

# 基于本质矩阵的摄像机自标定方法

江泽涛 吴文欢

(南昌航空大学教育部无损检测重点实验室,南昌 330063) (南昌航空大学信息工程学院,南昌 330063)

**摘要** 本质矩阵描述了在摄像机内参数矩阵已知的条件下的对极几何关系,是归一化图像坐标下的基础矩阵。鉴于本质矩阵具有两相等的非零奇异值,提出了一种基于本质矩阵的自标定方法,该方法首先利用本质矩阵这个特性来构造目标函数,考虑到传统非线性优化算法的诸多不足,最后用粒子群优化算法来求解。实验结果表明,该方法精度较高、鲁棒性较强,是一种简单而有效的自标定方法。

**关键词** 摄像机自标定 本质矩阵 基础矩阵 粒子群算法

中图法分类号: TP391 文献标志码: A 文章编号: 1006-8961(2010)04-565-05

## An Essential Matrix-based Camera Self-Calibration Method

JIANG Zetao, WU Wenhuan

(Key Laboratory of Nondestructive Test, Ministry of Education, Nanchang Hangkong University, Nanchang 330063)

(School of Information Engineering, Nanchang Hangkong University, Nanchang, 330063)

**Abstract** The essential matrix, which is a fundamental matrix described in the normalized image coordinate, represents the epipolar geometry relation under the condition of known camera intrinsic parameters. It is well known, two non-zero singular values of the essential matrix must be equal. Therefore, according to the intrinsic property, an essential matrix-based self-calibration approach is for proposed the first time in this paper. First, the objective function is constructed by the intrinsic property of the essential matrix. Second, the particle swarm optimization is used to solve the objective function considering the drawbacks of traditional optimization algorithms. Analytical results show that the proposed method is not only highly accurate but also robust. Consequently, it is a simple but valid self-calibration method.

**Keywords** intrinsic parameters, essential matrix, fundamental matrix, PSO

## 0 引言

摄像机标定是从2维图像获取3维信息必不可少的步骤,被广泛应用于航空航天、遥感图像、机器视觉、虚拟现实、3维重建等领域<sup>[1]</sup>。广义上摄像机标定可以分为两种:传统标定方法和自标定方法。传统标定方法需要使用经过精密加工的标定块,通过建立标定块上3维坐标已知的点与其图像点间的对应关系来计算摄像机的内外参数,该方法的优点

在于可以获得较高的精度,但标定过程费时费力,不适合在线标定和不可能使用标定块的场合,而自标定不需要标定物,仅利用从图像序列中得到的约束关系来进行标定。

20世纪90年代初,Maybank等人提出了基于Kruppa方程的自标定方法<sup>[2]</sup>,由于Kruppa方程是二次非线性约束,所以求其解比较困难。尽管运用代数几何知识或同伦连续算法可以直接求解Kruppa方程,但是当图像数目增加时,可能解的个数呈指数增长,使得直接求解失去意义。Hartley首次利

基金项目:国家自然科学基金项目(60673055,60973096);国家航空基金项目(2007zc56003);江西省自然科学基金项目(2008GZS0033)

收稿日期:2008-10-16;改回日期:2009-03-13

第一作者简介:江泽涛(1961—),男,教授,博士生导师。2006年获西北工业大学计算机应用技术专业博士学位。主要研究领域为图像处理模式识别和计算机视觉等。E-mail: zetaojiang@yahoo.com.cn

用本质矩阵拥有两相等非零奇异值这一特性得出两个关于摄像机内参数的约束条件,然后利用这些约束进行自标定,但他所提算法只能基于两幅视图并且最多只能求出两个内参数<sup>[3]</sup>。文献[4]则改进了Hartley的算法,提出了一种简单的适用于多幅视图摄像机标定技术,可以标定出所有摄像机内参数。但是该算法是用传统的非线性优化算法来最小化由奇异值所构造的代价函数,并且其初始值是根据经验所得,而没有给出一个有效初始化的方法,从而更加容易陷入局部最优。

由于粒子群优化算法是目前比较流行的一种优化方法,具有原理简单、易于实现、不需要初始值并能以较大概率来求得全局最优解的优点,已经在许多优化问题中得到成功应用,所以针对传统优化方法的缺点,本文提出了利用粒子群优化算法最优化由奇异值构造的代价函数从而实现摄像机自标定的方法。

## 1 基于本质矩阵的自标定

### 1.1 摄像机参数的本质矩阵

假设摄像机的模型是常用的针孔模型,因此从3维空间点  $X = (x, y, z, 1)^T$  到2维图像点  $m = (u, v, 1)^T$  的成像关系可以表示为  $m \cong K [R | t] X$ 。其中

$$K = \begin{bmatrix} f_u & 0 & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

其中,  $K$  是摄像机的内参数矩阵,  $R, t$  是摄像机坐标系相对于世界坐标系的旋转矩阵与平移向量,符号“ $\cong$ ”表示在相差一个比例因子意义下的相等。

如果已知内参数矩阵  $K$ , 用它的逆作用于图像点  $m$  得到点  $\tilde{m} = K^{-1}m = [R | t]X$ , 则  $\tilde{m}$  就是图像点在归一化坐标下的表示。而本质矩阵就是归一化图像坐标下的基础矩阵的特殊形式。由于本质矩阵比基础矩阵少两个自由度,所以在基础矩阵已知的条件下,可以构成对摄像机内参数的两个约束。它与基础矩阵之间的关系是

$$E = K^T F K \quad (1)$$

既然基础矩阵  $F$  的秩为2以及内参数矩阵  $K$  是满秩,那么本质矩阵  $E$  的秩显然为2,所以本质矩阵  $E$  有个奇异值为0,而文献[3]还证明了本质矩阵  $E$  的其他两个非零奇异值必须相等,并且这两个条件构成了任意3阶方阵是本质矩阵的充分必要条件。

### 1.2 基于本质矩阵的标定方法

假定图像序列中有  $M$  幅视图,并且用相连的两幅视图得到一个基础矩阵,那么可以得到  $M-1$  个基础矩阵,对应地有  $M-1$  个本质矩阵。对于第  $i$  个本质矩阵  $E_i$ ,假定它两个非零的奇异值分别为  $\sigma_{i1}, \sigma_{i2}$ ,其中  $\sigma_{i1} \geq \sigma_{i2}$ ,在理想情况下,  $\sigma_{i1} = \sigma_{i2}$ ,只要找到某个  $K$  使得以上奇异值相等或尽可能地近似相等,那么就实现了摄像机自标定。为了实现自标定,构造如下目标函数:

$$fitness = \sum_{i=1}^{M-1} \frac{\omega_i}{\sum_{i=1}^{M-1} \omega_i} \left(1 - \frac{\sigma_{i1}}{\sigma_{i2}}\right) \quad (2)$$

式中,  $\omega_i$  是加权因子,在  $[0, 1]$  范围之内,表示估计第  $i$  个基础矩阵时的可信度,等于在计算该基础矩阵时用到的对应点个数。这个目标函数中有4个变量  $f_u, f_v, u_0, v_0$ ,可以通过最小化这个目标函数使其趋于0进而求得各个摄像机内参数。因此,目标函数的最小化计算是成功的关键,下面探讨用粒子群优化算法来完成这个最小化过程。

## 2 自标定算法

### 2.1 粒子群优化算法

粒子群优化算法(PSO)是Kennedy和Eberhart于1995年共同提出的一种基于群体智能的进化算法<sup>[5-7]</sup>。PSO初始化为一群随机粒子(随机解),然后通过迭代找到最优解。在每一次迭代中,粒子通过跟踪两个“极值”来更新自己:一个是粒子本身所找到的最优解,称为个体极值(pBest);另一个极值是整个种群目前找到的最优解,这个极值就是全局极值(gBest)。每个粒子都有由目标函数决定的适应值,将其位置坐标代入目标函数可以得到它的适应度值,进而通过所求的适应值来衡量该粒子的优劣。并且每个粒子还有一个速度矢量,决定着它在解空间飞行的方向和距离。每个粒子根据当前所找的gBest与其自身的pBest来更新速度和位置,经过迭代搜索,算法最终可得到问题的最优解。

假设在一个  $D$  维的搜索空间中,有  $N$  个粒子组成一个群体,其中第  $t$  代第  $i$  个粒子的位置坐标表示为向量

$$x_i^t = (x_{i1}, x_{i2}, \dots, x_{id}, \dots, x_{iD})^T$$

其速度可表示为

$$\mathbf{v}_i^t = (\nu_{i1}^t, \nu_{i2}^t, \dots, \nu_{id}^t, \dots, \nu_{id}^t)^T$$

个体最优位置表示为

$$\mathbf{p}_i^t = (p_{i1}^t, p_{i2}^t, \dots, p_{id}^t, \dots, p_{id}^t)^T$$

第  $t$  代种群的全局最优表示为

$$\mathbf{p}_g^t = (p_{g1}^t, p_{g2}^t, \dots, p_{gd}^t, \dots, p_{gd}^t)^T$$

对于第  $t+1$  代第  $i$  个粒子,其速度和位置的第  $d$  维分量,根据如下公式迭代更新:

$$\mathbf{v}_{id}^{t+1} = \omega \mathbf{v}_{id}^t + c_1 r_1 (\mathbf{p}_{id}^t - \mathbf{x}_{id}^t) + c_2 r_2 (\mathbf{p}_{gd}^t - \mathbf{x}_{id}^t) \quad (3)$$

$$\mathbf{x}_{id}^{t+1} = \mathbf{x}_{id}^t + \mathbf{v}_{id}^{t+1} \quad (4)$$

式中,  $\omega$  是非负数,称为动量系数,控制着前一速度对当前速度的影响。 $\omega$  值越大,全局寻优能力越强,局部寻优能力越弱;反之,则局部寻优能力增强,而全局寻优能力减弱。通常将迭代开始时的惯性权重  $\omega$  设置较大,并在迭代过程中逐渐减小。加速度常数  $c_1, c_2$  用来调整粒子自身经验与群体经验在其运动中所起作用的权重,通常  $c_1 = c_2 = 2$ 。 $r_1$  与  $r_2$  为(0,1)范围内均匀分布的随机数。一般需要设定粒子速度每一维分量的上限  $\mathbf{v}_{\max}$ ,即当  $|\mathbf{v}_{id}| > \mathbf{v}_{\max}$  时,取  $|\mathbf{v}_{id}| = \mathbf{v}_{\max}$ 。

### 2.2 自标定算法

以下是用粒子群优化算法进行优化的摄像机自标定步骤:

- 1) 检测出每幅视图中的兴趣点;
- 2) 匹配每相连两幅视图中的兴趣点;
- 3) 用文献[8]的方法计算每相连两幅视图间的基础矩阵;
- 4) 构造式(2)中的目标函数  $fitness$ ;
- 5) 设定群体规模,初始化所有粒子,即在允许范围内对群体中粒子的位置和速度进行随机初始化;
- 6) 把每个粒子的位置代入目标函数  $fitness$  中,

计算出每个粒子的适应值,并与它经历过的最好位置  $pBest$  进行比较,进而更新粒子当前的最好位置  $pBest$ ;

7) 在粒子群中,选择拥有最好适应值的粒子作为当前群体全局极值  $gBest$ ;

8) 根据式(3)和(4)分别对每个粒子的位置和速度进行更新;

9) 检查终止条件(当  $fitness$  小于误差容忍限度或到达最大迭代次数),若上述条件满足,终止迭代,输出  $gBest$  作为所要求的摄像机内参数,否则返回第6)步。

## 3 实验结果

### 3.1 模拟数据实验

在模拟实验中,摄像机内参数设置为  $f_u = f_v = 1\ 000, u_0 = 300, v_0 = 400$ 。让此模拟摄像机对着随机分布在空间中球心位于原点的单位球内的50个点拍摄并连续地改变其外参数,从而得到10幅模拟图像。由于在实际应用中,提取兴趣点必定会引入噪声,所以在模拟实验中向图像点加上了零均值高斯噪声,噪声水平从0.0像素以0.2像素的步长依次增加到2.0个像素,并且在每种噪声水平下分别进行100次独立实验。为了评价本文方法的性能,在此考察每个内参数的相对误差,即  $\varepsilon_p = \frac{1}{Np} \sum_{i=1}^N |\bar{p}_i - p|$ ,其中  $N$  表示实验次数,  $p \in (f_u, f_v, u_0, v_0)$  表示某个参数的真值,  $\bar{p}_i \in (f_{ui}, f_{vi}, u_{0i}, v_{0i})$  表示某个参数在第  $i$  次实验中得到的值。图1表示模拟实验中摄像机内参数相对误差随噪声变化的情况。

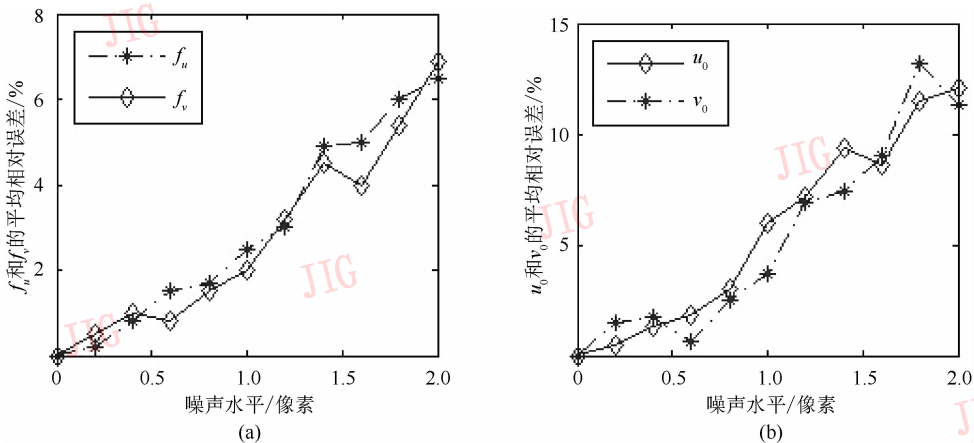


图1 4个内参数在不同噪声水平下的相对误差

Fig. 1 The relative errors of the four intrinsic parameters under different noise level

从实验结果可以看出,在噪声水平为 1.0 像素时,  $f_u, f_v$  的相对误差在 2% 左右,  $u_0, v_0$  的相对误差也不到 5%,而在噪声比较大的情形下(如 2.0 个像素),相对误差仍然比较小,标定结果还是令人满意的,这说明本文方法标定精度较高,鲁棒性也较强。

### 3.2 真实图像实验

为了说明本文方法的有效性,为此采用与文献[9-12]中相同的图像序列,对同一摄像机进行标定,然后与文献中的标定结果进行对比。采用从 <http://www.robots.ox.ac.uk/~vgg/data/data-mview.html> 下载的 Valbonne church 图像序列,图像分辨率为  $512 \times 768$ ,选取其中 9 幅作为测试图片,图 2 是其中的一幅。首先用 SIFT 对图像进行特征提取和匹配<sup>[13]</sup>,手工选取正确的匹配点;然后用 8 点算法求解基础矩阵,构造出目标函数;最后用粒子群优化算法求解。



图 2 Valbonne church 的图像帧

Fig. 2 A frame image of Valbonne church sequence

表 1 对本文方法得出的结果与文献[9-12]给出的结果做了比较。这里采取与文献[11-12]相同的做法,把文献[9-10]的结果当做参考标准,因为他们是比较经典的基于 Kruppa 方程的自标定算法,有着较高的精度和鲁棒性,重构效果也相当好。从表 1 可以看出,相比于文献[11-12],本文所得结果与标准结果更接近,这进一步证实了本文方法的精确性和有效性。

本文利用 PSO 寻优时一般可以在 600 代内快速且准确地收敛到最优值,对于奔腾 4、2.80 GHz 的

表 1 不同算法的标定结果对比

Tab. 1 Estimation results of the proposed method and other methods

	$f_u$	$f_v$	$u_0$	$v_0$
文献[9]	682.84	682.843	255.999	383.999
文献[10]	679.285	681.345	258.802	383.188
文献[11]	667	693	241	398
文献[12]	619	699	234	372
本文方法	670.457	679.323	248.648	392.561

处理器,大概耗时在 20 s 内,这和特征点提取与匹配耗时长达几分钟相比,本文方法的自标定时间消耗并不算长。

## 4 结论

提出了一种基于本质矩阵的摄像机自标定方法。本质矩阵不仅描述摄像机对之间的位置关系,还可以对摄像机内参数构成约束。本文利用本质矩阵具有两相等的非零奇异值这一约束条件来构造目标函数。考虑到传统的优化方法很依赖于初始值,并且容易陷入局部最优,所以利用粒子群优化算法对这个目标函数进行最小化,进而求得各个摄像机内参数。

在粒子群优化的过程中,根据一般相机的实际情况,主点范围取在以图像中心点为中心、大小为图像大小的  $1/5$  的矩形范围内,两等效焦距的初始范围可以根据 Kruppa 方程求得,并以 500 的长度进行选取。当两焦距的长度比约为 1,并且主点接近图像中心点时,所得结果就是最终所求的内参数,最后也可以用重构效果来检验。模拟数据实验和真实图像实验充分证明了提出的方法精度较高,鲁棒性较强,是一种有效的、易实施的自标定方法。

### 参考文献 (References)

- [1] Fusiello A. Uncalibrated Euclidean reconstruction: A review[J]. *Image and Vision Computing*, 2000, 18(6-7): 555-563.
- [2] Maybank S J, Faugeras O D. A theory of self-calibration of a moving camera[J]. *International Journal of Computer Vision*, 1992, 8(2): 123-151.
- [3] Hartley R. Estimation of relative camera position for uncalibrated cameras[C]//*Proceedings of the 2nd European Conference on Computer Vision*. London, UK: Springer-Verlag, 1992: 579-587.

- [4] Mendonca P, Cipolla R. A simple technique for self-calibration [C]. In the Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE Press, 1999: 112-116.
- [5] Kennedy J, Eberhart R. Particle swarm optimization [C]// Proceedings of the IEEE International Conference on Neural Networks. Piscataway, NJ, USA: IEEE Press, 1995: 1942-1948.
- [6] Kennedy J, Eberhart R. A new optimizer using particle swarm theory [C]// Proceedings of the 6th International Symposium on Micro Machine and Human Science. Piscataway, NJ, USA: IEEE Press, 1995: 39-43.
- [7] Shi Y, Eberhart R. A modified particle swarm optimizer [C]// Proceedings of the IEEE International Conference on Evolutionary Computation. Piscataway, NJ, USA: IEEE Press, 1998: 69-73.
- [8] Hartley R. In defence of the eight-point algorithm [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(6): 580-593.
- [9] Lourakis L, Deriche R. Camera Self-Calibration Using the svd of the Fundamental Matrix: Technical Report 3748 [R]. Sophia Antipolis, France: The France National Institute for Research in Computer Science and Control, 1999.
- [10] Zeller C, Faugeras O D. Camera Self-Calibration from Video Sequences: the Kruppa Equations Revisited: Research Report 2793 [R]. Sophia Antipolis, France: The France National Institute for Research in Computer Science and Control, 1996.
- [11] Hated A, Boufama B. Camera self-calibration from bivariate polynomials derived from Kruppa's equations [J]. Pattern Recognition, 2008, 41(8): 2484-2492.
- [12] Fusiello A, Benedetti A, Farenzena M, et al. Globally convergent autocalibration using interval analysis [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(12): 1633-1683.
- [13] Lowe D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2): 91-110.