

基于软指派算法的相机位姿估计研究

杜广胜, 丛明*, 刘毅, 刘冬

大连理工大学机械工程学院, 辽宁 大连 116024

摘要 针对未知特征点匹配关系下的单目相机位姿估计问题, 提出了一种基于软指派算法的相机位姿估计算法。结合了正交迭代算法和软指派算法, 以三维(3D)和二维(2D)特征点的物方空间共线性误差作为确定特征点之间匹配关系的指派依据和计算相机位姿的目标函数值, 通过迭代方式确定特征点的匹配关系和相机的位姿。其不仅能处理 3D/2D 特征点一一对应的情况, 而且能够处理同时存在遮挡的 3D 特征点和错误的 2D 特征点的情况。合成图像的实验结果表明: 在含有图像噪声、遮挡的 3D 特征点和错误的 2D 特征点情况下, 该算法的成功率 > 82%, 真实图像实验对算法的性能进行了验证。

关键词 机器视觉; 位姿估计; 正交迭代算法; 指派算法

中图分类号 TP391.41 **文献标识码** A

doi: 10.3788/LOP55.091501

Camera Pose Estimation Based on Softassign Algorithm

Du Guangsheng, Cong Ming*, Liu Yi, Liu Dong

School of Mechanical Engineering, Dalian University of Technology, Dalian, Liaoning 116024, China

Abstract Aiming at the problem of monocular camera pose estimation under the matching relations of unknown feature points, a new camera pose estimation algorithm is proposed based on the Softassign algorithm. Based on the advantages of orthogonal iteration and Softassign algorithm, the collinear error of the object space in three-dimensional (3D) feature points and two-dimensional (2D) feature points is used as the basis for determining the matching relationship between feature points and the objective function value for calculating the camera pose. An iteratively way determines the matching relationship of feature points and the pose of the camera. It can handle the case where the 3D/2D feature points are in one-to-one correspondence, and the case which occlude 3D feature points and erroneous 2D features. The results of the synthetic image experiments show that the success rate of this algorithm is over 82% in one case with image noise, occlusion and clutter. The experiment of real image can verify the property of algorithm.

Key words machine vision; pose estimation; orthogonal iteration; assignment algorithm

OCIS codes 150.1135; 120.0280; 110.3925; 100.4995

1 引言

相机位姿求解问题是机器视觉领域的基础研究内容, 在相机内参数已知的情况下, 根据一组二维(2D)特征点和对应的三维(3D)特征点, 求解相机的位姿, 又称为透视 n 点成像 (P_nP) 问题。成熟的解决方法如解析算法^[1-2]、迭代算法^[3-5]和全局优化算法^[6]等已广泛应用于手眼标定、物体追踪、自主巡航、增强现实等技术^[7]。在机器人自动化装配等新领域中, 相机位姿估计所用到 3D/2D 特征点匹配关

系未知, 而且包含无对应特征点的 2D 特征点(2D 伪特征点)和 3D 特征点(3D 遮挡特征点), 这类问题称之为同时确定相机位姿和特征点匹配关系问题^[8]。

该问题的解决方案包括假设检验法和迭代算法。随机采样一致性(RANSAC)算法^[9]是一种假设检验法, 其原理是随机假设一小部分 2D/3D 特征点相对应, 采用线性变换方法求解相机的位姿, 将求解的位姿结果应用于其他特征点以检验是否正确。该算法的缺点是搜索速度慢, 计算复杂度高, 容易返

收稿日期: 2018-02-23; 修回日期: 2018-03-23; 录用日期: 2018-04-02

基金项目: 国家自然科学基金(51575078)

* E-mail: congm@dlut.edu.cn

回局部最优值。SoftPOSIT 算法^[10-11]是一种确定相机位姿和特征点匹配关系的优化算法,将迭代求解 PnP 问题的比例正交投影迭代变换(POSIT)算法^[4]和软匹配(Softassign)算法^[12]结合为全局优化算法,该算法的可靠度高,但缺点是 POSIT 算法的局限性导致目标到相机的距离较近时搜索结果不收敛。其他学者将 SoftPOSIT 算法与遗传算法相结合^[8],将初始值作为个体,在种群迭代过程中,以 SoftPOSIT 算法的目标函数作为适应度函数,通过遗传算法的筛选、交叉和变异机制,最终得到最优的结果,该算法改进了初始值的给定方法,但是在特征点数量较低时收敛效果不明显。SoftSI 算法^[13]是将基于形状的迭代算法(SI)和 Softassign 算法结合起来,通过奇异值分解得到相机的位姿。SoftSI 算法避免了姿态模糊,但其时间复杂度高,不能处理存在 3D 错误特征点和 2D 伪特征点的情况,应用范围有限。

基于运动质点系的相机位姿估计(PSKPE)算法^[14-15]是将共线性误差^[5]和匹配矩阵^[16]相结合,用于特征点对应关系未知的相机位姿估计,不仅可以应用于 2D/3D 特征点一一对应问题,还可以应用于存在 3D 遮挡特征点和 2D 伪特征点的情况,但是当二者同时存在时,算法无法得到正确解,此外,该算法搜索速度慢,得到的相机位姿精度低。其他方法则采用了增加传感器的方式进行测量,精度有较大的

提升^[17-19]。

在机器人自动化装配中,为了提高定位精度,工件尽量占满相机的视野,因而距离相机较近,而 POSIT 算法应用的前提是工件在相机坐标系 z 轴方向的厚度远小于其在 z 轴方向的深度,当工件与相机较近时,SoftPOSIT 算法收敛缓慢,成功率低,而其他算法耗时长,难以满足时间要求,因此有必要对现有算法进行改进。在文献^[14-15]的基础上,本文提出了改进的基于软指派算法的相机位姿估计算法,将 Softassign 算法与正交迭代(OI)算法融合在一起,在不增加算法时间复杂度的情况下,可以同时存在 3D 遮挡特征点和 2D 伪特征点的相机位姿估计问题,并对提出的算法进行了实验验证和分析。

2 算法描述

2.1 物方空间共线性误差

如图 1 所示,假设世界坐标系 W -XYZ 下存在一点 P^w ,其坐标表示为 $P^w = (X^w, Y^w, Z^w)^T$,经过旋转和平移之后,对应相机坐标系 O_c - $X_c Y_c Z_c$ 下的点 P^c ,坐标表示为 $P^c = (X^c, Y^c, Z^c)^T$, P^c 和 P^w 之间的转换关系为

$$P^c = rP^w + t, \quad (1)$$

式中: r 为旋转矩阵,是 3×3 单位正交阵,满足 $r^T r = I$; $t = (t_x, t_y, t_z)^T$ 为平移向量。

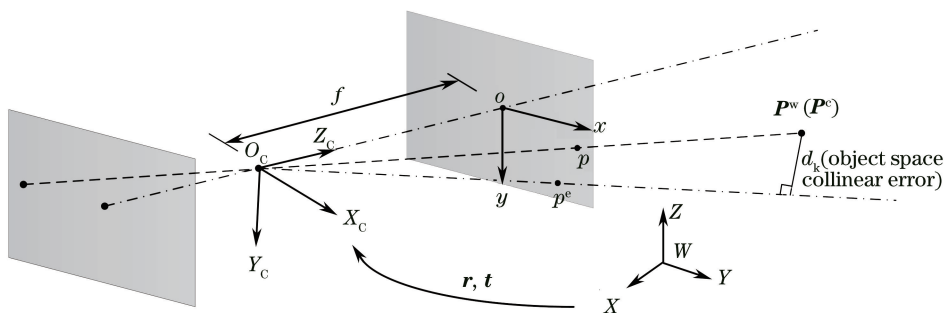


图 1 物方空间共线性误差

Fig. 1 Object space collinear error

假定相机的焦距为 f ,则点 P^c 投影到图像物理坐标系 o - xy 上的点为 p ,表示为 $p = (x_0, y_0)^T$,由于噪声和误差, P^c 实际投影到图像上的点为 $p^e = (x_0^e, y_0^e)^T$,与理想图像点 p 不重合,则点 P^c 到视线 $\overline{O_c p^e}$ 的距离可以表示为

$$d^2 = \|(V - I)P^c\|^2 = \|(V - I)(rP^w + t)\|^2, \quad (2)$$

式中 d 即为点 P^c 与图像点 p^e 之间的物方空间共

线性误差^[5], V 为视线 $\overline{O_c p^e}$ 的投影矩阵:

$$V = \frac{vv^T}{v^T v}, v = (X_0^e, Y_0^e, f)^T. \quad (3)$$

2.2 相机位姿估计算法

已知一组由 L 个 3D 特征点 P_k^w 组成的点集,在 W -XYZ 下的坐标表示为 $\{P_k^w = (X_k^w, Y_k^w, Z_k^w)^T, k = 1, 2, \dots, L\}$,和一组由 N 个 2D 特征点 p_i 组成的点集,在图像物理坐标系 o - xy 下的坐标表示为

$\{p_i = (x_i, y_i)^T, i=1, 2, \dots, N\}$, 特征点之间的匹配关系未知, 由(1)式得相机坐标系 $O_C X_C Y_C Z_C$ 下特征点集 $\{P_k^c = (X_k^c, Y_k^c, Z_k^c)^T, k=1, 2, \dots, L\}$, 3D 特征点 P_k^w 与 2D 特征点 p_i 之间的共线性误差可以根据(2)式表示为

$$d_{ik}^2 = \|(V_i - I)(RP_k^w + T)\|^2, \quad (4)$$

式中 V_i 为视线 $\overline{O_C p_i}$ 的投影矩阵, 由(3)式求得, R 为旋转矩阵, $T = [T_x \ T_y \ T_z]^T$ 为平移矩阵。

在特征点匹配关系未知情况下, 必须在确定特征点之间的匹配关系之后进行相机位姿估计, 根据(4)式得到的结果, 采用 Softassign 算法对 3D/2D 特征点进行匹配。构建距离矩阵 $D = [d_{ik}^2]_{N \times L}$ 和匹配概率矩阵 $M = [m_{ik}]_{(N+1) \times (L+1)}$, 匹配概率矩阵 M 第 $N+1$ 行和第 $L+1$ 列由松弛元素组成, 其中 m_{ik} 代表 p_i 与 P_k^c 的匹配权重, 是一个连续变量, 取值范围为 $m_{ik} \in [0, 1]$ 。该匹配概率矩阵具有以下特点: 当 $m_{ik} = 0$ 时, p_i 与 P_k^c 完全不匹配; 当 $m_{ik} = 1$ 时, p_i 与 P_k^c 完全匹配; 当 $m_{i(L+1)} = 1$ 时, 标记 p_i 为 2D 伪特征点; 当 $m_{(N+1)k} = 1$ 时, 标记 P_k^c 为遮挡 3D 特征点, 使得匹配概率矩阵能够处理 3D 遮挡特征点与图像 2D 伪特征点同时出现的情况。同时确定相机位姿和特征点匹配关系问题转化为求解优化函数最小值:

$$\min E = \sum_{i=1}^N \sum_{k=1}^L m_{ik} d_{ik}^2. \quad (5)$$

当给定距离矩阵 D 时, 匹配概率矩阵 M 可以由算法 1 求得: 1) 取 N, L 中较大的值, 赋值给 n_{\max} , 令 $S_{\text{scale}} = 1/n_{\max}$; 2) 令 $m_{ik} = S_{\text{scale}}, 1 \leq i \leq N+1, 1 \leq k \leq L+1$; 3) 令 $m_{ik} = S_{\text{scale}} \cdot \exp[-\beta(d_{ik}^2 - \alpha)], 1 \leq i \leq N, 1 \leq k \leq L$; 4) 将匹配矩阵的元素归一化, 使得每一行每一列元素之和为 1, 即 $\sum_{i=1}^{N+1} m_{ik} = 1, \sum_{k=1}^{L+1} m_{ik} = 1$, 得到匹配概率矩阵 M 。

算法 1 中, β 为模拟退火参数, 在迭代过程中逐渐增大; α 为误差阈值, 避免了匹配概率矩阵的松弛元素全部退化为 1。相机位姿的估计过程即在视线 $\overline{O_C p_i}$ 对匹配点 P_k^c 的引力作用下, P_k^c 向视线 $\overline{O_C p_i}$ 靠近的过程, 当 M 逐渐退化为置换矩阵时, (5) 式优化函数达到最小值, 从而得到相机的位姿。

将所提出的确定相机位姿和特征点匹配关系算法标记为 SoftOI 算法, 算法的步骤如下。

1) 将 L 个质点视作质点系, 设质心为 P_c^w , 质心坐标为

$$P_c^w = \frac{1}{L} \sum_{k=1}^L P_k^w. \quad (6)$$

2) 给定相机位姿估计的初始值 R_0 和 T_0 , 令 $R = R_0, T = T_0$, 将世界坐标系的原点平移到质心 P_c^w 处, 新世界坐标系下的 3D 特征点集表示为 $\{P_k^w = (X_k^w, Y_k^w, Z_k^w), k=1, 2, \dots, L\}$, 设新世界坐标系转换到相机坐标系的平移向量为 T' , 则

$$P_k^c = RP_k^w + T', \quad (7)$$

$$T = T' - RP_c^w. \quad (8)$$

3) 设定图像噪声 σ, τ , 匹配参数 $\lambda = [\lambda_n]_{1 \times 300}$, 迭代代数 n , 迭代系数 β_{update} 和 β_{final} , 令 $\alpha = \sigma^2 + 1, n = 1$ 。

4) 计算视线 $\overline{O_C p_i}$ 对点 P_k^c 的引力 $F_{ik} = (V_i - I)(RP_k^w + T')$ 、距离矩阵 D 和匹配概率矩阵 M 。

5) 根据每个 3D 特征点受到的合力 $F_k = \sum_{i=1}^N m_{ik} F_{ik}$, 求得引力对质心的主矩为

$$T_c = \sum_{k=1}^L r_k \times F_k, \quad (9)$$

式中 $r_k = P_k^c - T'$, 为 P_k^c 相对于质心的矢径。

6) 根据文献[14]的质点系运动原理得质点系的角加速度为

$$\alpha = (J_c)^{-1} T_c, \quad (10)$$

式中 J_c 为质点系相对于质心的转动惯量。

7) 将 R 的分向量分别绕 α 转过 $|\alpha|$ 弧度, 得到新的旋转矩阵 R 。

8) 计算 T' 关于旋转矩阵 R 的最优解为

$$T' = \left[\sum_{i=1}^N \sum_{k=1}^L m_{ik} (I - V_i) \right]^{-1} \left[\sum_{i=1}^N \sum_{k=1}^L m_{ik} (V_i - I) RP_k^w \right]. \quad (11)$$

9) 计算匹配矩阵 M 中特征点匹配对数 P_{match} 。如果 $P_{\text{match}} < \lambda_n \cdot L$, 则结束算法, 返回空值; 否则执行下一步。

10) 计算目标函数值 E 。如果 $E < \tau$ 或者 $\beta > \beta_{\text{final}}$ 超出了取值范围, 根据(8)式计算 T 值, 结束算法, 否则, 令 $\beta = \beta \cdot \beta_{\text{update}}, n = n + 1$, 转到步骤 4) 继续迭代计算。

给定的随机初始值 R_0 和 T_0 决定了 SoftOI 算法能否搜索到正确的相机位姿, 当判断 SoftOI 算法搜索结果不收敛时, 需要及时终止本次搜索以节约时间, 并重新给定随机初始值 R_0 和 T_0 进行搜索, 循环往复直至该算法搜索到正确的相机位姿。当 SoftOI 算法迭代到第 n 次, 特征点匹配对数

$P_{\text{match}} < \lambda_n \cdot L$ 时, 认定 SoftOI 算法搜索结果不收敛, 终止本次搜索。 $P_{\text{match}} < \lambda_n \cdot L$ 既作为是否是正确的相机位姿的判断条件, 又作为是否提前终止搜索的判断条件, 加快了算法的搜索速度。在算法迭代过程中, 特征点匹配对数的变化趋势是随着迭代次数的增加而逐渐增大, 因此匹配参数 λ_n 应该逐渐增大。根据实验结果得, λ_n 的最大值为 $\lambda_{\text{max}} = 0.5$ 时搜索效率最高。当 SoftOI 算法给定 η 次初始值后仍返回空结果, 认定相机位姿估计失败。

与文献[14-15]对比, 所提出的算法在匹配概率矩阵的确定、平移向量的计算和终止条件上做出了改进: 1) 匹配系数为连续变量, 保留了特征点之间的匹配可能性, 能够处理 3D 特征点被遮挡与图像 2D 伪特征点同时出现的情况; 2) 由更新后的旋转矩阵计算新的平移向量, 与上一次迭代的平移向量无关, 提高了迭代速度; 3) 迭代中判断匹配点对数是否满足预设比例, 以便及时终止当前搜索, 节约搜索时间。

3 实验与分析

3.1 合成图像实验

相机位姿估计算法的 2 个重要的指标是算法的成功率和所用时间, 因此将在此部分对上述指标进行探讨。采用计算机仿真的方式, 随机生成 3D/2D 特征点, 探讨特征点对的数量、图像噪声、2D 伪特征点比例和 3D 遮挡特征点比例对算法成功率和运行时间的影响。

合成图像实验根据已知的相机参数模拟产生随

机 3D 特征点和其对应的 2D 特征点, 同时添加图像噪声和干扰点, 仿真实验包括以下 5 个参数: σ 、 L 、 a 、 b 和 n_t 。其中, σ 为图像噪声的标准差, 单位为 pixel; L 为仿真实验中 3D 特征点的数量; a 为 3D 遮挡特征点在 3D 特征点中所占的比例; b 为 2D 伪特征点在 2D 特征点中所占的比例; n_t 则为上述 4 个参数组合中, 每次实验重复的次数, 此外, 设定每一次实验中随机搜索执行次数 η 最高为 10000。在计算机上创建的虚拟相机焦距为 760 pixel, 成像尺寸为 1600 pixel \times 1600 pixel, 每次实验中, 3D 特征点随机分布在相机坐标系中 $[-2, 2] \times [-2, 2] \times [4, 9]$ 的区域内, 3 个旋转角在 $[-90^\circ, 90^\circ]$ 之间随机产生, 初始平移向量的取值范围是 $[-5, 5] \times [-5, 5] \times [0, 10]$, 单位为 m。

相机位姿估计结果评估中, 位姿误差的定义为

$$p_{\text{err}} = \|\mathbf{R}_{\text{est}} \mathbf{T}_{\text{est}} - \mathbf{R}_{\text{ture}} \mathbf{T}_{\text{true}}\|, \quad (12)$$

式中 $\mathbf{R}_{\text{est}} \mathbf{T}_{\text{est}}$ 为算法搜索结果值, $\mathbf{R}_{\text{ture}} \mathbf{T}_{\text{ture}}$ 为真实值。

图 2 为合成图像实验的实验过程, 含有的 17 个 2D 特征点和 3D 特征点中, 包括 2 个 3D 遮挡特征点和 2D 伪特征点, 其中绿色圆圈代表 2D 特征点, 红点代表 3D 特征点, 连接线是为了方便观察而设置, 不影响结果。图 2(a) 是给定初始值后 3D 特征点重投影与 2D 特征点的分布情况, 图 2(b) 代表第 7 次迭代后的结果, 图 2(c) 则是最终的计算结果; 当红点位于绿色圆圈中, 代表 3D/2D 特征点对应, 反之, 则无对应点, 观察可得, 最终结果中包含 2 个 3D 遮挡特征点和 2D 伪特征点。

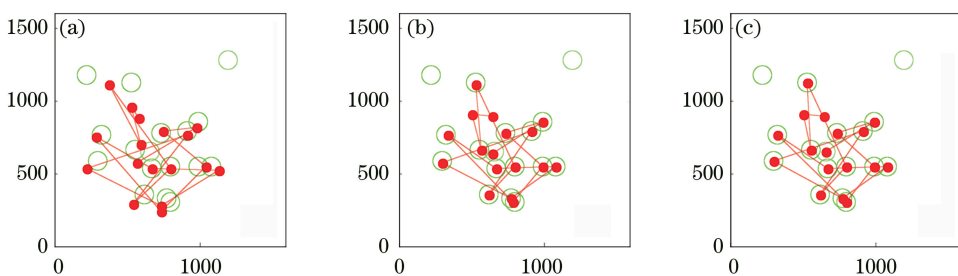


图 2 合成图像实验过程。(a) 第 1 次迭代; (b) 第 7 次迭代; (c) 最终结果

Fig. 2 Process of the synthetic experiment. (a) The 1st iteration; (b) the 7th iteration; (c) final result

图 3 为特征点数量、2D 伪特征点比例和 3D 遮挡特征点比例对算法成功率、算法运行时间和位姿估计误差等结果的影响图。实验中, 取图像噪声的标准差 $\sigma = 0.5$, 3D 特征点的数量为 $L \in \{10, 20, 30, 40, 50\}$, $a \in \{0, 0.2, 0.4\}$, $b \in \{0, 0.2, 0.4\}$, 实验次数为 $n_t = 100$ 。由图 3(a) 可以得出, 合成图像实验成功率在 92% 以上, 随着 3D 特征点数量增多, 成功率

逐渐升高; 随着 a 和 b 的值升高, 成功率逐渐下降。这是因为 3D 特征点数量越多, 相机位姿的迭代结果越准确, 由于 a 和 b 代表干扰点的比例, 干扰点的比例越大, 越容易陷入局部最优而返回空结果, 使成功率降低。由图 3(b) 分析可得, 算法的耗时随着特征点数量的增多而增大, 但 a 和 b 对算法耗时影响不大, 这是由于算法的时间复杂度与特征点的数量

相关,干扰点比例不影响特征点总数,因此不影响算法的时间复杂度。由图 3(c)可得,估计的相机位姿误差随着特征点的数量增大而呈减小趋势,这是因

为特征点的数量越多,图像噪声引起的随机误差之间相互抵消,因而位姿误差逐渐减小。

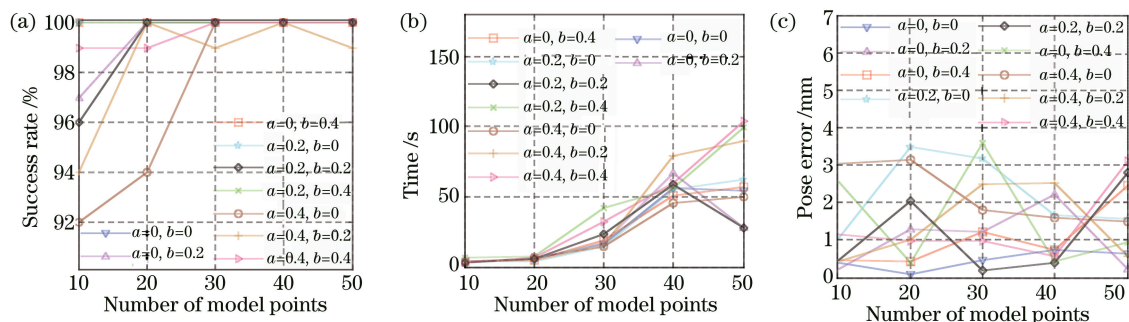


图 3 特征点的参数对相机位姿估计结果的影响。(a)成功率;(b)耗时;(c)位姿估计误差

Fig. 3 Effects of feature point parameters on the camera pose estimation. (a) Success rate; (b) time consuming; (c) pose estimation error

图 4 为图像噪声对相机位姿估计结果的影响图,其中 $\sigma \in \{0, 1, 2, 3, 4, 5\}$ 、 $L \in \{10, 20, 30, 40, 50\}$ 、 $a=b=0$ 、 $n_i=100$ 。由图 4(a)可得,当 3D 特征点数量不变时,随着图像噪声 σ 增大,成功率逐渐下降;同一噪声水平下,随着特征点数量增大,成功率先增大后减小,所有实验成功率 $> 82\%$ 。该实验结果表明图像噪声对算法的成功率影响较大,图像噪声越大,算法的成功率越低。这是因为噪声干扰了

特征点之间的匹配关系,从而导致成功率下降。由图 4(b)可得,图像噪声增加了算法的耗时。由图 4(c)可得,当特征点数量不变时,随着图像噪声增大,位姿估计误差逐渐增大;当图像噪声不变时,3D 特征点数量越多,位姿估计误差越小。这是由于图像噪声越大,估计的位姿越不准确,当 3D 特征点逐渐增多时,特征点之间的误差相互抵消,减弱了图像噪声的影响。

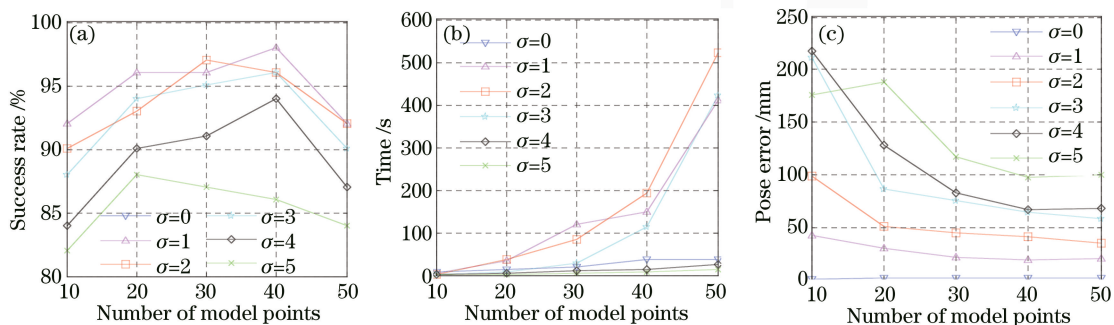


图 4 图像噪声对相机位姿估计结果的影响。(a)成功率;(b)耗时;(c)位姿估计误差

Fig. 4 Effects of pattern noise on the camera pose estimation. (a) Success rate; (b) time consuming; (c) pose estimation error

将所提出的 SoftOI 算法与文献[14]的 PSKPE 算法进行比较,令 $\sigma=2$ 、 $L \in \{10, 20, 30, 40, 50\}$ 、 $n_i=100$ 。由于 PSKPE 算法无法处理同时存在 3D 遮挡特征点和 2D 伪特征点的情况,则分别在 $a=b=0$ 、 $a=0$ $b=0.1$ 和 $a=0.1$ $b=0$ 情况下进行实验,PSKPE 算法的迭代次数最高为 20000。图 5 是实验对比结果,从图 5(a)可以得出,当 $a=b=0$ 时,SoftOI 算法的成功率低于 PSKPE 算法,其他情况下,SoftOI 算法的成功率与 PSKPE 算法基本持平

或者高于 PSKPE 算法;随着特征点数量增加,PSKPE 算法的成功率下降地更快一些。由图 5(b)、(c)可知,SoftOI 算法的运算时间较少,误差较大。改进的平移向量计算方式和迭代终止条件,使得 SoftOI 算法耗时减少,同时保证了较高的成功率,而 PSKPE 算法每次搜索的迭代次数均为 20000,在牺牲了时间的情况下,获得了较高的精度。而改进的匹配概率矩阵计算方式,使得 SoftOI 算法的应用范围更广。

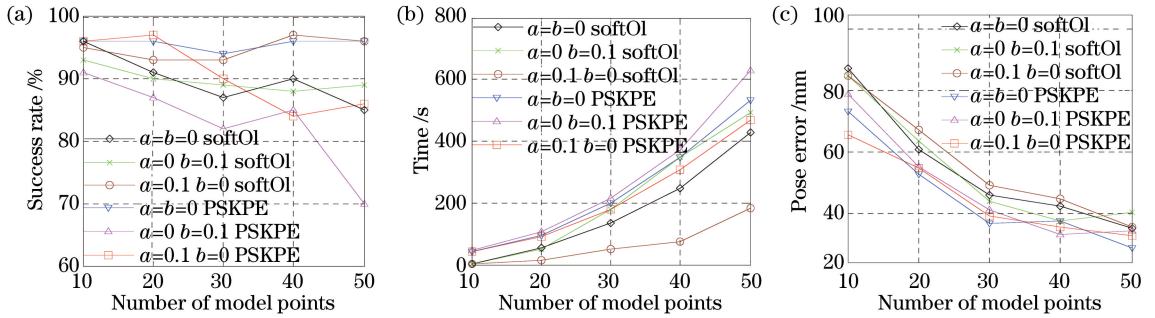


图 5 SoftOI 和 PSKPE 算法实验对比结果。(a)成功率;(b)所耗时间;(c)位姿估计误差

Fig. 5 Comparison results of SoftOI and PSKPE algorithms. (a) Success rate;

(b) time consuming; (c) pose estimation error

将所提出的 SoftOI 算法与文献 [10] 的 SoftPOSIT 算法进行更全面的对比。令 $\sigma=0, L \in \{10, 20, 30, 40, 50\}, a=b=0.1, n_t=100$, 在实验中引入相对距离 r_d 的概念, 即:

$$r_d = \frac{T_z}{d}, \quad (13)$$

式中 T_z 为平移矩阵中的第三个元素, d 为 3D 特征点质点系的直径, $d = \frac{2}{L} \sum_{k=1}^L (\mathbf{P}_k^w - \mathbf{P}_c^w)$ 。取 $r_d \in \{1, 3, 5, 7\}$, 对两种算法进行实验。

图 6 所示为 SoftOI 与 SoftPOSIT 算法的实验对比结果。由图 6(a)可得, 实验中 SoftOI 算法成功率高于 SoftPOSIT 算法的成功率, 在特征点数量不变的情况下, r_d 对 SoftOI 算法的成功率影响较

小, 而 r_d 越小, SoftPOSIT 的成功率越低; 当特征点数量较少时, SoftPOSIT 算法成功率较低。由图 6(b)可知, 在 L 和 r_d 值一定的情况下, SoftOI 算法的搜索时间要小于 SoftPOSIT 算法的搜索时间; 随着 r_d 增大, SoftOI 算法的搜索时间增加, SoftPOSIT 算法的搜索时间减少。由图 6(c)可知, 实验条件相同时, SoftOI 算法的位姿估计误差小于 SoftPOSIT 算法的位姿估计误差。综上所述, SoftOI 算法在成功率、耗时和位姿误差上要优于 SoftPOSIT 算法, 这是因为 OI 算法无应用的限制条件, 对近场 (r_d 较小) 情况下的相机位姿估计更迅速准确, 而 POSIT 算法在近场情况下收敛速度慢或者不收敛, 导致成功率低, 耗时长; 在终止条件上, SoftOI 算法采用的空间误差比 SoftPOSIT 算法采用的图像误差更精确, 因而位姿误差更低。

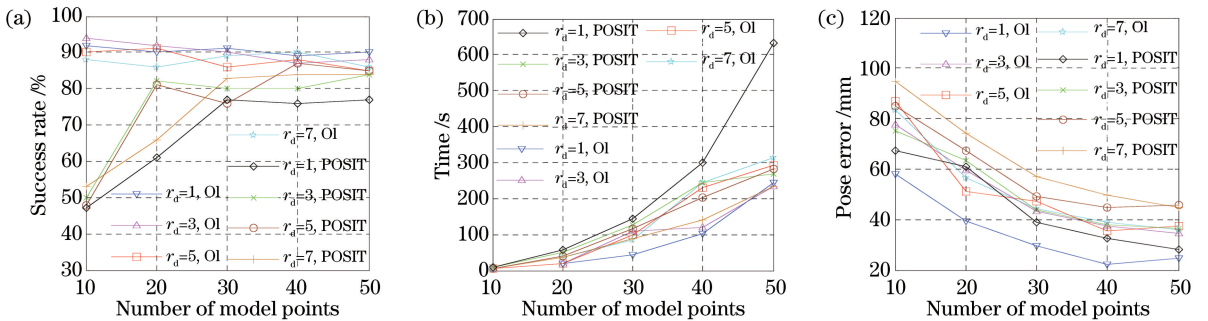


图 6 SoftOI(OI)和 SoftPOSIT(POSIT)算法实验对比结果。(a)成功率;(b)所耗时间;(c)位姿估计误差

Fig. 6 Comparison results of SoftOI(OI) and SoftPOSIT(POSIT) algorithms. (a) Success rate;

(b) time consuming; (c) pose estimation error

3.2 真实图像实验

实验采用真实图像对算法进行验证, 图像由焦距为 1391 pixel 的相机对真实物体拍照所得, 图像尺寸为 1920 pixel \times 1080 pixel, 主点位于 (971.850 pixel, 520.210 pixel), 噪声的标准差取 $\sigma=1$, 其他参数与数字仿真实验相同。

实验中被拍摄物体为一 T 型装配体, 如图 7 所

示。从图像中选取 10 个点作为 2D 特征点, 如图 7(a)所示; 从装配体中选取 11 个角点作为 3D 特征点, 为了更直观的展示, 将 3D 特征点重投影到图像上, 如图 7(b)、(c)所示。为了验证算法的稳定性, 特征点中包含了 3 个 3D 遮挡特征点和 2 个 2D 伪特征点。图 7(b)为算法迭代到第 7 次的重投影效果, 图 7(c)为算法最终结果下的重投影效果。3D

特征点的重投影点逐渐靠近图像中相匹配的点,最终完全重合。将一虚拟长方体投影到装配体上,如图7(d)所示,虚拟长方体与物体及周围的环境相符合,说明相机位姿估计误差较小。该实验运行次数为20,成功次数为19,成功率为95%,搜索时间平均为2.67 s,实验证明,该算法能在真实图像中快速准确地获得相机的位姿。

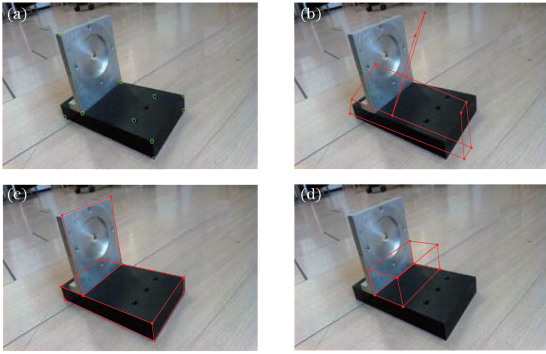


图7 真实图像的相机位姿估计过程。(a) 2D特征点设置;(b)算法运行至第7次时的结果;(c)最终的3D特征点重投影结果;(d)虚拟长方体的投影结果

Fig. 7 Camera pose estimation process of real image.

(a) Configuration of 2D feature points; (b) result after the 7th iteration; (c) final reprojection result of the 3D feature points; (d) projection result of virtual cube

4 结 论

针对3D/2D特征点对应关系未知的情况,提出了一种结合了Softassign算法和OI算法的相机位姿估计算法,通过3D/2D特征点之间的共线性误差确定特征点的匹配关系,从而计算相机的位姿。实验表明:该算法在不增加算法复杂度的情况下能够处理同时存在3D遮挡特征点和2D伪特征点的情况;加入了图像噪声后,成功率 $>82\%$ 。与其他算法对比,该算法的成功率高、相机位姿估计误差小,在被测物体距离相机较近的场景中,估计相机位姿成功率高,适用于机器人装配等领域。未来将在存在图像噪声环境下提高算法的成功率方面做进一步深入研究,并将该算法应用到机器人装配领域中。

参 考 文 献

[1] Vincenzani L, Francese M N, Pascual F. EPnP: an accurate $O(n)$ solution to the PnP problem [J]. International Journal of Computer Vision. 2009, 81(2): 155-166.

[2] Hesch J A, Roumeliotis S I. A direct least-squares

(DLS) method for PnP [C] // Proceedings of IEEE International Conference on Computer Vision, 2011: 383-390.

- [3] Dai D K, Wang X S, Hu C S, *et al.* Camera calibration and attitude measurement technology based on astronomical observation [J]. Acta Optica Sinica, 2012, 32(3): 0312005.
戴东凯, 王省书, 胡春生, 等. 基于天文观测的相机标定及姿态测量技术研究 [J]. 光学学报, 2012, 32(3): 0312005.
- [4] Oberkempf D, DeMenthon D F, Davis L S. Iterative pose estimation using coplanar feature points [J]. Computer Vision and Image Understanding, 1996, 63(3): 495-511.
- [5] Lu C P, Hager G D, Mjolsness E. Fast and globally convergent pose estimation from video images [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(6): 610-622.
- [6] Kahl F, Henrion D. Globally optimal estimates for geometric reconstruction problems [J]. International Journal of Computer Vision, 2007, 74(1): 3-15.
- [7] Yue X K, Wu Y Y, Wu K Z. A novel binocular OI (orthogonal iteration) fusion algorithm better for estimation of spacecraft position and attitude [J]. Journal of Northwestern Polytechnical University, 2011, 29(4): 559-563.
岳晓奎, 武媛媛, 吴侃之. 基于视觉信息的航天器位姿估计迭代算法 [J]. 西北工业大学学报, 2011, 29(4): 559-563.
- [8] Yang H W, Wang F, Li Z, *et al.* Simultaneous pose and correspondence estimation based on genetic algorithm [J]. International Journal of Distributed Sensor Networks, 2015, 11(11): 828241.
- [9] Fischler M A, Bolles R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography [J]. Readings in Computer Vision, 1987: 726-740.
- [10] David P, DeMenthon D, Duraiswami R, *et al.* SoftPOSIT: simultaneous pose and correspondence determination [J]. International Journal of Computer Vision, 2004, 59(3): 259-284.
- [11] Shi J F, Ulrich S. Soft POSIT enhancements for monocular camera spacecraft pose estimation [C] // Proceedings of IEEE 21st International Conference on Methods and Models in Automation and Robotics, 2016: 30-35.
- [12] Gold S, Rangarajan A. A graduated assignment

- algorithm for graph matching[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996, 18(4): 377-388.
- [13] Zhou H Y, Zhang T, Lu W N. Vision-based pose estimation from points with unknown correspondences[J]. IEEE Transactions on Image Processing, 2014, 23(8): 3468-3477.
- [14] Chen P, Hu G D, Yan G R. Camera pose estimation algorithm based on particle system kinematics [J]. Acta Optica Sinica, 2014, 34(1): 0112002.
陈鹏, 胡广大, 闫桂荣. 基于质点系运动原理的像机位姿估计算法 [J]. 光学学报, 2014, 34(1): 0112002.
- [15] Ugurdag H F, Gören S, Canbay F. Gravitational pose estimation [J]. Computers & Electrical Engineering, 2010, 36(6): 1165-1180.
- [16] Ugurdag H F, Goren S, Canbay F. Correspondenceless pose estimation from a single 2D image using classical mechanics[C] // Proceedings of IEEE 23rd International Symposium on Computer and Information Sciences, 2008: 1-6.
- [17] He F Y, Lin J R, Gao Y, *et al.* Optimized pose measurement system combining monocular vision with inclinometer sensors [J]. Acta Optica Sinica, 2016, 36(12): 1215002.
何斐彦, 林嘉睿, 高扬, 等. 单目视觉与倾角仪组合优化的位姿测量系统 [J]. 光学学报, 2016, 36(12): 1215002.
- [18] Guan R F, Yang L H, Wang L J, *et al.* Position and attitude precision measurement of spatial objects based on orthogonal cylindrical imaging [J]. Acta Optica Sinica, 2016, 36(11): 1112003.
关瑞芬, 杨凌辉, 王丽君, 等. 基于正交柱面成像的空间物体位姿精密测量 [J]. 光学学报, 2016, 36(11): 1112003.
- [19] Zhang L L, Zhang W, Hu C H, *et al.* Mobile videometrics method based on relay camera [J]. Acta Optica Sinica, 2017, 37(4): 0412001.
张林龙, 张伟, 胡昌华, 等. 基于传递像机的移动摄像测量方法 [J]. 光学学报, 2017, 37(4): 0412001.