

文章编号: 0253-2239(2004)05-628-5

一种无相机标定的立体图像对校正新方法*

曾吉勇 苏显渝

(四川大学光电科学技术系, 成都 610064)

摘要: 双目立体视觉是光学被动三维测量的主要方法。为了快速、准确地寻找对应点,通常要将立体图像对进行校正,消除垂直视差。提出了一种无需标定相机的立体图像对校正方法。该方法从基本矩阵计算初始透视投影、旋转和竖直平移变换矩阵,然后以对应点坐标为基础对这些变换矩阵进行优化计算,从而有效地避免了优化计算的局部最小值,而且不过分依赖基本矩阵的计算精度。通过对提出的方法和完全基于基本矩阵的图像校正方法以及无需基本矩阵计算的图像校正方法进行实验比较,结果表明提出的方法图像校正速度快,能有效地消除垂直视差,而产生的图像变形较小。

关键词: 计算机视觉; 图像校正; 立体图像对; 投影校正; 基本矩阵

中图分类号: TN863 文献标识码: A

A New Rectification Method of Stereo Pairs Without Calibration for Cameras

Zeng Jiyong Su Xianyu

(Department of Optoelectronics, Sichuan University, Chengdu 610064)

(Received 10 March 2003; revised 23 May 2003)

Abstract: The binocular stereovision is a primary approach to realize optical passive 3-D measurement. To search corresponding points fast and accurately, the stereo pairs must be rectified. A method is presented to rectify stereo pairs without calibration for cameras. In this method the initial values of perspective projective matrix; rotative matrix and translative matrix are calculated from the fundamental matrix, than these matrices are optimized based on the coordinate value of corresponding points. This method can avoid the local least value in the optimal calculation effectively, and the dependence on the calculation precision of the fundamental matrix excessively. By the experiment of image rectification to compare this method put forward this here with two typical methods that one depend on fundamental matrix calculation and the other dispenses with fundamental matrix calculation, it is shown that this method can rectify stereo pairs fast and accurately, but bring little image distortion.

Key words: computer vision; image rectification; stereo pairs; projective rectification; fundamental matrix

1 引 言

光学三维传感分为主动三维传感和被动三维传感^[1~3],双目立体视觉是光学被动三维传感的主要方法。立体图像对是相机在不同的位置对同一场景的

记录,同一物点在两个成像平面内的像点称为对应点,从立体图像对中对对应点的视差可以提取出场景的距离信息。由于立体图像对存在外极几何约束^[4],对应点可以在相应的外极线上寻找(如图 1 所示)。理想的立体图像对中对对应点只有水平视差,不存在垂直视差,理想立体图像对的匹配只需沿图像的水平扫描线搜寻对应点,因此理想立体图像对的匹配速度较快,能消除匹配的多义性。立体图像对的校正就是将一般的立体图像对变换为理想立体图像对的过程^[5]。

* 国家自然科学基金(60177001)资助课题。

E-mail: zengjiyong7@163.com

收稿日期:2003-03-10;收到修改稿日期:2003-05-23

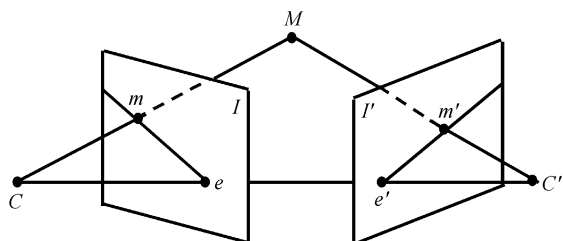


Fig. 1 Epipolar geometry between a pair of stereo images

在立体视觉中,图像对校正是相当重要的,可以在相机标定^[6]和无相机标定^[7,8]的情况下进行。由于无相机标定的立体视觉^[9,10]具有更大的适应性,已成为立体视觉的重要研究领域。无相机标定的立体图像对校正的前提是已知一系列对应点的坐标。因此在图像对校正前须分别在图像对中通过兴趣点提取算子提取出兴趣点集,然后用特征匹配方法建立两个兴趣点集中兴趣点的一一对应关系,获得对应点集。在无相机标定的立体图像对校正算法中,Loop等^[7]提出了基于基本矩阵的图像校正方法,该方法过分依赖于基本矩阵的精确求解,而对对应点计算基本矩阵还没有较为稳定的算法,因此这种图像校正方法很容易崩溃,而且没有较好地校正垂直视差。Francesco等^[8]提出了一种无需基本矩阵计算的图像校正方法,该方法只依赖于对应点的坐标,而无视基本矩阵的作用,校正过程中非线性优化计算的初始值选取缺乏可信度,采用金字塔结构的优化过程计算量很大,是一种不稳定的方法。

本文提出了一种无需相机标定的立体图像对平面投影校正方法。本方法由基本矩阵计算透视投影、旋转和竖直平移变换矩阵的初始值,然后以对应点坐标为基础对这些变换矩阵进行优化计算,校正中还考虑了图像变形的消除。由于利用了基本矩阵的指导作用和对对应点坐标的决定性作用,本方法有效地避免了非线性优化计算的局部最小值,而且不过分依赖基本矩阵的计算,是一种速度快、精度高的立体图像对校正方法。

2 图像对平面投影校正的数学描述

2.1 图像对平面投影校正

设 \mathbf{H} 、 \mathbf{H}' 分别为图像 I 、 I' 的平面投影变换矩阵, m 、 m' 为对应点,变换后图像对应点为 \bar{m} 、 \bar{m}' ,有

$$\bar{m} = \mathbf{H}m, \quad (1)$$

$$\bar{m}' = \mathbf{H}'m', \quad (2)$$

校正前图像对的基本矩阵为 \mathbf{F} ,外极方程为^[1]

$$m^T \mathbf{F} m = 0, \quad (3)$$

校正后图像对的外极点位于水平方向上的无穷远处,基本矩阵为

$$\bar{\mathbf{F}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \quad (4)$$

校正后图像对的外极方程为

$$(\mathbf{H}'m')^T \bar{\mathbf{F}} (\mathbf{H}m) = 0. \quad (5)$$

图像平面投影校正的任务是计算 3×3 投影变换矩阵 \mathbf{H} 、 \mathbf{H}' ,满足(5)式的投影变换矩阵 \mathbf{H} 、 \mathbf{H}' 用于图像对的变换得到校正后的图像对具有理想立体图像对的特征,但满足(5)式的变换 \mathbf{H} 、 \mathbf{H}' 不是唯一的,所以图像对的校正是一个多解的问题。评价一个图像对校正算法的好坏要看它是否符合以下原则: i) 对应点垂直视差趋于 0; ii) 变换后的立体图像对变形较小; iii) 算法稳定,速度快。

2.2 平面投影变换矩阵

在图像对的校正过程中,图像的平面投影变换可以分解为 4 个过程: i) 透视投影变换,将外极点变换至无穷远处; ii) 旋转变换,使外极线位于水平方向; iii) 竖直平移,使对应点的纵坐标相等,二幅图像中只要一幅平移即可; iv) 水平缩放,使图像变形最小,并不影响竖直方向的校正结果。因此平面投影变换矩阵可表示为 $\mathbf{H} = \mathbf{H}_s \mathbf{H}_t \mathbf{H}_r \mathbf{H}_p$,其中 \mathbf{H}_p 为透视投影矩阵, \mathbf{H}_r 为旋转矩阵, \mathbf{H}_t 为竖直平移矩阵, \mathbf{H}_s 为水平缩放矩阵。

1) 透视投影变换

$$\mathbf{H}_p = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \tau_a & \tau_b & 1 \end{bmatrix}, \quad (6)$$

2) 旋转变换

$$\mathbf{H}_r = \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (7)$$

3) 竖直平移

$$\mathbf{H}_t = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{bmatrix}, \quad (8)$$

4) 水平缩放

$$\mathbf{H}_s = \begin{bmatrix} s_a & s_b & s_c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (9)$$

图像校正的关键在于通过透视投影、旋转和竖直平移获得外极线位于水平方向且已水平校正的图

像对。因此图像对的校正精度可以用校正后图像对对应点纵向绝对视差的平均值衡量,完全校正的图像对其纵向绝对视差平均值为 0。透视投影、旋转和竖直平移决定了校正精度。如果二幅图像中只有第一幅图像在竖直方向上平移,则第一幅图像的平面投影变换的前三个过程为

$$\mathbf{H}_{\text{tp}} = \mathbf{H}_t \mathbf{H}_r \mathbf{H}_p = \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ w_a t - \sin \theta & w_b t + \cos \theta & t \\ w_a & w_b & 1 \end{bmatrix}, \quad (10)$$

第二幅图像的平面投影变换前三个过程为

$$\mathbf{H}'_{\text{tp}} = \mathbf{H}'_r \mathbf{H}'_p = \begin{bmatrix} \cos \theta' & \sin \theta' & 0 \\ -\sin \theta' & \cos \theta' & 0 \\ w'_a & w'_b & 1 \end{bmatrix}. \quad (11)$$

$$f_i(\mathbf{H}'_{\text{tp}}, \mathbf{H}_{\text{tp}}) = (\mathbf{H}'_{\text{tp}} m'_i)^T \bar{\mathbf{F}}(\mathbf{H}_{\text{tp}} m_i) = (w'_a x'_i + w'_b y'_i + 1)[(t w_a - \sin \theta) x_i + (t w_b + \cos \theta) y_i + t] - (w_a x_i + w_b y_i + 1)(y'_i \cos \theta' - x'_i \sin \theta').$$

已知多于 7 对对应点的坐标,用非线性最小二乘法求解 (12) 式,可以计算出使评价函数 $\cos t(\mathbf{H}_{\text{tp}}, \mathbf{H}'_{\text{tp}})$ 最小的 $\theta, w_a, w_b, t, \theta', w'_a, w'_b$, 确定变换矩阵 \mathbf{H}_{tp} 和 \mathbf{H}'_{tp} 。

3.2 水平缩放矩阵的计算

水平缩放只是为了确保校正后图像变形较小,故式(9)中可令 $S_c = 0$, 即

$$\mathbf{H}_s = \begin{bmatrix} s_a & s_b & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

假设图像的宽度为 w , 高度为 h 。选取图像边缘的四个中点(如图 2)所示。

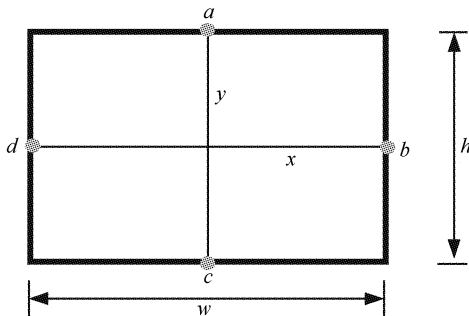


Fig. 2 The four midpoints on edges selected to calculate the horizontal zoom matrix

$$a = [w/2 \ 0 \ 1], \quad b = [w \ h/2 \ 1], \\ c = [w/2 \ h \ 1], \quad d = [0 \ h/2 \ 1],$$

四个边缘中点经过透视投影、旋转和竖直平移

3 无相机标定的立体图像对校正方法

3.1 透视投影、旋转和竖直平移矩阵的优化计算

通过透视投影、旋转和竖直平移将获得外极线位于水平方向,且已水平校正的理想图像对。因此确定变换矩阵 \mathbf{H}_{tp} 和 \mathbf{H}'_{tp} 是图像对校正的核心。

由校正后图像对的外极方程(5),得到评价函数

$$\cos t(\mathbf{H}_{\text{tp}}, \mathbf{H}'_{\text{tp}}) = \sum_{i=1}^N [(\mathbf{H}'_{\text{tp}} m'_i)^T \bar{\mathbf{F}}(\mathbf{H}_{\text{tp}} m_i)]^2, \quad (12)$$

其中 N 为图像对中参与计算的对应点点数。 m'_i, m_i 为对应点的齐次坐标。

$$m_i = [x_i \ y_i \ 1]^T, \quad m'_i = [x'_i \ y'_i \ 1]^T,$$

由(4)式、(10)式、(11)式、(12)式得

变换后(变换矩阵 \mathbf{H}_{tp})为 $a, \hat{b}, \hat{c}, \hat{d}$ 。

$$\text{二直线} \quad \begin{cases} x = \hat{b} - \hat{d} = [x_u \ x_v \ x_w] \\ y = \hat{c} - \hat{a} = [y_u \ y_v \ y_w] \end{cases},$$

图像变形较小可以从两个方面考虑

1) 二直线垂直,即

$$(\mathbf{H}_s x)^T (\mathbf{H}_s y) = 0. \quad (13)$$

2) 图像的纵横比不变

$$\frac{(\mathbf{H}_s x)^T (\mathbf{H}_s x)}{(\mathbf{H}_s y)^T (\mathbf{H}_s y)} = \frac{w^2}{h^2}, \quad (14)$$

解(13)式、(14)式得^[7]

$$\begin{cases} s_a = \frac{h^2 x_v^2 + w^2 y_v^2}{h w (x_v y_u - x_u y_v)}, \\ s_b = \frac{h^2 x_u x_v + w^2 y_u y_v}{-h w (x_v y_u - x_u y_v)}, \end{cases}$$

3.3 透视投影、旋转和竖直平移变换矩阵初始值的计算

用非线性最小二乘法求解(12)式确定变换矩阵 \mathbf{H}_{tp} 和 \mathbf{H}'_{tp} 的优化过程初始值选取十分重要。采用 Loop 等^[7]的方法由基本矩阵可以计算出透视投影变换的初始值 $w_{a0}, w_{b0}, w'_{a0}, w'_{b0}$ 和旋转变换矩阵的初始值 θ_0, θ'_0 , 在计算透视变换矩阵过程中保证了透视变形最小。竖直平移矩阵初始值的计算中,图像经过透视投影变换和旋转变换后外极线应为水平方向的平行线,为使对应的外极线在同一扫描线上,消除竖直方向的视差,应将图像对中的任一图像沿竖直方向平移。设图像 I, I' 经过透视投影变换和旋转变换

后的对应点的纵坐标为 $y_i, y'_i, i = 1, 2, \dots, N$ 。则图像 I 的平移量初始值 $t_0 = \frac{1}{n} \sum_{i=1}^n (y'_i - y_i)$ 。

4 实 验

用三种图像对校正方法进行了实验比较,这三种方法分别为:i) Loop 等^[7]提出的完全基于基本矩阵的图像校正方法;ii) Francesco 等^[8]提出的无需基本矩阵计算的图像校正方法;iii) 本文提出的用基本矩阵作为指导,以对应点坐标为基础的图像校正方法。实验选用三组图像对进行校正。i) 哈维头像对;ii) 桌面图像对。图像用 SONY DVCAM DSR-PDX10P 摄像机拍摄;iii) 花瓶图像对。图像用 Mintron EX1881 摄像机和圆刚图像采集卡拍摄。

表 1 比较了三种校正方法的校正精度、图像变

形和校正速度。图像校正精度用垂直视差衡量, Loop 方法校正后图像对的垂直视差最大,本文方法垂直视差最小。图像变形主要指透视变形和图像旋转, Francesco 方法图像变形较大,本文方法图像变形最小。一般来说非线性优化过程越少校正速度越快,因此可用非线性优化过程次数的多少来衡量图像校正快慢, Loop 方法用了一次解一元六次方程的非线性优化过程, Francesco 方法用了数十次解无约束非线性最小二乘法优化过程,本文的方法用了一次解一元六次方程的非线性优化过程和一次无约束非线性最小二乘法优化过程。可见本文方法校正精度高、图像变形小、校正速度较快。由于通过基本矩阵计算透视投影、旋转和竖直平移变换矩阵的初始值,因此本文的图像对校正方法对基本矩阵具有一定的依赖性,即基本矩阵的确定仍要有一定的精度。

Table 1 Compare among the three rectifying methods

	Images	Loop's method	Francesco's method	Author's method
Average vertical disparity / pixel	Head portrait	1.478	0.192	0.154
	Tabletop	0.194	0.118	0.119
	Vase	2.063	0.277	0.218
Image distortion	Head portrait	Middle	Large	Little
	Tabletop	Little	Large	Little
	Vase	Middle	Large	Little
Rectifying speed (the number of nonlinear optimization)	Head portrait	1	98	2
	Tabletop		28	
	Vase		50	

图 3 为桌面图像对及其外极线图,图 4 为用本文方法校正后的桌面图像对;图 5 为一座小山

的图像对及其外极线图,图 6 为用本文方法校正后的小山图像对。

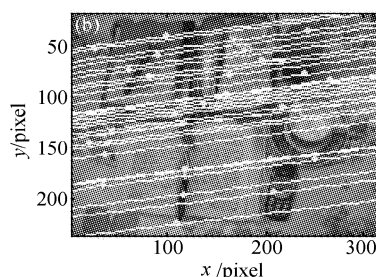
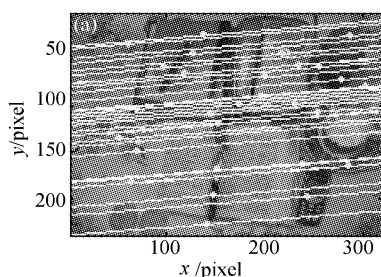


Fig. 3 A Pair of tabletop images and its epipolar line. (a) The left image, (b) the right image

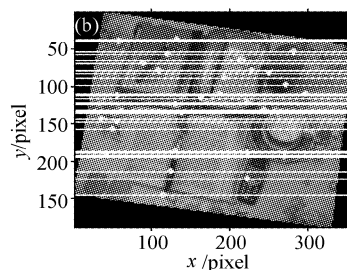
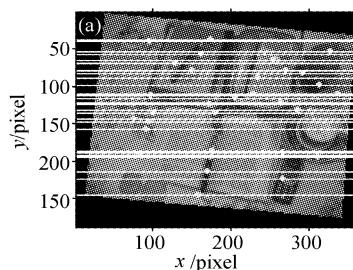


Fig. 4 A pair of rectified tabletop images. (a) The left image, (b) the right image

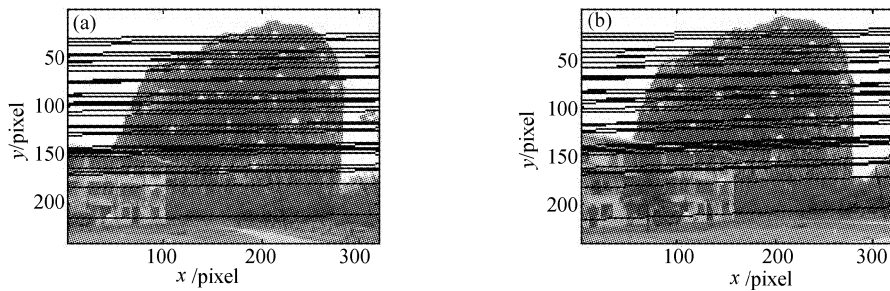


Fig. 5 A pair of hill images and its epipolar line. (a) The left image, (b) the right image

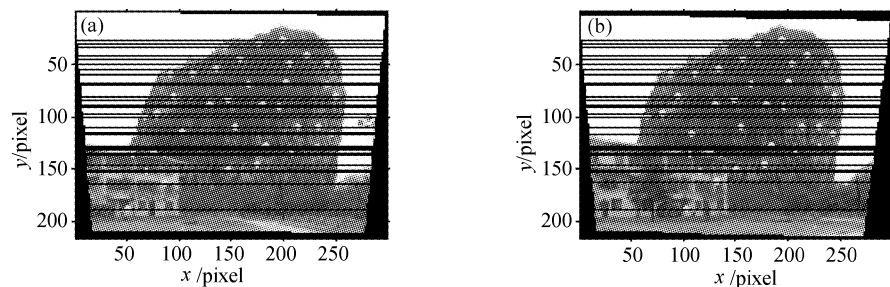


Fig. 6 A pair of rectified hill images. (a) The left image, (b) the right image

图 3、图 5 均用手持摄像机在二个不同位置拍摄,图 5 中图像对拍摄时摄像机移动了 8 m。图像校正时分别提取了 30 和 37 个对应点,从图 4 和图 6 可以看出校正后的图像中对对应点的垂直视差已基本消除。

结论 本文提出了一种无需相机标定的立体图像对校正方法,具有校正速度快、精度高的特点。用本文提出的方法和 Loop 等^[7]提出的完全基于基本矩阵的图像校正方法以及 Francesco 等^[8]提出的无需基本矩阵计算的图像校正方法进行了实验比较。实验表明本文的方法在平均垂直视差、图像变形和校正速度等方面具有明显的优越性。

参 考 文 献

- 1 Su Xianyu, Tan Songxin, Xiang Niquan *et al.*. Complex object shape measurement using FTP method. *Acta Optica Sinica* (光学学报), 1998, **18**(9):1128~1233 (in Chinese)
- 2 Zhang Qican, Su Xianyu. Dynamic liquid surface shape measurement. *Acta Optica Sinica* (光学学报), 2001, **21**

- (12):1506~1508 (in Chinese)
- 3 Li Wansong, Su Likun, Su Xianyu. Phase-measuring profilometry in big scale measureme. *Acta Optica Sinica* (光学学报), 2000, **20**(6):792~796 (in Chinese)
- 4 Faugeras O D. Stratification of 3-D vision: Projective, affine, and metric representations. *J. Opt. Soc. Am.* (A), 1995, **12**(3):465~484
- 5 Hartley R. Theory and practice of projective rectification. *Int. J. Comput. Vision*, 1999, **35**(2):1~16
- 6 Fusiello A, Trucco E, Verri A. A compact algorithm for rectification of stereo pairs. *Machine Vision Applications*, 2000, **12**(1):16~22
- 7 Loop C, Zhang Zhengyou. Computing rectifying homographies for stereo vision. *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, 1999. 125~131
- 8 Francesco I, Trucco E. Projective rectification without epipolar geometry. *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, 1999. 94~99
- 9 Faugeras O D. What can be seen in three dimensions with an uncalibrated stereo rig? *Proc. 2nd European Conf. on Computer Vision*, Santa Margherita Ligure, 1992. 563~578
- 10 Hartley R, Gupta T, Chang T. Stereo from uncalibrated cameras. *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, Champaign, 1992. 761~764