

Pablo Carneiro Elias

**Calibração e Posicionamento de Câmera
Utilizando Fotos e Modelos de Edificações**

Dissertação de Mestrado

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-graduação em Informática do Departamento de Informática da PUC-Rio

Orientador: Prof. Marcelo Gattass

Rio de Janeiro
Abril de 2008

Pablo Carneiro Elias

Calibração e Posicionamento de Câmera Utilizando Fotos e Modelos de Edificações

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-graduação em Informática do Departamento de Informática do Centro Técnico Científico da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Prof. Marcelo Gattass

Orientador

Departamento de Informática — PUC-Rio

Prof. Waldemar Celes

Departamento de Informática — PUC-Rio

Prof. Alberto Barbosa Raposo

Departamento de Informática — PUC-Rio

Prof. Asla Medeiros e Sá

TecGraf PUC-Rio

Prof. Paulo Cezar Pinto de Carvalho

IMPA

Prof. José Eugênio Leal

Coordenador Setorial do Centro Técnico Científico — PUC-Rio

Rio de Janeiro, 28 de Abril de 2008

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Pablo Carneiro Elias

Graduou-se em Bacharelado em Informática na PUC-Rio em 2005. Desde 2004 trabalha no laboratório de Computação Gráfica da universidade (TecGraf) desenvolvendo sistemas de realidade virtual e visualização científica.

Ficha Catalográfica

Elias, Pablo Carneiro

Calibração e Posicionamento de Câmera Utilizando Fotos e Modelos de Edificações / Pablo Carneiro Elias; orientador: Marcelo Gattass. — Rio de Janeiro : PUC-Rio, Departamento de Informática, 2008.

v., 103 f: il. ; 29,7 cm

1. Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Informática.

Inclui referências bibliográficas.

1. Informática – Tese. 2. Calibração de câmera. 3. Pontos de fuga. 4. Realidade Aumentada. I. Gattass, Marcelo. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. III. Título.

CDD: 004

À minha mulher, Cíntia, por estar sempre ao meu lado e a meu pai por
prover sempre tudo que precisei.

Agradecimentos

A meu orientador e professor Marcelo Gattass, por estar sempre disposto a me atender e ajudar, por suas idéias valiosas, sua incrível boa vontade e por toda vasta experiência que permitiu compartilhar comigo.

À amiga e professora, a doutora Asla Sá, por sua enorme ajuda e boa vontade e também por suas importantes idéias e colaborações para esta dissertação.

Ao professor Paulo Cezar, por suas sugestões e idéias que também foram fundamentais para a realização desta pesquisa.

À minha esposa que, além de tudo, dedicou muito tempo e paciência à leitura desta dissertação e efetuou diversas correções ortográficas e melhorias fundamentais relativas ao texto.

A toda equipe do Tecgraf, por seu companheirismo e auxílio, especialmente a Alberto Raposo, cuja gestão humana e inteligente foi decisiva para que este trabalho pudesse ser concebido.

Aos amigos da equipe do grupo de Realidade Virtual por me proporcionarem bons momentos e produtivas discussões.

Resumo

Elias, Pablo Carneiro; Gattass, Marcelo. **Calibração e Posicionamento de Câmera Utilizando Fotos e Modelos de Edificações**. Rio de Janeiro, 2008. 103p. Dissertação de Mestrado — Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

A reconstrução de câmera é um dos problemas fundamentais da visão computacional. Sistemas de software desta área utilizam modelos matemáticos de câmera, ou câmeras virtuais, para, por exemplo, interpretar e reconstruir a estrutura tridimensional de uma cena real a partir de fotos e vídeos digitais ou para produzir imagens sintéticas com aparência realista. As técnicas de reconstrução de câmera da visão computacional são aplicadas em conjunto com técnicas de realidade virtual para dar origem a novas aplicações chamadas de aplicações de realidade aumentada, que utilizam câmeras virtuais para combinar imagens reais e sintéticas em uma mesma foto digital. Dentre os diversos usos destes tipos de aplicação, este trabalho tem particular interesse naqueles que tratam de visitas aumentadas a edificações. Nestes casos, fotos de edificações — tipicamente de construções antigas ou ruínas — são reconstruídas a partir de modelos virtuais que são inseridos em meio a tais fotos digitais com a finalidade de habilitar a visão de como essas edificações eram em suas estruturas originais. Nesse contexto, este trabalho propõe um método semi-automático e eficiente para realizar tal reconstrução e registrar câmeras virtuais a partir de fotos reais e modelos computacionais de edificações, permitindo compará-los através de superposição direta e disponibilizando uma nova maneira de navegar de forma tridimensional por entre diversas fotos registradas. Tal método requer a participação do usuário, mas é projetado para ser simples e produtivo.

Palavras-chave

Calibração de câmera. Pontos de fuga. Realidade Aumentada.

Abstract

Elias, Pablo Carneiro; Gattass, Marcelo. **Camera Calibration and Positioning Using Photographs and Models of Buildings**. Rio de Janeiro, 2008. 103p. MSc Thesis — Department of Computer Science, Pontifícia Universidade Católica do Rio de Janeiro.

Camera reconstruction is one of the major problems in computer vision. Software systems in that field use mathematical camera models, or virtual cameras, for example, to interpret and reconstruct the tridimensional structure of a real scene from a set of digital pictures or videos or to produce synthetic images with realistic looking. Computer vision technics are applied together with virtual reality technics in order to originate new types of applications called augmented reality applications, which use virtual cameras to combine both real and synthetic images within a single digital image. Among the many uses of these types of applications, this work has particular interest in those concerning augmented visits to buildings. In these cases, pictures of buildings — typically old structures or ruins — are reconstructed from virtual models that are inserted into such pictures allowing one to have the vision of how those buildings were on their original appearance. Within this context, this work proposes a semi-automatic and efficient method to perform such reconstructions and to register virtual cameras from real pictures and models of buildings, allowing comparing them through direct superposing and making possible to navigate in a tridimensional fashion between the many registered pictures. Such method requires user interaction, but it is designed to be simple and productive.

Keywords

Camera calibration. Vanishing points. Augmented Reality.

Sumário

1	Introdução	14
2	Trabalhos Relacionados	22
2.1	Calibração de Câmera	22
2.2	Posicionamento de Câmera	28
2.3	Sistemas Fotogramétricos	29
3	O Método Proposto	31
3.1	Métodos para Calibração de Câmera	34
3.2	Extração das Arestas Estruturais do Modelo	53
3.3	Posicionamento Inicial de Câmera	58
3.4	Teste de Oclusão	59
3.5	Projeção da Geometria	60
3.6	Cálculo de Correspondências Modelo-Imagem	61
3.7	Cálculo de Posição e Orientação da Câmera	69
3.8	Limitações e Vantagens	75
4	3D-Pose: Uma Aplicação Exemplo	77
4.1	Descarte Manual de Arestas Estruturais	78
4.2	Posicionamento de Câmera	78
4.3	Cálculo de Correspondências	79
4.4	Informando Correlações Manualmente	79
4.5	Navegação Tridimensional entre Fotos	80
4.6	Visão Geral das Funcionalidades	82
5	Resultados	84
5.1	Comparação com o Método Tsai3D	84
5.2	Resultados Visuais	88
6	Conclusão	90
6.1	Trabalhos Futuros	92
	Glossário	94
	Referências Bibliográficas	95
A	Coordenadas Baricêntricas do Ortocentro de Um Triângulo	98
B	Minimização usando o Grupo $S0(3)$	101
B.1	O Grupo das Transformações de Corpo Rígido $SE(3)$	103

Lista de figuras

- 1.1 Aplicação de realidade aumentada. Marcadores planares são colocados sobre a mesa e utilizados para reconstruir câmeras virtuais que permitem inserir objetos sintéticos nas imagens capturadas pelas câmeras dos dispositivos móveis, como as locomotivas no exemplo. 15
- 1.2 Aplicação de realidade aumentada. O padrão planar mostrado impresso em folha de papel possui uma geometria conhecida. A partir dele é possível reconstruir uma câmera virtual para sintetizar objeto mostrado, no caso, um automóvel vermelho. O procedimento foi feito utilizando a biblioteca Artoolkit. 16
- 1.3 Exemplo de realidade aumentada aplicada a ruínas de edificações. Na imagem acima uma câmera virtual foi reconstruída a partir de marcadores eletrônicos para inserir o modelo do templo de Zeus, em Olympia, em sua aparência original. Projeto Archeoguide. 17
- 1.4 Exemplo de aplicação de realidade aumentada não interativa que utiliza marcações feitas pelo usuário para recuperar, dentre outras coisas, a câmera que gerou a imagem. Na imagem acima, as arestas em verde são marcações feitas por um usuário. A imagem foi extraída do trabalho feito por Paul E. Debevec em 1996 em sua tese de doutorado. 18
- 2.1 Processo de formação de pontos de fuga. A figura **a** ilustra que pontos igualmente espaçados no mundo são projetados para pontos cada vez mais próximos na imagem. Um ponto muito distante na cena é projetado em um ponto chamado ponto de fuga, representado pelo ponto v na reta x ou v' no plano de projeção da câmera. A figura **b** mostra que um ponto de fuga corresponde a uma direção no mundo construída a partir de um vetor que vai do centro de projeção da câmera até esse ponto na imagem. 25
- 2.2 A figura ilustra marcações de segmentos no espaço da imagem $(s_{x1}, s_{x2}, s_{y1}, s_{y2}, s_{z1}, s_{z2})$ que correspondem a segmentos paralelos às direções principais da cena (x, y, z) (espaço do mundo). Para cada direção existem dois segmentos cuja interseção é exatamente o ponto de fuga. 26
- 3.1 Diagrama esquemático do funcionamento geral do método proposto. 34
- 3.2 Etapa de calibração vista em mais detalhes. Caso não haja na imagem de entrada informações EXIF que permitam a calibração automática da câmera, ela ocorre manualmente utilizando-se marcações na própria imagem para se obter pontos de fuga e, então, calcular a matriz de calibração K . 35
- 3.3 Modelo de câmera *pinhole*. Nesse modelo um ponto P do mundo é projetado para um ponto p' do plano focal através de um mapeamento simples obtido por semelhança de triângulos. 36
- 3.4 Sistema de coordenadas da imagem (x, y) e da câmera (x_{cam}, y_{cam}) . 37

3.5	Mapeamento entre coordenadas da imagem (em pixels) e coordenadas do plano de projeção da câmera. A posição final no plano depende de um ponto adotado como origem (geralmente, mas não necessariamente, o centro do plano, o ponto principal) e do tamanho de um pixel na horizontal e na vertical.	39
3.6	Simbologia atribuída aos parâmetros da câmera que são mapeados pelas tags EXIF. O significado de cada símbolo e suas respectivas chaves EXIF são mostradas nas tabelas 3.1 e 3.2.	41
3.7	Interseção entre diversas retas. O ponto de interseção é calculado utilizando-se mínimos quadrados.	45
3.8	Cálculo de pontos de fuga por múltiplas retas utilizando uma parametrização mais robusta.	45
3.9	Interpretação geométrica da distância focal e do centro da imagem, que coincide com o ortocentro do triângulo formado pelos pontos de fuga.	49
3.10	Descarte de arestas contidas em planos. Cada aresta que for adjacente a dois triângulos coplanares (aresta <i>a</i> pontilhada) é descartada.	55
3.11	Processo de agregação de arestas colineares consecutivas.	56
3.12	Arestas curtas que pertencem à malha geométrica de um modelo feito no 3D Studio Max® (arestas que compõem o telhado). Tal modelo contém milhares de arestas pequenas que chegam a 0.1% do tamanho da maior aresta e podem ser eliminadas durante o processo de extração.	57
3.13	Arestas selecionadas para uso (em amarelo) na etapa de detecção de relações entre modelo e imagem.	58
3.14	Modelo criado por ferramenta de modelagem comercial exibido de duas formas: em modo <i>wireframe</i> , onde as arestas dos triângulos que o compõem aparecem; e em modo normal, onde somente um subconjunto dessas arestas aparecem.	59
3.15		63
3.16	Cálculo da região efetiva de interesse dentro de um ROI.	65
3.17	Tratamento de interseções entre diferentes ROIs.	68
3.18	Representação de uma reta tridimensional $\mathbf{r}_i \in \mathbb{R}^6$, pelos vetores $\hat{\mathbf{v}} \in \mathbb{R}^3, \mathbf{d} \in \mathbb{R}^3$, escrita no sistema de coordenadas da câmera.	71
3.19	Representação da reta $\mathbf{r}_i \in \mathbb{R}^6$ e da normal \mathbf{m} nos espaços da câmera e do mundo.	72
3.20	A medição do erro entre o segmento medido na imagem e o segmento que representa a linha do modelo-base projetada. Os coeficientes m_x, m_y, m_z são os coeficientes da normal (equações 3-23 e 3-24).	74
4.1	Linhas-guias para manipulação do modelo virtual.	78
4.2	Diálogo de configuração de arestas. É possível selecionar o algoritmo de detecção (PPHT ou HT) e configurar valores de threshold para cada um deles. A imagem usada como referência para as configurações é a selecionada como corrente.	79
4.3	Cálculo automático de correlações entre modelo e imagem a partir do posicionamento de câmera inicial.	80

4.4	Representação dos vetores \mathbf{d} , $\hat{\mathbf{v}}$, \mathbf{t}_i e \mathbf{t}_a usados para calcular a classificação de distância a partir de uma direção.	82
4.5	Barras de Ferramentas do sistema. As principais funcionalidades são indicadas pelos números. Os círculos vermelhos indicam as funcionalidades mais importantes.	82
5.1	Esquema ilustrativo do teste com dados sintéticos. Os símbolos $\langle R_0, T_0 \rangle$ significam respectivamente a rotação e a translação iniciais fornecidas como entrada para o método de ajuste proposto.	85
5.2	Testes de reconstrução de câmera com dados sintéticos.	86
5.3	Teste de reconstrução de câmera com dados reais.	87
5.4	Imagem real da aplicação 3D-Pose. Casamento entre imagem e modelo de uma edificação.	88
5.5	Imagem real da aplicação 3D-Pose. Casamento entre imagem e modelo de uma edificação, com linhas estruturais do modelo mostradas em azul.	89
5.6	Comparação entre o modelo virtual e uma imagem do Convento. A câmera que gerou a imagem foi utilizada para posicionar o modelo exatamente no mesmo ponto de vista da imagem, permitindo sobrepor-los. Pode-se verificar visualmente qual parte da estrutura do modelo é afetada pela rachadura mostrada pela foto.	89
5.7	Casamento entre o modelo virtual e foto do convento.	89
A.1	Ilustração do ortocentro (H) de um triângulo qualquer. Embora o triângulo utilizado para ilustração na figura seja isósceles, a formulação proposta para calcular as coordenadas baricêntricas do ortocentro é geral.	98

Lista de tabelas

3.1	Parâmetros de câmera que são obtidos diretamente das <i>tags</i> EXIF (vide Figura 3.6).	42
3.2	Descrição dos parâmetros que são calculados a partir das <i>tags</i> EXIF.	42
5.1	Coordenadas dos vértices do cubo virtual nos espaços do mundo e da janela, utilizadas como entrada para o método T3D (tabela da esquerda) e lista de arestas utilizadas para o MP. A resolução da imagem de entrada é de 961×633 (Figura 5.2).	86
5.2	Testes com dados sintéticos onde as correspondências são dadas e são ideais. A tabela mostra uma comparação entre o método de Tsai 3D (T3D), o método proposto (MP) e os dados exatos da cena virtual.	86
5.3	Coordenadas dos vértices do cubo real nos espaços do mundo e da janela, utilizadas como entrada para o método T3D (tabela da esquerda) e lista de arestas utilizadas para o MP. A resolução da imagem de entrada para o caso real é de 3888×2592 (Figura 5.3).	87
5.4	Tabela de comparação entre o método de Tsai 3D (T3D) e o método proposto (MP) utilizando dados reais. Neste teste foi possível comparar a calibração por pontos de fuga com a calibração por EXIF tags, já que a foto de entrada foi obtida por uma câmera real. A medida em pixels obtidas dessas tags são baseadas no tamanho real de um pixel e a distância focal em milímetros é a distância focal real da câmera que capturou a foto. A unidade <i>px</i> utilizada é uma abreviação para <i>pixels</i> .	88

*Quando alguém pergunta “Qual o caminho?”.
O Zen simplesmente responde “Caminhe”.*

Ditado Budista, *Autor Desconhecido.*