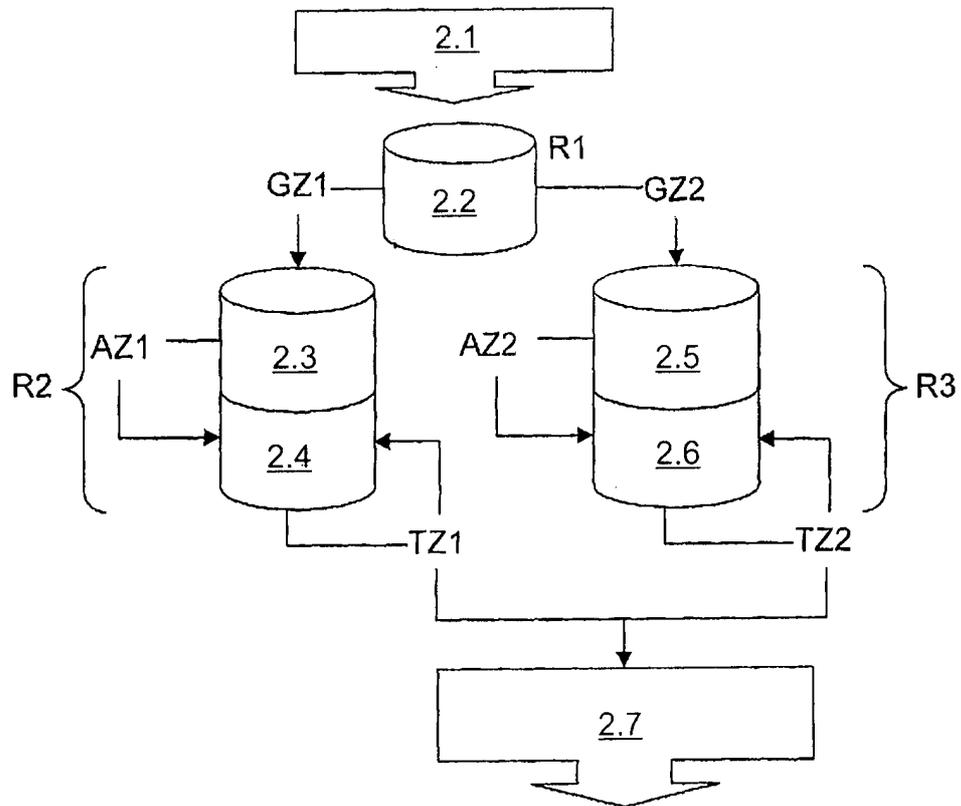


FIGURA 2

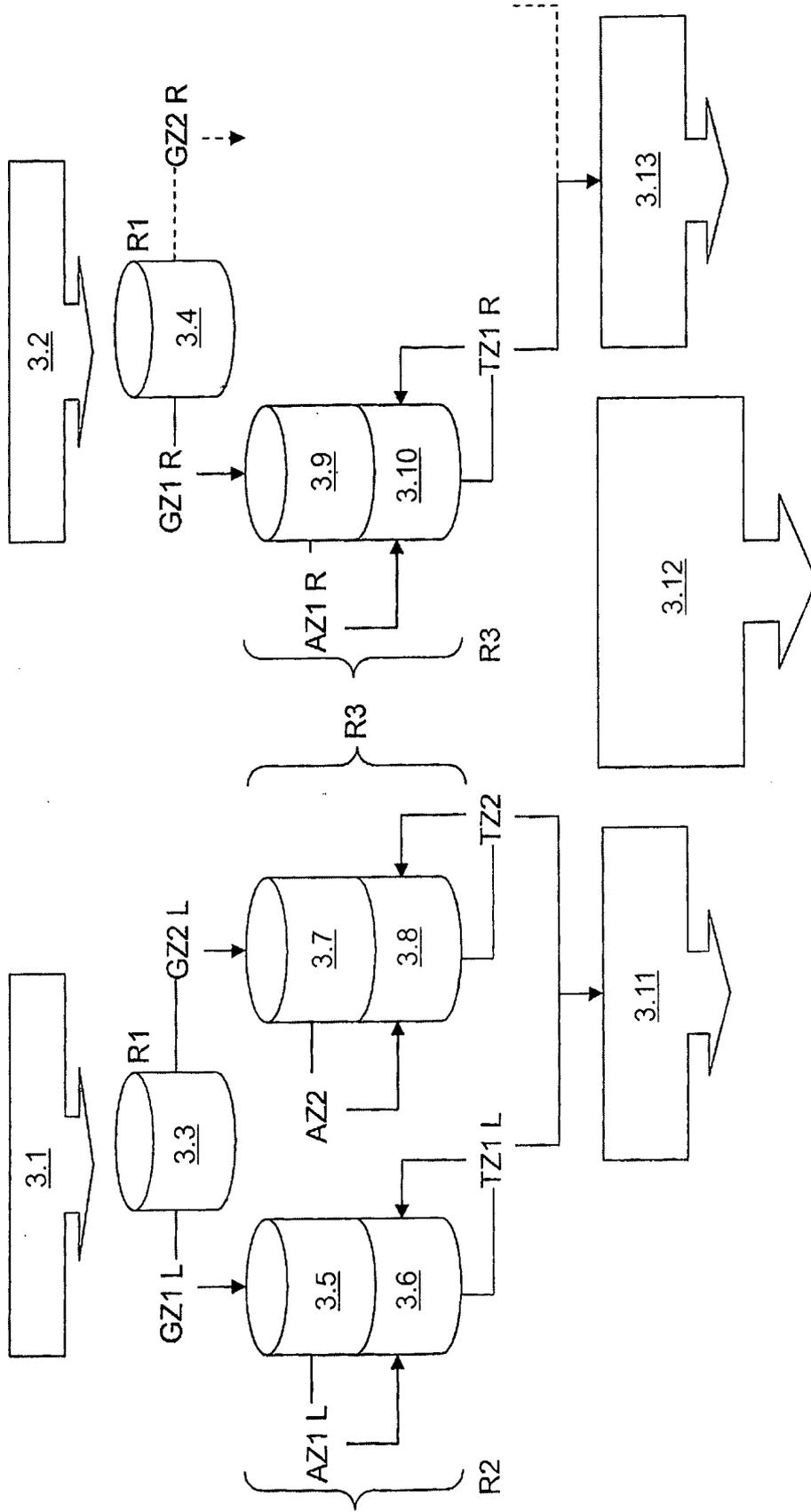


FIGURA 3

RESUMO

“MÉTODO E SISTEMA DE RECONHECIMENTO E RASTREAMENTO, EM TEMPO REAL, DA POSIÇÃO OCULAR DE MÚLTIPLOS USUÁRIOS”, que se refere a um método e a um arranjo de circuito para a

5 detecção e rastreamento isenta de contato de posições de olhos ou pupilas de múltiplos observadores em tempo real. Os dados de entrada consistem de uma seqüência de quadros digitais de vídeo. O método mencionado compreende os seguintes passos: a ação conjunta de um recurso de localização de faces para a

10 detecção de faces, seguido por um recurso de localização de olhos para a detecção de regiões de olhos, e um recurso de rastreamento de olhos para detectar e rastrear pontos de referência de olhos. O objetivo da presente invenção é determinar a posição dos olhos, dentro de uma rotina hierarquicamente organizada, que tem como objetivo reduzir gradualmente a região de procura, iniciando com o quadro completo de vídeo (VF) para a região da face alvo (GZ) e

15 subsequentemente para a região alvo dos olhos (AZ). Além disso, um recurso ou um grupo de recursos é executado em cada um dos casos em uma unidade dedicada de computação, enquanto processos separados estão sendo executados em paralelo. A invenção é usada para rastrear as posições dos olhos de usuários de *displays* autoestereoscópicos e em holografia de vídeo.