

Sistemas de Rastreamento de Baixo Custo para Desktops

Manuel E. Loaiza, Gustavo N. Wagner, Thiago A. Bastos,
Alberto B. Raposo, Marcelo Gattass

Grupo de Tecnologia em Computação Gráfica (Tecgraf)
Depto. de Informática – PUC-Rio

R. Marquês de São Vicente, 225, Gávea – 22453-900, Rio de Janeiro – RJ – Brazil

{manuel, gustavow, tbastos, abraposo, mgattass}@tecgraf.puc-rio.br

Abstract. *With the development of VR (Virtual Reality) technologies, several input and output devices have been created. There is a category of interaction devices that is becoming important due to the increasing availability of digital cameras. This kind of optical device uses the image captured by a set of digital cameras to track the position of objects and to provide positioning and orientation data for RV systems. This paper presents the development of two low-cost optical tracking devices with six degrees of freedom for desktop applications. The first device uses three webcams mounted on a box and detects the movements of a white marker over its black background. An alternative approach is also presented, using a single camera that, when moved by the user over a mark, works as a 3D mouse. Finally, the use of these devices in VR applications is presented.*

Resumo. *Com o desenvolvimento das tecnologias de RV (Realidade Virtual), diversos tipos de dispositivos de interação e de saída foram surgindo. Existe uma categoria de dispositivos de interação que está se destacando devido à crescente disponibilidade de câmeras digitais. Este tipo de dispositivo óptico utiliza a imagem capturada por um conjunto de câmeras digitais para rastrear a posição de um ou mais objetos e, com isso, fornecer dados de posicionamento e orientação para os sistemas de RV. Este artigo apresenta a implementação de dois dispositivos ópticos de rastreamento de baixo custo, com seis graus de liberdade, para aplicações desktop. O primeiro dispositivo usa três câmeras web montadas sobre uma caixa, detectando o movimento de um marcador branco sobre seu fundo preto. Também é apresentada uma abordagem alternativa que usa uma única câmera que, movimentada pelo usuário sobre um padrão, funciona como um mouse tridimensional. Finalmente, é mostrado o uso desses dispositivos em aplicações de RV.*

1. Introdução

Uma das conseqüências do advento da RV foi a necessidade de se redefinir o paradigma de interação humano-computador. O sistema tradicional mouse-teclado-monitor foi substituído por dispositivos que permitiram maior imersão do usuário no ambiente virtual e o manuseio de todas as potencialidades dessa tecnologia. O modo como os participantes interagem com o sistema de RV influencia enormemente suas experiências no ambiente virtual, facilitando seu uso, aumentando a sensação de imersão e ampliando a variedade de ações que podem ser tomadas pelo usuário.

Este novo paradigma de interação tem como uma conseqüência direta a criação de novos dispositivos de hardware, dirigidos a dar suporte a eventos de interação que se originam da experiência de trabalhar e interagir em um espaço 3D. São dispositivos especializados e orientados a dar suporte a estas novas formas de interação, também permitindo dar certo grau de imersão ao usuário dentro das aplicações 3D com as quais interage. Como este tipo de hardware não é planejado para ser de uso geral, sua construção normalmente tem um alto custo de produção, o que o leva a ficar confinado a laboratórios de pesquisa especializada com alto nível de investimento na área de RV. Este trabalho faz parte de uma linha de pesquisa que procura uma nova opção de dispositivo de interação para RV, que seja fácil de montar e tenha um baixo custo para ser construído.

Um importante dispositivo de interação em RV é o rastreador de posição, que pode ser utilizado para acompanhar a posição do corpo e os movimentos do usuário, assim como a posição de outros objetos sendo por ele utilizados. Existe uma variedade de dispositivos de rastreamento, com seis graus de liberdade, cada um utilizando uma tecnologia diferente, entre eles, os eletromagnéticos, mecânicos, acústicos, inerciais e ópticos. Entretanto, esses dispositivos ainda são extremamente caros, e apenas alguns deles são voltados para sistemas desktop (a maioria é direcionada para sistemas de projeção imersiva, como as CAVes). Neste artigo, são apresentadas experiências na construção e implementação de dois dispositivos ópticos de manipulação em 3D de baixo custo.

Para garantir o baixo custo de implementação, ambos dispositivos foram planejados para funcionar com câmeras de preço acessível (como câmeras web), executar em PCs com configurações “normais” e funcionar em quaisquer condições de luz.

O objetivo dessa pesquisa é criar dispositivos de rastreamento acessíveis para qualquer pessoa interessada em experimentar a possibilidade de interação e semi-imersão em aplicações 3D de RV. Adicionalmente, esse tipo de ferramenta permite a interação com as aplicações de RV em suas próprias mesas de trabalho, sem a necessidade de se deslocar para salas ou laboratórios com recursos especiais.

Na seção seguinte, alguns trabalhos relacionados são apresentados e o primeiro dispositivo de rastreamento é descrito. Na seção 3, uma abordagem alternativa é apresentada e, na seção 4, algumas aplicações que usam esses sistemas são mostradas.

2. O dispositivo da caixa

O dispositivo da caixa foi projetado com o objetivo de capturar e simular, em essência, a mesma arquitetura e funcionamento que apresentam os dispositivos de rastreamento óptico comerciais, utilizados para o rastreamento de áreas amplas que usam a luz infravermelha como componente elementar dentro de seu funcionamento. Estes dispositivos comerciais têm em sua arquitetura certas características em comum:

- Uma área física bem definida, onde se centra o campo de visão das câmeras e se faz o rastreamento dos marcadores definidos.
- Um conjunto de câmeras, que são colocadas a uma altura moderada da área de rastreamento para não atrapalhar a movimentação que o usuário possa fazer dentro desse espaço.
- Um conjunto de marcadores, que na maioria das vezes é esférico, e que serão os alvos no processo de rastreamento.

As características acima são refletidas no projeto criado para o dispositivo proposto. Porém, a maioria dos dispositivos ópticos para rastreamento do usuário em ambientes RV usa luz infravermelha para destacar os marcadores do ambiente. Dispositivos experimentais [Schoenfelder 2004] [Graf 2004] e produtos comerciais como o Vicon Motion System [Vicon 2004] tradicionalmente usam marcadores reflexivos especiais, fontes de luz infravermelha e câmeras capazes de registrar esse tipo de luz, funcionando em ambientes livres de quaisquer fontes de ruído infravermelho, como lâmpadas incandescentes ou luz solar.

Como um dos objetivos dos dispositivos de interação desenvolvidos nesse trabalho é evitar o uso de equipamentos especializados e permitir que os dispositivos funcionem em qualquer condição de luz, foi necessário eliminar o uso de luz infravermelha (ou outras formas especiais de luz como luz negra, também usada em dispositivos ópticos [Kim 2004]) no processo de rastreamento.

A luz infravermelha é usada em dispositivos de rastreamento como uma maneira de distinguir os marcadores de outros objetos no espaço físico de rastreamento. No dispositivo desenvolvido, esse resultado é obtido por meio do contraste entre as cores do fundo e dos marcadores.



Figura 1. Dispositivo da caixa

No dispositivo desenvolvido (Figura 1), câmeras web são montadas em um suporte e a área de trabalho é definida por uma caixa de $0.5 \times 0.5 \times 0.5\text{m}$, permitindo que o dispositivo seja posicionado ao lado de um PC, em ambientes de trabalho comuns. Para permitir o contraste necessário para o processo de rastreamento, o fundo da caixa é preto e os marcadores são brancos (o oposto é possível, mas as sombras criadas pelo braço do usuário sobre um fundo branco podem atrapalhar o processo de detecção).

Além da caixa com fundo preto, o dispositivo é montado com:

- Três câmeras web, duas Logitech QuickCam 4000 com sensor CCD e uma Creative NX Pro com sensor CMOS.
- Duas lâmpadas, usadas em situações de pouca iluminação.
- Um suporte de alumínio para as câmeras web.
- Esferas de isopor brancas, com 25mm de diâmetro, montadas em uma forma triangular por meio de arames.

O desenho do dispositivo permite o livre movimento dos marcadores dentro da caixa. Esses marcadores são pequenas esferas brancas, agrupadas em uma forma triangular, sendo rastreadas individualmente ou como um objeto único (Figura 2).

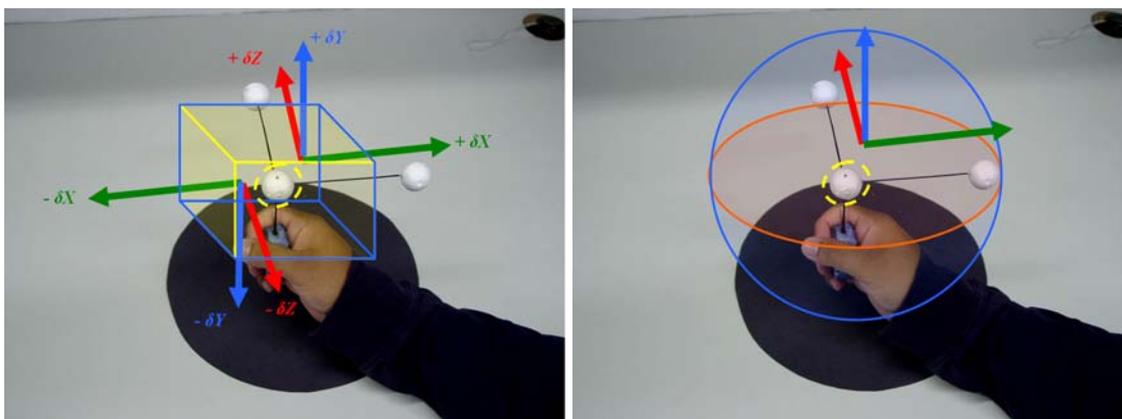


Figura 2. Objeto triangular rastreado e os seis graus de liberdade

As câmeras são conectadas a portas USB em um único PC. Os *streams* de vídeo são capturados com resolução de 320×240 . Componentes móveis no suporte de alumínio permitem que o usuário adapte o campo de visão das câmeras de acordo com as características da aplicação a ser usada e do tamanho do ambiente de trabalho disponível.

Algumas características adicionais também foram levadas em conta na construção do dispositivo, em razão do uso e tipo de aplicações ao qual está dirigido seu projeto físico:

- O desenho relativamente compacto permite colocar o dispositivo dentro um ambiente desktop comum.
- As dimensões do espaço de trabalho foram projetadas para suportar o rastreamento da mão do usuário, permitindo liberdade de movimentos na hora

do rastreamento dos marcadores, os quais poderão estar colados na mão ou serem agrupados em padrões específicos, que também podem ser controlados pela mão do usuário.

2.1. Processo de rastreamento

Com a finalidade de recuperar a posição e orientação dos marcadores enquanto eles se movimentam dentro do espaço de rastreamento, é necessário definir o conjunto de processos, implementados com base em técnicas de visão computacional para dar apoio às tarefas de captura, identificação e reconstrução 3D dos marcadores. Este conjunto de processos pode ser resgatado em trabalhos como [Kutulakos 1998], [Dorfmüller-Ulhaas 2001] e [Ribo 2001], que descrevem alguns sub-processos base, implementados na maioria dos sistemas de rastreamento ópticos. Este conjunto de processos pode se resumir no seguinte fluxo:

- Captura e processamento das imagens do *stream* de vídeo.
- Calibração das câmeras.
- Correlação e identificação dos marcadores específicos nas diferentes imagens de vídeo.
- Reconstrução das posições 3D dos marcadores.

O primeiro passo (captura e processamento) é implementado convertendo as imagens capturadas para escala de cinzas e depois aplicando filtros gaussianos e de limiarização nas mesmas. A Figura 3 ilustra o resultado após a aplicação desses filtros. Nota-se, na mesma figura, que a mão e o braço do usuário também se destacam do fundo preto. Portanto, se faz necessário extrair apenas os contornos circulares dos pontos de referência a serem rastreados e, com esta informação, calcular o centro dessas áreas circulares. A estratégia para achar o centro das áreas definidas pelos contornos circulares na imagem binária é aproximar o centro como sendo a media aritmética das coordenadas 2D dos pontos que formam o contorno das áreas brancas que representam os marcadores.



Figura 3. Imagens binárias capturadas pelas 3 câmeras. No alto à direita está a imagem original capturada pela câmera do meio.

O segundo subprocesso na implementação do sistema de rastreamento óptico é a calibração de cada câmera web utilizada. A calibração tem como objetivo encontrar uma medida de relacionamento entre o mundo 3D, ou seja, o espaço físico que é visto pela câmera, e o plano da imagem 2D que se vê na tela do computador visualizando o espaço físico capturado. Esta medida de relacionamento é expressa como uma transformação que inclui uma matriz de rotação e um vetor de translação que nos permitam levar coordenadas do mundo 3D físico – medidas tomadas em milímetros – para posições 3D do sistema de coordenadas internas definidas para cada câmera web – medidas em escala livre – e, em um último passo, transformar as coordenadas 3D expressas no sistema de cada câmera em posições 2D no plano da imagem que é desenhado no computador – medidas expressas em pixels. O plano da imagem está posicionado em uma distância f (distância focal expressa em coordenadas do sistema da câmera).

Para a implementação do dispositivo proposto decidiu-se utilizar o método TSAI não-coplanar [Tsai 1986], [Tsai 1987] que se baseia na captura de um padrão 3D conhecido, do qual são conhecidas as medidas físicas de certos pontos de referência. Este método é descrito na Figura 4.

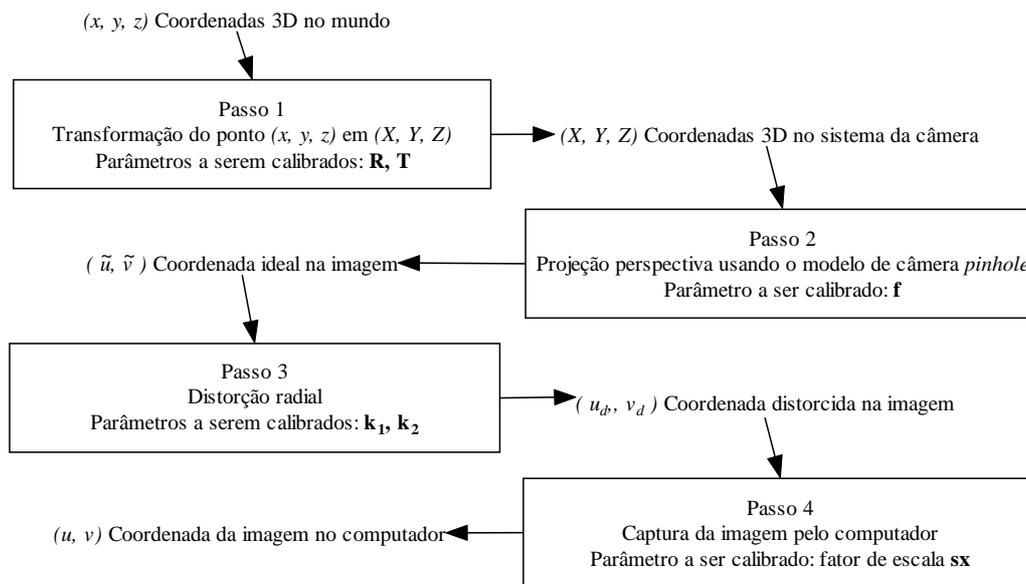


Figura 4. Descrição da transformação de coordenadas no mundo em coordenadas da imagem.

O terceiro subprocesso na implementação do sistema de rastreamento óptico é a correlação entre os diferentes marcadores que aparecem nas três imagens das câmeras web. Um caso intuitivo, é quando há apenas um marcador nas imagens. Nesse caso, a estratégia seria só pegar a coordenada 2D do único ponto em contraste que se consegue ver em cada imagem das câmeras. Porém até mesmo neste caso pode acontecer o problema de ruído, que pode criar um falso marcador nas imagens, precisando ser tratado por técnicas de filtragem. O caso mais complexo é quando há mais de um marcador nas imagens e, com isto, é necessário fazer uma correlação entre “n” marcadores detectados nas câmeras. Em outras palavras, precisa-se definir qual é cada

marcador de uma imagem base em relação aos que aparecem nas imagens das outras câmeras.

Para solucionar este problema de correlação entre os diferentes marcadores que aparecem nas imagens, utilizaremos conceitos da geometria epipolar para fazer uma correlação inicial entre os marcadores detectados nas imagens. Segundo a geometria epipolar, as linhas epipolares projetadas devem cair sobre os marcadores que aparecem nas outras câmeras web, e a precisão no cálculo das linhas epipolares depende só da quantidade de pontos de amostra com os quais se criem as matrizes fundamentais. Adicionalmente, foram implementadas heurísticas auxiliares orientadas a resolver casos especiais que acontecem no momento de capturar e que podem gerar problemas na correlação, tais como o problema de perda de marcadores por oclusão na imagem de alguma câmera. Em casos de oclusão entre os marcadores, também é usado o filtro de Kalman [Welch 2001] para estimar a posição do marcador e evitar a interrupção do processo de rastreamento.

Finalmente, a última etapa na implementação do sistema de rastreamento óptico é a reconstrução 3D dos diferentes marcadores que aparecem nas imagens das câmeras web. O método escolhido para esse passo é descrito em [Ribo 2001]. Para fazer a reconstrução toma-se como base os pontos em coordenadas da imagem que vêm da etapa de calibração e as tabelas de correlação criadas na etapa seguinte. O resultado final serão pontos 3D que são escolhidos como correspondentes aos marcadores físicos rastreados. O resultado final da reconstrução 3D será usado na etapa seguinte para rastrear os movimentos do conjunto de marcadores como um objeto único, a partir do qual se faz a recuperação dos 6 graus de liberdade que este objeto tem dentro do espaço de trabalho.

2.2. Eventos de entrada

O passo final no processo de rastreamento é gerar eventos que permitam ao usuário interagir com as aplicações de RV. Atualmente o dispositivo da caixa é usado em três configurações: para interação com 3 graus de liberdade, para interação com 6 graus de liberdade e para interação com 3 graus de liberdade com abertura dos dedos.

Na configuração de 6 graus de liberdade, a interação ocorre como em um mouse 3D, como o Spaceball 5000 [3dconnexion 2004]. A posição e orientação rastreadas dos marcadores são enviadas para a aplicação, onde podem ser usadas para movimentar objetos ou navegar dentro da cena virtual. Para obtermos a informação de orientação, é necessário o uso de pelo menos 3 marcadores.

A configuração de 3 graus de liberdade funciona de maneira semelhante à anterior, mas sem a captura da orientação dos marcadores. Nesse caso, a posição do marcador pode ser obtida usando-se apenas um marcador.

A última configuração expande a captura da posição do marcador para permitir a captura da abertura e do fechamento de um par de dedos. Ao invés de um único marcador, dois marcadores são usados nas pontas dos dedos indicador e polegar do usuário. A abertura dos dedos é calculada a partir da distância entre os dois marcadores.

Nas duas configurações em que dois ou mais marcadores são usados, a posição transmitida para a aplicação deve ser definida. Na configuração com 6 graus de

liberdade, esta é a posição da esfera que fica no ângulo reto do triângulo (Figura 2). Na configuração com 3 graus de liberdade com abertura, essa posição é calculada a partir da média da posição das outras duas bolinhas.

As informações de posição e orientação podem ser interpretadas de duas maneiras bem distintas. Na primeira, chamada de mapeamento absoluto, cada posição e orientação possível dentro da caixa são mapeadas para uma posição e orientação no mundo virtual. Na outra maneira, definimos uma região neutra na caixa, dentro da qual nenhum evento é gerado. Quando o dispositivo deixa essa posição neutra, eventos são gerados para movimentar o objeto virtual na direção assumida. O objetivo da região neutra é criar um volume onde pequenos movimentos involuntários da mão do usuário não geram eventos inconvenientes. A orientação dos marcadores também respeita uma área neutra, definida por um ângulo máximo a partir de uma orientação inicial pré-definida. Posições assumidas fora dessa área neutra geram eventos de rotação que manipulam o ambiente virtual.

2.3. Resultados experimentais

Na implementação do processo de rastreamento óptico, os sub-processos de calibração de câmera, correlação e reconstrução dos marcadores foram submetidos a testes com o objetivo de conhecer a precisão dos parâmetros e elementos calculados por eles.

No processo de calibração, se mediu a precisão das matrizes calculadas para cada câmera. Os testes foram feitos projetando os pontos da amostra de pontos 3D, usadas no processo de calibração, sobre as imagens 2D de cada câmera utilizando as matrizes calculadas. Então, se mediu a diferença entre as coordenadas 2D calculadas e as coordenadas 2D válidas que foram obtidas na etapa de captura da amostra de pontos. Obteve-se um erro médio de ± 1.8 pixels, entre as coordenadas válidas e as calculadas. Esta precisão é suficiente para nossa experiência.

No processo de correlação é importante analisar a precisão da matriz fundamental calculada, que relaciona cada par de câmeras. O teste realizado consistiu em projetar cada linha epipolar gerada pelos pontos capturados na câmera pivô sobre os planos da imagem das câmeras subordinadas (da esquerda e direita). Para cada linha se calcula a distância do ponto correspondente capturado na câmera subordinada analisada até a linha projetada. O erro neste caso foi definido pela distância média dos pontos às linhas. O erro obtido foi de ± 0.9 pixels. Esta foi uma taxa de erro baixa, bastante útil para garantir a precisão da correlação.

No processo de reconstrução foi testada a precisão das posições 3D calculadas para cada marcador. De novo se utilizaram as coordenadas 3D da amostra de pontos usada no processo de calibração para fazer o teste. Foi calculada a distância média entre as coordenadas 3D conhecidas da amostra de pontos de calibração, e as coordenadas 3D calculadas com o algoritmo de reconstrução aplicado sobre as coordenadas 2D conhecidas também para a amostra de pontos da calibração. O resultado foi um erro médio de ± 8.0 mm.

Finalmente, outro dado importante é que o processo de rastreamento óptico manteve uma taxa de captura maior que 20fps (quadros por segundo), que é o mínimo

permitido para se ter um rastreamento constante em tempo real, evitando assim tremores ou inconsistências. A taxa original de captura das câmeras é de 30fps a uma resolução de 320×240. Após a execução recorrente das etapas de correlação e reconstrução, que são constantemente executadas dentro do processo de rastreamento óptico, a taxa média obtida foi de 25fps. Esta é uma boa taxa de desempenho especialmente considerando que todo o processo é executado em uma só CPU, que é a mesma ao qual as três câmeras estão conectadas.

Estes resultados mostraram que os parâmetros calculados e métodos usados dão uma boa precisão ao processo de rastreamento, refletindo isto no desempenho global do dispositivo.

3. O dispositivo com câmera livre

Em outro dispositivo estudado (Figura 5), a câmera assume um novo papel. Ao invés de ser posicionada ao redor da área de manipulação, ela é manipulada pelo usuário, como se fosse uma *wand* ou um mouse espacial. Sua posição e orientação, são calculadas da imagem que a câmera obtém de um grupo de padrões de referência posicionados na área de trabalho.

Uma abordagem similar é implementada em [Piekarski 2004], onde múltiplas câmeras e marcadores de referência são usados para rastrear a posição aproximada do usuário dentro de um prédio. A maior diferença desse trabalho se deve ao fato dele trabalhar em uma escala muito maior e, conseqüentemente, precisar de menos precisão.

3.1. Hardware e processo

O dispositivo usa uma única câmera web (atualmente, uma Logitech QuickCam Pro 4000) e um conjunto de padrões de referência impressos em papel comum (Figura 5).

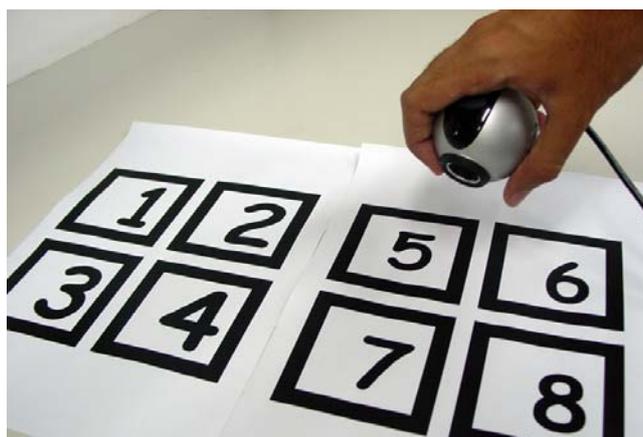


Figura 5. Câmera e padrões de referência

O processo de detecção da posição e orientação da câmera é muito similar ao usado no ARToolKit [ARToolKit 2004], em que cada padrão de referência é usado para se obter uma matriz que leva pontos das coordenadas locais do padrão para coordenadas locais da câmera. Essa matriz, que funciona exatamente como a matriz *modelview* do

OpenGL é usada no ARToolKit para desenhar objetos sobre a imagem capturada, criando a ilusão de que estão realmente posicionados sobre os padrões.

O dispositivo de rastreamento desenvolvido obtém essa matriz que, em sua essência, relaciona a posição do padrão com a posição da câmera no mundo real, nos seguintes passos:

- Uma imagem do espaço de trabalho é capturada pela câmera
- A imagem é processada e todos os possíveis padrões de referência são encontrados. Esse passo procura todas as regiões escuras da imagem e seleciona aquelas em que suas fronteiras formam polígonos com quatro lados.
- Em seguida, cada possível padrão encontrado é comparado com uma imagem previamente armazenada. Essa comparação permite que falsos padrões sejam ignorados e que tipos distintos de padrões sejam identificados.
- Pontos chave em cada padrão encontrado são usados para a calibração com o método de TSAI coplanar [Tsai 1987]. Atualmente, apenas os quatro vértices do quadrilátero externo do padrão são usados. Os 8 pontos necessários para a calibração com o TSAI coplanar são encontrados pela interpolação dos 4 pontos conhecidos, levando em consideração a perspectiva da imagem capturada.
- A posição detectada é filtrada com o filtro de Kalman e o resultado é transmitido para a aplicação.

3.2. Eventos de Entrada

A conversão da posição e orientação da câmera para a cena virtual é direta, sendo necessárias apenas correções de escala ou orientação que podem ser feitas com transformações elementares.

Existem duas formas de interpretar os movimentos realizados quando a câmera tem sua visão bloqueada ou não consegue capturar nenhum padrão de referência. Na forma mais simples, assim que a posição da câmera é obtida novamente, o movimento realizado pela câmera é transmitido abruptamente para a cena virtual, podendo causar falhas visuais. Na outra forma, a movimentação realizada durante esse período é ignorada. Isso evita as falhas visuais e ainda permite um efeito interessante, que permite estender a área útil do dispositivo indefinidamente, se necessário. Como o dispositivo ignora movimentos feitos quando nenhum padrão de referência é encontrado, podemos obstruir a visão da câmera e movê-la de volta ao centro da área de trabalho, para continuar um movimento, de maneira semelhante a quando erguemos um mouse quando este chega ao final do *mouse pad*.

4. Aplicações

4.1. Navegação de um robô submarino

A primeira aplicação RV onde o dispositivo da caixa foi extensivamente testado consistiu na navegação de um robô submarino. Esta aplicação usa o dispositivo com a configuração de 6 graus de liberdade, mencionada na seção 2.2. A aplicação (Figura 6) tem dois componentes:



Figura 6: Aplicação do Robô Submarino

- O primeiro é o cenário virtual que representa um campo de exploração petrolífera no fundo do mar, onde se tem como objetivo navegar em volta das diferentes estruturas distribuídas no fundo.
- O segundo é um robô submarino, que é o objeto utilizado pelo usuário para navegar no cenário virtual.

O dispositivo tem como objetivo permitir que o usuário controle o robô submarino da forma mais natural possível, isto é, ver como os movimentos feitos com sua mão, diretamente refletidos sobre a posição dos marcadores, em diferentes direções e fazendo giros, têm uma correspondência com os movimentos de translação e rotação que o robô faz enquanto navega no cenário virtual.

Aqui é importante citar como medida qualitativa de desempenho o grau de dificuldade na adaptação do usuário ao dispositivo para a realização da tarefa de navegação e controle de um objeto virtual dentro do cenário 3D onde interage. A precisão da reconstrução e a uniformidade do rastreamento se vêem refletidas na forma suave e coerente dos movimentos que o robô faz, isto em razão de diminuir possíveis saltos quando os marcadores têm problemas como oclusão.

Outras aplicações onde o dispositivo foi testado foram um “Pintor 3D” e um “Braço de Robô Virtual”, este último também utilizado em um ambiente de exploração petrolífera submarina. Estas duas aplicações são interessantes de apresentar porque mostram uma outra estratégia de interpretação e emissão de eventos para o dispositivo da caixa, que está baseada diretamente no rastreamento de marcadores dentro do espaço absoluto da caixa que compõe o dispositivo. A idéia básica é fazer um escalamento das dimensões do espaço de rastreamento (volume da caixa preta) para dimensões do mundo virtual, com a finalidade de recuperar posições 3D de dentro da caixa do dispositivo, que terão uma posição 3D correspondente no mundo virtual.

4.2. Braço de robô virtual

O simulador de braço de robô virtual tem como objetivo controlar a movimentação de um braço de robô com 3 graus de liberdade e, adicionalmente, controlar o evento de

abrir e fechar a garra que tem o robô na ponta de seu braço. A aplicação simula cenários de operações submarinas usando dinâmica de corpo-rígido, visão estéreo e som 3D. Para fazer isto, a posição da ponta do braço robô é controlada com a movimentação da mão do usuário dentro do espaço de rastreamento. Para isto, se extrai a posição 3D média entre os marcadores colocados em uma luva na mão do usuário e, em todo instante da movimentação da mão, calcula-se o valor médio entre as coordenadas 3D dos marcadores (Figura 7). A partir da posição da ponta do braço, rastreada pelo dispositivo, as posições das demais juntas do robô são calculadas por cinemática inversa.

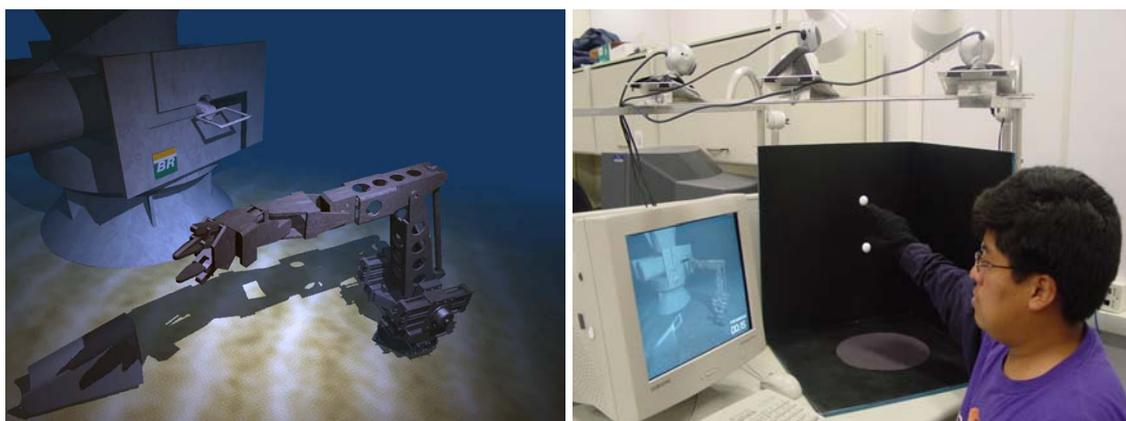


Figura 7. Aplicação do Braço Robô Virtual controlada pela mão do usuário

Os eventos de abrir e fechar a garra também estão ligados ao movimento da mão do usuário. Utilizamos os dois marcadores colocados nos dedos indicador e polegar e, fazendo uma analogia com a movimentação dos dedos, se detecta se eles chegam a juntar, até ter uma distância mínima entre os marcadores representando que a garra está se fechando, ou se eles se separam representando que a garra está se abrindo.

Esta adaptação de um evento adicional mostra a flexibilidade que o sistema de rastreamento do dispositivo dá ao usuário para criar novos eventos dependendo da tarefa que se queira realizar dentro da aplicação 3D. Nessa aplicação, o usuário opera uma réplica virtual de um braço manipulador real montado em um veículo operado remotamente, usado para operações submarinas em águas profundas.

Como esta aplicação é baseada no ViRAL [Bastos et al. 2004], um framework de RV bastante flexível, não foi muito problemático testar e integrar diferentes dispositivos de entrada. Na tentativa de se utilizar um Spaceball 5000, que tem fácil utilização e preço acessível, a manipulação do braço não se mostrou tão intuitiva. Claramente, os melhores resultados vieram de dispositivos que podem rastrear a posição da mão do usuário. Os dados obtidos dessa posição podem ser usados, com cinemática inversa, para reconstruir a posição de cada junta do braço virtual e permitir métodos mais intuitivos de interação.

Geralmente, dispositivos de rastreamento não são tão acessíveis como um mouse 3D, apesar de algumas soluções para espaços limitados terem preços razoáveis. Nessa categoria, o Ascension's Flock of Birds [Ascension 2004] pareceu ser uma boa escolha. Foi usado um sensor amarrado ao pulso do usuário e resultados satisfatórios foram

obtidos. Mesmo assim, por ser um dispositivo magnético de rastreamento com fio, o Flock of Birds tem suas desvantagens, especialmente para o uso em desktops.

O nosso dispositivo de rastreamento da caixa pareceu ser uma boa alternativa para essa aplicação. Além de ser barato, ele pode operar na maior parte dos ambientes e não requer fios presos ao usuário. Além disso, pôde-se customizar o dispositivo para rastrear, além da posição do marcador, a abertura dos dedos do usuário, o que nos permitiu controlar a abertura da garra do braço manipulador no mesmo dispositivo, o que não é possível no Flock of Birds.

4.3. Pintor 3D

O pintor 3D é mostrado na Figura 8. Ele é composto por um cenário 3D com latas de tinta de diferentes cores e um pincel flutuante que será o objeto controlado pelo usuário.



Figura 8. Pintor 3D.

O objetivo desta aplicação é fazer desenhos livres 3D no cenário virtual e ela foi implementada usando os dois dispositivos apresentados neste artigo. A utilidade desses dispositivos está na facilidade que o usuário tem para realizar os desenhos utilizando movimentos da sua mão. Para isto, no caso do dispositivo da caixa, somente precisamos colocar dois marcadores na mão do usuário e fazer o rastreamento dentro do espaço definido na caixa preta. A posição média entre os marcadores reconstruídos em coordenadas 3D serão as que guiarão a movimentação do pincel pelo cenário virtual. Nesta abordagem, só estamos resgatando movimentos de translação absoluta dos marcadores, utilizando para isto só três graus de liberdade. No segundo dispositivo, a própria câmera, movida pela mão do usuário, representa o pincel virtual.

As vantagens que este tipo de adaptação apresenta é a sensação de imersão e liberdade que o usuário tem enquanto desenha livremente com sua mão dentro do cenário virtual. Além disso, o usuário percebe como seus movimentos são diretamente representados no mundo virtual. A interação entre o usuário e a aplicação é natural, não precisando de maior treinamento para fazer uso dos dispositivos e interagir com a aplicação.

5. Conclusão e trabalhos futuros

Os dois dispositivos apresentados nesse trabalho podem ser alternativas viáveis para os dispositivos comerciais existentes, pelo menos para aplicações em que precisão milimétrica não seja necessária. O primeiro dispositivo funciona com três câmeras estáticas e um conjunto de marcadores dentro de um espaço restrito. O segundo usa uma câmera livre e marcadores de referência, impressos em papel. Enquanto o segundo dispositivo usava técnicas que permitiam menos precisão que o primeiro, ele se justifica por ser menor e mais fácil de ser transportado e montado.

Os exemplos apresentados mostram que é possível construir dispositivos de rastreamento óptico de baixo custo usando tecnologia acessível. Esses dispositivos, em muitos casos, são mais intuitivos para manipulação de ambientes virtuais do que soluções tradicionais, como joysticks ou mouses 3D.

No futuro, pretende-se aumentar o número de marcadores suportados para permitir o uso de duas mãos na mesma caixa e tentar reconhecer automaticamente diferentes configurações de marcadores no dispositivo da caixa. Também pretende-se aumentar o número de marcadores de referência rastreados para aumentar a área efetiva do dispositivo com câmera livre. Finalmente, é também importante estudar técnicas de interação que tirem vantagem dos dispositivos apresentados nas aplicações de RV.

Alguns vídeos mostrando as aplicações e dispositivos apresentados neste artigo são encontrados em <http://www.tecgraf.puc-rio.br/~manuel/videos/>

Agradecimentos

A pesquisa em Realidade Virtual do Tecgraf/PUC-Rio é apoiada primordialmente pelo CENPES/PETROBRAS, pela FINEP e RNP (projeto Rede GIGA).

Referências

- 3D Connexion - www.3dconnexion.com. Acesso em março/2004.
- ARToolKit. <http://www.hitl.washington.edu/artoolkit/>. Acesso em março/2004
- Ascension Technology Corporation- www.ascension-tech.com. Acesso em março/2004.
- Bastos, T. A., Silva, R. J. M., Raposo, A. B., and Gattass, M. (2004). “ViRAL: Um Framework para o Desenvolvimento de Aplicações de Realidade Virtual”, VII Symposium on Virtual Reality (SVR), pp. 51–62.
- Dorfmueller-Ulhaas, K., Schmalstieg, D. (2001) “Finger Tracking for Interaction in Augmented Environments”, IEEE and ACM International Symposium on Augmented Reality (ISAR'01), pp. 55-66.
- Graf, H., Koch, M., Stork, A. (2004) “Cyberstilo, Towards an Ergonomic and Aesthetic Wireless 3D-Pen”, Beyond Wand and Glove Based Interaction, IEEE VR 2004 Workshop, pp. 51-54.
- Kim, H., Fellner, D.W. (2004) “Interaction with Hand Gesture for a Back-Projection Wall”, Computer Graphics International (CGI'04), pp. 395-402.

- Kutulakos, K., Vallino, J. (1998) "Calibration-Free Augmented Reality", IEEE Transactions on Visualization and Computer Graphics, 4(1), pp. 1-20.
- Piekarski, W., Avery, B., Thomas, B.H., Malbezin, P. (2004) "Integrated Head and Hand Tracking for Indoor and Outdoor Augmented Reality", IEEE VR 2004, pp.11-18.
- Ribo, M., Pinz, A., Fuhrmann, A.L. (2001) "A New Optical Tracking System for Virtual and Augmented Reality Applications", IEEE Instrumentation and Measurement Technology Conference, vol. 3, pp. 1932-1936.
- Schoenfelder, R., Maegerlein, A., Regenbrecht, H. (2004) "TACTool: Freehand Interaction With Directed Tactile Feedback", Beyond Wand and Glove Based Interaction, IEEE VR 2004 Workshop, pp. 13-15.
- Tsai, R.Y. (1986) "An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision", IEEE Conf. on Computer Vision and Pattern Recognition, pp. 364-374.
- Tsai, R. Y. (1987) "A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses", IEEE Journal of Robotics and Automation, RA-3(4), pp. 323-344.
- Vicon Motion Systems Ltd. - www.vicon.com. Acesso em março/2004.
- Welch, G. Bishop, G. (2001) "An Introduction to the Kalman Filter", Course 8, SIGGRAPH.