

Um Método Projetivo para Cálculo de Dimensões de Caixas em Tempo Real

Leandro A. F. Fernandes¹, Manuel M. Oliveira (Orientador)¹

¹ Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Caixa Postal 15.064 – 91501-970 – Porto Alegre – RS – Brasil

{laffernandes, oliveira}@inf.ufrgs.br

Abstract. This work describes an accurate method for computing the dimensions of boxes and the uncertainty associated to the measurements of each of their edges directly from single perspective projection images in real-time. The approach is based on projective geometry and computes the box dimensions using data extracted from the box silhouette and from the projection of two parallel laser beams on one of the imaged faces of the box. The effectiveness of the proposed approach is demonstrated by automatically computing the dimensions of boxes using a scanner prototype that implements the proposed algorithms.

Resumo. Este trabalho descreve um método acurado para o cálculo das dimensões de caixas e da incerteza associada a medição de cada uma de suas aresta diretamente a partir de uma única imagem em tempo real. A técnica é baseada em geometria projetiva e realiza as medições usando informações extraídas da silhueta da caixa de interesse e da projeção de dois feixes de laser paralelos sobre uma das faces visíveis da caixa. A eficiência da técnica proposta é demonstrada a partir construção do protótipo de um scanner que implementa os algoritmos descritos neste trabalho e pelo seu uso na medição de diversas caixas.

1. Introdução

A obtenção de medidas (e.g., dimensões, volume, etc.) de objetos tridimensionais a partir de imagens encontra aplicações em diversas áreas como controle de qualidade, vigilância, análises forenses, estimativa do custo, planejamento de armazenagem e transporte de mercadorias. Infelizmente, o problema de realizar medições de objetos geometricamente complexos diretamente sobre imagens é bastante difícil. Caso se acrescente a restrição de que a medição deva ser realizada utilizando uma única imagem, o problema não apresenta solução em geral. Como será demonstrado neste trabalho, soluções podem ser obtidas para o caso de objetos com geometrias regulares, como é o caso de paralelepípedos. De fato, a possibilidade de estimar as dimensões de paralelepípedos a partir de imagens é de enorme utilidade para diversas empresas como companhias aéreas, correios e empresas de armazenagem, entre outros que manipulam uma grande quantidade de caixas em suas operações diárias. Atualmente, tais companhias adotam caixas de tamanhos padronizados, enquanto caixas de dimensões arbitrárias precisam ser medidas manualmente. A padronização pode comprometer o uso otimizado do espaço de armazenamento e medições manuais tendem a tornar o fluxo de atividades mais lento. A solução ideal para o problema está no uso de caixas que melhor se ajustem ao seu conteúdo e cujas dimensões possam ser obtidas de forma rápida e confiável.

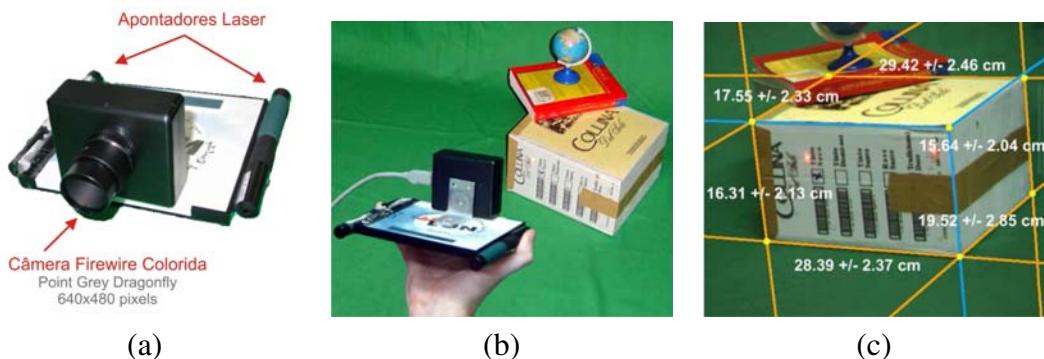


Figura 1. (a) Protótipo do scanner para o cálculo das dimensões de caixas a partir de imagens. (b) O usuário seleciona a caixa de interesse apontando feixes de laser sobre uma de suas faces. (c) As medidas das arestas são estimadas para cada quadro do vídeo, em tempo real, bem como suas respectivas incertezas.

Este trabalho apresenta um método automático para o cálculo das dimensões de caixas a partir de uma única imagem obtida em projeção perspectiva. O foco do trabalho está na automatização do processo de cálculo das dimensões de caixas em tempo real e na análise do erro resultante deste processo. A abordagem faz uso de informações extraídas da silhueta da caixa de interesse e pode ser aplicada até mesmo quando apenas duas de suas faces são visíveis, ou quando a caixa encontra-se parcialmente oculta por outros objetos na cena (Figura 1, b). Para lidar com o problema da ambigüidade projetiva (*i.e.*, causado pela perda da informação de profundidade no momento que a imagem é capturada), dois feixes de laser paralelos são projetados sobre uma das faces da caixa. A técnica de medição proposta é demonstrada a partir da construção de um protótipo do scanner para cálculo de dimensões de caixas (Figura 1, a) e de seu uso em diversas medições em tempo real.

As principais contribuições do trabalho incluem:

- Um algoritmo para o cálculo das dimensões de caixas de forma automática e em tempo real;
- Um algoritmo para a extração da silhueta da caixa de interesse na presença de oclusão parcial de suas arestas;
- Um modelo estatístico para detecção do fundo da cena sob diferentes condições de iluminação;
- Um esquema de votação eficiente para a transformada de Hough; e
- Uma derivação de como estimar, em tempo real, o erro associado às medidas obtidas para cada uma das arestas da caixa (Figura 1, c).

É importante comentar que, devido a restrições de espaço, este artigo não apresenta uma descrição detalhada dos algoritmos propostos ou dos resultados obtidos. O trabalho na íntegra [Fernandes 2006], bem como publicações provenientes do mesmo [Fernandes et al. 2005, Fernandes et al. 2006, Fernandes and Oliveira 2006], são encontrados em <http://www.inf.ufrgs.br/~laffernandes/boxdimensions.html>. Para melhor compreensão, recomenda-se que o vídeo presente nesta página seja assistido.

2. Trabalhos Relacionados

Em geral, técnicas que extraem medidas de objetos tridimensionais a partir de imagens dependem da interação de um usuário e não visam performance em tempo real. Por exem-

plo, Criminisi et al. [Criminisi et al. 1999] apresentaram uma técnica para fazer medições em 3D a partir de uma única imagem em projeção perspectiva. Caso a distância entre algum elemento da cena e o plano de referência seja conhecida, a técnica permite calcular distâncias entre pontos quaisquer e o plano de referência. Caso tal distância não seja conhecida, as dimensões calculadas ficam sujeitas a um fator de escala. Nesta técnica, o usuário é responsável pela identificação de planos na imagem e da distância conhecida.

Em um trabalho mais próximo ao nosso, Lu [Lu 2000] descreve um método para calcular as dimensões de caixas a partir de uma única imagem em tons de cinza. Para simplificar a tarefa, Lu assume que as imagens são adquiridas a partir de projeção paralela ortográfica e que três faces da caixa são visíveis simultaneamente. Além disso, caso a distância entre a câmera e a caixa não seja informada, as dimensões calculadas ficam sujeitas a um fator de escala. A necessidade de conhecer a distância entre a câmera e a caixa é um fator bastante restritivo no método proposto por Lu. Além disso, cuidados especiais são requeridos a fim de distinguir as arestas da caixa de linhas na textura da mesma, fazendo com que o método não seja executado em tempo real.

Nossa abordagem calcula as dimensões de caixas a partir de uma única imagem adquirida com projeção perspectiva, produzindo reconstrução métrica em tempo real e de maneira completamente automática. O método proposto pode ser aplicado a caixas com texturas arbitrárias, pode ser usado quando apenas duas faces da caixa são visíveis, ou mesmo quando as arestas da caixa alvo estão parcialmente ocultas por outros objetos.

3. Calculando as Dimensões de Caixas

Dada uma caixa, suas dimensões são calculadas a partir de sua silhueta e da projeção de dois feixes de laser paralelos sobre uma das faces visualizadas pela câmera. O processo completo pode ser dividido em três estágios: (i) encontrar as marcações laser e os vértices da caixa na imagem; (ii) eliminar a ambigüidade projetiva presente em imagens obtidas por meio de projeção perspectiva; e (iii) calcular as posições em 3D dos vértices da caixa e, delas, as dimensões das arestas. A fim de simplificar o problema, as caixas são modeladas como paralelepípedos.

3.1. Encontrando as marcações laser e os vértices da caixa

As posições das marcações laser na imagem são identificadas buscando os pontos com maior valor de luminância. Encontrando as marcações laser, temos uma pista da localização da caixa na imagem. Falta agora identificar seus limites (silhueta). A fim de evitar a influência da textura na identificação das arestas que definem a silhueta, um fundo de cor conhecida é adotado (*e.g.*, verde, no caso da Figura 1, b). Note que apesar de todos os pixels de fundo apresentarem a mesma cor, seus tons tendem a variar de pixel para pixel em função das condições de iluminação. Uma vez removido o fundo de cena, o sistema gera uma imagem binária que classifica os pixels como sendo de fundo ou como sendo de objetos em cena (entre eles a caixa de interesse). É importante comentar que a identificação do fundo baseia-se em um modelo estatístico que funciona sob diferentes condições de iluminação. Este modelo é uma das contribuições originais deste trabalho.

Após a remoção dos pixels de fundo, uma busca que inicia em uma das marcações laser identifica a silhueta dos objetos em cena que sobrepõe a silhueta da caixa alvo. Desta forma, *a cena pode conter um número arbitrário de objetos*, desde que a silhueta projetada

por eles não oclua completamente nenhuma aresta da caixa alvo. Uma vez descartados os fragmentos de silhueta que pertencem a objetos sobrepostos e identificada a silhueta da caixa de interesse, um novo e eficiente esquema de votação para a transformada de Hough 2D (outra contribuição original deste trabalho) é aplicado a fim de encontrar as linhas suporte das arestas que definem a silhueta da caixa (linhas alaranjadas na Figura 1, c).

Finalmente, as coordenadas 2D dos vértices na imagem são obtidas pela intersecção de pares de linhas suporte adjacentes. No caso em que três faces da caixa são visíveis, elas são delimitadas por um vértice interno (intersecção das linhas azuis na Figura 1, c). A linha suporte de uma aresta interna (azul) é calculada como sendo a linha que conecta seu ponto de fuga correspondente e o vértice na silhueta que é mais próximo deste ponto de fuga. O conjunto completo de arestas e vértices identificados *automaticamente* pela técnica é ilustrado na Figura 1 (c).

3.2. Eliminando a ambigüidade projetiva

Antes de poder re-projetar os vértices da caixa de volta para o espaço 3D, é preciso eliminar a ambigüidade projetiva. Isto é resolvido usando a distância entre as projeções das duas marcações laser sobre uma face da caixa, medida no espaço 3D. Esta distância é obtida usando as coordenadas 2D das marcações laser na imagem, a distância conhecida entre os feixes de laser e o vetor normal da face que contém as marcações. O vetor normal de uma dada face é calculado a partir da linha de fuga desta face [Hartley and Zisserman 2000]. Finalmente, a equação do plano que contém as marcações laser é calculada usando seu vetor normal e a posição em 3D de uma das marcações (calculada assumindo um modelo de câmera *pinhole*).

3.3. Calculando a posição 3D dos vértices da caixa

De posse da equação do plano que contém uma das faces da caixa, a posição em 3D dos vértices desta face é calculada como a intersecção deste plano com raios que partem do centro de projeção da câmera e passam pelos vértices identificados na imagem. Visto que toda face visível compartilha algum vértice com as outras faces, a equação dos planos das outras faces também pode ser obtida, permitindo a recuperação das coordenadas 3D de todos os vértices visíveis.

4. Resultados

O método descrito foi validado utilizando-se um scanner construído para este fim e que consiste de três módulos: uma câmera de vídeo conectada a um PC, dois apontadores laser e um módulo de software. A câmera foi montada sobre uma caixa plástica e os apontadores laser foram alinhados e colados nas laterais dessa caixa (Figura 1, a). Os feixes são (quase) perpendiculares ao plano de imagem e encontram-se separados por uma distância conhecida. O software foi escrito em C++.

Com o auxílio do scanner foram realizados experimentos e análises estatísticas a fim de verificar a confiabilidade das medições realizadas pelo sistema. Além da análise estatística, uma derivação analítica da propagação de erros introduzidos pelos dados de entrada ao longo da cadeia computacional permite a estimativa do erro nas medidas obtidas para cada imagem em tempo real. As validações estatísticas e experimentais mostraram que a abordagem proposta é acurada e precisa. O erro médio relativo em uma

dimensão computada é de 3,75% e está relacionado à geometria das caixas, que de alguma forma diferem de um paralelepípedo devido a imperfeições (e.g., faces e cantos curvados, assimetrias, etc.).

Nos experimentos, as imagens das caixas foram adquiridas a uma distância que variou de 1,7 a 3,0 metros em relação à câmera, sendo que o scanner foi desenvolvido para cobrir uma faixa de 1,3 a 4,0 metros de distância. Para uma cena típica, como a cena exibida na Figura 1, o sistema processa o vídeo e calcula as dimensões das caixas a uma taxa de 29 quadros por segundo (qps). Quando o esquema de votação proposto para a transformada de Hough é substituído pelo esquema de votação tradicional, a taxa de processamento cai para apenas 10 qps. Isso ilustra a eficiência do processo de votação proposto. Essas medições foram realizadas em um computador com processador de 2,8 GHz.

5. Conclusões

Neste trabalho foi apresentado um método completamente automático para o cálculo de dimensões de caixas em tempo real a partir de uma única imagem em projeção perspectiva. Além do método em si, as contribuições do trabalho incluem um algoritmo para a extração da silhueta da caixa de interesse, mesmo com a sobreposição parcial de suas arestas; um método de votação eficiente para a transformada de Hough; e um modelo estatístico para a identificação do fundo de cena sob diferentes condições de iluminação. Métodos estatísticos foram utilizados na verificação da acurácia e precisão do sistema e a teoria de propagação de erros foi utilizada a fim de estimar a incerteza nas medições realizadas sobre cada imagem. A técnica proposta é capaz de estimar as dimensões de dezenas de caixas por segundo.

Acreditamos que as idéias apresentadas permitam a otimização de diversos procedimentos que atualmente são baseados em medições manuais de caixas.

Referências

- [Criminisi et al. 1999] Criminisi, A., Reid, I., and Zisserman, A. (1999). Single view metrology. In *Proc. of the 7th IEEE ICCV-99*, pages 434–441, Kerkyra, Greece.
- [Fernandes 2006] Fernandes, L. A. F. (2006). Um método projetivo para cálculo de dimensões de caixas em tempo real. Master's thesis, UFRGS, Porto Alegre, Brasil.
- [Fernandes and Oliveira 2006] Fernandes, L. A. F. and Oliveira, M. M. (2006). A scanner for computing box dimensions in real time. ACM SIGGRAPH 2006 Conference Abstracts and Applications. Available in the conference CD-ROM (ISBN 1-59593-366-2).
- [Fernandes et al. 2005] Fernandes, L. A. F., Oliveira, M. M., da Silva, R., and Crespo, G. J. (2005). Computing box dimensions from single perspective images in real time. In *Proc. of the XVIII SIBGRAPI*, pages 155–162, Natal, Brazil. IEEE Computer Society.
- [Fernandes et al. 2006] Fernandes, L. A. F., Oliveira, M. M., da Silva, R., and Crespo, G. J. (2006). A fast and accurate approach for computing the dimensions of boxes from single perspective images. *Journal of the Brazilian Computer Society*, 12(2):19–30.
- [Hartley and Zisserman 2000] Hartley, R. I. and Zisserman, A. (2000). *Multiple view geometry in computer vision*. Cambridge University Press, Cambridge, UK.
- [Lu 2000] Lu, K. (2000). Box dimension finding from a single gray-scale image. Master's thesis, SUNY Stony Brook, New York.