

基于二阶锥规划的运动控制相机轨迹跟踪研究^①

陈明伟^{②*} ** *** 朱登明^{*} 毛天露^{*} 王兆其^{*}

(^{*} 中国科学院计算技术研究所 前瞻研究实验室 北京 100190)

(^{**} 昆明理工大学 昆明 650500)

(^{***} 中国科学院研究生院 北京 100049)

摘 要 提出了一种运动控制相机轨迹跟踪方法,以解决传统基于图像特征匹配方法难以处理的存在稀疏纹理和动态物体视频图像的问题。针对运动控制中现有无特定参照物手眼标定算法精度较低的问题,对螺旋运动理论进行扩展,用于降低输入相机参数中的噪声水平,并提出了一种矩阵张量乘积的手眼标定方程表现形式,运用二阶锥规划和松弛变量约束对其优化。模拟实验和真实场景实验表明,与现有方法相比,这种方法有效提高了无特定参照物模式的手眼标定精度及其对应的运动控制相机轨迹跟踪精度。视频图像虚实融合效果的对比验证了该方法的有效性。

关键词 相机轨迹跟踪,手眼标定,数据筛选,矩阵张量乘积,二阶锥规划

0 引 言

基于视频图像序列的相机运动分析是计算机视觉领域里的重要技术^[1,2],其基于特征点检测和匹配算法,如 KLT-Tracker^[3] 特征追踪方法或 SIFT^[4] 匹配法,计算各相机位置的空间运动矩阵,并运用捆绑调整(bundle adjustment)法优化相机矩阵。这类方法的优点是高精度全自动追踪,且只需要一部普通的数码相机。但其严重依赖于拍摄场景的纹理信息,当场景图像中存在稀疏纹理(图像区域颜色图案单一,如白色墙壁等)、透射(如玻璃)、反射(如光滑金属)或动态物体时,容易导致求解中断和失败。

在机器人运动控制领域,机器手臂上固定安置摄像机,通过视觉技术给机器人提供眼的功能,实现手臂操作的精确定位,广泛运用于汽车制造、智能感知、医疗手术等领域。其核心技术是分析摄像机相对于机器人末端的固定不变的未知三维空间变换矩阵。这类技术称为机器人手眼标定(hand-eye calibration)。只要求解出手眼标定矩阵,就能通过控制柜读取新的拍摄视频对应的机器位置信息,计算出摄像机的运动轨迹参数。

手眼标定中的输入相机矩阵的获取方式,根据不同的拍摄条件分为两类:一类是拍摄特定的标定参照物(如打印的棋盘图纸),其相机矩阵参数运用经典的标定算法^[5] 求出。大部分现有手眼标定算法都基于此,代表方法有 Tsai 等^[6] 的线性二分法、Chen^[7] 的螺旋轴对称理论、Daniilidis 等^[8] 的对偶四元数法、Strobl 等^[9] 的基于物理度量的最大可能性误差分析法、毛等^[10] 的旋转方程的四元数解法、Zhao 等^[11] 的全局凸规划解法等;另一类不需要特定参照物,直接从场景图像中由运动恢复结构(structure from motion, SFM)方法获取。此类方法与第一类相比,拍摄简单,但因缺少场景实际尺度信息,相机平移参数中多存在一个未知的缩放因子 λ ,解法有 Andreff 等^[12] 的 SVD 分解法, Lofberg^[13] 的对偶四元数法扩展和非线性法、Heller 等^[14] 的二阶锥规划的全局优化法等。受特征匹配噪声和缩放因子影响,第二类方法精度略低于第一类基于参照物的手眼标定方法。

本文基于自主研发的一套运动控制设备,提出一种无需特定参照物的手眼标定优化方法。针对此类方法中 SFM 求解的相机参数存在较大噪声的问题,提出一种基于螺旋运动(screw motion)性质的筛

① 国家自然科学基金(61173067),国家自然科学基金-广东联合基金(U0935003)资助项目。

② 男,1977 年出生,博士生;研究方向:计算机视觉与虚拟现实;联系人, E-mail: chenmingwei@ict.ac.cn
(收稿日期:2012-10-13)

选策略,用于降低输入的相机运动数据噪声,并进一步提出矩阵张量乘积(kronecker product)的手眼标定方程表述形式,运用二阶锥规划(second order cone programming, SOCP)和松弛变量(slack variances)约束方法优化标定矩阵。实验结果表明,现有以SFM作为相机输入的手眼标定方法相比,本文方法有效提高了手眼标定精度,可达到与基于特定标定参照物的方法相近的精度水平。该方法已实际应用于运动控制摄影领域的相机轨迹跟踪,通过虚实融合效果的逼真性,验证了其有效性。

1 手眼标定方程

运动控制中的核心技术手眼标定方程具体描述为 $A_{12}X = XB_{12}$,其中 $A_{12} = A_2 * A_1^{-1}$ 是已知的两个相机运动位置间的变换矩阵,而 $B_{12} = B_2^{-1} * B_1$ 是已知的对应的两个机器末端位置间的变换矩阵, X 是待求的相机镜头中心到机器人末端的变换矩阵。这些矩阵都由 3×3 的旋转 R 和 3×1 的平移 t 组成:

$$M_i = \begin{bmatrix} R_i & t_i \\ 0_3 & 1 \end{bmatrix}, M \in \{A, B, X\} \quad (1)$$

我们的运动控制拍摄系统示意如图1所示,其中 cam_i 是待求的相机轨迹矩阵。拍摄位置对应的机器末端矩阵 B_i 的值由机器控制柜读取并计算出,相机矩阵 A_i 的值则运用相机标定算法获取。求解标定方程获得 X 后,拍摄一段新的视频图像,读取视频对应的机器轨迹参数 B'_i ,如果以机器坐标系为原点,相机运动矩阵 $cam_i = B'_i X$,在我们的实验中,以第一个拍摄位置 A_0 为坐标系原点, B_0 是其对应的机器末端位置,则拍摄视频对应的相机运动矩阵为

$$cam'_i = A_i = XB_i^{-1}B_0XA_0 = XB_i^{-1}B_0X \quad (2)$$

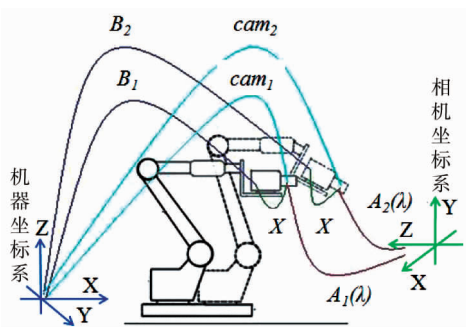


图1 手眼标定和相机轨迹示意图

1.1 基于特定参照物的手眼标定方程

当拍摄特定的参照物,如均匀分布且平整的棋盘图打印纸,运用经典标定算法^[5]可获取精确的相机运动矩阵 A 。标定方程拆分为两个子方程的形式:

$$R_a R_x = R_x R_b \quad (3)$$

$$R_a t_x + t_a = R_x t_b + t_x \quad (4)$$

其中 R_i 和 t_i 表示式(1)对应的旋转矩阵和平移矩阵。此方程可运用Strobl等^[9]的算法求解出精确解。

1.2 基于运动恢复结构的手眼标定方程

当不使用特定参照物,直接拍摄任意具有丰富纹理信息的场景时,可运用SFM方法求解相机运动矩阵,但因缺乏拍摄场景的实际尺度信息,求解出的相机运动矩阵与真实值之间存在一个未知的缩放因子 λ ,式(4)变为

$$R_a t_x + \lambda t_a = R_x t_b + t_x \quad (5)$$

与基于标定棋盘图的算法相比,此方法精度较低,主要原因是SFM方法检测和匹配的特征点依赖于场景图像的纹理信息,而纹理稀疏图像区域容易产生误匹配点对。因此提高此类方法手眼标定精度的途径之一是建立有效的相机噪声筛选方法。

2 运动恢复结构相机数据筛选

相对于相机矩阵数据,经过精确校正过的机器人装置,其从控制柜中读取的机器末端运动矩阵数据的精度和噪声水平与拍摄位置无关,不会随着运动姿态发生改变。而螺旋运动性质建立了两者的关系,且与待求的标定矩阵无关,因此我们以螺旋运动性质为基础,提出一种相机参数噪声筛选策略。

2.1 扩展的螺旋运动性质

运动矩阵中的分量旋转 R 可以表示为轴和旋转角的形式 (l, θ) ,罗德里格斯描述了这种转换关系:

$$R = I + \sin(\theta)\hat{l} + (1 - \cos(\theta))\hat{l} \times \hat{l} \quad (6)$$

Chasles理论^[7]描述螺旋运动性质为:一个刚体变换可用旋转和平移矩阵 $(R, \hat{t})$ 表示,也可用三个螺旋量表示:围绕不经过原点的轴 l 、旋转角 θ ,以及平移量 $\hat{t}$ 在轴 l 上的投影 d :

$$\theta = \|r\|, l = r/\|r\|, d = l^T \hat{t} \quad (7)$$

其中 r, l 和 θ 由式(6)和(7)求取。刚体中的任何位置都可用上式表示,并且任意一个位置做刚体运动,其运动前后的对应位置 (a, b) 具有如下性质:

$$\theta_a = \theta_b, d_a = d_b \quad (8)$$

对于图 1 所示的运动控制摄影系统,机器人末端的云台与摄像机固定不变,在拍摄过程中始终保持刚体不变性,因此相机矩阵 $A_{12}(R_a, t_a)$ 和对应的机器末端阵 $B_{21}(R_b, t_b)$, 均可用上式描述。

然而对于 SFM 方法获取相机矩阵 A'_{12} , 虽然在旋转分量上保持旋转角相等 $\theta'_a = \theta_b$, 但平移分量存在一个待定的缩放因子 λ , 导致投影不再相等, λ 由投影关系表示为

$$\lambda = d_b/d'_a = (\vec{l}_b^T \vec{t}_b) / (\vec{l}_a^T \vec{t}'_a) \quad (9)$$

因噪声数据的存在, λ 不能简单地由上式取均值解。我们根据上述螺旋运动性质提出一种筛选策略, 用于降低相机输入数据中的噪声。

2.2 相机参数噪声筛选策略

我们可以对比两者数据, 剔除出现在高斯峰值部分的相机运动数据噪声。提出以下螺旋运动约束条件, 用于筛选相机数据噪声:

(a) 基于旋转角相等 $\theta'_a = \theta_b$ 性质, 可在旋转分量上筛选相机输入参数:

$$A''_a = \{\theta''_a \mid \|\theta'_a - \theta_b\| < \delta_1, \theta''_a \in \theta'_a\} \quad (10)$$

(b) λ 由式(9)解出, 但因相机参数存在噪声, 求解出的 λ 值存在较大的误差。我们运用 RANSAC 方法, 迭代消除平移分量上噪声较大的数据, 筛选函数为

$$\|R_a t_x + \lambda t_a - R_b t_b + t_x\| < \delta_2 \quad (11)$$

(c) 式(9)中当 $\vec{l} = 0$ 时无解。为了避免退化情况, 我们设置阈值加以限制:

$$\{\|\arccos(\vec{l} \cdot \vec{t}) - \frac{\pi}{2}\| > \delta_3\} \quad (12)$$

算法 1: 对输入的 N 组相机运动数据和机器运动数据依次运用上面 3 个约束不等式, 筛选去掉相机矩阵中处在高斯峰值部分的噪声数据 A_i 和对应的机器末端数据 B_i 。其中 $\delta_{1,2,3}$ 是数据筛选的设定阈值, $\delta_{1,2}$ 值在实验中取对应表达式数据集差异幅度的 80%, δ_3 取 5 度。筛选后的数据用于手眼标定方程的输入。

3 手眼标定方程优化

3.1 二阶锥规划

计算机视觉领域中, Kahl 等人^[15]运用无穷范数 (infinite norm) 的最小最大化 (min-max) 的形式描述基于特征匹配的三维重构、单应矩阵分析、角度三角化等问题:

— 842 —

$$\min_x \max_i \frac{\|F_i X - b_i\|_2}{c_i^T X + d_i}, c_i^T X + d_i > 0 \quad (13)$$

其中 x 是需要求解的变量, F, b, c 和 d 是已知的数据集, 优化目的是最小化数据集中的最大误差。

最小最大化问题常用二阶锥规划 (second order cone programming, SOCP) 方法迭代求解, 其形式如下:

给定 γ

是否有解 X

$$\text{使得 } \frac{\|F_i X - b_i\|_2}{c_i^T X + d_i} < \gamma, c_i^T X + d_i > 0 \quad (14)$$

其算法思想是对于初始给定的一个误差阈值 γ , 求解式(14), 如果有可行解, 则减少 γ 的值, 如果没有可行解, 则增加 γ 的值。 γ 的范围由给定的误差上下界值限定, 迭代求解的过程中不断更新上下界值, 直到上下界的差值小于一定阈值为止。

3.2 矩张量乘积

一个 $n \times m$ 的矩阵 A 和一个 $p \times q$ 的矩阵 B 的张量乘积定义为

$$A \otimes B = [a_{ij} B] = \begin{bmatrix} a_{11} B & \cdots & a_{1n} B \\ \vdots & \ddots & \vdots \\ a_{m1} B & \cdots & a_{mn} B \end{bmatrix} \quad (15)$$

一个 $n \times m$ 矩阵 A 的一维形式定义为:

$$\text{Vec}(A) = [a_{11} a_{12} \cdots a_{mn}]^T \quad (16)$$

引理 1:

$$AVB^T = F \Leftrightarrow (A \otimes B) \text{Vec}(V) = \text{Vec}(F)$$

$$\text{vec}(AB) = (A \otimes I) \text{Vec}(B)$$

$$\text{vec}(BA) = (I \otimes A) \text{Vec}(B) \quad (17)$$

其中 I 是单位矢量, 其维度与 B 矩阵相同。

3.3 手眼标定矩张量表示和二阶锥规划优化

我们把标定式(3)写成张量乘积形式如下:

$$\begin{aligned} R_a R_x R_b^T &= R_x \Leftrightarrow (R_a \otimes R_b) \text{Vec}(R_x) \\ &= \text{Vec}(R_x) \Leftrightarrow [(R_a \otimes R_b) - I] [\text{vec}(R_x)] = 0 \end{aligned} \quad (18)$$

公式(5)写成张量乘积形式如下:

$$\begin{aligned} \text{vec}(R_x t_b) &= (I_3 \otimes t_b) \text{vec}(R_x) \Leftrightarrow \\ &[(R_a - I) t_a I_3 \otimes t_b] \begin{bmatrix} t_x \\ \lambda \\ \text{vec}(R_x) \end{bmatrix} = 0 \end{aligned} \quad (19)$$

合并式(18)和(19)得到:

$$\begin{bmatrix} 0 & 0 & (R_a \otimes R_b) - I \\ (R_a - I) t_a & I_3 \otimes t_b & \end{bmatrix} \begin{bmatrix} t_x \\ \lambda \\ \text{vec}(R_x) \end{bmatrix} = 0 \quad (20)$$

上式是 $AX=0$ 形式的方程,然而运用常用的线性 SVD 分解法求出的 X 不能保证是最优解。因此我们把方程改写为无穷范数式(13)的形式,其中 $c_i^T=0, d_i=1$:

$$\min_x \max_i \|A_i X\| < \gamma, \gamma > 0 \quad (21)$$

为了避免出现零解,增加约束 $0 < \eta < \|A_i X\|$, 并运用二阶锥规划方法对其求解,如果有可行解 X_i , 则对应误差为

$$e_i = \max \|A_i X\| \quad (22)$$

二阶锥规划求解算法流程如下:

算法 2: 二阶锥规划手眼标定算法

输入: 相机参数 A_i , 机器参数 B_i ;

设置上下界: $\text{low}=0, \text{high}=1$;

设置最大误差: $\text{maxError} = \text{无穷大}$;

1. repeat

2. $\gamma = (\text{low} + \text{high})/2$;

3. 求解式(21)是否有可行解

4. If (有可行解 && $e_i < \text{maxError}$)

{ $\text{high} = \gamma$; $\text{maxError} = e_i$; }

5. Else { $\text{low} = \gamma$; $\text{high} = \text{high} * 2$; }

6. Until ($\text{high}-\text{low} < 1e-8$)

3.4 松弛变量约束

无穷范数形式的二阶锥规划方法可以在现有的噪声下,经过不断迭代尝试获得最大噪声数据的最小优化解。然而如果某一些数据的噪声较大时,必然影响到整体标定方程的精度。Olsson 等^[16]提出了一种运用松弛变量删除噪声数据的方法,应用于基于图像的三维建模方法,我们将其应用到手眼标定领域,对式(21)增加松弛变量约束 s_i , 变为如下形式:

$$\min_{s_i, X} \sum_i s_i \text{ s.t. } \|A_i X\| \leq \gamma + s_i, \gamma \geq 0 \quad (23)$$

在给定阈值 $\gamma > 0$ 的前提下,我们运用线性规划方法求解数据集中每个数据对应的 s_i 值:

$$\|A_i X\| = \gamma + s_i, \gamma \geq 0 \quad (24)$$

如果 $s_i < 0$, 则说明其对应的是噪声较小的数据,反之则对应噪声较大的数据。当数据筛选阈值设置为大于零的 μ 时,则删除所有 $s_i > \mu$ 对应的输入数据 A_i 和 B_i 。优化后的输入数据再运用无穷范数的二阶锥规划方法(算法 2)求解式(21)的最小最大化解。

算法 1 剔除的是处在高斯峰值部分的相机噪声

数据,松弛变量约束是在算法 1 处理后,进一步剔除噪声较大的相机和对应位置的机器数据。

4 相机轨迹获取和图像虚实融合

我们以一台德国 KUKA(KR 110)机器人为基础,自主研发一套运动控制摄影系统,在机器手臂末端加装两个扩展轴,用于安放摄像机。

基于一套自主研发的模拟编辑控制和通讯控制软件系统,控制硬件设备执行设定姿态的视频拍摄,整个相机轨迹获取流程如图 2 所示,其主要步骤为:(1)运用模拟控制软件围绕无特定参照物的普通场景设定十几个机器拍摄姿态,采集标定图像(此时图像具有丰富的纹理信息)和机器数据;(2)运用我们提出的手眼标定算法,求解手眼标定矩阵;(3)控制装置拍摄需要获取相机轨迹的场景视频(此时视频图像可以是任意纹理、拍摄姿态和长度),并读取拍摄轨迹对应的机器末端位置矩阵;(4)由求解出的手眼标定矩阵和机器轨迹矩阵计算出相机轨迹矩阵,见式(2);(5)编写生成建模动画软件(如 3DMAX 等)的数据脚本,在动画软件中导入相机轨迹数据、图像视频和虚拟物体,运用建模动画软件的渲染功能进行图像的虚实融合。

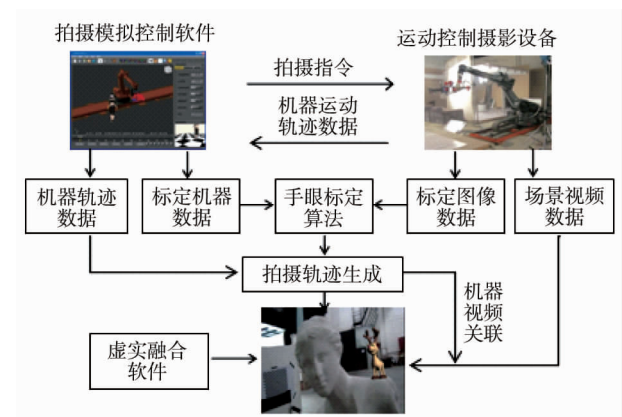


图 2 基于运动控制的轨迹获取和虚实融合流程图

5 实验

5.1 模拟实验

我们在一台 CPU 双核 2.4GHz, 内存 4G 的机器和 Matlab 环境中,设计模拟手眼标定的各种数据,并在其中用 Monte Carlo 方法随机生成原始数据和

高斯噪声数据,并对两者数据叠加,用于对比我们提出的方法和现有方法的标定精度。用“OUR”代表我们的方法,“SL”代表二分线性最小二乘法^[6],“SDQ”代表旋转矩阵先用二分法求出,平移和缩放因子用对偶四元素法^[8]求解。实验中加入高斯噪声的数据生成方式:Monte Carlo 方法随机产生标定矩阵 X ,和摄像机坐标系到机器坐标系矩阵 Z ,以及多个不同位置的机器矩阵 B ,并计算出相机矩阵 $A_i = ZB_iX^{-1}$ 。加入均方差为(0.01—0.1)的 10 组不同的高斯噪声到 A 矩阵上,并随机生成缩放因子 λ ,模拟出存在缩放因子的相机矩阵 $A' = A(\lambda)$ 。其中 A 、 X 和 Z 旋转矩阵的 Roll-Pitch-Yaw 三个旋转角生成范围为(−180,180)度, A 和 X 平移分量的生成范围为(−100,100), Z 的平移分量设定为[200,200,200]。缩放因子的生成范围为(0.001,1000)。上述标定数据每次随机生成 20 组,共生成 100 次。

标定误差的衡量标准如下: $e_t = \|\hat{t} - t\| / \|t\|$ 和 $e_r = \|\hat{R} - R\|$, 其中($\hat{t}$, $\hat{R}$)是运用本文提出的方法和现有方法计算的标定结果, (t , R)表示上述随机生成的理论生成值。100 次标定结果的误差均值的实验比较如图 3 和图 4 所示,图中横

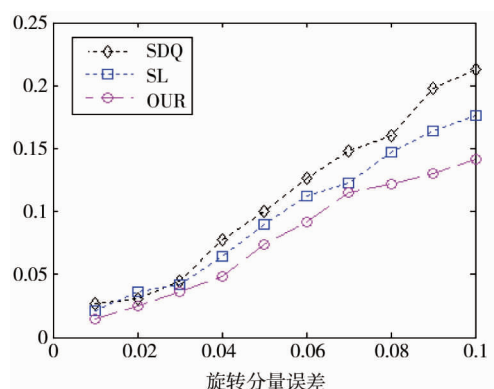


图 3 不同噪声水平下旋转分量误差比较

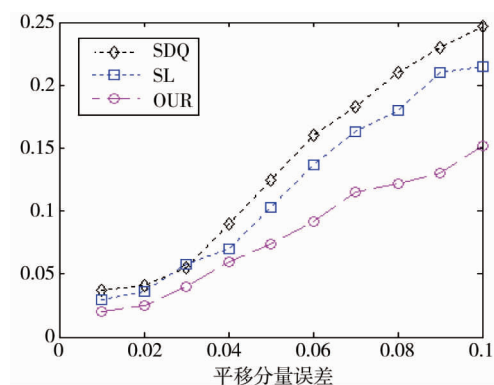


图 4 不同噪声水平下平移分量误差比较

坐标表示加入的高斯噪声水平,而纵坐标表示标定的误差水平。从图 3 和图 4 的模拟实验结果在旋转分量和平移分量上的误差看出,与“SDQ”和“SL”的方法相比,我们提出的方法求解出误差更小的手眼标定方程解。

5.2 真实实验

真实实验通过 VC++ 编写模拟控制软件进行测试,如图 2 所示。在机器人拍摄装置上固定放置一台佳能 SX240 HS 数码相机。由式(2)可以看出运动控制相机轨迹跟踪的精度取决于手眼标定的精度。我们通过模拟控制程序设定机器姿态,拍摄分辨率为 1920×1080 像素的 12 张无特定参照物场景图像,如图 5 左图所示,运用本文算法和现有的“SL”和“SDQ”方法求解无参照物手眼标定矩阵。并在相同位置拍摄打印的棋盘图纸张,如图 5 右图所示,运用 Strobl 等的手眼标定软件 DLR CalLab^[17] 计算标定结果,对比我们提出的方法与这些方法在真实环境中的标定结果精度。



无标定参照物场景用于提出的算法(左图),棋盘图用于 DLR CalLab(右图),两图像序列拍摄位置相同,分辨率均为 1920×1080 像素

图 5 相机矩阵参照物

相机运动矩阵运用 SFM 算法,如“Bundler”^[18],计算各个相机位置参数 A_i 和三维点云。同时读取对应位置的机器人九轴参数并计算出机器矩阵 B_i ,运用提出的数据筛选方法优化相机数据,剔除噪声较大的数据。

由于此时并不知道所用设备的手眼标定矩阵真实值,因此采用相对误差函数作为衡量手眼标定矩阵精度的方式:

$$E_R = \frac{1}{N} \sum \|R_a R_x - R_x R_b\|^2$$

$$E_t = \frac{1}{N} \frac{\sum \| (R_a - I)\vec{t}_x - R_x \vec{t}_b + \vec{t}_a \|^2}{\sum \| R_x \vec{t}_b - \vec{t}_a \|^2} \quad (25)$$

我们提出的方法与“SL”方法和“SDQ”方法,以及基于棋盘图的 DLR CalLab 软件计算的标定精度

的比较结果如表 1 所示。

实验结果表明,我们提出的基于螺旋运动的相机数据筛选和二阶锥规划算法,可以有效地剔除相机参数噪声并获得更精确的优化解。比“SL”方法和“SDQ”方法更精确,甚至标定精度还略高于基于棋盘图的 DLR CalLab 软件。即在简化标定设施的同时,提高了标定精度。

表 1 真实实验数据标定误差比较

标定算法	E_R	E_t
SL(SFM)	0.0057	0.072
SDQ(SFM)	0.0061	0.069
Our Method(SFM)	0.0030	0.056
DLR(棋盘图)	0.0031	0.055

经过手眼标定校正后,我们拍摄两段场景视频(分辨率 1920×1080 像素,帧率 25,视频解析出 1727 幅图),分别拍摄纹理信息丰富和稀疏纹理的场景用于测试提出方法的相机轨迹求解精度。

第一段特征丰富的视频如图 6 所示。对其运用鲁棒的特征匹配跟踪方法(如 Boujou 软件)追踪的效果如图 6 右图所示;分别运用我们提出算法、DLR CalLab 软件、和“SDQ”方法计算出的手眼标定矩阵以及对应的相机轨迹,并分别与特征匹配方法的结果进行对比,如图 7 所示。其中实线段表示特征匹配的相机轨迹,虚线段表示上述几种运动控制方法的相机轨迹。从图中可以看出,我们提出的方法能很好的与特征匹配方法的轨迹重合,与棋盘图的方法 DLR CalLab 软件的结果近似,优于 SFM 方法“SDQ”的结果。

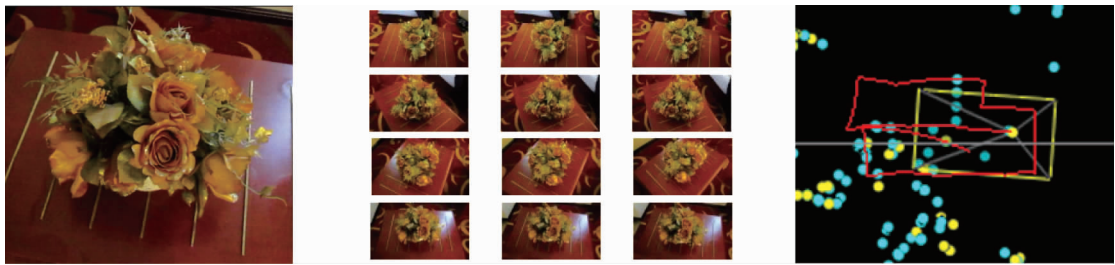
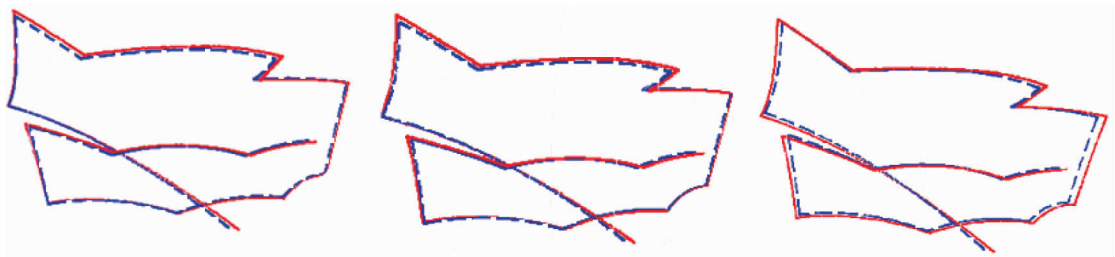


图 6 特征匹配跟踪方法对纹理丰富图像获取的相机轨迹



(实线段:特征匹配结果;虚线段:运动控制结果:左图提出的算法、中图 DLR 软件、右图 SDQ 方法)

图 7 运动控制方法与特征匹配方法的相机轨迹对比

第二段纹理稀疏的视频如图 8 所示,从图上看,因场景中人物是动态的,而且背景是纯白色,特征匹配的方法难以有效求解此类场景的相机轨迹,而我们的方法与场景的纹理信息无关,依然能获得精确的相机轨迹,因机器运动的轨迹路线不变,因此获取的相机轨迹数据与第一段视频一样。

我们根据拍摄的时刻信息建立相机轨迹和视频图像的对应关系,把生成的相机轨迹写成 3dmax 软件的脚本并导入其中,并在 3dmax 中导入对应的视频图像和虚拟的物体模型“马鹿”,调节视图视点为相机轨迹视点,并在相机视点下调节虚拟物体的空

间位置,使其位于场景图像中黑色椅子对应的三维空间处,应用 3dmax 的渲染工具把虚拟物体从相机各帧视点的方向投影到每一帧图像上,形成虚实融合的混合图像,如图 9 所示,其中左上图是相机轨迹在 3dmax 中的表示。从图中可看出在不同视角拍摄的图像,都能与空间中位置固定不变的虚拟物体对应,验证了求解相机轨迹的有效性和精确性。

6 结 论

本文提出一种基于运动控制的相机轨迹追踪方



图 8 纹理稀疏和存在动态物体图像的特征检测匹配效果

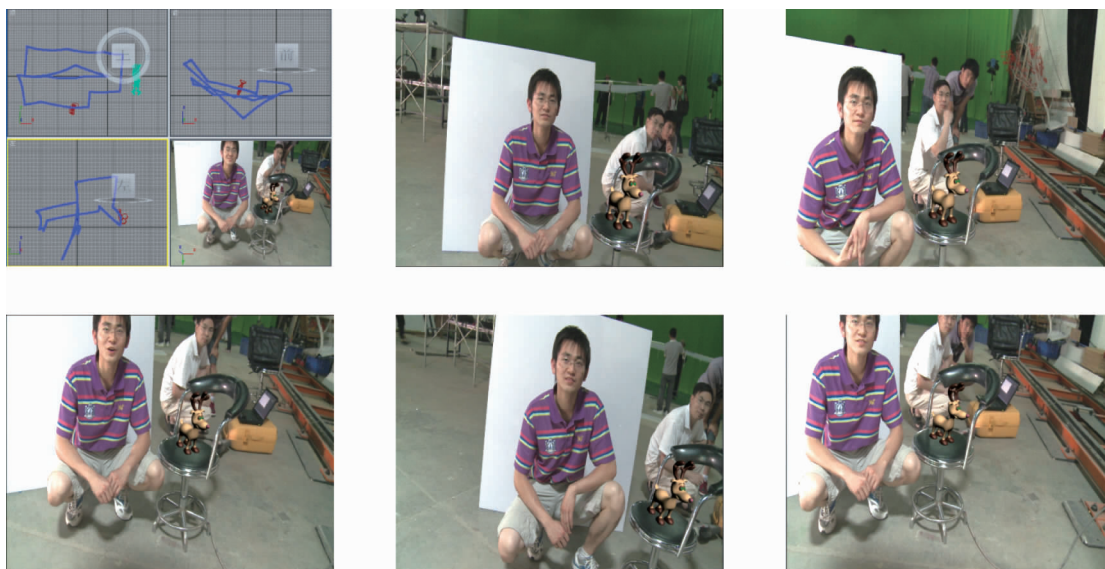


图 9 基于提出算法求解的相机轨迹和虚实融合效果

法,有效解决了传统特征匹配的轨迹跟踪方法难以处理纹理稀疏和动态物体的场景视频图像的问题。我们提出的算法,首先基于螺旋运动理论提出数据筛选策略,用于剔除相机参数中噪声较大的数据,然后基于二阶锥规划和松弛变量约束求解出手眼标定矩阵在一定噪声下的最优解。通过模拟和真实实验与现有方法的比较,验证了提出的算法在不使用特定标定物时具有较高的标定精度。并运用机器人通讯控制技术,读取机器轨迹参数,由正向动力学计算出机器末端的运动矩阵,从而得到拍摄视频对应的相机轨迹。导入数据到虚实合成软件中,通过虚拟物体和真实图像场景的良好融合效果,验证了提出方法的有效性。

参考文献

- [1] Thormahlen T, Hasler N, Michael W, et al. Merging of feature tracks for camera motion estimation from video. In: 5th European Conference on Visual Media Production, London, British, 2008. 1-8
- [2] Hengel A, Dick A, Thormahlen T, et al. VideoTrace: rapid interactive scene modelling from video. *ACM Transactions on Graphics*, 2007, 26(3): 86-94
- [3] Shi J B, Tomasi C. Good Features to Track. In: IEEE Conference on Computer Vision and Pattern Recognition, Seattle, USA, 1994. 593-600
- [4] Lowe D G. Distinctive image features from scale invariant keypoints. *International Journal of Computer Vision*, 2004, 60(2): 91-110
- [5] Zhang Z. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(11): 1330-1334
- [6] Tsai R Y, Lenz R K. A new technique for fully autonomous and efficient 3D robotics hand-eye calibration. *IEEE Transactions on Robotics and Automation*, 1989, 5(3): 345-358
- [7] Chen H. A screw-motion approach to uniqueness analysis of head-eye geometry. In: Proceedings of the IEEE Conference on Vision and Pattern Recognition, Hawaii, USA, 1991. 43-630
- [8] Konstantinos D. Hand eye calibration using dual quaternions. *The International Journal of Robotics Research*, 1999, 18(3): 286-298

- [9] Strobl K H, Hirzinger G. More accurate camera and hand-eye calibrations with unknown grid pattern dimensions. In: Proceedings of the IEEE International Conference on Robotics and Automation, Pasadena, USA, 2008. 1398-1405
- [10] 毛剑飞, 邵黄芳, 蒋莉等. 求解方程 $RaRx = RxRb$ 的四元数几何研究. 中国图象图形学报, 2010, 15(6): 951-957
- [11] Zhao Z J. Hand-eye calibration using convex optimization. In: IEEE International Conference on Robotics and Automation, Shanghai, China, 2011. 2947-2952
- [12] Andreff N, Horaud R, Espiau B. Robot hand-eye calibration using structure from motion. *Journal of Robotics Research*, 2001, 20: 228-248
- [13] Lofberg J. YALMIP: A toolbox for modeling and optimization in MATLAB. <http://users.isy.liu.se/johanl/yalmip>; Lofberg, 2012
- [14] Heller J, Havlena M, Sugimoto A, *et al.* Structure-from-motion based hand-eye calibration using L_∞ minimization. In: IEEE Conference on Computer Vision and Pattern Recognition, Rhode Island, USA, 2011. 1173-1178
- [15] Kahl F, Hartley R. Multiple view geometry under the L_∞ norm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, 30: 1603-1617
- [16] Olsson E, Hartley R. Outlier Removal Using Duality. In: Computer Vision and Pattern Recognition Conference, Alaska, USA, 2008. 1450-1457
- [17] Strobl K H. The DLR camera calibration toolbox software. http://www.dlr.de/rm/desktop_default.aspx/tabid-3925; DLR, 2010
- [18] Snavely N. Bundler: structure from motion (SfM) for unordered image collections. <http://phototour.cs.washington.edu/bundler>; University of Washington, 2008

Motion control camera tracking based on second cone programming

Chen Mingwei * ** ***, Zhu Dengming*, Mao Tianlu*, Wang zhaoqi*

(* Institute of Computing Technology; Chinese Academy of Sciences, Beijing 100190)

(** Kunming University of Science and Technology, Kunming 650500)

(*** University of Chinese Academy of Sciences, Beijing 100049)

Abstract

A motion control based camera tracking method is proposed to solve the problem of existence of sparse textures and dynamic object images, which is hard to deal with when using a conventional image feature detector based camera tracking method. To resolve the problem of hand eye calibration in motion control, which has less accurate result when it does not use calibration pattern, a screw motion based camera data selection strategy is proposed to reduce the input camera motions noise. The hand eye calibration function is changed by kronecker product and solved by second cone programming and slack variances constrain methods. Comparing with existing methods in a simulated and a real environment, the proposed method shows the validity in no calibration pattern sense, and it is applied to virtual reality field.

Key words: camera tracking; hand eye calibration; data selection; kronecker product; second order cone programming