

非合作纹理目标单目位姿计算

冯肖维, 谢安安, 肖健梅, 王锡淮

引用本文:

冯肖维, 谢安安, 肖健梅, 等. 非合作纹理目标单目位姿计算[J]. *光学精密工程*, 2020, 28(8): 1775–1784.

FENG Xiao-wei, Xie An-an, Xiao Jian-mei, et al. Monocular pose calculation of non-cooperation textured object[J]. *Optics and Precision Engineering*, 2020, 28(8): 1775–1784.

在线阅读 View online: <https://doi.org/10.3788/OPE.20202808.1775>

您可能感兴趣的其他文章

Articles you may be interested in

空间非合作旋转目标的模型重建与位姿优化

Model reconstruction and pose optimization of non-cooperative rotating space target

光学精密工程. 2019, 27(8): 1854–1862 <https://doi.org/10.3788/OPE.20192708.1854>

二维合作目标的单相机空间位姿测量方法

Methods of monocular pose measurement based on planar objects

光学精密工程. 2017, 25(1): 274–280 <https://doi.org/10.3788/OPE.20172501.0274>

非合作旋转目标闭环检测与位姿优化

Closed-loop detection and pose optimization of non-cooperation rotating targets

光学精密工程. 2017, 25(4): 1036–1043 <https://doi.org/10.3788/OPE.20172504.1036>

工业机器人单目视觉对准技术研究

Research on industrial robot alignment technique with monocular vision

光学精密工程. 2018, 26(3): 733–741 <https://doi.org/10.3788/OPE.20182603.0733>

基于多直线对应和加权最小二乘的位姿估计

Pose estimation based on multiple line hypothesis and iteratively reweighted least squares

光学精密工程. 2015, 23(6): 1722–1731 <https://doi.org/10.3788/OPE.20152306.1722>

文章编号 1004-924X(2020)08-1775-10

非合作纹理目标单目位姿计算

冯肖维*, 谢安安, 肖健梅, 王锡淮
(上海海事大学 电气自动化系, 上海 201306)

摘要:为了实时测量非合作纹理运动目标相对位姿,提出一种单目视觉同步建模与位姿计算方法。选择具有良好特征共视关系与分布的模型帧,增量式恢复目标三维共视特征模型,实现非合作目标的合作化,并依靠运动预测模型实现基于特征跟踪的位姿计算。根据模型对应网格的拓扑关系估计目标表面未知区域特征三维信息,结合网格面法向场进行局部光束平差优化,同时利用闭环优化抑制尺度漂移,在减少特征模型恢复误差的同时提高位姿计算精度。实验结果表明,在非结构化环境中能够在线恢复目标三维信息,实现目标相对位姿准确计算,特征平均重投影误差小于 1.5 pixel,位姿计算平均绝对误差为 4.29 mm 和 1.54°,平均计算时间小于 120 ms,为基于单目视觉的在线三维感知与测量建模提供技术支持。

关键词:单目视觉;非合作目标;位姿计算;特征建模;非线性优化

中图分类号:TP391 **文献标识码:**A **doi:**10.3788/OPE.20202808.1775

Monocular pose calculation of non-cooperation textured object

FENG Xiao-wei*, Xie An-an, Xiao Jian-mei, Wang Xi-huai

(Department of Electrical Automation, Shanghai Maritime University, Shanghai 201306, China)

* Corresponding author, E-mail: xwfeng1982@163.com

Abstract: To measure the relative pose of moving non-cooperation textured objects in real time, a monocular simultaneous modeling and pose calculation method was proposed. A 3D covisibility model was incrementally constructed with frames containing the highest covisibility of features and best distribution to achieve cooperation between non-cooperation objects. Subsequently, the relative pose of the object was calculated via feature tracking by the motion prediction model. The mesh of the model was used to restore the 3D information of feature points that were distributed in an unknown area of the object surface. To reduce model error and improve the accuracy of pose estimation, bundle adjustment optimization was performed using a facet normal field, and the scale drift was reduced using closed-loop optimization. Experiments show that the method is a real-time online system that can recover 3D information of object in unstructured environments and accurately estimate relative poses in unstructured environments to provide technical support for 3D sensing and measurement modeling based on monocular vision. The mean reprojection error (MRE) of the features using the proposed method is less than 1.5 pixels, and the mean absolute error (MAE) of pose calculation is 4.29 mm and 1.54° while the average time consumption is less than 120 ms.

收稿日期:2020-03-11;修订日期:2020-04-17.

基金项目:国家自然科学基金资助项目(No. 61503241, No. 61801287)

Key words: monocular vision; non-cooperation objects; pose calculation; feature modelling; nonlinear optimization

1 引言

非合作目标位姿测量广泛应用于机器人环境感知、空间交会对接和虚拟现实等领域^[1-4]。由于无法事先获得目标物体外形信息,使非合作目标位姿计算变得复杂与困难,必须首先恢复目标表面三维信息(如纹理特征、外形轮廓等),然后利用这些信息提供的空间约束关系解算位姿^[5]。

单目视觉系统无法直接获得景深信息,只能利用目标与相机间的相对运动来恢复场景三维结构(Structure From Motion, SFM)^[6]。而单目位姿计算不仅要恢复目标三维结构,还需要实时精确估计相机与目标之间的相对位姿。Segal^[7]首先恢复目标三维信息,然后利用扩展卡尔曼滤波算法迭代求解位姿,由于需要事先恢复目标的三维结构,无法实时在线应用。在一些非结构化应用场景中,由于无法事先恢复目标完整外形信息,只能根据逐渐恢复的目标三维表面结构信息来确定目标与相机间的相对位姿,此过程类似于同时定位与地图构建(Simultaneous Localization and Mapping, SLAM)^[8]。Augenstein^[9]利用粒子滤波实现目标同步三维重建与位姿估计,依靠粒子滤波的多假设估计提高算法抗干扰能力。为了实现目标特征和相对位姿的联合估计,前述方法需要滤波器对每帧图像进行处理,将时间消耗在处理具有极少新增信息的冗余图像帧中,且忽略了累积线性化误差。为此, Klein^[4]和 Mouragnon^[10]提出基于关键帧的单目视觉方法,仅使用选定的图像帧重构三维信息,因此可以执行更费时但更准确的非线性优化。Strasdat 等人证明在相同的计算成本下,基于关键帧的优化技术比滤波法更为准确^[11]。

在单目视觉 SLAM 中,环境中的所有特征都可以用于解算位姿,当镜头移动时,连续图像帧之间存在丰富的共视特征,保证视觉 SLAM 具有较高稳定性与抗干扰能力。但在指定目标的位姿计算中,目标只占图像的局部区域,提取的特征数量有限,使得空间约束关系变少,算法的稳定性变差,特别是当背景特征丰富时,会进一步给目标特

征辨识造成困难,导致计算失败。此外,单目视觉具有尺度漂移问题,目标模型在增量式恢复过程中会产生累积误差,进而影响位姿计算的精度。Mur-Artal 等^[8]利用关键帧思想增量式恢复环境三维共视特征地图,并依靠回环检测与位姿图优化保证地图尺度一致性,实现相机相对非结构环境的位姿计算,但是无法针对具体目标计算相对位姿。在此基础上,刘宗明等^[12]实现非合作目标基于关键帧的纯旋转位姿测量,由于没有建立目标模型,无法分离前景与背景,因此该方法只能在结构化环境中使用。并且前述方法在优化时都只考虑成像模型约束,而没有考虑特征点与目标轮廓之间的几何关系。

非合作目标位姿计算的关键是如何在线恢复目标外形信息、抑制累积误差以及避免环境干扰。本文选择具有良好共视特性的图像帧——模型帧——构建共视特征模型,实现非合作纹理目标三维信息的合作化。利用网格模型对目标表面轮廓及其拓补关系进行估计,增量式恢复目标表面未知区域特征点与非共视特征点三维信息,增加模型包含特征的数量。并利用网格法向场引导特征模型的优化调整,减少由于网格近似造成的特征恢复误差,从而提高目标位姿的计算精度。在此基础上引入模型帧闭环优化,抑制尺度漂移形成的累积误差。依靠运动预测模型跟踪共视特征实现目标相对位姿的实时计算。

2 目标特征模型恢复

2.1 共视特征模型

纹理目标表面具有丰富的特征点信息(如 ORB, SIFT, SURF 等^[13]),不受物体表面形状的限制,且提取速度较快,可以满足特征匹配的准确性与位姿计算的实时性需求。本文根据不同时刻 k 获得的模型帧序列 $\{I_m^k\}$, 实现模型帧中共视 ORB 特征 $\{p^{i,k}\}$ 三维坐标 $\{P_w^k\}$ 与目标位姿的同步计算。如果利用初始模型帧 I_m^0 构建世界坐标系 \mathbf{o}_w , 将后序模型帧的位姿表示为该模型帧相对世界坐标系 \mathbf{o}_w 的位姿变换矩阵 $\{T_{m,w}^k\}$, 则目标的共视特征模型可以表示为:

$$M = \{ \{T_{m,w}^k\}, \{p^{i,k}\}, \{P_w^i\} \}. \quad (1)$$

考虑实时性与尽可能多的引入视差,模型帧在选择时需要满足如下条件:

- (1)模型帧的间隔大于 15 个连续图像帧;
- (2)模型帧必须包含 $\alpha_m=40$ 个以上目标表面特征点,从而提供尽可能多的约束条件,便于三维信息的恢复与优化;
- (3)新增加的模型帧与某个已有模型帧之间共视特征的比率要小于 $\beta_m=85\%$,使邻近模型帧之间具有较大的视觉变化,提高匹配特征之间的视差。

由于使用了共视特征模型,位姿计算与模型优化可以在局部区域进行,而与图像帧率无关。同时,为了控制共视特征模型 M 中模型帧的总体数量,每增加新的模型帧都会伴随冗余帧检测,剔除模型中特征匹配率高于 $\gamma_m=90\%$ 的模型帧,使得在增量式建模过程中可以灵活地扩展目标模型,当连续观测到目标表面相同区域时,模型无需被重复构建。

2.2 模型初始化

根据 2 帧图像之间的对极几何恢复目标与相机间的初始相对位姿^[14]。首先,通过交互手段在初始图像帧中选择目标的大致区域,检测任意邻近 2 帧图像之间匹配的 ORB 特征对;然后,为了防止计算出现退化,根据匹配的 ORB 特征点是否位于同一个平面上,分别选用基础矩阵模型或单应性矩阵模型解算 2 帧图像间相对位姿变换矩阵 $T_{m,w}^1$,增加初始化过程的鲁棒性;最后,根据初始化成功时 2 帧图像中的第 1 帧图像作为初始模型帧 I_m^0 构建世界坐标系 o_w ,此时初始模型帧的位姿 $T_{m,w}^0$ 为单位矩阵,并将第 2 帧图像作为第 1 帧模型帧 I_m^1 。 $T_{m,w}^1$ 解算成功后,根据三角测量法^[14]恢复模型帧 I_m^0 和 I_m^1 中匹配 ORB 特征的三维信息 $\{P_w^i\}$;若解算不成功,则重新选择 2 帧图像重复上述过程,直至位姿变换 $T_{m,w}^1$ 解算成功。

2.3 三维信息恢复

如果目标为刚体,单目视觉系统可以根据不同时刻图像帧中二维特征点间匹配关系,利用三角测量法恢复三维信息^[14],如图 1 中 P_{match}^{int} 和 P_{match}^{ext} 。而目标表面没有共视关系的特征点,本文借助已恢复特征点估计其三维坐标,如 P_{new}^{int} 和 P_{new}^{ext} 。

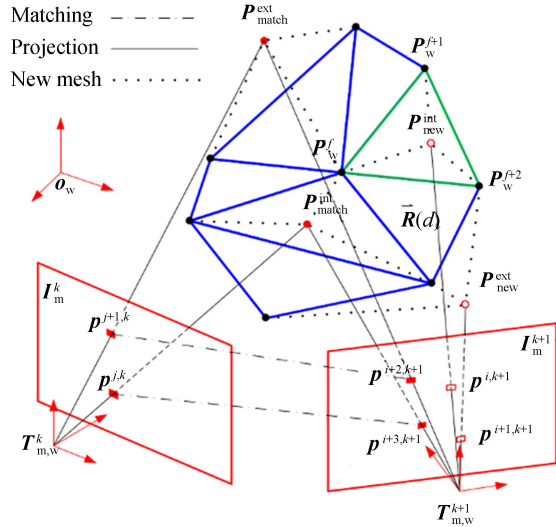


图 1 特征点三维重构

Fig. 1 3D reconstruction of feature points

根据模型 M 中已恢复的三维特征点 $\{P_w^i\}$,可以获得目标表面轮廓的近似三角网格模型 F ,如图 1 所示。如果 $o_{m,w}^{k+1}$ 为模型帧 I_m^{k+1} 的光心, r 为射线 $R(d)$ 的单位向量,则该射线上任意一点可表示为:

$$R(d) = o_{m,w}^{k+1} + d \cdot r. \quad (2)$$

如果网格模型 F 中某三角面 f 的顶点为 $\{P_w^f, P_w^{f+1}, P_w^{f+2}\}$,则 f 所在平面内任意点的坐标 $P(u, v)$ 可以表示为:

$$P(u, v) = (1 - u - v) \cdot P_w^f + u \cdot P_w^{f+1} + v \cdot P_w^{f+2}, \quad (3)$$

则模型帧 I_m^{k+1} 上某 ORB 特征 $p^{i,k+1}$ 的三维坐标估计值可以通过射线 $R(d)$ 与三角形 f 所在平面的交点获得,即:

$$o_{m,w}^{k+1} + d \cdot r = (1 - u - v) \cdot P_w^f + u \cdot P_w^{f+1} + v \cdot P_w^{f+2}. \quad (4)$$

如果 $u \geq 0, v \geq 0$,且 $u + v \leq 1$,则 $P(u, v)$ 位于三角形 f 内,如点 P_{new}^{int} ;否则,位于三角形外,如点 P_{new}^{ext} 。将计算出的深度信息 d 代入公式(2)即可获得交点坐标,从而恢复模型帧中未匹配的新增特征点三维信息。

2.4 特征点的选择

为了只恢复目标表面的特征,本文利用网格模型 F 的拓扑关系对位于网格外的特征点进行筛选,避免背景特征加入共视特征模型。

如果新恢复的特征点不在网格中任何一个三

角面内(如图 1 中 $\mathbf{P}_{\text{new}}^{\text{ext}}$),即不在已恢复共视模型区域内,计算 $\mathbf{P}_{\text{new}}^{\text{ext}}$ 与最近的三角面 $f_c(\mathbf{c}_{f_c}, \mathbf{n}_{f_c})$ 的垂直偏离程度:

$$D_c = \mathbf{n}_{f_c} \cdot (\mathbf{P}_{\text{new}}^{\text{ext}} - \mathbf{c}_{f_c}), \quad (5)$$

其中 $\mathbf{n}_{f_c}, \mathbf{c}_{f_c}$ 分别为 f_c 的法向量与中心点,可以通过 f_c 的 3 个顶点计算得到。若 D_c 大于某个阈值 ϵ ,则认为 $\mathbf{P}_{\text{new}}^{\text{ext}}$ 不为目标上的点。同时,假设目标表面轮廓在局部区域连续,则网格面只能对其邻近的未知区域形状进行拓扑估计,因此需要剔除离中心 \mathbf{c}_{f_c} 距离较远的特征点。此外,本文只保留至少能被 3 个模型帧观测到的特征点,在提高模型共视关系的同时避免信息冗余。

2.5 模型法向约束优化

由于三角网格模型是实际物体表面的一个近似逼近,因此根据网格模型重投影恢复的特征三维信息含有误差,影响位姿计算的精度。当共视特征模型被更新,即有新的模型帧被加入时,本文所述方法会对共视模型中与新加入模型帧具有共视关系的区域进行优化调整。

假设镜头无畸变,相机的内参矩阵为 \mathbf{K} ,则成像模型 π 可以表示为:

$$s^{i,k} \mathbf{p}^{i,k} = \pi(\mathbf{T}_{m,w}^k, \mathbf{P}_w^i) = \mathbf{K} \mathbf{T}_{m,w}^k \mathbf{P}_w^i, \quad (6)$$

其中 $s^{i,k}$ 为尺度系数,取默认值 $s^{i,k} = 1$,则第 i 个特征点对第 k 个模型帧的重投影误差为:

$$\mathbf{e}_{i,k} = \mathbf{p}^{i,k} - \pi(\mathbf{T}_{m,w}^k, \mathbf{P}_w^i). \quad (7)$$

如果将新模型帧及共视模型帧的位姿和对应的特征点作为优化初值 $\{\tilde{\mathbf{T}}_{m,w}^k\}$ 和 $\{\tilde{\mathbf{P}}_w^i\}$,根据重投影误差可以建立基于投影约束的模型优化目标函数:

$$E_1(\{\tilde{\mathbf{T}}_{m,w}^k\}, \{\tilde{\mathbf{P}}_w^i\}) = \sum_k \sum_i \rho_H(\mathbf{e}_{i,k}^T \cdot \boldsymbol{\Omega}_{i,k}^{-1} \cdot \mathbf{e}_{i,k}), \quad (8)$$

其中: ρ_H 为 Huber 鲁棒核函数, $\boldsymbol{\Omega}_{i,k} = (s^{i,k})^2 \cdot \mathbf{I}_{2 \times 2}$ 为特征点 \mathbf{P}_w^i 在模型帧 I_m^k 上的重投影误差协方差矩阵。

目标表面三维特征除了和相机之间具有投影约束关系外,还应该位于目标表面轮廓上,即特征点 $\{\mathbf{P}_w^i\}$ 的位置优化还应该满足目标外形约束^[15]。网格中三角面是目标局部区域切平面的近似,因此其法向也是对应目标表面区域法向的近似。为了使特征点在优化迭代过程中收敛到真实目标表面,本文利用网格面法向场引导三维特征点位置的更新。

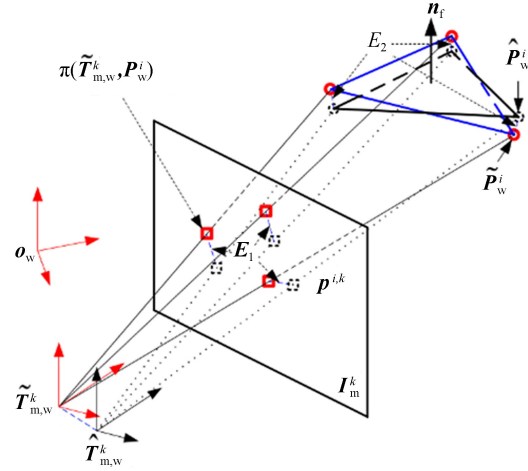


图 2 特征模型局部优化

Fig. 2 Local optimization of feature model

如图 2 所示,目标表面三角网格模型 F 中局部区域 \mathbb{N}_f 的某个三角面片 $f \in \mathbb{N}_f$ 的法向量为 \mathbf{n}_f 。则根据 \mathbb{N}_f 中每个三角面片的法向量与其 3 条边之间的正交关系,定义如下优化正则项:

$$E_2(\{\mathbf{P}_w^i\}) = \sum_{f \in \mathbb{N}_f} \sum_{(m,n) \in \partial f} [\mathbf{n}_f \cdot (\mathbf{P}_w^m - \mathbf{P}_w^n)]^2, \quad (9)$$

其中 ∂f 表示面 f 对应的边。则最终的基于面法向场约束的模型局部优化表示为:

$$\operatorname{argmin}_{\{\tilde{\mathbf{T}}_{m,w}^k\}, \{\tilde{\mathbf{P}}_w^i\}} [E_1(\{\tilde{\mathbf{T}}_{m,w}^k\}, \{\tilde{\mathbf{P}}_w^i\}) + \theta E_2(\{\mathbf{P}_w^i\})], \quad (10)$$

其中参数 θ 用于控制投影约束优化与模型法向约束优化之间的平衡。本文取 $\theta = 1$,使重投影误差与法向调整误差具有相同重要性。本文使用 Levenberg-Marquardt 法进行非线性优化求解。

2.6 模型闭环尺度优化

如果相机再次采集到目标表面某个已建模区域,则可以通过模型中已建立的模型帧调整当前模型帧的尺度。

首先利用特征匹配在共视特征模型 M 中搜索当前模型帧 I_m^c 对应的闭环模型帧 I_m^l ,并从 2 个模型帧之间 m 组具有共视关系的 ORB 特征点对 $\mathbf{p}^{i,c} \leftrightarrow \mathbf{p}^{j,l}$ 中任意选择 3 组计算这 2 个模型帧之间的相似变换矩阵初值^[8],并根据 2 帧之间的对称转换重投影误差 \mathbf{e}_c 和 \mathbf{e}_l 构建目标函数优化 $\mathbf{S}_{m,m}^{c,l}$:

$$\mathbf{e}_c = \mathbf{p}^{i,c} - \pi_c(\mathbf{S}_{m,m}^{c,l}, \mathbf{P}_w^i), \quad (11)$$

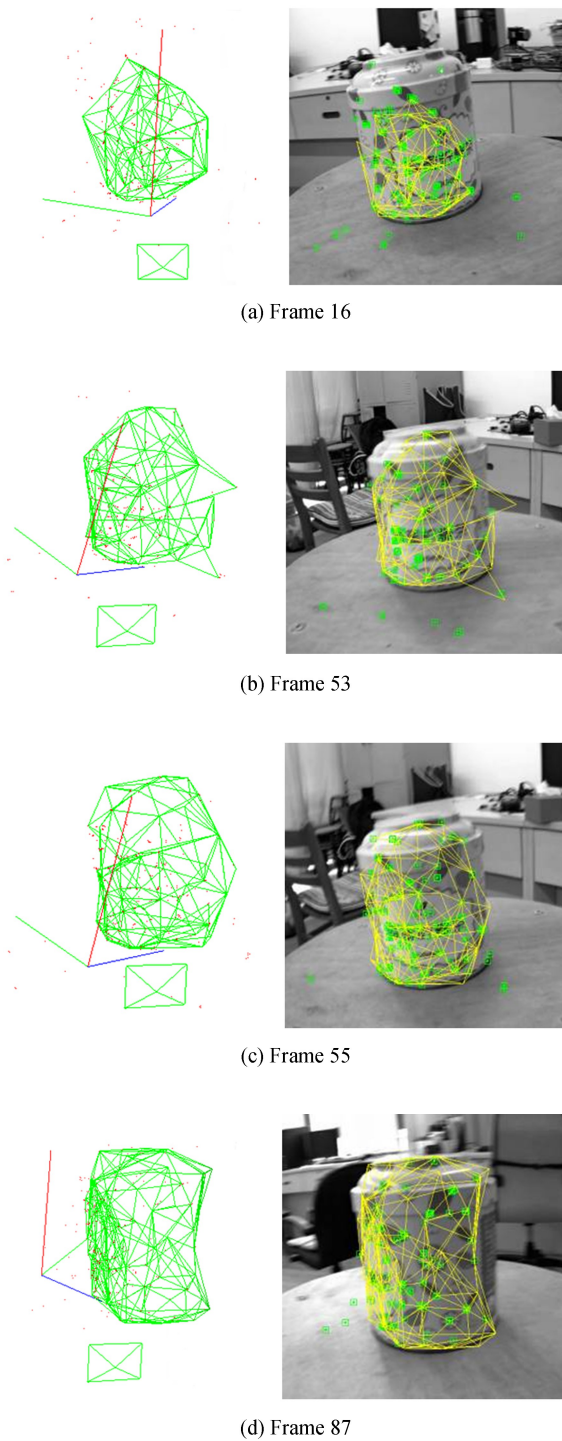


图 4 圆筒增量式建模结果

Fig. 4 Incremental modelling results of cylinder

同时,将左侧三维网格模型根据计算出的目标位姿重投影到右侧对应二维图像帧上,右图中绿色小方框为提取的 ORB 特征。随着计算的推进,重投影网格与目标轮廓保持重合,说明本文所述方法可以对非合作目标进行在线同步建模与位姿

计算。此外,图 4(b)中有 2 个背景特征点被误加入模型,使模型产生了畸变,随着视角的变化,新信息的到来,干扰特征点被辨识,并从模型中剔除,如图 4(c)所示,说明本文所述方法的特征选择机制可以有效区分目标与背景特征,从而保证只有目标表面的特征用于位置计算。

图 5 显示了图 4 所示建模过程中,模型帧的数量随时间变化的情况。本文所述方法的冗余帧回收机制使模型帧的数量保持在一个稳定的数量上,而不会随时间持续增长。同时由于目标识别机制,背景的变化并不会使新模型帧被加入,保证本文方法可以长时间高效运行。图中柱状图表示不同时刻共视特征模型中每个模型帧的共视模型帧的平均数量,可以保持在 5~10 frame 之间,说明共视特征模型可以提供足够的特征进行位姿计算。

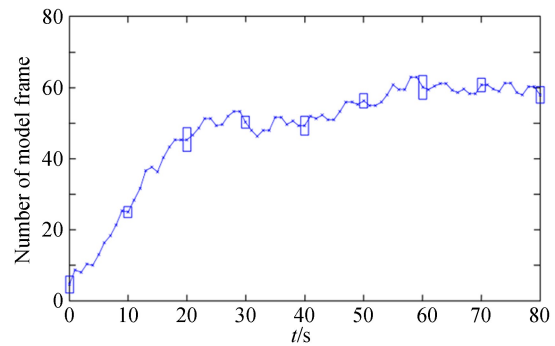


图 5 模型帧数量及其共视关系的变化

Fig. 5 Evolution of the number of model frames and covisibility

为了直观显示模型优化的效果,图 6 为目标三维网格模型在优化过程中的变化情况。可以发现,随着特征点不断优化,网格模型逐渐逼近真实目标轮廓,说明在建模过程中由于网格近似造成的特征恢复误差可以得到有效抑制。

图 7 显示了插入第 40 个模型帧后,模型优化前后特征模型恢复与特征重投影的结果(彩图见期刊电子版)。每幅图从左到右分别为特征模型与对应的真实目标外形线框模型、当前帧提取的特征点(绿色圆圈)和根据目标当前位姿得到的反投影特征点(红色交叉)及目标线框模型、特征重投影放大视图。图 7(a)是未进行模型优化时的结果,由于初始估计的位姿与恢复的目标共视特

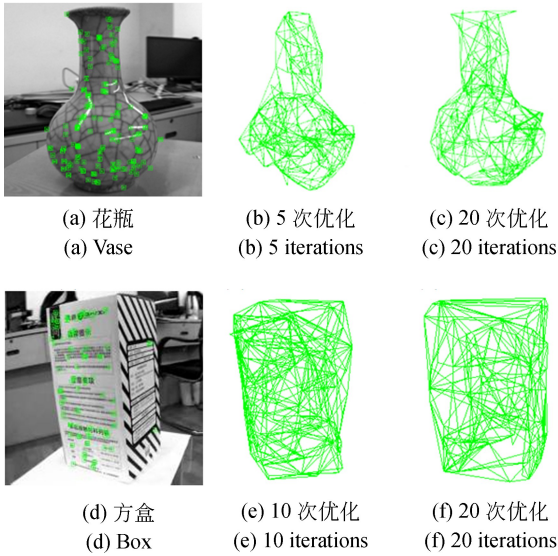


图 6 目标网格模型优化结果

Fig. 6 Optimization results of target mesh model

征模型存在误差,重投影的特征与线框模型都与实际值具有较大偏差。图 7(b)是模型经过无法向约束优化后的结果,通过共视特征点与位姿的优化,重投影误差得到明显改善,但是特征模型并没有收敛到目标真实轮廓。图 7(c)是经过本文所述法向约束优化后的结果,重投影误差得到改善的同时,共视特征模型也得到优化,特征点朝着目标真实表面轮廓调整。

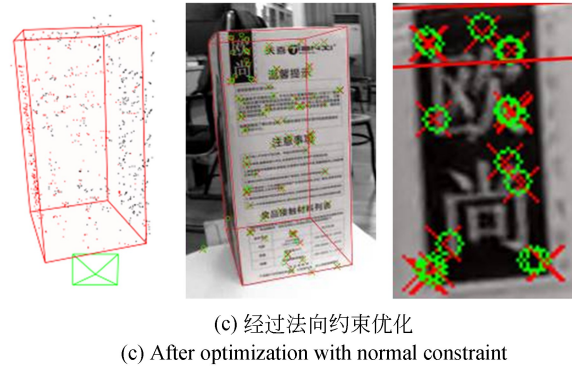
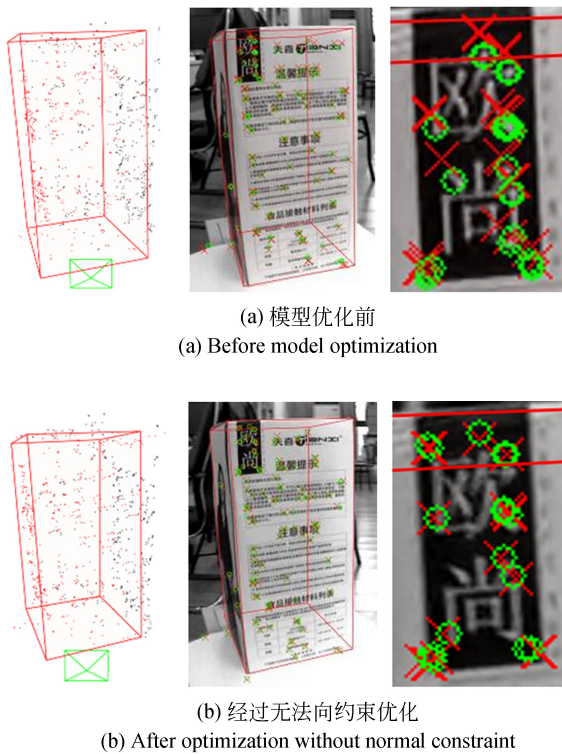


图 7 特征模型与重投影误差比较

Fig. 7 Comparison of feature model and reprojection error

为了定量评价模型优化对特征恢复精度的影响,求取不同时刻图 7(c)中方盒目标每个面 s_j (n_j, c_j) ($1 \leq j \leq 6$) 上恢复的三维特征点 $P_w^i \in s_j$ 的偏离误差的均方根值:

$$D_{RMS} = \sqrt{\sum_{j=1}^6 \sum_{i \in s_j} [n_j \cdot (P_w^i - c_j)]^2 / M}, \quad (20)$$

其中 n_j, c_j 分别为通过 $\{P_w^i\} \in s_j$ 计算得到 s_j 的法向量与中心点, M 为模型中所有特征点的数量。由于无法恢复模型的真实尺度,因此对各个时刻 D_{RMS} 值进行归一化处理,如图 8 中蓝色粗柱状线(彩图见期刊电子版)。同时,为了评价法向约束对模型恢复和位姿计算精度的影响,图 8 显示了不同时刻特征点重投影误差的均方根值:

$$E_{RMS}^k = \sqrt{\sum_i e_{i,k}^2 / N^k}, \quad (21)$$

其中 N^k 为 k 时刻图像帧中重投影特征的数量。图 8 中红色和紫色细柱状线分别对应无法向约束模型优化(图 8(b))和法向约束模型优化(图 8(c))的结果。当特征模型趋于稳定时,经过本文所述法向约束模型优化后的特征重投影误差 RMS 值小于 1.5 pixel,优于无法向约束模型优化的结果。从图 8 中可以发现,随着新模型帧加入以及伴随着的模型优化,三维特征模型的偏离程度 D_{RMS} 不断下降,共视特征模型对于目标外形估计精度逐步提高;同时特征重投影误差 E_{RMS}^k 随 D_{RMS} 变小,说明模型质量的提高可以有效增加位姿计算的精度。

为了比较不同方法位姿计算精度,将多面体目标固定于高精度 6 自由度云台上做空间运动,

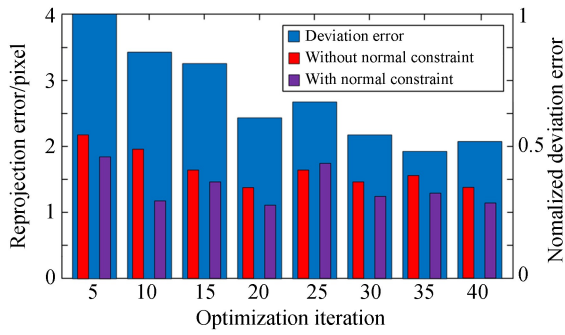


图 8 特征重投影误差与偏离误差随优化的变化

Fig. 8 Evolution of reprojection error and normalized deviation error w. r. t. optimization

利用固定于支架上的相机对目标进行连续拍摄,如图 11 所示。由于相机和目标之间的相对位姿真值未知,而云台提供的平移距离和旋转角度只能反映目标姿态的变化量,所以根据测量位姿的变化量来衡量位姿计算精度。

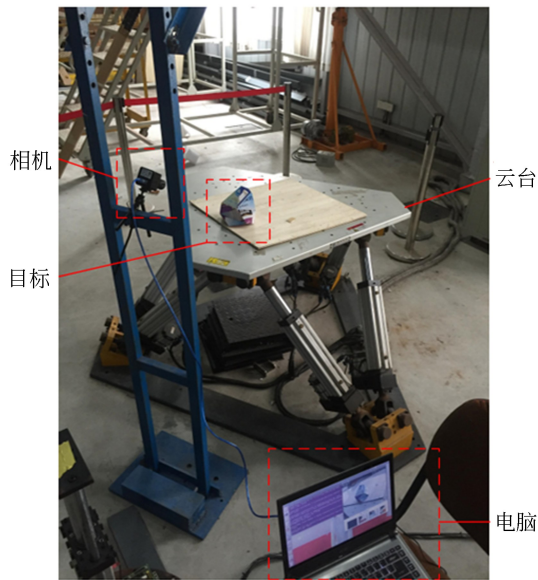


图 9 精度分析实验平台

Fig. 9 Experiment platform of precision analysis

图 10 显示了,目标做 20 次位姿变换,平移距离变化量和旋转变化量的绝对和相对测量误差。将目标平移量测量值与对应云台平移量相比较,利用该比值对测量值进行尺度恢复后与云台实际平移量进行比较,得到平移误差。而旋转误差根据式(22)计算^[16]:

$$\Delta R = \max_{k=1}^3 \{ \arccos(\mathbf{r}_{\text{true}}^k \cdot \mathbf{r}^k) \times 180/\pi \}, \quad (22)$$

其中 $\mathbf{r}_{\text{true}}^k$ 和 \mathbf{r}^k 分别为云台旋转矩阵和测量所得旋转矩阵的第 k 列。在位姿 9 处,目标发生了较大的旋转运动,扩展卡尔曼滤波(Extended Kalman Filter, EKF)法受到线性化误差和高斯分布的影响,特征或位姿估计误差较大,导致特征匹配失败,从而使得 EKF 方法发散,目标位姿跟踪失败;而粒子滤波(Particle Filter, PF)的多假设机制可以让位姿计算从失败中恢复,但状态分布的采样精度影响了位姿计算精度;基于关键帧的优化法(Bundle Adjustment, BA)由于同时考虑多个特征的新息,具有较强的稳定性和较高的位姿计算精度;而本文所述基于目标特征建模的位姿计算在具备关键帧优化法鲁棒性的同时,得益于特征模型法向约束优化,本文所述方法的精度总体优于其他 3 种方法,位姿测量的绝对误差小于 7.47 mm 和 2.63°,平均绝对误差(MAE)为 4.29 mm 和 1.54°,均方根误差(RMSE)为 4.50 mm 和 1.63°。

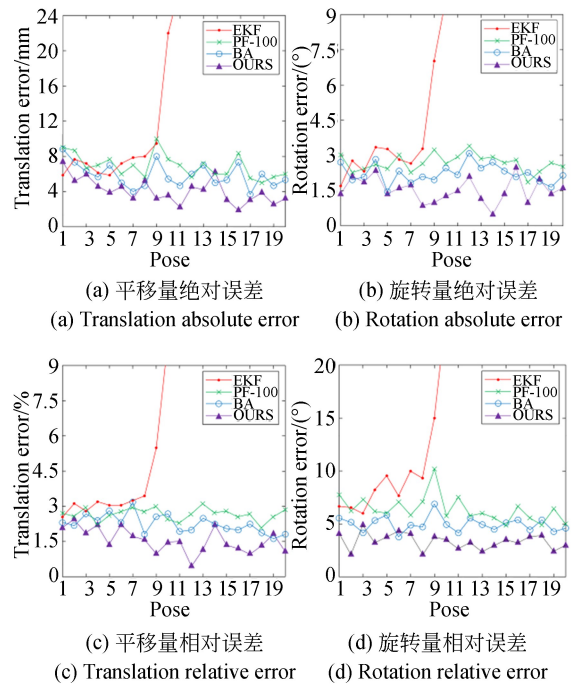


图 10 回环优化对于共视特征模型的影响

Fig. 10 Influence of loop detection on covisibility

为了分析算法的时效性,对程序中各个功能模块的运行时间进行了统计,如表 1 所示。由于建模过程中网格数量是逐渐增长的,实验中等待

模型帧数量稳定后再进行统计。从表中看出,位姿计算率维持在 8 Hz,4 个模块中最耗时的是网格构建,由于各模块采用多线程机制,因此其他 3 个模块对位姿计算的实时性影响较小。网格构建与模型优化过程中,当有新的模型帧插入时,局部优化会被终止,因此耗时变化较大。

表 1 位姿计算时效分析

Tab.1 Time efficiency analysis of pose calculation(ms)

功能模块	中值	均值	标准差
位姿计算	131.89	119.95	8.93
增量建模	29.52	31.04	11.91
网格构建	5 779.73	12 126.37	9 553.75
模型优化	2 614.51	2 980.40	791.28

参考文献:

- [1] LI J, BESADA J A, BERNARDOS A M. A novel system for object pose estimation using fused vision and inertial data[J]. *Information Fusion*, 2016, 33:15-28.
- [2] SHARMA S, D'AMICO S. Comparative assessment of techniques for initial pose estimation using monocular vision[J]. *Acta Astronautica*, 2016, 123:435-445.
- [3] WANG X, YU H, FENG D. Pose estimation in runway end safety area using geometry structure features[J]. *The Aeronautical Journal*, 2016, 120(1226): 675-691.
- [4] KLEIN G, MURRAY D. Parallel tracking and mapping for small AR workspaces[C]. *International Symposium on Mixed and Augmented Reality*, Nara, Japan, 2007: 225-234.
- [5] 周婧,高印寒,刘长英,等. 基于自适应算法的单目视觉系统的姿态解算[J]. *光学精密工程*, 2012, 20(12): 2796-2803.

5 结 论

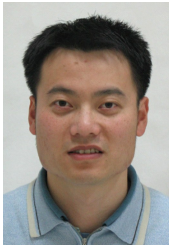
本文根据机器视觉和计算机图形学相关理论,实现了非合作纹理目标在线同步建模与位姿计算。从连续图像帧序列中选择具有良好特征共视关系的图像帧构建基于 ORB 特征的共视特征模型,实现目标的合作化。利用网格模型指导目标未知区域特征点的增量式恢复,避免背景特征干扰和目标整体建模。在共视模型的局部重投影优化中引入法向约束,提高特征模型的恢复质量,并对模型进行闭环尺度优化,减少累积误差的同时提高位姿计算精度。实验结果表明,本文所述方法能够对非合作纹理目标进行实时的特征建模与位姿计算,为基于单目视觉的三维感知与测量建模提供一种有效方法。

在今后的研究中,可以考虑引入基于像素深度匹配的光流法进行非纹理目标的位姿计算;并利用如直线等特征施加更加稳定的空间约束关系,进一步提高算法的精度与稳定性。

- 2012, 20(12): 2796-2803(in Chinese).
- [6] CRANDALL D J, OWENS A, SNAVELY N, *et al.*. SfM with MRFs: Discrete-continuous optimization for large-scale structure from motion[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(12): 2841-2853.
- [7] SEGAL S, CARMÍ A, GURFIL P. Vision-based relative state estimation of non-cooperative spacecraft under modeling uncertainty[C]. *Aerospace Conference*, MT, USA, 2011: 1-8.
- [8] MUR-ARTAL R, MONTIEL J M, TARDOS J D. ORB-SLAM: a Versatile and Accurate Monocular SLAM System[J]. *IEEE Transactions on Robotics*, 2015, 31(5): 1147-1163.
- [9] AUGENSTEIN S, ROCK S M. Simultaneous estimation of target pose and 3-D shape using the FastSLAM algorithm [C]. *Proceedings of the AIAA Guidance, Navigation and Control Conference*, Chicago, USA, 2009: 1-15.
- [10] MOURAGON E, DEKEYSER F, SAYD P, *et al.*. Real time localization and 3D reconstruction [C]. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, New York,

- USA, 2006: 363-370.
- [11] STRASDAT H, MONTIEL J M M, DAVISON A J. Visual SLAM: Why filter[J]. *Image and Vision Computing*, 2012, 30(2): 65-77.
- [12] 刘宗明,张宇,卢山,等. 非合作旋转目标闭环检测与位姿优化[J]. *光学精密工程*, 2017, 25(4): 504-511.
LIU Z M, ZHANG Y, LU SH, *et al.*. Closed-loop detection and pose optimization of non-cooperation rotating targets[J]. *Opt. Precision Eng.*, 2017, 25(4): 504-511(in Chinese).
- [13] RUBLEE E, RABAUDE V, KONOLIGE K, *et al.*. ORB: an efficient alternative to SIFT or SURF[C]. *IEEE International Conference on Computer Vision, Barcelona, Spain*, 2011: 2564-2571.
- [14] HARTLEY R, ZISSERMAN A. *Multiple View Geometry in Computer Vision, 2nd ed*[M]. UK: Cambridge University Press, 2004.
- [15] 冯肖维,姜晨,何敏,等. 三维距离图像基于特征估计的自适应平滑[J]. *光学精密工程*, 2019, 27(12): 2693-2701.
FENG X W, JIANG CH, HE M, *et al.*. Adaptive smoothing for three-dimensional range image based on feature estimation[J]. *Opt. Precision Eng.*, 2019, 27(12): 2693-2701(in Chinese).
- [16] LI S, XU C, XIE M. A robust $o(n)$ solution to the perspective-n-point problem[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2012, 34(7): 1444-1450.

作者简介:



冯肖维(1982—),男,上海人,博士,讲师,2004年、2007年、2011年于上海大学分别获得学士、硕士、博士学位,主要从事非线性系统控制、机器人技术、机器视觉等方面的研究。E-mail: xw-feng1982@163.com



谢安安(1995—),男,江西吉安人,硕士研究生,2017年于南昌航空大学获得学士学位,主要研究方向为机器视觉与图像处理等。E-mail: xiean1972@qq.com