

Referências Bibliográficas

- [Alvarez, 2002] Alvarez, B. S. V. (2002). "Edição Tridimensional de Fotografias Arquitetônicas". Master's thesis, Pontifícia Universidade Católica do Rio de Janeiro, PUC, Rio de Janeiro, Brazil. 2.1.1, 2
- [Braga, 2007] Braga, I. F. (2007). *Realidade Aumentada em Museus: As Batalhas do Museu Nacional de Belas Artes, RJ*. PhD thesis, COPPE/UFRJ. 1
- [Canny, 1986] Canny, J. (1986). "A Computational Approach to Edge Detection". pages 679–698. 3.6.1
- [Chou, 2000] Chou, G. T.-S. (2000). *Large-Scale 3D Reconstruction: A Triangulation Based Approach*. PhD thesis, MIT. 1
- [Cipolla et al., 1999a] Cipolla, R.; Robertson, D.; and Boyer, E. (1999a). "Camera calibration from vanishing points in images of architectural scenes". *BMVC. British Machine Vision Conference*. 2.1.1, 2.2, 3.1.4
- [Cipolla et al., 1999b] Cipolla, R.; Robertson, D.; and Boyer, E. (1999b). "Photo-Builder - 3D Models of Architectural Scenes from Uncalibrated Images". *IEEE International Conference on Multimedia Computing and Systems*, 1, pp. 25 to 31. 2.1.1, 2.2, 3.1.4
- [Debevec, 1996] Debevec, P. E. (1996). *Modeling and Rendering Architecture from Photographs*. PhD thesis, University of California at Berkeley. 2.3, 3.7.1, 6.1
- [Delage et al., 2007] Delage, E.; Lee, H.; ; and Ng, A. Y. (2007). "Automatic Single-Image 3d Reconstructions of Indoor Manhattan World Scenes". *Springer Tracts in Advanced Robotics*. 1
- [Fischer and Bolles, 1981] Fischer, M. A. and Bolles, R. C. (1981). "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography". *Communications of the ACM*, 6, pp. 381–395. 1, 2.3, 3.6
- [Gamba et al., 1996] Gamba, P.; Mecocci, A.; and Salvatore, U. (1996). "Vanishing Point Detection By A Voting Scheme". *IEEE International Conference on Image Processing*, pages 301–304. 2.1.1, 6.1

- [Hartley and Zisserman, 2003] Hartley, R. and Zisserman, A. (2003). *Multiple View Geometry in Computer Vision*. Cambridge University, 2nd edition. 2.1, 2.1.1, 2.1.1, 2.2, 3.1.2, 3, 3.1.4
- [J.Matas et al., 2004] J.Matas; C.Galambos; and J.Kittler (2004). "Progressive Probabilistic Hough Transform". *British Machine Vision Conference*. 3.6.1
- [Kang, 2000] Kang, S. B. (2000). "Radial Distortion Snakes". Technical report, Microsoft Research at Microsoft Corporation. 2.1
- [Lowe, 2004] Lowe, D. G. (2004). "Distinctive Image Features from Scale-Invariant Keypoints". *International Journal of Computer Vision*. 2.1, 2.3
- [Luong and Faugeras, 1999] Luong, Q. and Faugeras, O. (1999). "Camera Calibration, Scene Motion and Structure recovery from point correspondences and fundamental matrices". Technical report, INRIA, 06902 Sophia-Antipolis, France. 2.1
- [McLean and Kotturi, 1995] McLean, G. and Kotturi, D. (1995). "Vanishing Point Detection by Line Clustering". *IEEE Transactions On Pattern Analysis And Machine Intelligence*. 2.1.1, 6.1
- [Pajdla et al., 1997] Pajdla, T.; Werner, T.; and Hlaváč, V. (1997). "Correcting Radial Lens Distortion Without Knowledge of 3-D Structure". Technical report, Czech Technical University. 2.1
- [Raposo et al., 2004] Raposo, A. B.; Szenberg, F.; Gattass, M.; and Celes, W. (2004). "Visão Estereoscópica, Realidade Virtual, Realidade Aumentada e Colaboração". Technical report, Tecgraf-Grupo de Tecnologias em Computação Gráfica, Departamento de Informática, PUC-Rio, Rio de Janeiro, Brasil. 1
- [Rother, 1999] Rother, C. (1999). "A new Approach for Vanishing Point Detection in Architectural Environments". *Image and Vision Computing*, 20(9-10), pp. 647-655. 2.1.1, 6.1
- [Ruggiero and da Rocha Lopes, 1996] Ruggiero, M. A. G. and da Rocha Lopes, V. L. (1996). *Cálculo Numérico - Aspectos Teóricos e Computacionais*. UNICAMP, 2 edition. 3.7.2
- [Saxena et al., 2007] Saxena, A.; Sun, M.; and Ng, A. Y. (2007). "Learning 3-D Scene Structure from a Single Still Image". *IEEE International Conference on Computer Vision and Pattern Recognition*. 1

- [Schaffalitzky and Zisserman, 2000] Schaffalitzky, F. and Zisserman, A. (2000). "Planar Grouping for Automatic Detection of Vanishing Lines and Points". Technical report, Department of Engineering Science, University of Oxford. 2.1.1, 6.1
- [Snavely et al., 2006] Snavely, N.; Seitz, S. M.; and Szeliski, R. (2006). "Photo Tourism Exploring Photo Collections in 3D.". Technical report, Microsoft at Microsoft Research, Redmond WA 98052, USA. 2.1.2, 2.3, 6.1
- [Taylor and Kriegman, 1994] Taylor, C. J. and Kriegman, D. J. (1994). "Minimization on the Lie group $SO(3)$ and Related Manifolds". *IEEE Computer Graphics and Applications*. 3.7.1, B
- [Taylor and Kriegman, 1995] Taylor, C. J. and Kriegman, D. J. (1995). "Structure and Motion from Line Segments in Multiple Images". *IEEE Transactions On Pattern Analysis And Machine Intelligence*, 17(11), pp. 1021–1032. 2.2, 2.3, 3.6.2, 3.7, 3.7.1, 3.7.1, 3.7.1, 3.7.2, 3.7.2
- [Trucco and Verri, 1998] Trucco, E. and Verri, A. (1998). *Introductory Techniques for 3-D Computer Vision*. Prentice Hall, Inc. 3.6.1
- [Tsai, 1986] Tsai, R. Y. (1986). "An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision". *IEEE International Conference on Computer Vision and Pattern Recognition*. 2.1, 3.7, 5.1
- [Vlahakis et al., 2002] Vlahakis, V.; Ioannidis, N.; Karigiannis, J.; Tsotros, M.; Gounaris, M.; Stricker, D.; Gleue, T.; Daehne, P.; and Almeida, L. (2002). "Archeoguide: An Augmented Reality Guide for Archaeological Sites". *IEEE Computer Graphics and Applications*. 1
- [Zhang, 1998] Zhang, Z. (1998). "A Flexible New Technique for Camera Calibration". Technical report, Microsoft Corporation at Microsoft Research. 2.1, 3.1.4

Glossário

Renderizar Processo de transformar a representação de um modelo de dados para a forma de uma imagem digital.

Feature Sinônimo para *característica* de determinado objeto, em inglês.

A Coordenadas Baricêntricas do Ortocentro de Um Triângulo

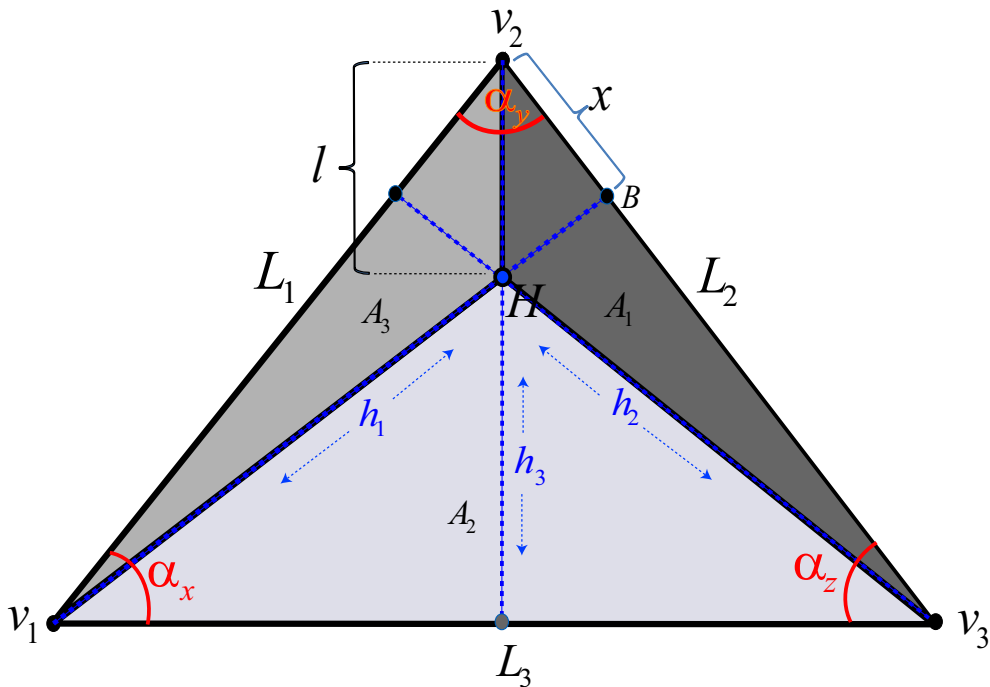


Figura A.1: Ilustração do ortocentro (**H**) de um triângulo qualquer. Embora o triângulo utilizado para ilustração na figura seja isósceles, a formulação proposta para calcular as coordenadas baricêntricas do ortocentro é geral.

Adota-se aqui a notação de $\Delta(V_1, V_2, V_3)$ para designar a área de um triângulo formado por vértices V_1, V_2 e V_3 .

Utilizando a Figura A.1 como referência, são derivadas as coordenadas baricêntricas (r_x, r_y, r_z) do ortocentro H de um triângulo qualquer, dado que

$$\begin{aligned} r_x &= \Delta(V_2, H, V_3) / \Delta(V_1, V_2, V_3) \\ r_y &= \Delta(V_1, H, V_3) / \Delta(V_1, V_2, V_3) \\ r_z &= \Delta(V_1, H, V_2) / \Delta(V_1, V_2, V_3) \end{aligned}$$

Pela Área de um Triângulo Qualquer, vem que

$$\begin{aligned}\Delta(V_1, V_2, B) &= \frac{1}{2} \sin(90 - \alpha_y) h_1 L_1 = \frac{x h_1}{2} \\ x &= \sin(90 - \alpha_y) L_1 = \cos \alpha_y L_1\end{aligned}\quad (\text{A-1})$$

usando então o valor de x calculado, temos:

$$\begin{aligned}\cos(90 - \alpha_z) &= \frac{x}{l} = \frac{\cos \alpha_y L_1}{l} \\ l \sin \alpha_z &= \cos \alpha_y L_1 \\ l &= \frac{\cos \alpha_y L_1}{\sin \alpha_z}\end{aligned}\quad (\text{A-2})$$

calculando a área do triângulo oposto ao vértice V_1 , vem que:

$$\begin{aligned}\Delta(V_2, H, V_3) &= \frac{1}{2} \sin(90 - \alpha_z) (l L_2) = \\ &= \frac{(\cos \alpha_z \cos \alpha_y) L_1 L_2}{2 \sin \alpha_z}\end{aligned}\quad (\text{A-3})$$

Como temos

$$r_x = \Delta(V_2, H, V_3) / \Delta(V_1, V_2, V_3)$$

e, novamente pela Área de um Triângulo Qualquer,

$$\Delta(V_1, V_2, V_3) = \frac{\sin \alpha_y (L_1 L_2)}{2}\quad (\text{A-4})$$

podemos obter r_x :

$$r_x = \frac{(\cos \alpha_z \cos \alpha_y) (L_1 L_2)}{2 \sin \alpha_z} = \frac{(\cos \alpha_z \cos \alpha_y)}{\sin \alpha_z \sin \alpha_y} \frac{\sin \alpha_y (L_1 L_2)}{2}$$

$$\mathbf{r}_x = \cot \alpha_z \cot \alpha_y\quad (\text{A-5})$$

O mesmo cálculo pode ser feito para se obter as coordenadas baricêntricas r_y e r_z .

Assim, tem-se

$$r_x = \cot \alpha_z \cot \alpha_y$$

$$r_y = \cot \alpha_z \cot \alpha_x$$

$$r_z = \cot \alpha_x \cot \alpha_y$$

Lembrando que as coordenadas baricêntricas são homogêneas, logo, as coordenadas r_x, r_y, r_z fazem parte da classe de equivalência (wr_x, wr_y, wr_z) .

B Minimização usando o Grupo $SO(3)$

Muitos problemas na área de visão computacional e na robótica envolvem a busca por uma matriz de rotação ótima dentro do conjunto das rotações rígidas $SO(3) = \{R \in \mathbb{R}^{3 \times 3} : R^t R = I, \det(R) = 1\}$ [Taylor and Kriegman, 1994]. Por exemplo, o problema de determinar a posição de câmera em relação a um conjunto conhecido de marcações na imagem, pode ser escrito na forma de uma função-objetivo

$$O : SO(3) \rightarrow \mathbb{R}$$

que mede o quão bem a matriz R ajusta os dados observados. Muitos destes problemas não têm uma fórmula matemática fechada, o que significa que métodos numéricos são normalmente aplicados para calcular as soluções. No entanto, para aplicar tais técnicas de otimização, não se pode simplesmente construir um conjunto de parâmetros utilizando, por exemplo, ângulos de Euler, pois o próprio espaço $SO(3)$ não é isomórfico ao espaço vetorial \mathbb{R}^3 , ou seja, não existe uma função $\mathbb{R}^3 \rightarrow SO(3)$. Isso significa que uma definição ruim dos parâmetros pode causar singularidades ou anomalias em vários pontos do espaço desses parâmetros. *Tais anomalias podem causar sérios problemas quanto utilizados em métodos de gradiente-descente baseados em procedimentos de minimização.*

Contudo, se o processo de minimização acontece em torno de uma rotação inicial conhecida é possível abandonar a parametrização global em detrimento de uma estratégia local utilizando um atlas de parametrização local. Isso equivale a construir um mapeamento contínuo, inversível e diferenciável em torno de cada ponto R_0 do espaço $SO(3)$ (que pode ser visto como uma superfície contínua em torno do ponto R_0). Tal mapeamento é expresso como:

$$R(\omega) = R_0 \exp\{J(\omega)\}, \quad \omega \in \mathbb{R}^3, \sqrt{\omega^t \omega} < \pi \quad (\text{B-1})$$

onde $\omega = (\omega_x, \omega_y, \omega_z)^t$, \exp é o operador exponencial para matrizes e $J(\omega)$ é o operador simétrico de rotação (*skew symmetric operator*) $J : \mathbb{R}^3 \rightarrow SO(3)$

dado por:

$$J(\omega) = \begin{pmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{pmatrix}$$

O operador exponencial de uma matriz X é definido como:

$$e^X = \sum_{k=0}^{\infty} \frac{X^k}{k!}.$$

A equação B-1 define um mapeamento isomórfico um-para-um com a esfera aberta $\sqrt{\omega^t \omega} < \pi$, o que permite então criar uma função $\mathbb{R}^3 \rightarrow SO(3)$ já que dentro deste intervalo não há ambiguidades.

Quando se expande a expressão da equação B-1 utilizando a definição de operador exponencial para matrizes, pode-se perceber que ela equivale a uma aproximação por série de Taylor em torno de um ponto conhecido R_0 , para valores muito pequenos de $\sqrt{\omega^t \omega}$. Essa é a idéia principal por trás desse método, que permite definir então a função-objetivo O em termos de sua parametrização local em torno de um ponto inicial:

$$O(R(\omega)) \approx O(R_0) + \mathbf{q}^t \omega + \omega^t \mathbf{H} \omega \quad (\text{B-2})$$

onde \mathbf{q} e \mathbf{H} representam respectivamente o gradiente e a matriz Hessiana da função-objetivo avaliados no ponto $\omega = 0$, que corresponde a matriz de rotação R_0 .

A B-2 representa uma aproximação quadrática por expansão da série de Taylor, e pode ser utilizada para calcular o passo que leva a um nível mais baixo na curva de erro:

$$\omega_s = -\mathbf{H}^{-1} \mathbf{g}$$

Esse passo incremental pode ser aplicado para calcular uma nova matriz de rotação $R = R_0 \exp \{J(\omega_s)\}$ que alinha melhor os pontos observados e os pontos do espaço. Aplicando sucessivamente este processo, pode-se chegar a um erro mínimo estipulado que é representado pela norma do vetor gradiente \mathbf{g} . Quando $\|\mathbf{g}\|$ se aproxima de um valor mínimo de parada, a função-objetivo O foi minimizada.

B.1**O Grupo das Transformações de Corpo Rígido $SE(3)$**

Elementos do grupo das transformações de corpo rígido $SE(3)$ — utilizado para minimizar a função-objetivo da Seção 3.7.1 — são geralmente representados pelo par $\langle R, T \rangle$ onde $R \in SO(3)$ e $T \in \mathbb{R}^3$. Utilizando a técnica recém-mostrada para minimização utilizando o grupo das rotações, podemos também criar uma parametrização local em $S : \mathbb{R}^6 \rightarrow SE(3)$ que mapeia elementos de uma esfera aberta no \mathbb{R}^6 em uma vizinhança em torno de um ponto $\langle R_0, T_0 \rangle$ nesse *manifold*:

$$S(\omega, t) = \langle R_0 \exp\{J(\omega)\}, T_0 + t \rangle \quad \omega, t \in \mathbb{R}^3 \quad (\text{B-3})$$

onde a dupla $\langle \omega, t \rangle$ pode ser vista como um elemento de uma região aberta no \mathbb{R}^6 .