

文章编号 1004-924X(2017)11-2958-10

基于旋转投影二进制描述符的空间目标位姿估计

邹 瑜¹, 梁 斌^{1*}, 王学谦², 张涛¹, 宋靖雁^{1,2}

(1. 清华大学 自动化系, 北京 100084;

2. 清华大学 深圳研究生院, 深圳 518055)

摘要: 为了实现基于点云的空间目标相对位姿快速估计, 提出一种旋转投影二进制描述符(BRoPH)。该描述符首先建立特征点处的局部参考坐标系, 然后通过旋转投影局部点云生成不同视角下的密度图像块和深度图像块, 最后根据图像块生成特征点的多尺度二进制字符串。针对位姿估计对实时性的要求, 在分析 BRoPH Hamming 距离分布的基础上, 提出了基于 Hamming 距离阈值的特征匹配策略, 用于剔除潜在的错误配对, 加快位姿估计收敛速度。最后, 在基于局部特征描述符位姿估计框架下分别与 SHOT 描述符和 FPFH 描述符进行了比较。结果表明: BRoPH 描述符在仅需要 SHOT 和 FPFH 平均内存 1/80 的基础上, 得到了远高于 SHOT 和 FPFH 的平均位姿估计精度, 其平均姿态误差小于 0.1° , 平均位置误差小于 $1/180 R$ 。此外, 基于 Hamming 距离阈值的特征匹配策略使得 BRoPH 的位姿粗估计速度加快了 7 倍, 总体位姿估计频率超过 7 Hz, 比 SHOT 和 FPFH 分别快 3~6.8 倍。该方法具有占用内存小、计算速度快、位姿估计精度高和抗干扰能力强等优点, 满足基于点云的空间目标位姿估计实时性要求。

关键词: 空间目标, 位姿估计, 点云配准, 旋转投影直方图, 二进制描述符

中图分类号: TP391.41; TP242.6 **文献标识码:** A **doi:** 10.3788/OPE.20172511.2958

Spacetarget pose estimation based on binary rotational projection histogram

ZOU Yu¹, LIANG Bin^{1*}, WANG Xue-qian², ZHANG Tao¹, SONG Jing-yan^{1,2}

(1. Department of Automation, Tsinghua University, Beijing 100084, China;

2. Graduate School at Shenzhen, Tsinghua University, Shenzhen 518055, China)

* Corresponding author, E-mail: bliang@tsinghua.edu.cn

Abstract: To estimate quickly the relative pose of a space target based on point cloud, a Binary Rotational Projection Histogram (BRoPH) feature descriptor was proposed. Firstly, a Local Reference Frame (LRF) for the feature point was established; Then, the density and depth images were generated under different views by rotationally projecting the local surface of feature point, Finally, the multi-scale binary string of the feature point was produced based on the images. For implementing the pose estimation of space target in real time, a Hamming distance threshold based feature matching strategy was proposed further to exclude false matching pairs to accelerate coarse pose estimation procedure. The comparison experiments were performed with SHOT descriptor and

收稿日期: 2017-04-21; 修订日期: 2017-06-20.

基金项目: 广东省自然科学基金资助项目 (No. 2014A030310318, No. 2015A030313881); 深圳基础研究资金资助项目 (No. JCYJ20140509172959962, No. JCYJ20160301153317415, No. JCYJ20160301100921349)

FPFH descriptor. The results demonstrate that BRoPH achieves an accurate pose estimation with only about 1/80 average memory cost of SHOT and FPFH descriptors. The average attitude error of BRoPH is under 0.1° , and its average translation error is less than 1/180R. Besides, the Hamming distance threshold based feature matching strategy speeds up the subsequent RANSAC by 7 times, and the overall pose estimation frequency exceeds 7 Hz, which is 3 to 6.8 times faster than those of SHOT and FPFH descriptors respectively. The proposed feature descriptor is compact and efficient, and the pose estimation method is accurate and robust for the requirements of space target pose estimation.

Key words: space target; pose estimation; point cloud register; rotational projection histogram; binary descriptor

1 引 言

作为航天器在轨临近操作的必要前提^[1],传统的位姿(位置与姿态)估计大多基于光学相机采集的二维图像,通过基于立体三维重建^[2]或是引入约束^[3]来估计目标航天器的位姿。然而,太空中光照条件恶劣,二维图像的采集容易受到光照影响,抗干扰能力有限。近年来,随着三维成像技术的发展及星载硬件能力的提高,激光雷达等深度传感器在航天任务中的应用越来越频繁^[4-5]。相比传统的二维图像,深度传感器采集的三维点云包含了更多的信息(距离)且测量结果具有不受尺度、旋转和光照等因素影响的优点,因此从点云提取的位姿抗干扰能力强且精度高,深度传感器也有望成为未来航天测量中的常备手段^[4]。

基于点云的位姿估计本质上是一个点云配准问题,即给定模型和场景点云,求解两者之间的相对变换关系,使得两者重叠部分精准地拼合在一起。点云配准通常先建立两个点云之间的对应关系,估计一个粗匹配,再通过迭代最近点(Iterative Closest Point, ICP)算法^[6]进行精调^[7]。粗匹配方法主要可以分为两类^[8]:一类是基于全局特征,将位姿估计转化为不同视角下的点云特征搜索问题^[9];另一类是基于局部特征,通过提取点云的关键特征并匹配来建立对应关系,进而估计粗位姿。前者需要离线训练及点云分割等预处理,并且对自遮挡、杂波等干扰敏感,而后者则对部分重叠等干扰鲁棒性较强,因此,本文主要考虑基于局部特征的位姿估计。

目前三维视觉领域已有许多局部特征描述符,如 SI^[10]、FPFH^[11]、SHOT^[12]、USC^[13]和

RoPS^[14]等。基于这些局部特征的位姿估计方法有 3 方面的局限^[15]:有的特征描述符对噪声敏感,如 SI 仅使用点云的法向量信息;有的则仅适用于特定三维模型,如 RoPS 只适于三维网格模型;已有局部特征描述符都是浮点类型的,大多存在计算慢、占用内存大的缺点,如 USC 的一个特征描述符就需要占用 7 840 bit,其受星载硬件水平的限制,难以在轨应用;现有浮点类型局部特征描述符基本上通过欧氏距离进行搜索匹配^[7],考虑到深度传感器采集的点云通常规模较大,当参与匹配的点云数量较大时,基于欧氏距离的搜索匹配过程将非常耗时^[16],尤其是特征维度高的特征描述符,如 USC。这些局限性限制了它们在空间目标位姿估计任务中的应用。

本文提出一种新的基于旋转投影的二进制点云局部特征描述符(Binary Rotational Projection Histogram, BRoPH)。该描述符通过将特征点周围的局部点云旋转投影到不同的坐标平面得到一系列不同视角下的密度图像块和深度图像块,再对二维图像块进行二进制描述,进而得到三维点云的二进制描述符,该方法具有计算快和占用内存小的优点。在二进制描述符的基础上,进一步提出基于 Hamming 阈值的特征匹配方法,剔除大量潜在的错误配对,从而加快位姿粗估计的收敛速度,达到空间在轨临近操作任务中位姿估计的实时性要求。

2 BRoPH 特征描述符

由于三维点云不像二维图像那样排列有序,也不一定包含类似图像中亮度那样的可以直接比较的信息,因此不容易选择合适的采样模式和采

样点对来生成点云的二进制字符串,这也是目前三维点云领域几乎没有二进制描述符的原因。

本文提出的 BRoPH 核心思想是将特征点周围的局部点云旋转投影到不同的坐标平面得到一系列不同视角下的二维图像块,从而将复杂的三维点云二进制描述转换为对有限个二维图像块的二进制描述。整个算法包含 3 个主要步骤:建立局部参考坐标系;生成二维图像块;二进制描述图像块。

2.1 局部参考坐标系

局部参考坐标系 (Local Reference Frame, LRF) 是一个原点在特征点处且包含了特征点邻域空间信息的坐标系,相当于二维图像领域特征描述符的主方向。一个好的局部参考坐标系可以使特征对刚体变换保持不变性。

给定点云中的一个特征点 p 及特征描述半径 r , 距离点 p 范围小于 r 的所有点构成了特征点 p 的局部表面点云 $Q = \{q_i | \|p - q_i\| \leq r, 1 \leq i \leq N\}$ 。局部表面 Q 内的所有点都用于散点分布矩阵 M 的计算。为了提高计算效率,在计算 M 的过程中省去局部表面中心点的计算,而用特征点 p 来代替,并且为每个邻域点 q_i 赋以线性距离权重^[12]。散点分布矩阵 M 计算为:

$$M = \frac{1}{\sum_{i: d_i \leq r} (r - d_i)} \sum_{i: d_i \leq r} (r - d_i) (p - q_i) (p - q_i)^T, \quad (1)$$

其中 $d_i = \|p - q_i\|$ 是邻域点 q_i 与特征点 p 之间的欧氏距离。

得到 M 后, LRF 的 z 轴通过如下方式确定:

$$z(p) = \begin{cases} n(p), & \text{if } n(p) \cdot \sum_{i=1}^N (pq_i) \geq 0 \\ -n(p), & \text{otherwise} \end{cases}, \quad (2)$$

其中: $n(p)$ 是散点分布矩阵 M 的最小特征值对应的特征向量, $pq_i = q_i - p$ 代表点 p 到点 q_i 的向量。

x 轴和 y 轴的确定类似文献[17], 即将局部表面 Q 内的所有点都投影到与 z 轴垂直的平面上, 并计算每个邻域点 q_i 在该平面上的投影向量:

$$v_i = pq_i - (pq_i \cdot z(p)) \cdot z(p). \quad (3)$$

LRF 的 x 轴定义为所有投影向量的加权平均:

$$x(p) = \frac{\sum_{i=1}^N \omega_{n1} \omega_{z2} v_i}{\left\| \sum_{i=1}^N \omega_{n1} \omega_{z2} v_i \right\|}, \quad (4)$$

其中: $\omega_{n1} = (r - \|p - q_i\|)^2$ 为点 q_i 到点 p 的几何距离权重, $\omega_{z2} = (pq_i \cdot z(p))^2$ 为点 q_i 到投影平面的投影距离权重。

最终, LRF 的 y 轴根据右手坐标系确定:

$$y(p) = z(p) \times x(p). \quad (5)$$

其中 \times 代表向量叉乘。

2.2 二维图像块生成

从三维点云生成二维图像, 最简单的方式是将点云直接投影到某个平面上。然而直接投影会导致深度信息的丢失, 尤其是存在遮挡的情况下点云的细节会丢失更多。为了增强算法的鲁棒性, 本文一方面引入文献[14]和[18]中的旋转投影机制, 实现对局部点云的多视角描述, 另一方面在投影过程同时生成局部表面的密度图和深度图, 充分保留点云中的信息。具体图像块生成过程如下。

2.2.1 旋转投影

在建立特征点的 LRF 后, 将特征点 p 的邻域点 Q 分别绕 LRF 的第 j 个轴旋转 N_R 次 θ_i 角, 然后将旋转后的点云 $Q_j^{\theta_i}$ 投影到 LRF 的第 k 个坐标平面。其中 $1 \leq i \leq N_R$ 代表旋转次数, θ_i 代表每次旋转的角度且 $0 \leq N_R \theta_i < 90^\circ$, $j \in \{x, y, z\}$ 代表 LRF 不同坐标轴, $k \in \{xy, yz, zx\}$ 代表 LRF 的不同坐标平面。如此总共得到 $N_R \times 3_{\text{axes}} \times 3_{\text{planes}}$ 个投影。

2.2.2 网格化

将上述投影分别离散化为 $L \times L$ 大小的网格 ($L \geq 3$ 为奇数), 并将离散化的网格记为 G_{ijk} :

$$G_{ijk} = \text{Grid}(\text{Proj}(Q, \theta_i, j, k)). \quad (6)$$

2.2.3 标记网格点云

记落入不同网格 $G_{ijk}(m, n)$ 内的旋转局部点云 $Q_j^{\theta_i}$ 的子集为:

$$Q_{ijk}(m, n) = \{q_i | (q_i \in Q_j^{\theta_i}) \cap (\text{Proj}(q_i) \in G_{ijk}(m, n))\}, \quad (7)$$

其中: $1 \leq m, n \leq L$ 。

2.2.4 生成密度图像块 D

定义 D 为旋转局部点云 $Q_j^{\theta_i}$ 投影落入不同网格的点云的数量矩阵:

$$D(m, n) = \text{Count}(Q_{ijk}(m, n)). \quad (8)$$

2.2.5 生成深度图像块 I

定义 I 的各像素值为离视点最近的两个点的深度的平均值:

$$I(m, n) = \begin{cases} (d_1 + d_2)/2, \mathbf{D}(m, n) \geq 2 \\ d_1, \mathbf{D}(m, n) = 1 \\ 2r, \mathbf{D}(m, n) = 0 \end{cases}, \quad (9)$$

其中: $\{d_i = \text{Depth}(\mathbf{q}_i), d_i \leq d_j, i \leq j \mid \mathbf{q}_i \in \mathbf{Q}_{ijk}(m, n)\}$ 为投影落入对应网格内的点云子集的有序深度集合, r 为特征描述半径。

这样总共可以得到 $N_R \times 3_{\text{axes}} \times 3_{\text{planes}}$ 个不同视角下的密度图像块 $\{\mathbf{D}\}_{N_R \times 3 \times 3}$ 及深度图像块 $\{I\}_{N_R \times 3 \times 3}$, 即总共 $N_R \times 3 \times 3 \times 2$ 个二维图像块。

2.3 图像块的二进制描述

生成图像块以后, 接下来就是生成二进制串。考虑特征点邻域范围内的点云一般分布比较接近, 即密度图像块 \mathbf{D} 和深度图像块 \mathbf{I} 的各像素值可能比较接近, 因此先将生成的 \mathbf{D} 和 \mathbf{I} 归一化到 $[0, 1]$ 范围, 再使用 CS-LBP^[19] 来描述图像块中心。原始 CS-LBP 是全局二进制描述符且通常是固定半径的:

$$\text{CS-LBP}(x, y)_{R, N, T} = \sum_{i=0}^{(N/2)-1} s(n_i - n_{i+(N/2)}) 2^i, \quad (10)$$

其中:

$$s(x) = \begin{cases} 1, x > T \\ 0, \text{otherwise} \end{cases},$$

其中: n_i 对应于半径为 R 的圆上等分为 N 的第 i 个像素值, T 为二值化阈值, 设为 0.01。通常设置 $N = 8$, 即 CS-LBP 可以得到位长为 4 的二进制串。

本文将 CS-LBP 视为局部相关特征, 仅用于生成图像块中心(对应三维点云特征点 \mathbf{p}) 处的二进制字符串。同时, 为了增强描述符的区分性, 本文引入多尺度机制, 即对 $L \times L$ 大小的图像块, 分别使用不同圆半径 R 来生成图像块中心的 LBP 二进制串, 然后再将不同 R 下的二进制串串联到一起, 得到特征点 \mathbf{p} 的最终二进制描述符:

$$\text{CS-LBP}_{\text{Multi-Scale}} = \sum_R (2^{4(R-1)} \cdot \text{CS-LBP}_{R, 8, 0.01}), \quad (11)$$

其中 $1 \leq R \leq (L-1)/2$ 。

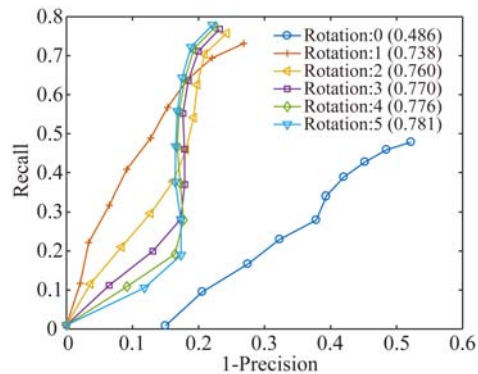
这样, 就可以得到特征点 \mathbf{p} 的 $N_R \times 3 \times 3 \times 2 \times 4 \times R_{\text{max}}$ 位二进制描述符。值得注意的是, 根据节 2.2 的二维图像块生成过程, 图像块的每一个像素值背后实际上都代表着一定数量的点云信息, 因此这种描述方式对噪声具有较强的鲁棒性。

2.4 参数分析

本节以斯坦福大学的 3D 模型库 (<http://graphics.stanford.edu/data/3Dscanrep/>) 中的 6 个模型点云作为模型, 对模型降采样 1/2 并加上均值为 0 标准差为 0.1 mr(Mesh Resolution, mr, 点云分辨率, 即各点离最近邻点距离的平均值) 的高斯噪声后得到的点云作为场景点云, 对不同参数下 BRoPH 的性能进行分析。实验过程参照文献[15]生成不同旋转次数 N_R 和不同图像块大小 L 下的 1-Precision vs. Recall 曲线 (PRC), 如无特殊说明, 默认图像块尺寸为 5, 旋转次数为 1。参数分析结果如图 1 所示, 图注括号中的数字代表对应曲线下的面积。

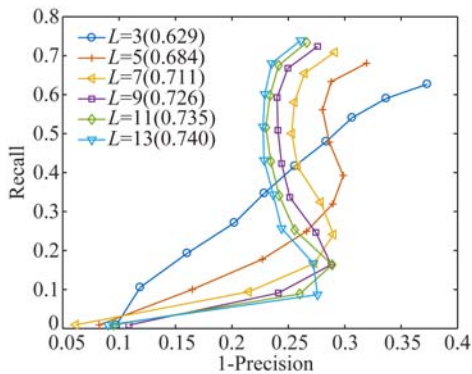
图 1(a) 为不同旋转次数下的 PRC 结果。可以看到, 随着旋转次数的增加, BRoPH 的性能逐渐提升, 尤其是从旋转 0 次到 1 次时, 性能得到大幅提升。这是由于旋转次数的增加可以带来更多不同视角的点云分布, 从而丰富 BRoPH 中的信息, 提升其区分性。旋转次数继续增加时, 由于大部分信息已经包含在 BRoPH 中, 因此, 进一步增加旋转次数所带来的性能提升不明显。图 1(b) 为不同图像块尺寸下的 PRC 结果。可以看到, 随着图像块尺寸的增加, BRoPH 的性能先提升再略微下降。这是由于在特征描述半径固定的前提下, 图像块尺寸越大, 图像块包含的点云细节信息越丰富, 相应地每个像素背后代表的点云也越少, 当图像块尺寸太大导致许多像素值非 0 即 1 时, 描述符的性能反而会恶化。

根据上述分析, 为了平衡算法的性能和效率, 在本文后续试验中, 设置 BRoPH 的旋转次数 N_R 为 1, 图像块尺寸大小 L 为 5。需要注意的是, 尽管旋转次数设为 1, 但视角数实际上达到了 9。



(a) 不同旋转次数

(a) Number of rotations



(b) 不同图像块尺寸
(b) Image patch size

图 1 不同参数下的性能分析

Fig 1 Performance under different parameters setting

3 基于点云局部特征的位姿估计

3.1 位姿估计流程

基于点云局部特征的位姿估计流程如图 2 所示。

3.1.1 点云获取

模型的三维点云可以通过设计的 CAD 或是在线重建^[20]得到,而场景三维点云则通过激光雷达或是其他深度传感器在线采集得到。模型的特征点检测和描述可以事先离线预处理。

3.1.2 特征点检测

特征点检测是特征描述与匹配的前提。考虑已有特征点检测方法存在检测慢或是特征数量偏少的局限性,本文采用均匀采样的结果作为特征点。

3.1.3 特征描述

特征描述是建立特征点对应关系的关键。在已有的点云局部特征描述符中,SHOT^[12]具有较

强的区分性、鲁棒性及计算效率^[15-16],而 FPFH^[11, 21]的紧凑型最强,本文将分别提取特征点处的 SHOT、FPFH 和 BRoPH 特征用于比较。

3.1.4 特征匹配

特征匹配直接影响位姿估计的结果。对于二进制特征描述符 BRoPH,采用暴力搜索结合交叉检验机制寻找 Hamming 距离最近的点作为配对;对浮点类型特征描述符 SHOT 和 FPFH,则采用 FLANN 搜索欧氏距离最近的点作为候选配对,而且只有最近点和次近点距离的比值(NNDR)小于一定阈值(本文设为 0.9)才认为是正确匹配。

3.1.5 错误匹配剔除

特征匹配结果中存在许多错误匹配,容易导致位姿粗估计误差太大。这里使用随机采样一致性算法(RANSAC)来剔除错误匹配。

3.1.6 位姿粗估计

位姿粗估计实际上和 RANSAC 过程同时进行。由于 3 个点即可确定一个位姿变换关系,因此在特征匹配结果中,随机选择 3 个点来计算位姿变换关系,根据设定的阈值判断其余点是内点还是外点,重复该过程直到达到最大迭代次数或设定阈值。最终选择最优的变换关系作为位姿粗估计的结果。

3.1.7 位姿精调

本文使用扩展 ICP (GICP, Generalized-ICP)^[22]来进行位姿精调。使用 RANSAC 得到的位姿粗估计结果作为 GICP 的初始位姿进行最近点迭代估计得到最终的位姿估计。

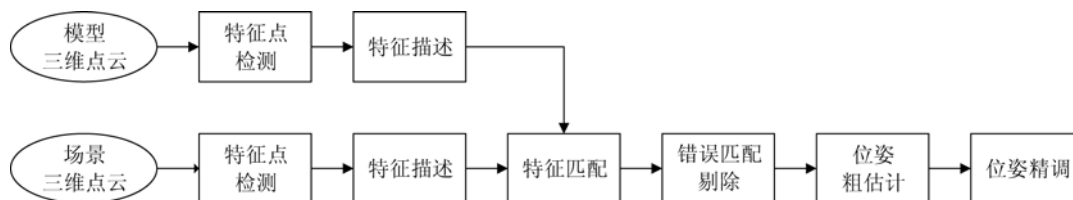


图 2 基于局部特征的位姿估计流程

Fig. 2 Flowchart of local feature descriptor based pose estimation

3.2 基于 Hamming 阈值的特征匹配策略

在特征匹配过程中,浮点类型描述符采用了 NNDR 技术来剔除潜在的错误配对。对于二进制特征描述符,其区分度一般比浮点类型描述符

低,这将导致特征匹配结果中包含更多的错误匹配。

根据 2.4 节的实验及已知的真实变换关系,可以得到特征匹配结果中的真实匹配点对(即配

对的点的物理距离小于一定阈值,本文设为 8 mr)和虚假匹配点对,统计其 Hamming 距离分布,结果如图 3 所示。

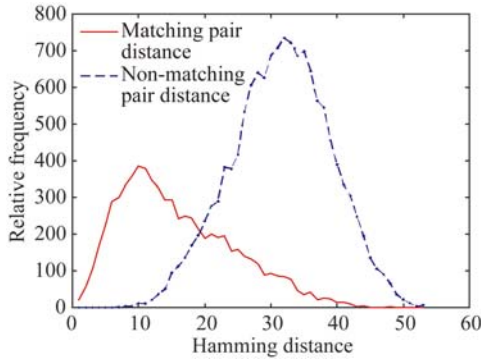


图 3 Hamming 距离分布

Fig. 3 Hamming distance distribution

可以看到,特征匹配结果中,真实匹配点对和虚假匹配点对的 Hamming 距离分布具有明显不同的峰值。这意味着可以在特征匹配过程中设置 Hamming 距离阈值,从而将部分虚假匹配剔除掉,并保留大多数真实匹配结果。这样带来的好处是特征匹配的点对数量大大减少,可以减少 RANSAC 的迭代次数并提高其收敛到正确位姿的概率。

4 实验与分析

4.1 实验条件设置

本文以小行星 Itokawa (<http://sbn.psi.edu/pds/resource/itokawashape.html>)和天宫一号为对象,采用文献[23]的技术来生成 100 组不同位置和姿态下的场景点云,其中小行星的观测距离范围为 [60, 1 050] m,平均观测距离为 405.7 m,天宫一号的观测距离范围为 [5, 60] m,平均观测距离为 20.5 m,场景序号越大表示平均观测距离越远。模型及部分场景点云如图 4 所示。为了增加对比,在生成的场景点云上添加均值为 0,标准差为 0.1 mr 的高斯噪声作为有噪声场景。

位姿估计过程中,SHOT 和 FPFH 均采用开源点云库(Point Cloud Library,PCL)^[24]中的默认参数,BRoPH 使用 2.4 节分析的参数,各特征描述符的基本情况总结如表 1 所示。其余参数中,特征描述半径为 15 mr,RANSAC 内点阈值为 10 mr,最大迭代次数为 50 000,GICP 的最大迭代次数为 100,迭代停止阈值为 10^{-6} 。所有实

验均在 Win7 64 位系统,开发环境为 VS2013,配置为英特尔 i7-4790@3.6GHz 8 核 CPU,16 GB 内存的台式机上进行。

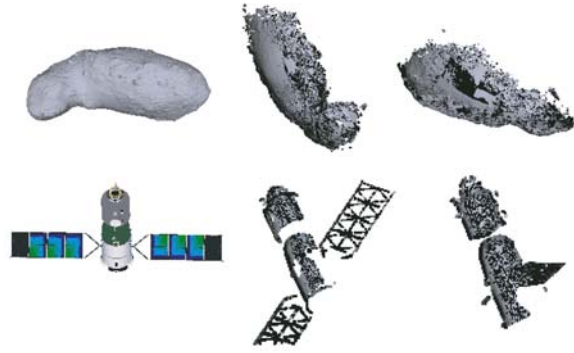


图 4 太空目标模型及部分场景点云图

Fig. 4 Model and scene examples of space target

表 1 参与实验的特征描述符

Tab. 1 Selected feature descriptors

特征描述符	维数	内存/bit	类型
SHOT ^[12]	352	1 408	浮点
FPFH ^[11]	33	132	浮点
BRoPH	144	18	二进制

4.2 错误匹配剔除策略分析

本节先分析 BRoPH 特征描述符的 Hamming 阈值特征匹配策略。实验以小行星 Itokawa 有噪声场景点云作为对象,在位姿估计的特征匹配环节中使用不同的 Hamming 距离阈值来剔除潜在的错误配对,并统计特征匹配得到的点对数,后续 RANSAC 耗时,以及最终位姿估计误差,结果如表 2 所示。

表 2 Hamming 距离阈值特征匹配

Tab. 2 Hamming distance based feature matching

Hamming 距离阈值	12	18	24	—
匹配点对	37	367	881	1078
T_{RANSAC}/ms	4.1	1.7	14.3	28.1
$\Delta\theta_x/(\circ)$	0.144 4	0.001 2	0.000 9	0.000 9
$\Delta\theta_y/(\circ)$	0.035 3	0.001 2	0.001 1	0.001 0
$\Delta\theta_z/(\circ)$	0.047 3	0.001 6	0.001 3	0.001 3
$\Delta T_x/m$	36.364 4	0.201 4	0.171 7	0.170 7
$\Delta T_y/m$	29.724 9	0.416 8	0.337 4	0.339 7
$\Delta T_z/m$	21.397 0	0.530 5	0.434 8	0.426 9

从表 2 中可以看到,随着 Hamming 距离阈值的变小,特征匹配得到的点对数量也逐渐减少,并且减少速度越来越快。相应地,位姿估计的误差也变得越来越大。

值得注意的是,随着特征匹配点对数量的减少,RANSAC 过程的时间先减少再增加。这是因为减少参与 RANSAC 的匹配点对会减少收敛所需的迭代次数,但是匹配点对太少导致无法收敛到给定阈值时,RANSAC 会一直迭代直到达到给定的最大迭代次数,因此时间反而变长了。为了平衡位姿估计精度及速度,后续实验设置 Hamming 距离阈值为 20。

4.3 位姿估计精度分析

根据 3.1 节的位姿估计框架,采用不同特征描述符对 Itokawa 和天宫一号的位姿进行估计,并计算位姿估计误差。直观起见,本文将位姿估计结果表达为 ZYX 顺序的欧拉角和三轴平移向量。

为方便阅读,这里仅显示 Itokawa 无噪声实验组位姿估计 Z 轴的平移误差,结果如图 5 所示。

从图 5 中可以看到,BRoPH 的平均误差最小,基本在 0 附近波动。随着场景序号增加(即平均观测距离增加),BRoPH 的位姿估计误差略有增加。SHOT 在大部分场景下的误差很小,但在少部分场景下的误差非常大,主要原因在于这些场景的点云较稀疏,特征匹配结果可靠性低,从而导致位姿粗估计错误,如场景 70(见图 5 标注),SHOT 特征匹配结果后只有 12 个特征点对,经过 RANSAC 后仅有 3 个特征点对,这导致了其平均位姿估计误差较大(见表 3)。FPFH 的位姿估计结

果非常不稳定,主要原因在于其区分度低,其结果呈现出随平均观测距离增加误差也增加的趋势。

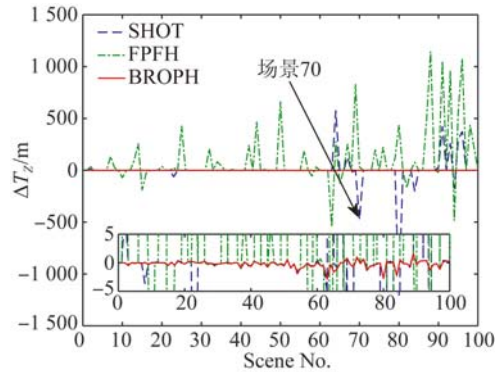


图 5 Itokawa 无噪声下 Z 轴平移误差

Fig. 5 Translation error of Z-axis of Itokawa without noise

进一步将两个空间目标在不同条件下的位姿估计平均误差统计为表 3。从表中可以看到,在所有实验条件下,BRoPH 的平均位姿估计精度都远高于 SHOT 和 FPFH。最重要的原因在于,SHOT 和 FPFH 在描述特征时都使用了点云的法向量信息,而法向量对噪声非常敏感,因而 SHOT 和 FPFH 的抗干扰能力都很差,这从没有噪声和加了噪声的结果对比可以得到验证。相反,BRoPH 在描述特征过程中使用的密度和深度信息都只和点云本身有关,因此对噪声的鲁棒性更强。SHOT 的平均位姿估计误差比 FPFH 小,主要原因是 SHOT 比 FPFH 的区分度更高,这和文献[15]的结果一致。

总体上,在各实验组下,基于 BRoPH 的位姿估计姿态误差小于 0.1° ,位置误差小于 $1/180R$ 。

表 3 位姿估计误差

Tab. 3 Error of pose estimation

模型	噪声水平/mr	描述符	$\Delta\theta_x/(\circ)$	$\Delta\theta_y/(\circ)$	$\Delta\theta_z/(\circ)$	$\Delta T_x/m$	$\Delta T_y/m$	$\Delta T_z/m$
小行星	0	SHOT	0.356 2	0.168 0	0.252 7	45.066 8	37.545 7	41.873 6
		FPFH	1.123 5	0.319 1	0.994 3	191.156 3	103.542 5	130.366 8
		BRoPH	0.000 9	0.001 1	0.001 3	0.155 5	0.334 2	0.430 0
	0.1	SHOT	2.651 3	0.740 4	1.304 4	264.293 9	310.769 1	198.553 8
		FPFH	2.303 4	0.832 4	1.693 3	303.809 9	268.084 6	244.918 1
		BRoPH	0.000 9	0.001 0	0.001 2	0.146 6	0.335 7	0.427 9
天宫一号	0	SHOT	0.476 9	0.133 0	0.259 2	1.686 9	3.144 5	2.109 6
		FPFH	2.502 1	0.791 8	1.369 8	11.380 5	17.279 2	12.553 6
		BRoPH	0.000 2	0.000 4	0.000 3	0.004 1	0.005 0	0.007 7
	0.1	SHOT	2.123 3	0.587 8	1.764 3	13.451 7	12.390 4	11.456 2
		FPFH	2.578 4	0.662 2	1.912 6	14.943 1	16.608 1	10.969 1
		BRoPH	0.003 8	0.003 2	0.006 1	0.004 5	0.032 8	0.093 5

4.4 位姿估计效率分析

本节仍然以小行星 Itokawa 有噪声实验组为对象,统计采用不同特征描述符时位姿估计各环节所用时间。该实验组中,Itokawa 模型点云包含 589 234 个点,场景点云平均包含 33 479 个点,特征点检测平均得到 1 819 个点,特征匹配得到的平均点对数及 RANSAC 后剩余点对(内点数)如表 4 所示。位姿估计各环节耗时统计结果如表 5 所示。

表 4 匹配点对数

Tab. 4 Number of matching pairs

	SHOT	FPFH	BRoPH
特征匹配点对	196	96	560
RANSAC 剩余点对	11	6	137

表 5 位姿估计各环节耗时

Tab. 5 Time for different procedure (ms)

	SHOT	FPFH	BRoPH
法向量估计	29.38	29.38	0.00
特征检测	13.73	13.73	13.73
特征描述	75.92	677.81	55.43
特征匹配	148.75	43.86	35.93
RANSAC	93.36	125.82	4.05
GICP	69.89	74.29	32.03
合计	431.04	964.89	141.17

从表 5 可以看到,总体而言,BRoPH 平均用时最少,每个场景点云位姿估计耗时约 141 ms,其次是 SHOT,用时约 431 ms,FPFH 最慢,需要约 965 ms。首先,SHOT 和 FPFH 需要估计点云

的法向量,而 BRoPH 不需要。特征检测过程三者一样。特征描述过程,FPFH 耗时最长,用了近 677 ms,而 SHOT 和 BRoPH 耗时相当。特征匹配过程中,尽管 BRoPH 用的是暴力搜索,但因为采用 Hamming 距离,匹配速度是最快的,而 FPFH 因为特征维度较低,且采用 FLANN 加速搜索,匹配过程也很快,SHOT 则因为特征维度高又用欧式距离匹配,即便采用了 FLANN 加速搜索也依然很慢。RANSAC 过程 SHOT 和 FPFH 耗时都比 BRoPH 明显要长,主要因为在有噪声情形下,SHOT 和 FPFH 的特征匹配结果较差,RANSAC 收敛速度慢,而 BRoPH 对噪声不敏感,且基于 Hamming 阈值的特征匹配减少了近一半的特征匹配点对,相比不使用 Hamming 阈值的特征匹配(见表 2),RANSAC 收敛速度提高了近 7 倍。至于 GICP,其收敛速度主要取决于位姿粗估计的精度,基于 SHOT、FPFH 和 BRoPH 的位姿估计在位姿精调环节的平均迭代次数分别为 11.3、13.1 和 3.2。总体上,基于 BRoPH 的位姿估计频率超过了 7 Hz,比 SHOT 快 3 倍,比 FPFH 快 6.8 倍。

4.5 位姿估计结果

为了进一步验证算法的有效性,本节将基于 BRoPH 描述子的位姿估计算法应用于小行星 Itokawa、天宫一号及斯坦福大学的 Bunny 模型的部分场景,得到模型与场景点云的配准结果如图 6 所示,其中左列为位姿粗估计结果,右列为位姿精调后的结果。可以看到,不论是有噪声的合成场景还是通过激光扫描仪得到的真实场景,模型与场景点云在位姿粗估计后的配准效果较好,这充分说明了基于 BRoPH 位姿估计算法的有效性。

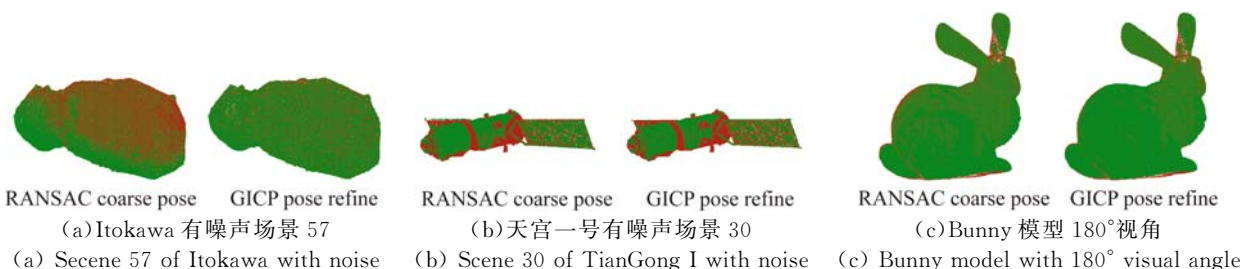


图 6 部分位姿估计结果,其中红色为模型点云,绿色为场景点云

Fig. 6 Pose estimation result, red indicates model clouds while green indicates scene clouds

5 结 论

本文针对空间目标位姿估计需求提出一种旋转投影二进制特征描述符 BRoPH, 并且在分析其性能的基础上, 提出一种基于 Hamming 距离阈值的特征匹配策略, 可以剔除大量错误配对, 加快位姿估计的速度。BRoPH 只占用 SHOT 和 FPFH 平均内存的 1/80 而达到了平均姿态误差

小于 0.1° , 平均位置误差小于 $1/180R$ 的精度。基于 Hamming 距离阈值的特征匹配策略使得 BRoPH 的位姿粗估计速度加快了 7 倍, 总体位姿估计频率超过 7Hz, 比 SHOT 和 FPFH 分别快 3~6.8 倍。本文提出的方法具有占用内存小、计算速度快、位姿估计精度高和抗干扰能力强等优点, 可以满足基于点云的空间目标位姿估计在轨实时性要求。

参考文献:

- [1] 梁斌, 杜晓东, 李成, 等. 空间机器人非合作航天器在轨服务研究进展[J]. 机器人, 2012, 34(2): 242-256.
LIANG B, DU X D, LI CH, *et al.*. Advances in space robot on-orbit servicing for non-cooperative spacecraft [J]. *Robot*, 2012, 34(2): 242-256. (in Chinese)
- [2] 徐文福, 刘宇, 梁斌, 等. 非合作航天器的相对位姿测量[J]. 光学精密工程, 2009, 17(7):1570-1581.
XU W F, LIU Y, LIANG B, *et al.*. Measurement of relative poses between two non-cooperative spacecrafts [J]. *Opt. Precision Eng.*, 2009, 17(7):1570-1581. (in Chinese)
- [3] 苗锡奎, 朱枫, 丁庆海, 等. 基于星箭对接环部件的飞行器单目视觉位姿测量方法[J]. 光学学报, 2013, 33(4):0412006.
MIAO X K, ZHU F, DIGN Q H, *et al.*. Monocular vision pose measurement based on docking ring component [J]. *Acta Optica Sinica*, 2013, 33(4): 0412006. (in Chinese)
- [4] 梁斌, 何英, 邹瑜, 等. ToF 相机在空间非合作目标近距离测量中的应用[J]. 宇航学报, 2016, 36(9):1080-1088.
LIANG B, HE Y, ZOU Y, *et al.*. Application of Time-of-Flight camera for relative measurement of non-cooperative target in close range [J]. *Journal of Astronautics*, 2016, 37(9):1080-1088. (in Chinese)
- [5] WOODS J O, CHRISTIAN J A. Lidar-based relative navigation with respect to non-cooperative objects [J]. *Acta Astronautica*, 2016, 126:298-311.
- [6] BESL P J, MCKAY N D. Method for registration of 3-D shapes [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 1992, 14(2): 239-256.
- [7] 王欣, 张明明, 于晓, 等. 应用改进迭代最近点方法的点云数据配准[J]. 光学精密工程, 2012, 20(9):2068-2077.
WANG X, ZHANG M M, YU X, *et al.*. Point cloud registration based on improved iterative closest point method [J]. *Opt. Precision Eng.*, 2012, 20(9):2068-2077. (in Chinese)
- [8] ALDOMA A, MARTON Z C, TOMBARI F, *et al.*. Tutorial: point cloud library: three-dimensional object recognition and 6 dof pose estimation [J]. *IEEE Robotics & Automation Magazine*, 2012, 19(3):80-91.
- [9] RHODES A, KIM E, CHRISTIAN J A, *et al.*. LIDAR-based relative navigation of non-cooperative objects using point cloud descriptors [C]. *AIAA/AAS Astrodynamics Specialist Conference*, AIAA, 2016.
- [10] JOHNSON A E, HEBERT M. Using spin images for efficient object recognition in cluttered 3D scenes [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1999, 21(5):433-449.
- [11] RUSU R B, BLODOW N, BEETZ M. Fast point feature histograms (FPFH) for 3D registration [C]. *IEEE International Conference on Robotics and Automation*, IEEE, 2009.
- [12] SALT I S, TOMBARI F, STEFANO L D. SHOT: Unique signatures of histograms for surface and texture description [J]. *Computer Vision and Image Understanding*, 2014, 125:251-264.
- [13] TOMBARI F, SALT I S, STEFANO L D. Unique shape context for 3D data description [C]. *Proceedings of the ACM Workshop on 3D Object Retrieval*, ACM, 2010.
- [14] GUO Y L, SOHEL F, BENNAMOUN M, *et al.*. Rotational projection statistics for 3D local surface description and object recognition [J]. *In-*

- International Journal of Computer Vision*, 2013, 105(1):63-86.
- [15] GUO Y L, BENNAMOUN M, SOHEL F, *et al.*. A comprehensive performance evaluation of 3D local feature descriptors [J]. *International Journal of Computer Vision*, 2016, 116(1):66-89.
- [16] GARCIA-GARCIA A, ORTS-ESCOLANO S, GARCIA-RODRIGUEZ J, *et al.*. Interactive 3D object recognition pipeline on mobile GPGPU computing platforms using low-cost RGB-D sensors [J]. *Journal of Real-Time Image Processing*, 2016:1-20.
- [17] YANG J Q, ZHANG Q, XIAO Y, *et al.*. TOLDI: An effective and robust approach for 3D local shape description [J]. *Pattern Recognition*, 2017, 65:175-187.
- [18] 郭裕兰, 万建伟, 鲁敏, 等. 激光雷达目标三维姿态估计[J]. *光学精密工程*, 2012, 20(4):843-850.
GUO Y L, WAN J W, LU M, *et al.*. Three dimensional orientation estimation for ladar target [J]. *Opt. Precision Eng.*, 2012, 20(4):843-850. (in Chinese)
- [19] HEIKKILÄ M, PIETIKÄINEN M, SCHMID C. Description of interest regions with center-symmetric local binary patterns [C]. *Computer Vision, Graphics and Image Processing, Indian Conference, ICVGIP*, 2006.
- [20] SCHNITZER F, JANSCHKE K, WILLICH G. Experimental results for image-based geometrical reconstruction for spacecraft rendezvous navigation with unknown and uncooperative target spacecraft [C]. *IEEE International Conference on Intelligent Robots and Systems*, IEEE, 2012, 5040-5045.
- [21] 沈荣华, 李卓嘉, 杨成, 等. 基于法向特征直方图的点云配准算法[J]. *光学精密工程*, 2015, 23(10Z):591-598.
SHENY H, LI ZH J, YANG CH, *et al.*. Point cloud registration with normal feature histogram [J]. *Opt. Precision Eng.*, 2015, 23(10Z):591-598. (in Chinese)
- [22] SEGAL A V, HAEHNEL D, THRUN S. Generalized-ICP [C]. *Proceedings of Robotics, Science and Systems*, 2009.
- [23] KATZ S, TAL A, BASRI R. Direct visibility of point sets [J]. *ACM Transactions on Graphics*, 2007, 26(3):1276377-1276407.
- [24] RUSU R B, COUSINS S. 3D is here: Point Cloud Library (PCL) [C]. *IEEE International Conference on Robotics and Automation*, IEEE, 2011.

作者简介:



邹 瑜(1988—),男,博士研究生,2010年于清华大学获得学士学位,主要从事空间机器人测量及计算机视觉方面的研究。E-mail: zouyu.yoyo@gmail.com



梁 斌(1968—),男,博士生导师,教授,1994年于清华大学获得博士学位,主要从事空间机器人控制、导航与制导以及视觉测量等方面的研究。E-mail: bliang@tsinghua.edu.cn