

Laboratório VISGRAF

Instituto de Matemática Pura e Aplicada

Rastreamento de Movimento com Marcadores Óticos

Sergio Biasi

Luiz Velho, Bruno Feijo (orientadores)

Technical Report

TR-04-08

Relatório Técnico

April - 04 - April

The contents of this report are the sole responsibility of the authors.
O conteúdo do presente relatório é de única responsabilidade dos autores.



Departamento de Informática

Proposta de Dissertação de Mestrado

Versão de 26.02.2004

Sergio Coutinho de Biasi
Candidato a Mestre em Informática

**“Rastreamento de
Movimento
com Marcadores Óticos
Retrorefletores”**

Professor Orientador : Bruno Feijó (PUC-Rio)

Professor Co-Orientador : Luiz Velho (IMPA)

1. Apresentação

Venho através deste documento, como candidato a mestre pelo departamento de informática da Pontifícia Universidade Católica do Rio de Janeiro, apresentar minha proposta para dissertação de mestrado.

Minha dissertação se situa na área de visão computacional, um campo importante da computação gráfica. Pretendo apresentar nela a descrição da teoria, projeto e construção de um sistema de rastreamento de movimento baseado na análise digital por computador de imagens obtidas com o auxílio de câmeras de vídeo retroiluminadas e marcadores ótico retrorefletores. Acredito ser este um tema de substancial relevância acadêmica e prática e que abrirá caminho para várias linhas de aprofundamento de pesquisa.

2. A Área de Visão Computacional

A área de visão computacional sempre foi uma das mais importantes da computação gráfica, contendo um grande número de problemas teóricos importantes assim como uma extensa variedade de aplicações tecnológicas de alto interesse prático. Infelizmente, essa é também uma das áreas com os requisitos mais fortes em termos de capacidade computacional, especialmente para as aplicações em tempo real. Por esse motivo, enquanto nos últimos cinquenta anos outras áreas da ciência da computação passaram por grandes avanços e sínteses teóricas no aspecto acadêmico e se traduziram em aplicações acessíveis às massas no campo tecnológico, a visão computacional permaneceu mais ou menos restrita em seus progressos, prejudicada pela dificuldade de realização de implementações práticas.

Em tempos recentes (em particular nas últimas duas décadas), porém, assistimos a uma explosão sem precedentes do poder computacional bruto disponível para todos os tipos de aplicações, associada a um barateamento e disponibilidade quase universal também sem precedentes desses recursos. Concomitantemente, observamos um salto tecnológico e de mercado similar na área de equipamentos de captura de vídeo.

A associação desses dois fatos permitiu o desenvolvimento e a aplicação de um grande número de técnicas antes tecnologicamente inviáveis na área de visão computacional. Inicialmente com a disponibilidade e baixo custo de scanners de qualidade cada vez maior, seguida pela revolução das câmeras de fotografia digital, e finalmente entrando numa era em que a captura de vídeo digital está se tornando uma tecnologia madura e universalmente acessível, foi se tornando cada vez mais fácil para o pesquisador assim como para o industrial experimentar com novas aplicações envolvendo a interpretação de imagens com o uso de computadores.

Dessa forma a área de visão computacional é um dos campos de pesquisa que tem crescido mais rapidamente nos últimos anos, e que tem gerado mais aplicações inovadoras na área tecnológica. É particularmente relevante ressaltar que numa época em que a interação com os computadores se torna cada vez mais freqüente e necessária, a visão computacional contém a promessa de abrir toda uma vasta área de interfaces mais eficazes, eficientes e intuitivas entre o mundo humano e o dos computadores. Os computadores se tornarão capazes de ver e interpretar o que nós mesmos vemos e interpretamos, e isso muito provavelmente terá implicações revolucionárias tanto em termos tecnológicos e científicos como sociais.

Citamos a seguir alguns exemplos (de forma alguma exaustivos) de aplicações da visão computacional que já hoje em dia estão tendo aplicação significativa, em alguns casos com conseqüências já revolucionárias :

- Controle de qualidade em linhas de produção. Controle automatizado de tamanho, perfil, continuidade e diversas outras características de peças produzidas em série que antes deviam ser verificadas por um operador humano.
- Reconhecimento automático de caracteres e textos em imagens digitais planas. Conversão para formato digital com indexação automática de conteúdo e referências cruzadas de enormes massas de documentos em papel, incluindo décadas de artigos científicos cuja consulta em busca de pesquisa pré-existente foi facilitada em ordens de grandeza.
- Reconhecimento automático de marcas e caracteres em fotos digitais. Reconhecimento automático de placas de veículos em trânsito, permitindo um novo paradigma em termos de controle de infrações e de veículos roubados.
- Robótica. Robôs que conseguem reagir a modificações do ambiente e operar corretamente sem a necessidade de uma programação prévia milimetricamente rigorosa de todos os movimentos que devem realizar.
- Reconhecimento de faces, reconhecimento de íris, reconhecimento de impressões digitais, reconhecimento de formato das mãos e outros. Vários métodos de identificação e autenticação para aplicações de segurança são baseados em informações biométricas como essas. Todos esses métodos citados já são comercialmente utilizados, e existe uma possibilidade concreta de que venham a substituir ou complementar, no dia-a-dia, os métodos usuais de autenticação (objetos, senhas).
- Fotografia 3D. A aquisição da estrutura tridimensional e da textura da superfície de objetos físicos é uma tecnologia que já está em uso atualmente, e que possui um grande número de aplicações, entre elas a geração posterior de imagens sintéticas foto realistas de um objeto a partir de qualquer ponto de vista.
- Interfaces de usuário. Os movimentos corporais, o foco de visão e a expressão do usuário podem ser usados como parte de um sistema de interface com programas de computador. Já existem várias aplicações concretas desse tipo de tecnologia, algumas com aplicação comercial bem sucedida em larga escala, como por exemplo na indústria de jogos e em realidade virtual.
- Realidade aumentada. Consiste na modificação de cenas reais com o uso de computador, por exemplo acrescentando legendas ou informações, ou realçando, acrescentando ou retirando objetos. Se feita em tempo real pode ser uma forma poderosíssima de interface, por exemplo para um piloto de jato localizar seus alvos.
- Aquisição automática de alvos. É uma tecnologia que permitiu, por exemplo, a existência da geração mais moderna mísseis inteligentes capazes de “achar” o alvo que devem atingir procurando-o ativamente em seu campo de visão.

3. O Problema da Captura de Movimento

Entre todas as aplicações que a revolução de visão computacional está permitindo, uma das mais excitantes e que ainda está praticamente na infância, tanto em termos acadêmicos como tecnológicos, é a captura de movimento.

Por captura de movimento entendemos a operação de um sistema de permita de alguma forma determinar automaticamente, utilizando sensores adequados, uma sucessão temporal de posições e/ou a orientações de pontos ou estruturas selecionadas que apresentem movimento em um ambiente tridimensional real. Essa tecnologia apresenta diversos usos, principalmente nas áreas de animação de modelos, interface com usuários, realidade virtual e realidade aumentada. A captura de movimento, porém, é uma vasta área em si mesma, e muitos outros usos são possíveis e estão emergindo (por exemplo em robótica) à medida que nossa tecnologia progride.

Para realizar o rastreamento, vários tipos de sistemas têm sido experimentados, estando entre os mais populares os que utilizam sensores mecânicos, sensores magnéticos, sensores ultrassônicos e sensores óticos.

Recentemente, a tecnologia que vem mostrando maior eficácia juntamente com menor custo e menor dificuldade de realização é a que utiliza como sensores câmeras de vídeo retroiluminadas e como alvo de rastreamento marcadores óticos retrorefletores. Com essas características, essa combinação tem se mostrado verdadeiramente ganhadora, e tem sido empregada comercialmente com grande sucesso em contextos altamente competitivos (como por exemplo na indústria cinematográfica).

Em sua forma mais geral, a captura de movimento é um problema prático bastante complexo, que leva a dificuldades teóricas para as quais freqüentemente não se conhece (ou não há) uma solução estável, robusta ou computacionalmente viável. Adicionalmente, ou talvez exatamente por esse motivo, os vários estágios da captura de movimento não foram suficientemente estudados do ponto de vista acadêmico. Não existe uma visão geral bem formulada do assunto, uma linguagem universalmente aceita para modelagem e determinação de invariantes e restrições, nem consenso sobre a melhor forma de realizar certas operações básicas.

A questão é exacerbada pelo fato de que apesar das dificuldades teóricas a área apresenta grande utilidade prática, levando diversas empresas a desenvolver e vender sistemas proprietários baseados em soluções parciais desenvolvidas privadamente e cujos detalhes são freqüentemente considerado segredo comercial.

Em resumo, nos deparamos com uma situação na o conhecimento da área é incompleto, parcial e fragmentado e adicionalmente grande parte dele não está academicamente disponível. Diante disso, um dos objetivos do presente plano de pesquisa é resgatar e unificar (mesmo que parcialmente) esse conhecimento, e materializá-lo num sistema de hardware e software academicamente disponível (tanto em termos de informações técnicas como de custo) para uso em pesquisa.



Captura de movimento para animação de modelos geométricos



Captura de movimento em tempo real para realidade virtual interativa



Captura de movimento para estudos biométricos

4. Marcadores Óticos Retrorefletores

O rastreamento de movimento através da análise por computador de imagens capturadas com o uso de câmeras de vídeo vem sendo tentado há muito tempo, com resultados que inicialmente foram de utilidade limitada.

A idéia básica consiste em capturar sucessivas imagens do ambiente onde se encontram os alvos de rastreamento com pelo menos duas câmeras de vídeo, identificar no plano de cada imagem as posições dos objetos através do uso de processamento por software, e então deduzir, por triangulação (e conhecimento dos parâmetros das câmeras utilizadas) as sucessivas posições dos objetos rastreados no mundo físico tridimensional.

A maior dificuldade com essa forma de realizar o rastreamento está em realizar a correspondência entre as imagens bidimensionais capturadas e os modelos tridimensionais que se tem do ambiente, ou seja, em localizar com precisão de forma eficiente e robusta nas imagens os objetos a serem rastreados.

Essa dificuldade advém, em um primeiro momento, do fato de que qualquer contexto minimamente complexo no mundo real apresenta inúmeras texturas, sombras, variações de iluminação e os próprios objetos rastreados podem apresentar elasticidade, plasticidade e um grande número de graus de liberdade.

Se desejamos, por exemplo, rastrear os movimentos do braço de uma pessoa, precisamos acompanhar exatamente onde está o pulso a cada instante. Porém, dada uma imagem de um braço, especialmente considerando que ela pode ser apresentada a cada momento sob diferentes orientações, é muito custoso e, ainda pior, pouco confiável (pelo menos com os métodos disponíveis atualmente) determinar automaticamente, por processamento de imagens, a posição do pulso (ou de qualquer outro ponto fixo não artificial de referência) no braço.

Dessa forma, o rastreamento baseado exclusivamente no processamento de imagens até o presente momento não resultou em sistemas de grande aplicabilidade prática para o problema de rastreamento de movimentos tridimensionais.

Para facilitar a análise das imagens capturadas, porém, há um recurso que desde cedo se mostrou de grande utilidade : o uso de marcadores.

Tais marcadores consistem em padrões facilmente reconhecíveis que são afixados em pontos conhecidos dos objetos a serem rastreados. Dessa forma, artificialmente são criados “pontos de referência” nas imagens a serem processados que facilitam enormemente sua interpretação.

Os marcadores podem ser de vários tipos, desde adesivos bidimensionais contendo imagens e/ou cores conhecidas até soluções mais sofisticadas, como marcadores esféricos de alta refletividade.

Um grande salto ocorreu, porém, com a introdução dos marcadores óticos retrorefletores. Tais marcadores, independente de seu formato, são recobertos por uma camada de material retrorefletor, isto é, que apresenta a propriedade de refletir a iluminação que recebe exatamente de volta na direção da qual veio. É importante ressaltar que isso é uma propriedade que nenhum material comum exhibe espontaneamente; a maioria dos objetos exhibe propriedades refletoras que são uma combinação de componentes difusas (não privilegiam nenhuma direção em particular) com especulares (privilegiam a direção simétrica, com relação à normal, ao ângulo de incidência).

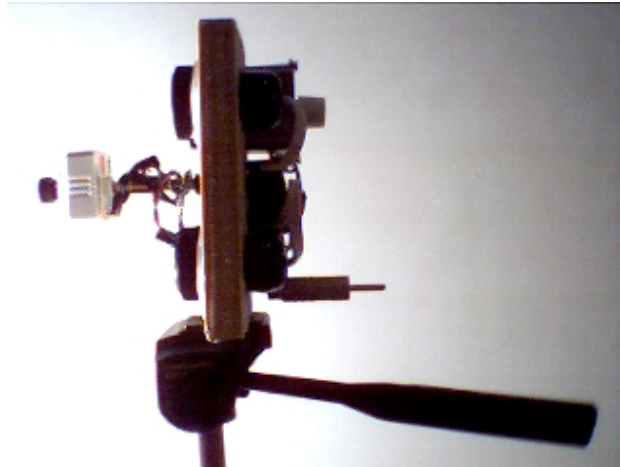
O desenvolvimento de materiais retrorefletores já é uma área razoavelmente madura, mas em anos recentes sua aplicabilidade a praticamente qualquer tipo de superfície, incluindo tecidos, e o aumento da eficiência da retroreflexão assim como a queda no custo de produção tornaram o uso de tais materiais uma opção cada vez mais atraente.

Para utilizarmos tais marcadores para rastreamento, fixamos marcadores em pontos-chave dos objetos a serem rastreados e então acrescentamos, exatamente alinhada com cada câmera utilizada para captura de imagens, uma fonte de iluminação direcional (por exemplo, uma matrix de LEDs de alta emissão). Chama-se a isso de uma fonte de iluminação retroiluminada.

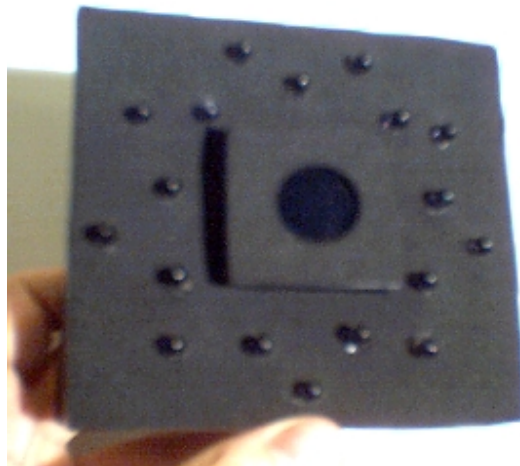
O resultado é que cada marcador devolverá a cada câmera um reflexo de grande intensidade, correspondente à fonte de luz alinhada com a câmera em questão. Com suficiente iluminação orientada, os marcadores devolverão reflexos tão intensos que facilmente superarão a intensidade de quaisquer outras regiões da cena. Dessa forma, os marcadores serão facilmente detectáveis e localizáveis, independentemente das texturas que o cercam e dentro de uma grande gama de variações das condições de iluminação ambiente.

Resta ao sistema então apenas uma segunda dificuldade, que surge em qualquer tecnologia baseada em marcadores, que é fazer a correspondência entre os marcadores rastreados e os objetos rastreados, isto é, em deduzir qual o movimento de uma estrutura a partir do conhecimento das sucessivas posições de uma “nuvem” de pontos a ela fixados. Caso utilizemos marcadores diferenciados para cada ponto da estrutura, esse problema tem solução imensamente facilitada. Em grande parte dos casos, porém, em particular no caso de marcadores óticos, é comum a utilização de marcadores idênticos para todos os pontos rastreados, caso em que uma análise mais complexa dos dados é necessária para inferir a correspondência entre os marcadores e o modelo.

A implementação de um sistema desse tipo é o assunto de minha dissertação de mestrado, na qual desenvolvo e descrevo o equipamento básico para um sistema desse tipo e apresento uma biblioteca de software implementando as operações primitivas básicas para o rastreamento do movimento dos marcadores.



*Câmera infravermelha orientável retroiluminada para captura de movimento
(equipamento construído pelo candidato para sua dissertação de mestrado)*



*Iluminador infravermelho baseado em matriz de LEDs com filtro acoplado
(equipamento construído pelo candidato para sua dissertação de mestrado)*



*Marcador retrorefletor esférico de alta capacidade
(equipamento construído pelo candidato para sua dissertação de mestrado)*



As imagens acima foram capturadas de uma seqüência de vídeo filmado em infravermelho utilizando retroiluminação. O destaque do marcador retrorefletor, que apresenta uma resposta muito mais intensa que qualquer outra região da cena, pode ser facilmente percebido.

5. Objetivos da Dissertação

Nesta dissertação de mestrado, pretendo implementar um sistema de rastreamento de movimento utilizando câmeras de vídeo retroiluminadas e marcadores óticos retrorefletores. Isso inclui desenvolver tanto o sistema de hardware (câmeras, filtros, iluminadores, marcadores, sistemas de captura, procedimentos, etc...) quando o software necessário para operar o sistema.

Em particular, pretendo atingir os seguintes objetivos :

- Projetar e montar um sistema que permita capturar vídeo em forma digital de quatro pontos de vista simultaneamente com resolução espacial e temporal suficientes para um sistema de rastreamento de movimento.
- Projetar e construir os iluminadores, filtros e marcadores retrorefletores necessários.
- Escrever o software necessário para calibrar e operar esse sistema de hardware, para rastrear as posições dos marcadores nas imagens obtidas, e para a partir de tais dados inferir as posições tridimensionais de cada marcador no mundo real.
- Apresentar o software desenvolvido sob a forma de uma biblioteca que possa ser facilmente utilizada para escrever aplicações genéricas empregando esta tecnologia e que possa servir como base para sistemas acadêmicos mais sofisticados na área de rastreamento.
- Escrever um conjunto básicos de aplicações que demonstrem o uso dessa biblioteca, em particular :
 - Visualização das posições tridimensionais de marcadores num ambiente sintético gerado por computação gráfica.
 - Rastreamento da posição e da orientação de estruturas rígidas simples associadas a dois ou três marcadores.
 - Uso do rastreamento de estruturas rígidas simples como ferramenta de interface do usuário.
- Descrever as soluções e decisões de engenharia utilizadas no desenvolvimento do sistema e de seus componentes, discutindo as dificuldades encontradas e como reproduzir os resultados.
- Apresentar uma discussão acadêmica sobre o tema de rastreamento de movimento utilizando marcadores, abordando os principais conceitos envolvidos.

6. Sistema Proposto

Após considerável pesquisa e tentativas de implementação, o sistema proposto é correntemente o seguinte :

A aquisição de imagens se dará através de micro-câmeras industriais sensíveis ao infravermelho e com saída de vídeo composto. Tais câmeras utilizarão micro-lentes levemente grande-angulares cobertas por filtros bloqueadores de luz visível mas transparentes ao infravermelho próximo, e serão retroiluminadas por uma matriz incandescente de baixa intensidade calibrável pelo operador.

Para um sistema completo, serão utilizadas simultaneamente quatro câmeras, ligadas a uma placa de captura de vídeo através de um multiplexador de vídeo (“quad”).

A nível de software, o sistema será realizado em plataforma windows, utilizando para captura de vídeo as facilidades disponibilizados pelo SDK da biblioreca DirectX da Microsoft. O modelo WDM e as rotinas e filtros DirectShow disponibilizados no SDK são adequados ao sistema, já tendo sido usados com sucesso por várias aplicações similares.

A programação do software será feita em C/C++, em princípio utilizando algoritmos já bem conhecidos para calibração e segmentação.

7. Cronograma

Listo abaixo as tarefas envolvidas na realização da tese e o tempo estimado para sua realização. A parte de hardware encontra-se completamente realizada, portanto o valor apresentado é em retrospecto.

Parte 1 : Hardware

- (3 meses) Escolha e configuração de câmeras
- (1 mês) Seleção e montagem de filtros infravermelhos
- (1 sem) Determinação do equipamento para captura de vídeo
- (3 meses) Projeto e construção de iluminadores
- (2 meses) Projeto a manufatura de marcadores retrorefletores
- (1 sem) Projeto e montagem da estrutura de suporte físico do sistema

Parte 2 : Software

- (1 sem) Seleção de uma plataforma de software para captura de vídeo
- (1 mês) Determinação de uma solução para calibração de câmera
- (1 mês) Desenvolvimento de um sistema de rastreamento de marcadores em vídeo
- (1 mês) Realização da correspondência e da triangulação entre marcadores

Parte 3 : Aplicações

- (1 sem) Visualização das posições 3D dos marcadores em tempo real
- (1 sem) Rastreamento de estruturas simples
- (1 sem) Uso do rastreamento como ferramenta para interação

Parte 4 : Dissertação (2 meses)

- Descrição da área de rastreamento
- Justificativa da tecnologia escolhida
- Explicar as dificuldades encontradas e as soluções propostas
- Descrever o sistema construído
- Analisar resultados e comparar com outros sistemas

Para as tarefas ainda não realizadas, proponho o cronograma abaixo :

2004.1	Março	Abril	Maio	Junho	Julho
Software	■	■	■	■	
Aplicações				■	
Dissertação				■	■

Grande parte do tempo empregado até agora nesta dissertação foi utilizado com exaustiva pesquisa, seleção, construção e teste de sistemas de hardware. Diversas tecnologias foram experimentadas para cada componente. Vários tipos de câmeras, filtros, tecnologias de iluminação e marcadores foram examinados e por vezes implementados para que decidisse sobre sua viabilidade. A utilização de webcams, por exemplo, foi testada com diversos modelos, sendo finalmente descartada por fatores como baixa sensibilidade ao infravermelho, baixa qualidade de imagem e baixa taxa de captura. Meses foram gastos na construção de iluminadores, e vários protótipos de iluminadores com matriz de LED foram feitos. A solução para filtragem seletiva do infravermelho próximo também exigiu considerável experimentação até que uma solução satisfatória, que exigiu a confecção de filtros sob medida, foi adotada. Da mesma forma, os marcadores retrorefletores tiveram que ser confeccionados artesanalmente a partir de seus componentes básicos, tendo apenas a literatura da área como referência.

Assim sendo, o tempo inicialmente disponível para a realização da dissertação foi consumido sem que tenham ainda até o presente momento (fevereiro de 2004) sido atingidos os objetivos propostos. Com a continuação do trabalho, porém, e levando em conta as tarefas por realizar, considero realista considerar uma defesa de tese ao final de 2004.1.

7. Bibliografía Seleccionada

Miguel Ribo, Axel Pinz, Anton L. Fuhrmann, “A New Optical Tracking System for Virtual and Augmented Reality Applications”, IEEE Instrumentation and Measurement Technology Conference, Budapest, Hungary, May21-23, 2001.

Jianbo Shi and Carlo Tomasi, “Good Features to Track”, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pages 563-600, 1994

Jaeyong Chung, Namgyu Kim, Gerard Jounghyun Kim, Chan-Mo Park, “POSTRACK : A Low Cost Real-Time Motion Tracking System for VR Application”, Pohang University of Science and Technology, Korea

Klaus Dorfmueller, “Robust Tracking for Augmented Reality Using Retroreflective Markers”, Computers & Graphics / Elsevier, 23 (1999) 795-800

J.R.Parker, Algorithms for Image Processing and Computer Vision, Wiley Computer Publishing, 1997

Ramesh Jain, Rangachar Kasturi, Brian G. Schmuck, Machine Vision, McGraw-Hill, 1995