

Um sistema híbrido para rastreamento baseado em esferas retrorreflexivas e características do objeto rastreado

Lucas Teixeira, Manuel Loaiza, Alberto Raposo, Marcelo Gattass

*Tecgraf, Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro
{lucas, manuel, abraposo, mgattass}@tecgraf.puc-rio.br*

Abstract

The tracking of objects for Virtual and Augmented Reality applications normally considers either the use of markers or markerless techniques, based on characteristics of the object to be tracked. The approach based on markers is more established, but requires the insertion of these "artificial" elements in the tracked object. Markerless techniques, on the other hand, are still in an experimental phase, offering several challenges to be overcome. This paper presents a hybrid system that tries to get the best of both approaches. Invariant properties of retroreflexive spherical markers patterns are used to detect the markers in the object. The inclusion of these markers in known polygonal areas of the tracked objects helps the detection of intrinsic characteristics of them, providing more robustness to the tracking process. A case study using a construction in ruins mock-up is presented.

Resumo

O rastreamento de objetos para aplicações de Realidade Virtual e Aumentada normalmente considera ou o uso de marcadores ou técnicas sem marcadores, baseadas em características intrínsecas do objeto rastreado. A abordagem baseada em marcadores é mais estabelecida, mas tem a desvantagem de exigir a colocação desses elementos "artificiais" no objeto rastreado. Técnicas sem marcadores ainda estão em fase experimental, apresentando muitos desafios técnicos a serem superados. Este artigo apresenta um sistema híbrido que tenta obter o melhor dessas duas abordagens. Propriedades invariantes de padrões de marcadores esféricos retrorreflexivos são usadas para detecção dos mesmos. A inclusão de marcadores em áreas poligonais conhecidas do objeto rastreado ajuda a obter características intrínsecas do mesmo, garantindo maior robustez no rastreamento. É apresentado um

caso de estudo utilizando uma maquete de uma construção em ruínas.

1. Introdução

Os sistemas de rastreamento têm sido, por vários anos, ferramentas básicas e muito utilizadas na implementação de aplicações de realidade virtual e aumentada. O objetivo destes sistemas é realizar o rastreamento individual ou em grupo de diferentes objetos ou marcadores dentro de um cenário predefinido. Os marcadores podem variar desde marcadores artificiais que o usuário deseja colocar dentro da cena para representar pontos referenciais bem contrastados com resto de objetos da cena, até características próprias da cena, como cantos, arestas ou texturas. Entre as diversas tecnologias utilizadas para desenvolver estes dispositivos de rastreamento podemos citar a mecânica, magnética, sonora, óptica e híbrida. Entre elas, as tecnologias ópticas têm sido umas das mais utilizadas e desenvolvidas por conta de algumas vantagens que apresentam em relação às outras, tais como: seus sensores não precisam de fio, e são menos susceptíveis a ruído, permitindo o incremento do número de marcadores rastreados simultaneamente no mesmo cenário, sem que isto gere uma sobrecarga para o sistema, especialmente na parte de hardware.

Neste artigo é proposto um sistema de rastreamento óptico para aplicações de realidade aumentada que trabalha com marcadores sintéticos (padrões de marcadores) e pontos característicos próprios da estrutura do objeto real que desejamos rastrear e queremos enriquecer com informação virtual. Este sistema funciona como uma abordagem híbrida na utilização destes dois tipos de pontos referenciais. Inicialmente o sistema rastreia marcadores sintéticos tais como esferas retrorreflexivas colocadas em pontos bem definidos do objeto real. Esses pontos servem para calibrar e se obter uma posição inicial para a câmera, o

que permite colocar a informação virtual em coerência com o modelo real do objeto rastreado. Uma vez bem identificados estes pontos, é iniciado um processo de aprimoramento na extração e inclusão de outras características próprias da estrutura do objeto real, tais como cantos que ficaram distribuídos em áreas em volta dos marcadores sintéticos e cuja posição é bem conhecida no modelo do objeto real. A extração destas novas características será agilizada pelo fato de que onde elas deverão aparecer na imagem são áreas vizinhas às posições onde aparecem os marcadores sintéticos.

Esta nova abordagem tenta mostrar que este tipo de sistema híbrido pode ter um bom desempenho aproveitando as melhores características do rastreamento com padrões sintéticos assim como do rastreamento feito com base nas características próprias da estrutura do objeto real que será alvo do rastreamento.

Esse artigo prossegue organizado da seguinte forma. A seção 2 apresenta trabalhos relacionados. A seção 3 descreve a parte de hardware do sistema. A seção 4 descreve a extração e identificação dos marcadores, respectivamente. A seção 5 mostra como extrair características do objeto a partir da posição dos marcadores. A seção 6 apresenta um caso de estudo com uma aplicação dos algoritmos propostos usando uma maquete de uma construção em ruínas.

2. Trabalhos relacionados

Sistemas de rastreamento óptico utilizados na implementação de aplicações de realidade virtual e aumentada têm como tarefa principal a identificação de objetos, marcas ou estruturas específicas dentro de uma imagem, que são considerados os marcadores de rastreamento. Com esses marcadores de rastreamento, o sistema pode recuperar a geometria projetiva entre a câmera e a posição no mundo dos marcadores que aparecem na cena. O reconhecimento dos marcadores deve ser rápido e não ambíguo. Para tal efeito, diversas abordagens têm sido propostas, não só aprimorando o método de reconhecimento dos marcadores, mas também influenciando a definição de formato, cor, material e outras características visuais dos mesmos. Dentre as tecnologias mais comumente utilizadas podemos citar: marcadores fiduciais [7], marcadores retrorreflexivos [8, 9] e uma tendência chamada de “markerless” que se baseia no reconhecimento específico de características próprias da cena como: cantos, arestas, texturas e outros marcadores criados a partir da combinação das características antes

mencionadas, que são adaptadas para serem usadas como marcadores de rastreamento da cena [6, 10].

O processo de rastreamento de marcadores pode ser individual ou grupal (padrões de marcadores). Comumente, a tecnologia utilizada para o pareamento e identificação individual e grupal de padrões de marcadores se origina em pesquisas feitas na área de reconhecimento de padrões e paralelamente utilizadas na área de visão computacional. Muitas vezes técnicas robustas e bem desenvolvidas da área de reconhecimento de padrões são “exportadas” para utilização na área de visão computacional. Em [auto-referência omitida para revisão anônima] é proposto o uso de duas técnicas utilizadas em reconhecimento de padrões, que juntas são aplicadas na implementação de pareamento de padrões de marcadores em um sistema de rastreamento óptico. Uma dessas técnicas é a teoria sobre as propriedades projetivas não-variantes (projective invariant) descritas em [1], que será explorada na implementação do rastreamento híbrido aqui proposto, o qual consistirá no rastreamento de marcadores artificiais e feições características contidas dentro da imagem, que darão robustez a um sistema de realidade aumentada proposto. Outro sistema híbrido, que trabalha com marcadores e feições naturais, é descrito por Kanbara et al [4]. Esses autores usam um sistema de rastreamento mais complexo, com câmeras estéreo e um sensor inercial para auxiliar no rastreamento.

3. Hardware

O sistema óptico desenvolvido é composto basicamente por uma câmera colorida e marcadores dispersos pela objeto de interesse. Os marcadores precisam estar fixos em relação ao objeto, mas tanto a câmera como o objeto podem se movimentar.

3.1. Marcadores

Os marcadores ideais para esse sistema são os esféricos e que sejam capazes de ser rapidamente identificáveis na imagem.

O uso de marcadores esféricos é comumente usado para sistemas de rastreamento óptico [8, 9] porque a sua área projetada na imagem é comumente uma área circular. Calcular o centro do círculo é equivalente a calcular a projeção do centro da esfera sobre o plano da imagem. Isso porque o centro da esfera será projetado como sendo um ponto do raio que sai do centro da esfera, atravessa sua superfície, e intercepta perpendicularmente o plano da imagem. Esse ponto de intersecção ficará aproximadamente sobre o centro da

área circular que representa o marcador. Utilizamos então as coordenadas imagem do centro desta área circular para posicionar os marcadores dentro da imagem.

Outro ponto importante do marcador é sua fácil identificação na imagem. Para isso, o fenômeno da retrorreflexão é largamente utilizado. A retrorreflexão é a propriedade que alguns materiais compostos têm de, ao invés de fazer a difusão da luz, refletir a maioria dos raios luminosos na mesma direção em que os raios foram emitidos [5].

Esse tipo de marcador é produzido por poucas empresas no mundo. O uso de fita reflexiva sobre esferas, recurso normalmente utilizado, não tem o mesmo resultado de uma esfera pintada com tinta retrorreflexiva. A solução encontrada para produzir os marcadores foi pintar esferas de isopor com tinta látex branca e, com a tinta ainda úmida, cobrir a superfície com micro-esferas de vidro. Essas micro-esferas de vidro sobre tinta branca geram o fenômeno da retrorreflexão. As micro-esferas usadas são as mesmas usadas na tinta de demarcação de rodovias, sendo um produto barato e relativamente fácil de encontrar.

3.2. Iluminador e lente

Para poder extrair os marcadores da imagem com o menor processamento possível, em nossa aplicação utilizamos a abordagem de iluminação infravermelha passiva, cuja estrutura consiste em usar: marcadores pintados com tinta retrorreflexiva, uma câmera equipada com lente especial sem filtro IR, e um pequeno iluminador formado por 4 leds infravermelhos colocados perpendicularmente à lente da câmera. Isso permite que os marcadores sejam facilmente contrastados com o resto dos objetos na cena. Em especial, este contraste será facilmente visualizado no canal vermelho extraído da imagem.

O ambiente de rastreamento terá a restrição de só poder estar iluminado por luz fluorescente cuja emissão interna de luz infravermelha seja muito baixa e não gere iluminação excessiva que possa criar falsos marcadores na área de rastreamento.

4. Características projetivas invariantes

Uma vez que os marcadores retrorreflexivos são extraídos e identificados, eles podem ser agrupados em padrões específicos previamente definidos e treinados para seu posterior reconhecimento. O método utilizado para definir padrões de formatos e o reconhecimento individualmente destes padrões é a teoria de características projetivas invariantes.

A implementação das propriedades projetivas invariantes tem sido um tema discutido especialmente na área de reconhecimento de padrões [11]. Ela tem sua base na propriedade do cross ratio invariante na projeção em perspectiva. A propriedade de cross ratio define que se temos 4 pontos colineares (A,B,C,D), podemos definir um valor de cross ratio baseado nas distâncias destes pontos de acordo com a seguinte relação:

$$\text{Cross ratio (A,B,C,D)} = \frac{|AC|/|BC|}{|AD|/|BD|}$$

Esta propriedade do cross ratio é expandida não só para atingir padrões colineares de pontos, mas também padrões com marcadores co-planares. Neste caso, o cross ratio é obtido com base nas áreas dos triângulos gerados usando combinações dos vértices que compõem um padrão co-planar, como mostrado em [13]. Esta variante estendida das propriedades projetivas não-variantes do cross ratio aplicada sobre pontos co-planares foi bem desenvolvida até ficar não-variante a possíveis permutações na etiquetagem dos pontos que compõem os padrões co-planares. O trabalho de [12] mostra esta nova abordagem e testa a sua aplicação. O trabalho de [11] mostra outra extensão para gerar propriedades projetivas não-variantes para casos com mais de 5 pontos co-planares. Nos trabalhos acima, o identificador para a projetiva não-variante para um grupo de 5 pontos é definido como um vetor de 5 intervalos, com um valor mínimo e máximo para cada posição do vetor. Este vetor é obtido aplicando a técnica P2-Invariant sobre a amostra de 5 pontos, onde cada posição do vetor é relacionada a cada marcador que compõe o grupo de 5, o que segundo o autor permitirá um pareamento não só grupal, mas também individual para cada marcador que compõe grupo de 5. A única desvantagem é a necessidade de gerar todo o vetor de 5 intervalos para cada candidato de grupo de marcadores que quisesse ser pareado com um padrão específico.

Foi baseado nos resultados dos trabalhos acima descritos, que o algoritmo proposto definiu a configuração do padrão utilizado. Um padrão co-linear de 4 marcadores, mostrados na seção 3. O padrão tem duas características não-variantes à projeção em perspectiva, a co-linearidade e, por consequência, o cross ratio.

4.1. Colinearidade

A colinearidade é uma característica invariante à projeção perspectiva que é explorada na nossa

aplicação de rastreamento. Esta característica, além de ter um identificador único extraído com base na teoria das projetivas invariantes, pode gerar um formato específico que pode ser utilizado como um filtro pelo qual vários candidatos de “n” pontos na cena podem ser descartados se não cumprirem esta característica. Neste trabalho, o teste de colinearidade feito sobre padrões formados por 4 marcadores corretamente alinhados tem um custo computacional aproximadamente de $O(n)$. A idéia por trás do teste é:

1. Para cada grupo de 4 pontos, escolhemos os 2 primeiros pontos. Neste caso, os pontos não precisam estar ordenados.
2. Com os 2 primeiros pontos do conjunto de 4, calculamos a equação da reta que eles podem definir.
3. Para cada um dos 2 pontos restantes, calculamos a distância desses pontos para a reta.
4. Se as distâncias estiverem abaixo de um certo limiar, então o padrão é considerado válido.
5. Inicialmente, este primeiro limiar pode ser bem alto para permitir a aceitação de um bom número de fortes candidatos a padrões colineares. Num segundo teste, os padrões que passarem o primeiro teste de colinearidade são ordenados em relação ao eixo X primeiro e depois ao eixo Y, caso os pontos fiquem muito próximo nas coordenadas em X.
6. Neste segundo teste de colinearidade, padrões que realmente tenham um alto grau de colinearidade serão utilizados para o teste seguinte, que será o cálculo e comparação do valor da sua projetiva invariante.

4.2. Área envolvente

Uma segunda técnica utilizada para diminuir o custo computacional é a geração de uma caixa envolvente em volta da posição (no frame atual) onde são perfeitamente reconhecidos os padrões de rastreamento. Uma vez que um padrão específico é encontrado e validado, o sistema cria uma caixa envolvente com um certo dx e dy de largura e comprimento em volta do qual criaremos uma área de predição para que, num próximo frame, só aqueles marcadores que caíam dentro dessas respectivas áreas envolventes sejam considerados.

Este tipo de aprimoramento gera uma otimização e autonomia no processo de reconhecimento em tempo real.

5. Extração de características intrínsecas fortes

Características intrínsecas fortes de objetos, do ponto de vista do rastreamento, são características bem definidas, que podem ser usadas como pontos de referência no objeto para dar maior robustez no rastreamento.

No sistema desenvolvido, os marcadores são colocados em áreas poligonais do objeto passíveis de segmentação, para ser possível recuperar as posições dos pixels dos cantos dessas áreas poligonais. O processo de extração dessas características será explicado nas próximas seções.

5.1. Segmentação de regiões conhecidas sob os marcadores

As posições dos marcadores são conhecidas e as invariantes nos fornecem o identificador de cada marcador. Sendo assim, é possível usar um método de segmentação calibrado em pré-processamento. Foi usado o algoritmo Flood-Fill [2], principalmente pelo seu baixo custo computacional. Foi usada a sua versão que considera os 8 vizinhos de um pixel. O algoritmo itera sobre uma lista L de pixels (x,y) que inicialmente tem apenas a coordenada do pixel origem do Flood-Fill, que chamamos de $pxFFOr$. Nessa iteração o algoritmo retira o primeiro da lista L e visita todos os seus 8 vizinhos. Os vizinhos que atenderem ao critério de semelhança são marcados como pertencentes à região de $pxFFOr$ e inseridos no final da lista L. O algoritmo itera sobre essa lista L até não ter mais nenhum elemento na lista.

O critério de semelhança usado para segmentar as esferas brancas foi o pixel ter a cor entre 155 e 255 em todos os três canais de cor. Apesar desse critério ser bem amplo, podemos fazer isso se colocarmos os marcadores em lugares de bom contraste com o fundo. O critério de semelhança usado para segmentar a forma poligonal sob os marcadores é igual, apenas mudando os intervalos de cor aceito, para de 0 a 100 em todos os canais RGB. Os intervalos de cor foram escolhidos empiricamente. Essa parte do sistema pode ser substituída por qualquer técnica de segmentação que se adapte melhor ao objeto de interesse, desde que mantenha o custo computacional bastante reduzido.

Como queremos a segmentar por completo a região sob os marcadores, o resultado esperado é a

componente dos marcadores somada com a componente da região em torno dele, que faz parte da região conhecida. Para isso primeiramente aplicamos o algoritmo de segmentação de marcadores, depois substituímos o valor dos pixels dessa região segmentada por um valor que é aceito pelo critério de semelhança da segmentação da região poligonal conhecida. Aplicamos o algoritmo de segmentação final iniciando em $pxFFOr$ novamente. A resultante disso vai ser o que queremos, a soma da região sob o marcador com a da região dele, conforme ilustrado na Figura 1.

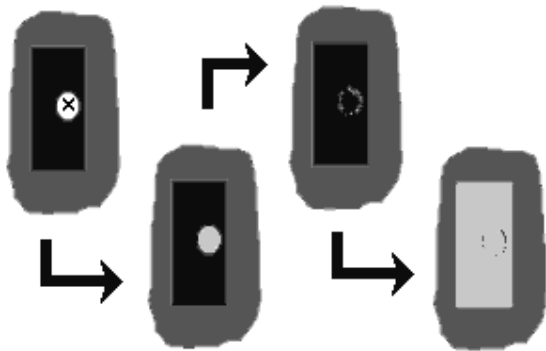


Figura 1. Representação do algoritmo de segmentação utilizado: a região poligonal conhecida em preto, o marcador em branco, o ponto detectado pelo sistema de rastreamento é o X sobre o marcador e em cinza claro os resultados das duas execuções do algoritmo de Flood-Fill.

5.2. Busca de um pixel do contorno externo da componente conexa gerada pelo algoritmo Flood-Fill

O algoritmo Flood-Fill tem como saída uma componente conexa, mas não garante a inexistência de buracos nessa componente. Portanto, a busca pelo contorno mais externo da área segmentada pode falhar se buscarmos apenas um pixel que seja vizinho de um pixel que não faça parte da componente conexa, pois esse pixel pode ser vizinho de um buraco interno do segmento, como os pontos indicados pelo quadrado pontilhado na Figura 2.

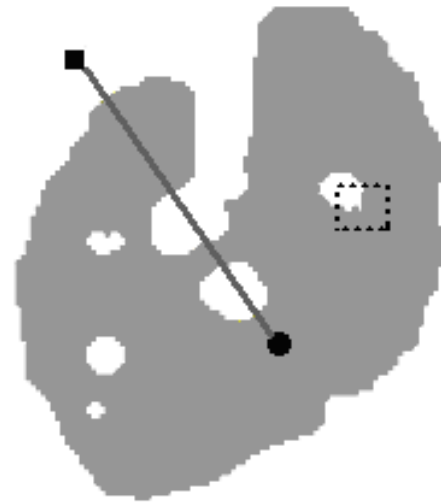


Figura 2. A componente conexa em cinza, o fundo em branco, um ponto qualquer pertencente à componente é o círculo preto, um ponto qualquer não pertencente à componente conexa é o quadrado preto.

Para o sistema, foi desenvolvido um algoritmo de busca do contorno externo de uma região segmentada (a componente conexa, que chamaremos a componente de C_c) na imagem. O primeiro passo do algoritmo é calcular uma caixa envolvente a C_c . Durante a execução do algoritmo de Flood-Fill é extremamente barato manter o valor do X_{min} e Y_{min} e o X_{max} e Y_{max} da C_c . Com esses quatro valores temos o retângulo envolvente alinhado com os eixos X e Y. Como esse retângulo envolve a C_c podemos afirmar que qualquer pixel fora desse retângulo está garantidamente fora da C_c . Pegamos um desses pixels fora do retângulo, que chamaremos de $pxExt$ (quadrado preto na Figura 2). Outro pixel importante para nosso algoritmo é um ponto qualquer pertencente a C_c . Podemos usar o pixel $pxFFOr$, que também é a posição detectada do marcador, pois ele sempre vai pertencer a C_c uma vez que é o valor desse pixel que usamos de referência para gerar a C_c . Esse pixel vai ser chamado de $pxInt$ (círculo preto na Figura 2). Se traçarmos um segmento de reta entre os pontos $pxInt$ e $pxExt$, esse segmento de reta pode passar por um ou mais pixels de contorno internos ou externos, mas garantidamente ele passa por um pixel que faz parte do contorno externo, pois o $pxExt$ é externo e para ele chegar a $pxInt$ ele precisa atravessar o contorno externo. Desses vários pixels de contorno precisamos de um que pertença ao contorno externo. Para isso, percorremos o segmento de reta a partir de $pxExt$ em direção a $pxInt$ até a primeira vez que encontrarmos um

pixel pertencente à C_c . Como iniciamos a busca em px_{Ext} que pertence à região externa à C_c então a primeira vez que encontrarmos um pixel de C_c será o encontro das duas regiões que nos interessa, ou seja, o contorno externo. Para bem definir um contorno, precisamos armazenar tanto o pixel de contorno pertencente à C_c , que chamaremos de px_{Cin} , como o pixel vizinho a ele pertencente à região externa, pois o px_{Cin} pode pertencer a mais de um contorno, caso exista um buraco interno do outro lado do pixel de contorno.

5.3. Critério local para reconhecimento de vértices de retângulos

Normalmente critérios globais são usados para determinar os vértices de um polígono [3], mas critérios dessa natureza são caros por ter que percorrer pelo menos toda área segmentada.

No sistema desenvolvido podemos tirar vantagem do fato que os vértices fazem parte do contorno, de modo que o espaço de busca pode ser bastante reduzido. Já que é possível caminhar sobre o contorno da região na qual procuramos os vértices, é suficiente então determinarmos um critério para, dado um ponto do contorno externo, saber se ele é ou não um dos vértices do retângulo. O critério que propomos supõe que seja conhecida uma estimativa da orientação do retângulo. Com isso, podemos segmentar o retângulo em 4 quadrantes de modo que cada vértice vai estar em um quadrante diferente. Depois rotacionamos o quadrante de modo que a reta A fique alinhada ao eixo y e a reta B com o eixo x. Para os quadrantes 1 e 3 rotacionamos mais 45° no sentido anti-horário e para os 2 e 4 rotacionamos mais 45° no sentido horário. Feito isso temos uma aproximação de parabolóide no espaço, que chamaremos de espaço rotacionado formado pelo eixo x' e y' , com ponto de máximo sendo uma boa aproximação do vértice que procuramos (linha tracejada na Figura 3). Como o contorno é discreto, podemos encontrar esse ponto de máximo testando, para cada pixel do contorno, se a reta paralela a x' que passa por esse ponto do contorno, passa em algum outro ponto dentro da região contornada. Para o vértice isso não ocorre. Isso seria caro se tivéssemos que testar, em todos os pontos dessa reta dentro da imagem, mas se a área retangular é razoavelmente bem determinada, o algoritmo Flood-Fill encontrará o retângulo com apenas pequenos problemas de locais como exemplificado na Figura 3 no quadrante 3 pela seta. Em pré-processamento podemos escolher um segmento de reta com tamanho razoável para escapar dos máximos locais analisando algumas imagens do

objeto rastreado. Esse tamanho tende a ser bem pequeno, como de 10 a 20 pontos para cada lado do ponto sendo testado. Uma boa otimização desse critério é começar pelos dois extremos do segmento e ir iterando em direção do ponto sendo testado, pois é mais provável que encontre nos extremos desse segmento de reta um ponto pertencente à região interna ao contorno, e então o ponto já pode ser descartado como vértice. Essa forma de decidir se o ponto é vértice ou não deixa uma ambigüidade, pois se percorremos todo o contorno não saberemos se o vértice é o do quadrante 1 ou 3, o mesmo acontecendo com o 2 e 4, já que o módulo da inclinação x' é o mesmo para esses pares quadrantes. Para resolver esse problema, só buscamos no contorno dentro do quadrante correspondente.

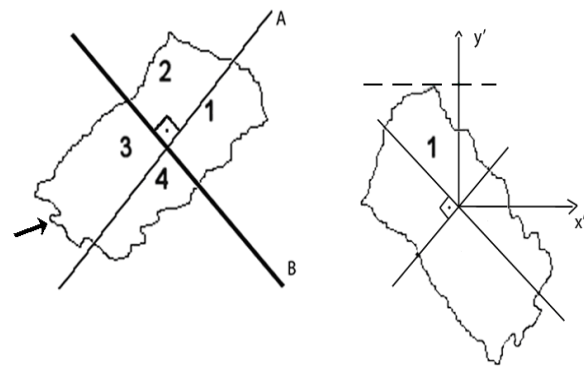


Figura 3. A imagem da esquerda é a área retangular com os eixos A e B que são a orientação do retângulo. A da direita mostra o espaço rotacionado

6. Caso de estudo

Para testar nossos algoritmos usamos uma maquete (Figura 5) das ruínas do convento São Boaventura (Figura 4), localizado em Itaboraí/RJ, construído em 1660. A realidade aumentada é uma boa forma de criar interatividade e aumentar o interesse pela história. A aplicação final tem como objetivo a reconstrução virtual do convento, em cima da imagem das ruínas.



Figura 4. Convento de São Boaventura



Figura. 5. Maquete do convento de São Boaventura. Em destaque estão 3 padrões colineares de marcadores.

6.1. Proposta da aplicação

A aplicação faz o rastreamento em tempo real de uma maquete com marcadores esféricos e retrorreflexivos presos a ela. Alguns desses marcadores estão em janelas retangulares. Como usamos a técnica de projetivas invariantes, temos um identificador único para cada marcador. Assim, caso o marcador for um que sabemos que está sobre uma janela retangular, extraímos os cantos dessa janela, para adicionar robustez ao rastreamento. Alguns outros critérios podem ser adicionados para decidir se adicionamos esses pontos para calibração como área mínima do retângulo na imagem e se o retângulo está muito cisalhado.

Após o rastreamento, aumentamos a imagem com partes das ruínas que não existem mais. A reconstrução do modelo original do convento está sendo objeto de pesquisa histórica e visitas a conventos similares.

A aplicação ,atualmente , é feita apenas para ambientes fechados, dadas as restrições de iluminação explicadas na seção 3.

6.2. Etapas da aplicação

A aplicação foi dividida em duas partes. A primeira é o treinamento. Nessa etapa, posicionamos a câmera em vários diferentes pontos de vista em frente à maquete e usamos essas imagens para calibrar o sistema. A segunda parte é o rastreamento, usando imagens capturadas com uma câmera em tempo real e renderizando a imagem aumentada correspondente.

6.2.1. Treinamento – Offline

A primeira etapa no treinamento é fazer um estudo sobre a iluminação e escolher o threshold adequado ao ambiente, para que os marcadores retrorreflexivos sejam bem detectados.

O segundo passo é calibrar as invariantes projetivas. Isso é feito usando as imagens dos vários pontos de vista com os marcadores já detectados e manualmente associando os marcadores correspondentes de cada padrão nos vários pontos de vista da maquete. Em cada imagem, a aplicação calcula o valor da invariante projetiva de cada padrão e armazena o mínimo e máximo para serem usados na etapa on-line.

Por fim, testamos o algoritmo de extração de características intrínsecas fortes, para determinar o tamanho mínimo do segmento usado no critério de busca pelos vértices descrito na seção 5.3 e quais marcadores associados aos padrões terão o retângulo ao redor detectado.

6.2.2 Rastreamento – On-line

O usuário altera o ponto de vista do objeto até que todos os padrões sejam detectados. Após isso, o sistema de rastreamento entra no modo de coerência temporal usando a caixa envolvente para o custo computacional ser adequado para tempo real. O reconhecimento dos padrões retorna a posição deste padrão na imagem e o identificador do marcador. Essa identificação é uma chave dupla composta pelo número do padrão ao qual ele pertence e qual a ordem dele dentro do padrão. Para os padrões que têm algum marcador com a área adjacente para identificar, calculamos a reta que melhor ajusta as posições dos 4 marcadores colineares do padrão pelo método dos mínimos quadrados. Como o padrão é colocado paralelo à base das janelas, com essa reta temos uma boa estimativa da orientação das mesmas. Usamos o algoritmo de segmentação da seção 5.1 para extrair a região conhecida sob o marcador e calculamos um ponto no contorno com o algoritmo da seção 5.2. Então percorremos o contorno e, para cada ponto, é testado o critério da seção 5.3 para saber se aquele ponto é um vértice. Por fim, usamos esses pontos para calcular a posição dos objetos virtuais nas coordenadas da imagem.

6.3 Testes e Resultados

Os testes foram feitos usando um vídeo feito com câmera fireware da Unibrain fire-i, o processador

utilizado foi Intel P4 3Ghz. A maquete foi instrumentada com 3 padrões colineares, sendo que 2 deles estão em algumas janelas, dando um total de 7 janelas sendo segmentadas (ver Figura 5).

A tabela 1 mostra os tempos médios obtidos para as diversas etapas do processo de rastreamento proposto. A tabela 2 relaciona o número pontos errados que foram detectados dentro da área envolvente com o tempo médio do cálculo das projetivas invariantes.

Tabela 1. Resultados médios do processo de rastreamento.

Passo	Tempo médio
carregar imagem	1.8 ms
Extraír os marcadores e calcular as invariantes no modo de coerência temporal	3.6 ms
Calcular os retângulos	2.7 ms por retângulo
total	24.3 ms

Tabela 2. Tempo de cálculo das projetivas invariantes em relação ao número de pontos errados dentro das áreas envolventes.

Número de pontos errados	Tempo médio
0	1.5 ms
1	1.9 ms
2	2.7ms
3	4.5 ms
4	7.3 ms
5	11.6ms

O tempo para extrair os marcadores e calcular as invariantes sem coerência temporal foi de 0.8 a 2.6 s

O sistema proposto apresentou bons resultados no rastreamento da maquete. O tempo médio é tempo real, sobrando tempo adicional para técnicas de estimação da posição da câmera e renderização de objetos virtuais. O normal foi ter até 3 pixels errados dentro das áreas envolventes, por isso o tempo médio foi baixo. Os valores maiores em geral são decorrentes de movimentos muito bruscos.

A Figura 6 mostra a visualização da maquete com a inserção de janelas virtuais.

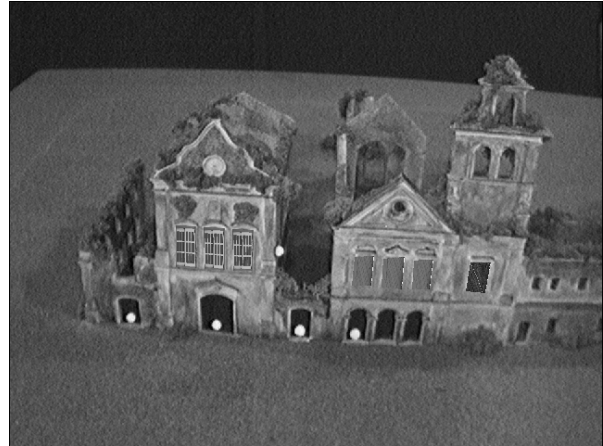


Figura 6 – Imagem aumentada com a renderização de fotos de janelas contemporâneas ao convento.

7. Conclusões e trabalhos futuros

Este trabalho apresentou um sistema de rastreamento híbrido, usando marcadores retrorreflexivos e características do objeto rastreado.

Um aspecto interessante do sistema é que, como os marcadores são discretos, é possível posicioná-los como foi feito nessa maquete, de forma que não ocupem grandes áreas na imagem, e que não façam oclusão de partes importantes da maquete. Com isso a imagem aumentada passa a poder mostrar mais partes da imagem real, sem prejudicar o rastreamento.

Como trabalho futuro, é necessário estudar uma maneira de reduzir o tempo para estimar as projetivas invariantes sem a coerência temporal e aumentar a robustez para permitir movimentos mais bruscos.

A aplicação apresentada é um protótipo de testes para o desafio de se utilizar realidade aumentada em ambientes abertos, sem o preparo do ambiente exigido quando se usam marcadores fiduciais

8. Agradecimentos

O Tecgraf/PUC-Rio é um grupo financiado principalmente pela Petrobrás. O convento São Boaventura se localiza na região do futuro COMPERJ (Complexo Petroquímico do Rio de Janeiro). Agradecemos ao Sr. Roberto Menegon pelo o apoio técnico no desenvolvimento das esferas retrorreflexivas.

9. Referências

[1] Suk, T., Flusser, J., "Point-based projective invariants", *Pattern. Recognition* (33),pp. 251-261,2000.

- [2] J. D. Foley, A. van Dam, S. K. Feiner, and J. F. Hughes, "Computer Graphics: Principles and Practice", Addison-Wesley, Reading, pp. 689-693, 1990.
- [3] Claudio Rosito Jung, Rodrigo Schramm, "Rectangle Detection based on a Windowed Hough Transform", *Proc. of SIBGRAPI*, pp. 113-120, 2004.
- [4] Masayuki Kanbara, Naokazu Yokoya, Haruo Takemura, "A Stereo Vision-Based Augmented Reality System with Marker and Natural Feature Tracking", *VSMM*, p. 455, 2001.
- [5] Helio Moreira and Roberto Menegon, "Sinalização Horizontal", São Paulo, 2003.
- [6] David G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, 60, 2, 2004, pp. 91-110.
- [7] Fiala, M., "ARTag, a fiducial marker system using digital techniques", *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol.2, no., pp. 590-596 vol. 2, 20-25 June 2005
- [8] Vicon Motion Systems - Optical Motion Capture Technology for Science, Engineering and Entertainment Industries. <http://www.vicon.com> – acessado em 11/2008
- [9] Advanced Realtime Tracking GmbH. <http://www.ar-tracking.de>, acessado em 11/2008.
- [10] Williams, B. and Klein, G. and Reid, I., "Real-time {SLAM} relocalisation", *Proc. International Conference on Computer Vision*, 2007.
- [11] Meer, P., Lenz, R., Ramakrishna, S.: Efficient Invariant Representations. *International Journal of Computer Vision* (26). pp. 137-152, 1998.
- [12] Meer, P., Lenz, R., Ramakrishna, S., "Correspondence of Coplanar Features Through p2 Invariant Representations", *Applications of Invariance in Computer Vision*, pp. 473-492, Springer - Verlag Press, 1993.
- [13] Suk, T., Flusser, J.: "The features for recognition of projectively deformed point sets", *Proceedings of the IEEE International Conference on Image Processing*, pp. 348-351, IEEE Press, Washington, USA, 1995.