

一种基于平面模板的单应矩阵求解方法

付海洋,牛连强,刘守琳

(沈阳工业大学 研究生学院,辽宁 沈阳 110023)

摘 要:在视觉测量、摄像机标定和三维重建等领域中,单应矩阵的估计都起着至关重要的作用。用平面模板代替传统的传感器进行视觉测量和摄像机标定,由于其成本低,操作快速简单等优点,已经得到了越来越广泛的应用。文中主要讨论了二维成像平面的透视投影模型,给出了一种最大似然估计法获取平面模板与成像平面的单应矩阵的求解过程,并用 Levenberg-Marquardt 算法对结果进行非线性优化。该方法简单、快速、易实现,在计算机视觉应用中具有一定的理论意义和应用价值。

关键词:单应矩阵;平面模板;最大似然估计;计算机视觉

中图分类号:TP391.41

文献标识码:A

文章编号:1673-629X(2010)04-0069-03

A Solving Homography Matrix Method Based on Planar Pattern

FU Hai-yang, NIU Lian-qiang, LIU Shou-lin

(Graduate School, Shenyang University of Technology, Shenyang 110023, China)

Abstract: Homography plays an important role in visual metrology, camera calibration and 3D reconstruction under both theoretical and practical standpoints. Planar pattern instead of the traditional sensor conduct visual metrology and camera calibration, because of its low cost, quick and simple operation, etc., has been more and more widely used. This paper discusses the model of 2D planar projective transform and proposes a solution procedure of acquiring planar pattern and the homography matrix of imaging planar using the method of maximum likelihood estimator and Levenberg-Marquardt algorithm with the results of nonlinear optimization. The method is simple, rapid and easy to accomplish, which also has the significances of theory and the value of application.

Key words: homography matrix; planar pattern; maximum likelihood estimator; computer vision

0 引言

用平面模板代替传统的传感器进行视觉测量和摄像机标定,由于其成本低,操作快速简单等优点,已经得到了越来越广泛的应用。而平面模板上的特征点和图像平面上的点的对应关系可以用单应矩阵来联系,所以,能否准确地估计单应矩阵直接影响测量的精度。

目前关于单应矩阵的估计有四类方法^[1]:代数方法、几何方法、鲁棒估计方法和统计估计方法。代数方法是基于代数距离的最小化方法,又称线性估计方法。几何方法是基于几何距离的最小化方法,该方法是非线性的,需要迭代计算。鲁棒估计方法如随机采样一致性(RANSAC)方法^[2],是从多个特征点中随机抽取4个特征点组成一组,计算出该组的单应矩阵,然后利

用平均几何配准误差来衡量匹配精度,通过筛选得出最优单应性关系。该方法对于测量数据包含错误数据点的情况亦能给出较好的估计,但实现起来比较复杂。统计估计方法^[3]是一种在概率模型下进行估计的方法,该方法在原理上比较简单,但在测量精度要求严格的情况下效果并不理想。

前两种方法是最基本的,简单、快速且易实现,所以应用也最为广泛。然而,通过研究和实际应用发现,在一些特殊的摄像机姿态下,这两种方法对噪声非常敏感,容易产生较大的误差。针对上述情况,文中给出了一种几何方法和统计估计方法相结合的单应矩阵估计方法。将所有对应点的 Mahalanobis 距离作为目标函数,利用最大似然估计法来获得单应矩阵,并用 Levenberg-Marquardt 算法对结果进行非线性优化,简化了计算过程,同时提高了单应矩阵的准确性。

1 计算机视觉中的坐标转换关系

在计算机视觉理论中通常用到三个坐标系:图像坐标系、摄像机坐标系和世界坐标系^[4]。

收稿日期:2009-08-25;修回日期:2009-11-04

基金项目:辽宁省科技攻关项目(2002216008)

作者简介:付海洋(1983-),男,辽宁葫芦岛人,硕士研究生,研究方向为科学计算可视化与可视化仿真;牛连强,教授,博士,研究方向为图形学、计算机辅助设计和科学计算可视化与可视化仿真。

图像坐标系是存储在计算机中的数字图像的坐标系,可以用两种方法表示:第一种是用像素单位来表示,图像坐标系的原点可以在图像的左上角,也可以在图像的左下角,这里规定原点位于图像的左下角。用 (u, v) 来表示图像坐标系中的某一坐标, u 轴水平向右, v 轴垂直向上, u 轴和 v 轴的单位都是像素;第二种是用物理单位来表示,比如使用毫米。设坐标系的原点位于 (u_0, v_0) 像素点上, x 轴水平向右, y 轴垂直向上,每一个像素在 x 轴和 y 轴方向上的物理尺寸为 d_x 和 d_y 。用 (x, y) 来表示图像坐标系中的某一坐标。图像坐标系的两种表示形式有如下关系:

$$u = \frac{x}{d_x} + u_0, v = \frac{y}{d_y} + v_0 \quad (1)$$

为方便以后使用,上式可以用矩阵形式表示为:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{d_x} & 0 & u_0 \\ 0 & \frac{1}{d_y} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (2)$$

摄像机坐标系是以摄像机镜头的光心为原点, X_c 轴和 Y_c 轴分别与图像坐标系的 x 轴和 y 轴平行,摄像机的光轴 Z_c 轴垂直于图像平面的三维直角坐标系。

平面模板上的特征点在图像上的成像关系可以用针孔模型近似表示,即任一点 P 在图像上的投影位置 p ,为光心 O 与 P 点的连线 OP 与图像平面上的交点,这种关系也称为透视投影,如图 1 所示,由比例关系有如下关系式:

$$x = \frac{fX_c}{Z_c}, y = \frac{fY_c}{Z_c} \quad (3)$$

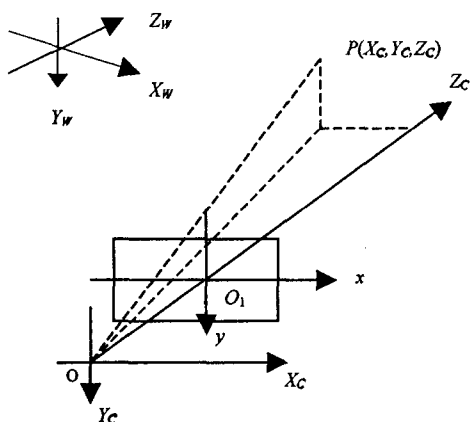


图 1 摄像机成像几何关系

则图像坐标系和摄像机坐标系有如下关系:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \frac{1}{Z_c} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \quad (4)$$

由式(2)和(4)可得:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{Z_c} \begin{bmatrix} \frac{1}{d_x} & 0 & u_0 \\ 0 & \frac{1}{d_y} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \quad (5)$$

世界坐标系是在环境中选择的一个基准坐标系,用来描述摄像机的位置,可以根据描述和计算方便等原则自由选取。

摄像机坐标系和世界坐标系的关系为:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R_{3 \times 3} & t_{3 \times 1} \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (6)$$

其中, $R_{3 \times 3} = (r_x, r_y, r_z)$ 为旋转矩阵, $t_{3 \times 1} = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}$ 为平移向量。

由式(5)和(6)可以得出世界坐标系到图像坐标系的一个线性变换:

$$km = A(R_{3 \times 3} t_{3 \times 1}) \tilde{M} \quad (7)$$

其中: k 为比例系数, $\tilde{m} = (u \ v \ 1)^T$ 为图像点的齐次坐标, $\tilde{M} = (X \ Y \ Z)^T$ 为平面模板上的点的齐次坐标。

2 单应矩阵的求解

2.1 使用最大似然估计法获得单应矩阵

最大似然估计法是一种统计方法,它用来求一个样本集的相关概率密度函数的参数。这个方法最早是由遗传学家及统计学家 Ronald Aylmer Fisher 在 1912 年开始使用的。似然即可能性,要在数学上实现最大似然估计法,首先要定义可能性: $\text{lik}(\theta) = f_D(x_1, \dots, x_n | \theta)$, 并且在 θ 的所有取值上,使这个函数最大化。这个使可能性最大的 θ 值即被称为 θ 的最大似然估计。其中: f_D 为概率密度函数, θ 为分布参数。利用最大似然估计法获取单应矩阵 H 的原理就是从平面模板中抽出若干特征点 M_1, M_2, \dots, M_n , 利用这些特征点可以找到一个关于单应矩阵 H 的估计,最大似然估计会寻找平面模板与成像平面的单应矩阵的最可能的值,即在所有可能的取值中,寻找一个取值使这个值的可能性最大化。

$$\text{令 } H = \lambda A(R \ t) \quad (8)$$

H 为单应矩阵, λ 为比例系数。

由式(7)和(8)可知,矩阵 H 实现了平面模板上的点和成像平面上的点的射影变换,称为单应矩阵^[5]。也就是说,平面模板上任一点 m 已知,根据矩阵 H 可求

出该点在物理坐标系下的坐标。矩阵 H 中包含了所有摄像机参数,这里给出 H 的求解过程,而不需要计算出各摄像机参数。单应矩阵 H 可使用最大似然估计法获取。

利用最大似然估计法获得单应矩阵的方法如下:

设 $\tilde{M}_i = (X_i Y_i Z_i 1)^T$ 为平面模板上的某一点, $\tilde{m}_i = (u_i v_i 1)^T$ 为对应的图像坐标系上的对应点。这里假设图像受到高斯噪声干扰,高斯噪声满足正太分布,其协方差矩阵为 Λ_{m_i} ,又由于各点的获取过程是独立的,因此,可以令 $\Lambda_{m_i} = \sigma^2 I$,即协方差矩阵为对角矩阵,且各对角元素相等。使用最大似然估计法获取矩阵 H 时,可以将所有对应点的 Mahalanobis 距离作为目标函数,目的就是要获取当距离取到最小值时的 H 矩阵。Mahalanobis 距离是由印度统计学家 P. C. Mahalanobis 提出的,表示数据的协方差距离。它是一种有效的计算两个未知样本集的相似度的方法。与欧式距离不同的是它考虑到各种特性之间的联系,并且是尺度无关的,即独立于测量尺度。欧氏距离与 Mahalanobis 距离相比的缺点在于,它将各变量之间的差别等同看待,这一点有时不能满足实际要求,各变量的不同属性对于区分各变量有着不同的重要性。Mahalanobis 距离矢量不仅在旋转变换和平移变换时不变,而且在尺度放大或缩小、仿射变换等情况下也保持不变^[6]。

$$\gamma^2 = \sum_i (m_i - \hat{m}_i)^T \Lambda_{m_i}^{-1} (m_i - \hat{m}_i) \quad (9)$$

其中: γ^2 为 Mahalanobis 距离的平方, $\hat{m}_i = \frac{1}{\tilde{h}_3^T \tilde{M}_i} \begin{bmatrix} \tilde{h}_1^T \tilde{M}_i \\ \tilde{h}_2^T \tilde{M}_i \end{bmatrix}$, \tilde{h}_i 为单应矩阵 H 的第 i 行。

显然,这里的最小化问题是一个非线性的最小二乘问题,可以使用 Levenberg - Marquardt 算法来解决,该算法相当于高斯牛顿法和梯度下降法的结合,具有很好的稳定性^[7]。

2.2 使用 Levenberg - Marquardt 算法进行优化

Levenberg - Marquardt 算法可用如下形式表示^[8]:

$$\begin{aligned} x_{(k+1)} &= x_k + \Delta x \\ \Delta x &= -[\nabla^2 E(x)]^{-1} \nabla E(x) \end{aligned} \quad (10)$$

其中: x_k 是 k 次迭代时权值和阈值组成的向量, $x_{(k+1)}$ 是新的权值和阈值组成的向量, $\nabla^2 E(x)$ 表示误差指标函数 $E(x)$ 的 Hessian 矩阵, $\nabla E(x)$ 表示梯度。

对于 Levenberg - Marquardt 算法,设定误差指标函数是关键步骤,这里可采用 γ^2 作为误差指标函数。

该算法用迭代法进行求解,所以对式(9)进行最小化之前还需为 H 矩阵设定一个初始值。

由式(7)和(8)可知,对于每一对应点对 i ,有如下关系:

$$Z_{ci} \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \begin{bmatrix} \bar{h}_1 \\ \bar{h}_2 \\ \bar{h}_3 \end{bmatrix} \begin{bmatrix} X_{wi} \\ Y_{wi} \\ Z_{wi} \\ 1 \end{bmatrix}, \text{ 即 } Z_{ci} \bar{m}_i = \begin{bmatrix} \bar{h}_1 \\ \bar{h}_2 \\ \bar{h}_3 \end{bmatrix} \bar{M}_i \quad (11)$$

从式(11)表示的两个方程中,可以消去 Z_{ci} ,即得

$$\begin{bmatrix} \bar{M}_i^T & 0^T & -u\bar{M}_i^T \\ 0^T & \bar{M}_i^T & -v\bar{M}_i^T \end{bmatrix} \begin{bmatrix} \bar{h}_1 \\ \bar{h}_2 \\ \bar{h}_3 \end{bmatrix} = 0 \quad (12)$$

设有 n 对对应点对,则可以得到 n 组类似式(12)的方程组,将它们进行组合后得到如下方程组:

$$\begin{bmatrix} \bar{M}_1^T & 0^T & -u\bar{M}_1^T \\ 0^T & \bar{M}_1^T & -v\bar{M}_1^T \\ \vdots & \vdots & \vdots \\ \bar{M}_n^T & 0^T & -u\bar{M}_n^T \\ 0^T & \bar{M}_n^T & -v\bar{M}_n^T \end{bmatrix} \begin{bmatrix} \bar{h}_1 \\ \bar{h}_2 \\ \bar{h}_3 \end{bmatrix} = 0 \quad (13)$$

由此,就可以求出 H 的初始值,然后再将其作为参数,使用 Levenberg - Marquardt 算法求得满足最大似然估计的单应矩阵 H 。

3 模拟试验结果

在模拟试验中,图像采用图2所示的棋盘格平面模板,经过处理的图像如图3所示。

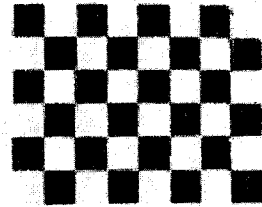


图2 棋盘格平面模板图像

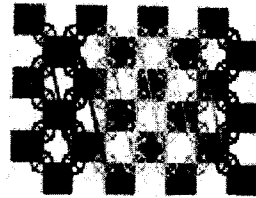


图3 经过图像处理后的模板图像

摄像机采用4个内参数,忽略畸变因子:

$f_u = 750, f_{0v} = 800, u_0 = 300, v_0 = 200$,由上述方法获得的单应矩阵结果如下:

$$H = \begin{bmatrix} 4.751 & -1.324 & 370.557 \\ -3.158 & 42.779 & 121.554 \\ -0.073 & 0.011 & 1.000 \end{bmatrix}$$

(下转第75页)

件的配置:

```
[macro-stdexten]
exten => s,1, Dial ($ {ARG1},20,t)
exten => s,n, Goto ($ {DIALSTATUS})
exten => s,n, hangup ()
exten => s,n (BUSY), voicemail (b $ {MACRO- EX-
TEN})
exten => s,n, hangup ()
exten => s,n (NOANSWER), voicemail (u $ {MACRO-
EXTEN})
exten => s,n, hangup ()
[local]
exten = 6601,1, macro (stdexten, SIP/6601)
```

3.3 CDR 电话详情记录

Asterisk 对每一个电话都产生一个记录文件,默认情形下,是以 CSV(comma separated value)格式保存在目录 /var/log/asterisk/cdr-csv 下。文件的字段主要有:Accountcode, Src, Dst, Dcontext, Clid, Channel, Duration 等等。

除了以 CSV 文件来存储电话记录,Asteriks 还提供了将记录存储到多种数据库的方式,用以满足需要存储更加详细信息的需求。这些数据库有 Mysql, pgsql, SQLite 以及 FreeTDS 等。

4 结束语

Asterisk 具有非常强的灵活性,而传统的 PBX 通常笨重且昂贵,面对 VOIP 不断发展的趋势,Asterisk 适应性强这一特点必将带来更多的应用前景,同时也

是主流的研究方向。Asterisk 的另外一个重要的价值在于,它不光可以让人们通过 VOIP 技术获得低价格长途通信,而是把声音看作是数据网络上的另一个应用程序,从而提供了开发更多新的应用的可能性。当然,目前 VOIP 还面临许多挑战,需要不断的努力使之更加完善^[7]。

创新之处:将 VOIP 技术与传统电话网有机地结合,产生出来许多有价值的附加应用,如:IVR, VOICEMAIL,CDR 等,用较低的成本完成了过去需要昂贵投资才能完成的企业电话功能^[8]。

参考文献:

- [1] 芦伟,曾文华.开源交换机 Asterisk 的应用[J].计算机系统应用,2007(2):80-83.
- [2] 杜能功,吴宇红.Asterisk 和 IAX 协议部署 VOIP 网络[J].西安邮电学院学报,2006(3):34-38.
- [3] 宋传旺,王静.基于软交换技术的 VoIP 系统方案设计[J].微计算机信息,2007(21):122-125.
- [4] 杜旭,沈强,余江.VoIP 网关实时网络传真的设计与实现[J].微计算机信息,2007(9):117-119.
- [5] 汤卓凡,栾杰.基于 Asterisk 的 VoIP 解决方案[J].计算机系统应用,2005(7):79-83.
- [6] 赵建涛,郑永静.基于软交换平台 Asterisk 的电力呼叫中心的设计与实现[J].电力科学与工程,2007(4):102-105.
- [7] 邓国栋,胡越明,白英彩.基于 Asterisk 的软交换运营平台的设计[J].计算机工程,2007(12):145-147.
- [8] 王继鹏,张维.基于 Asterisk 构架企业通信系统的研究[J].科学技术与工程,2007(5):738-741.

(上接第 71 页)

4 结束语

文中详细讨论了二维成像平面的透视投影模型,通过建立的图像坐标系、摄像机坐标系和世界坐标系推导出平面模板和成像平面之间的透视投影关系。给出一种利用最大似然估计法求取单应矩阵 H 的方法,并用 Mahalanobis 距离的平方作为误差指标函数,通过 Levenberg-Marquardt 算法对结果进行优化。实验证明,该方法简单、快速、易实现,在计算机视觉理论中有良好的应用价值。

参考文献:

- [1] 曾慧,邓小明,赵训坡,等.基于线对应的单应矩阵估计及其在视觉测量中的应用[J].自动化学报,2007,33(5):1-3.
- [2] 顾耀林,毛健.一种基于平面模板的虚实配准算法[J].

中国图像图形学报,2006,11(8):3-4.

- [3] Agarwal A, Jawahar C V, Narayanan P J. A survey of planar homography estimation techniques[R]. Hyderabad: International Institute of Information Technology, 2005:1-25.
- [4] 马颂德,张正友.计算机视觉——计算理论与算法基础[M].北京:科学出版社,1998.
- [5] Hartley R, Zisserman A. Multiple View Geometry in Computer Vision[M]. 2nd Edition. Cambridge: Cambridge University Press, 2003:87-129.
- [6] 宣国荣,柴佩琪.基于马氏距离矢量的计算机视觉不变量[J].模式识别与人工智能,1995,8(4):2-3.
- [7] Zhang Z Y. Flexible camera calibration by viewing a plane from unknown orientations[C]//Proceedings of the International Conference on Computer Vision (ICCV'99). USA: [s. n.], 1999:666-673.
- [8] 伏燕军,杨坤涛,邹文栋,等.基于 Levenberg-Marquardt 算法的图像拼接[J].激光杂志,2007,28(5):46-48.