

面向 AR 辅助维修的 3D 物体识别技术

王维, 张丹, 胡曾一震, 左敦稳
(南京航空航天大学 机电学院, 江苏 南京 210016)

摘 要:3D 物体识别与位姿估计是 AR 辅助维修信息增强的基础,目前的 3D 物体识别技术在识别准确率、运行实时性和位姿估计准确度上仍然存在不足。以物体识别的实时性和位姿估计的准确性为要求,研究基于纹理点云的局部参考系估计技术和点云全局特征描述符生成与匹配技术,提出融合形状与纹理信息的全局正交对象描述符及其 3D 物体识别算法,基于蜗轮蜗杆减速器的拆装维修场景进行实例验证。验证结果表明,3D 物体识别算法在 AR 辅助维修场景中具备实际可行性和真实有效性。

关键词:AR 辅助维修;局部参考系;点云描述符;3D 物体识别;位姿估计

中图分类号:TP391.41 **文献标志码:**A **文章编号:**1671-5276(2020)03-0155-04

3D Object Recognition Technology for AR-assisted Maintenance

WANG Wei, ZHANG Dan, HU Zeng yizhen, ZUO Dunwen

(College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

Abstract:3D object recognition and pose estimation are the basis of AR-assisted maintenance information enhancement. At present, the recognition accuracy and real-time and pose estimation accuracy shortage also exists in 3D object recognition technology. To meet the requirements of the target recognition real-time and pose estimation accuracy of scene recognition, this paper researches on the local reference frame estimation and descriptor generation and matching based on global feature of point cloud and proposes a fusion shape and texture global orthographic object descriptor and the 3D object recognition algorithm. Then, the algorithm is verified based on the disassembly and maintenance of the worm gear reducer. The verification results show that the algorithm above has practical feasibility and real validity in the AR-assisted maintenance real scene.

Keywords:AR-assisted maintenance; local reference frame; point cloud descriptor; 3D object recognition; pose estimation

0 引言

随着机械设备的集成度和复杂度越来越高,其相应的维修和维护也变得越来越困难,传统维修指导方式显然无法满足更高的要求,而基于增强现实技术的 AR 辅助维修成为一种令人期待的新型维修指导方式^[1]。

如同大多数的 AR 应用一样,3D 物体识别和定位技术一直以来都是限制其发展的主要羁绊,较早的 AR 辅助维修大多是基于对二维标记或二维图像的识别来实现信息增强的,如赵敏等人^[2]采用静态标识和动态标识相结合的方法,实现场景识别和信息增强;崔波等人^[3]采用基于图像匹配的算法,进行目标识别和信息增强;兰佳等人^[4]采用基于改进形状描述符等特征的立体视觉识别算法,实现了弱纹理工件的识别和抓取。基于二维标记的识别不仅污染场景,而且与基于二维图像的识别一样无法感知目标物体的空间位姿,而目前解决这一问题的最佳方案之一是采用基于三维数据的物体识别技术,其中最为普遍的是基于点云的 3D 物体识别技术。

基于点云的 3D 物体识别不仅可以给出类别,最重要的是可以给出空间位姿,关键在于点云的配准。点云的配

准可分为基于点云全局特征型和基于点云局部特征型,AR 辅助维修对目标识别时间效率和位姿估计准确度的要求是首要的,因此该类研究基于前者。基于不同方法的识别差异主要在描述符上,其中的 ESF^[5]以及近年出现的 GASD^[6]和 GOOD^[7]是目前实时性和准确性较好的几种全局描述符,而这三者中表现最好的是 GOOD,即全局正交对象描述符,其描述符计算时间仅为 1 ms 左右。

本文在 GOOD 描述符的基础上进行深度改进,提出一种名为 C-GOOD 的全局描述符以及基于该描述符的 3D 物体识别与定位算法。主要工作有:1) 改进关键技术,形成新的描述符生成方法。包括引入带距离权值的协方差矩阵,改进原有的局部参考系估计方法,以及在原有的 GOOD 描述符上加入纹理特征,形成新的描述符;2) 搭建基于蜗轮蜗杆减速器的实例验证维修场景,验证本文算法的实际可行性和具体识别效果。

1 识别算法框架

本文所提识别算法的整体框架如图 1 所示,主要分为离线阶段和在线阶段。

基金项目:国家自然科学基金(51705246)

第一作者简介:王维(1993—),男,陕西铜川人,硕士研究生,研究方向为数字化制造、增强现实。

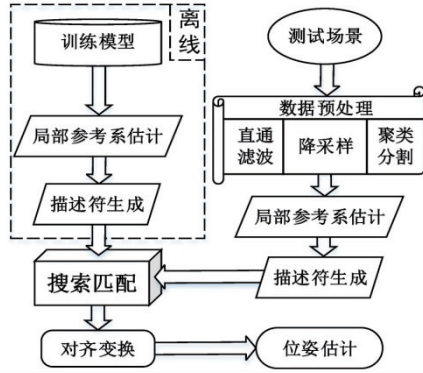


图1 3D物体识别算法整体框架

离线建库阶段,对所有训练对象点云模型的多个视角估计局部参考系并生成描述符,建立模板库。在线识别阶段,首先对测试场景进行预处理得到聚类目标,然后进行局部参考系估计和描述符生成;之后进行匹配搜索;匹配成功后将相应模型视角变换到场景目标即可获取场景目标初始位姿。

2 物体识别技术

2.1 局部参考系估计

在基于点云全局特征描述符的3D物体识别中,局部参考系是该类描述符具有唯一性、描述性和可重复性的基础,对于目标的位姿估计也不可或缺。常规的估计方法以主元分析法为基础,其关键在于构造归一化协方差矩阵:

$$M_c = \frac{1}{n} \sum_{i=1}^n (p_i - p_c) (p_i - p_c)^T \quad (1)$$

式中: p_i 为目标点云中的第*i*个三维点; n 为目标点云中的总点数; p_c 为目标点云质心。

由于局部参考系的估计依赖三维数据点,且本文所提算法的应用场景存在目标点云局部缺失、边缘突变、密度变化以及降采样等情况,因此本文对原有参考系估计方法加以改进,根据目标点云中三维数据点到目标点云质心的距离,设定一个与距离有关的贡献度权值,通过权值强化靠近质心三维数据点的作用、减弱远离质心三维数据点的作用。于是对式(1)进行改进即可得到新的协方差矩阵:

$$M_{CN} = \frac{1}{\sum_{i=1}^n (D - d_i)} \sum_{i=1}^n (D - d_i) (p_i - p_c) (p_i - p_c)^T \quad (2)$$

式中: p_i 、 i 、 n 和 p_c 与式(1)相同; D 为目标点云中距质心最远的三维点到质心的欧式距离; d_i 为目标点云中第*i*个三维数据点到质心的欧式距离。

对式(2)的协方差矩阵 M_{CN} 进行特征求解,可以得到一组特征向量 $[v_1, v_2, v_3]$,该向量即代表初始局部参考系,但此时的局部参考系具有方向二义性,本文通过引入三维点的空间分布以消除此方向二义性,其具体过程主要有:

1) 目标点云与初始局部参考系对齐

目标点云对齐时的变换矩阵由初始局部参考系定义,即

由向量 $[v_1, v_2, v_3]$ 定义,同时需要保证3个向量相互垂直。

2) 统计*x*和*y*坐标轴正负区域内的点数

分别统计对齐后的目标点云中位于*x*和*y*坐标轴正负区域的数据点个数。以*x*轴为例的具体统计规则为:

$$\begin{cases} N_x^+ = \{i: p_i \cdot x > \frac{d_b}{2}\} \\ N_x^- = \{i: p_i \cdot x < -\frac{d_b}{2}\} \end{cases} \quad (3)$$

式中: N_x^+ 和 N_x^- 分别表示对齐后的目标点云中位于*x*轴正负区域且在缓冲带之外的三维点总数; $p_i \cdot x$ 表示对齐后的目标点云中第*i*个三维点的*x*坐标值, d_b 为坐标轴附近的缓冲带宽度。

3) 确定*x*和*y*坐标轴方向标志

依据对齐后的目标点云在*x*和*y*坐标轴方向上的分布情况定义坐标轴方向标志。以*x*坐标轴为例的具体定义为:

$$S_x = \begin{cases} +1, & N_x^+ - N_x^- \geq N_T \\ -1, & \text{其他情况} \end{cases} \quad (4)$$

式中: N_x^+ 和 N_x^- 与式(3)相同,均为 ≥ 0 的自然数; S_x 表示*x*坐标轴的方向标志; N_T 为点数差异阈值。

4) 确定*z*坐标轴方向标志

对于*z*坐标轴的方向标志,只需要依据*x*和*y*坐标轴的方向标志即可确定,因为局部参考系的3个坐标轴之间有右手定则限制,所以*z*坐标轴的方向标志定义为:

$$S_z = \begin{cases} +1, & S_x \cdot S_y > 0 \\ -1, & \text{其他情况} \end{cases} \quad (5)$$

式中: S_x 、 S_y 和 S_z 分别表示局部参考系3个坐标轴的方向标志,三者取值均为+1或者-1。

根据初始局部参考系以及3个坐标轴的方向标志 S_x 、 S_y 和 S_z ,就可以确定唯一的最终局部参考系坐标轴的方向,消除歧义,获得唯一可重复的局部参考系。

2.2 描述符生成

针对全局正交对象描述符(GOOD)进行改进,通过添加纹理信息设计生成带纹理全局正交对象描述符(C-GOOD),其形成过程主要包含主体形状特征的生成、纹理信息的添加以及融合描述符的生成,具体分步细节为:

1) 局部参考系估计

此步骤在上一小节中已经说明。

2) 目标点云与最终局部参考系对齐

目标点云与局部参考系对齐时的变换矩阵由最终的唯一可重复局部参考系定义。

3) 确定3坐标轴方向投影描述符直方图

统计正交投影特征时,首先需要确定投影区域的大小和维度,假定对齐后的目标点云其最大绝对坐标值为 $C_{\max}$,则投影方格网阵的边长设置为 $L = 2C_{\max} + \delta$,而投影方格网阵维度*N*是本文描述符最为关键的参数,经过本文实验探索,其设置为10时可以达到识别效率和正确率的综合最佳,这里为了便于说明设置为5。活塞连杆在3坐标轴方向的方格网阵上投影效果如图2所示。

为了便于获取投影方格网阵相应区域的位置编号,在投影后需将投影方格网阵整体偏移至第1象限,2个方向偏移量均设置为投影方格网阵边长的一半。活塞连杆沿*z*轴方向

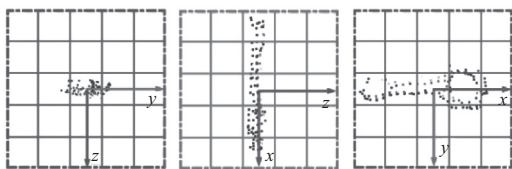


图2 活塞连杆点云沿3坐标轴方向上的投影

上的投影及其统计时方格网阵的偏移过程如图3所示。

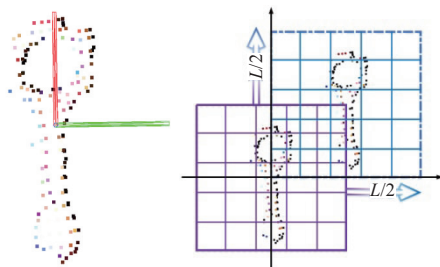


图3 活塞连杆沿z轴方向投影及投影方格网阵偏移

对于投影平面上的 N 维投影方格网阵,分别用 $row(\in \{0,1,2,\dots,n-1\})$ 和 $col(\in \{0,1,2,\dots,n-1\})$ 表示其区域方格所处的位置下标;对于投影点 p_i ,用 p_a 和 p_b 表示其在对应投影平面上的两个真实坐标,用 p_{row} 和 p_{col} 表示其将要投向的区域方格在方格网阵中的位置下标,那么根据投影方格网阵的边长 L 和维度 N ,即可计算出每一个投影点投向的区域方格位置横纵下标:

$$\begin{aligned} p_{row} &= \left\lfloor \frac{p_a + \frac{L}{2}}{\frac{L}{N}} \right\rfloor = \left\lfloor N \frac{p_a + \frac{L}{2}}{L} \right\rfloor \\ p_{col} &= \left\lfloor \frac{p_b + \frac{L}{2}}{\frac{L}{N}} \right\rfloor = \left\lfloor N \frac{p_b + \frac{L}{2}}{L} \right\rfloor \end{aligned} \quad (6)$$

将目标点云中的所有三维点都进行投影,并对每个区域方格进行点数统计,然后对投影矩阵进行归一化处理即可得到点云形状分布矩阵,将分布矩阵展开即为单个投影方向上的描述符直方图。对目标点云分别沿 x 、 y 、 z 3个坐标轴方向进行投影,即可得到3个对应于3个正交坐标轴方向的描述符直方图。

4) 形成主体形状描述符直方图

通过引入信息熵和方差对3)所得3个直方图进行排序拼接,消除排序歧义。首先以信息熵为第1排序标准,选出信息熵最大的排在首位;然后以方差为第2排序标准,在剩下的两直方图中选出方差较小的排在第2位;最后一个直方图排在末尾,由此便形成了一个具有 $3N^2$ 个区间的单一直方图,称为主体形状描述符直方图。

5) 添加纹理信息形成融合描述符

主体形状描述符只包含形状信息,不含纹理信息,因此其在物体识别上存在很大局限性。为了弥补不足,提高识别率和位姿估计准确度,本文在主体形状描述符基础上加入纹理信息,融合生成带纹理全局正交对象描述符。

带纹理全局正交对象描述符中纹理信息的添加,在主体形状描述符的投影点数统计时同步进行,具体来讲就是在统计主体形状描述符的投影点时,对相应点整合后的颜色信息进行统计,并在单个投影方格区域内,求取所有投影点整合后颜色信息的均值计入对应方格中,并在统计完成后与统计点数同步进行归一化处理。

对于每一个投影方向,纹理信息的加入方式和后续处理步骤都采用同样的方法,而对于纹理信息直方图的拼接只需与形状直方图保持同步即可。最后通过并联组合的方式将生成的整体纹理信息直方图和整体形状信息直方图合并,组合所得的形状纹理并行直方图即是最终的带纹理正交对象描述符(C-GOOD)直方图,其整体维度为 $2 \times 3N^2$ 。图4为活塞连杆的带纹理正交对象描述符直方图。

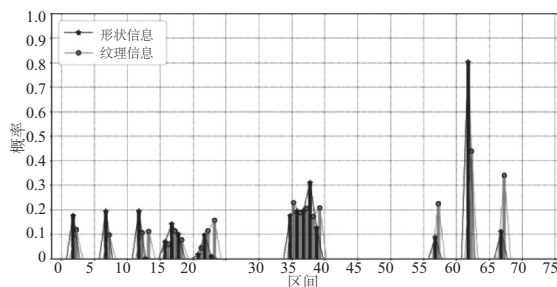


图4 带纹理正交对象描述符直方图

2.3 识别定位

本文所述物体识别的实质过程实为描述符的搜索匹配过程。在具体的匹配过程中,本文采用皮尔森相关系数来度量直方图描述符之间的相似性。当场景目标的点云描述符与训练对象点云描述匹配成功,则认为场景目标正确识别。

识别成功后,目标物的初始位姿估计具体过程为:

$$T_c = T_s^{-1} T_t = \begin{bmatrix} R_s & t_s \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} R_t & t_t \\ 0 & 1 \end{bmatrix} \quad (7)$$

式中: T_s 和 T_t 分别为场景目标和训练对象的局部参考系矩阵; T_c 为场景目标初始位姿。根据式(7)即可得到较为准确的初始位姿,但是若要得到更加精确的位姿,则可在牺牲一定实时性的基础上,使用ICP^[8]调优算法来实现。

3 场景实例验证

3.1 维修场景搭建

对于具体的维修场景,本文通过Kinect相机采集场景深度数据和彩色数据,并将深度图和彩色图进行融合,形成纹理点云对场景进行表达。本文以蜗轮蜗杆减速器的拆装维修为基础,搭建验证本文3D物体识别算法的维修场景,如图5(a)所示即为包含维修对象、采集设备以及处理设备的维修场景整体布置架构。

用于验证本文3D物体识别算法的维修场景以蜗轮蜗杆减速器为主要目标,通过对各个维修阶段相应场景中零部件的识别效果,来验证本文算法在实际维修场景中的表现。蜗轮蜗杆减速器的维修拆装场景如图5(b)所示,总共包括减速器底座、蜗轮、蜗轮轴、轴承端盖、套筒和减速器上盖6个零部件以及1个活动扳手。



图5 物体识别整体框架及维修场景纹理点云

3.2 实例识别验证

对于蜗轮蜗杆减速器的拆装维修,本文所提 3D 物体识别算法在具体识别的过程中,不仅可以给出目标的识别分类,同时会对目标的空间位姿进行估计。

对于零部件较多的某一维修场景,其识别过程中的两帧结果如图 6 所示。

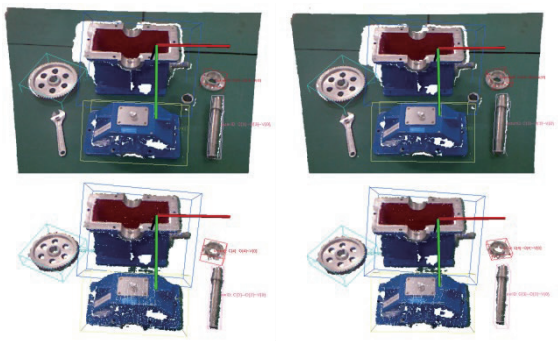


图6 维修场景的某两帧场景点云识别结果

在所有的点云数据帧中,5 个相对较大的目标物均能正确识别,只有相对较小的扳手和套筒因为实验中采集的数据精度太低导致无法成功识别。目标识别成功即可估计初始位姿,该维修场景下所有场景帧中,识别成功目标的平均初始位姿估计结果和所有场景帧的平均描述符计算匹配耗时如表 1 所示。

经过对蜗轮蜗杆减速器拆装维修场景的识别验证,可以明显看出本文所提算法在识别相对较大目标时效果显著,只是对于过小目标则相对较差,这主要是因为深度采集传感器的精度不足,无法获得较小目标的精确形状轮廓。同时在表 1 中不难发现,本文所提识别算法对单帧场景点云的识别耗时仅为 33.6 ms,基本达到 AR 辅助维修下实时识别目标物的要求,而且也给出了目标物的位置和姿态信息。

从总体来看,实例验证维修场景中的大多数目标都得到了成功识别,整体识别过程满足基本的实时性要求,并且给出了识别成功目标物的初始位姿。实例识别实验结果表明,本文所提 3D 物体识别算法在实际维修场景中具备实际可行性和真实有效性。

4 结语

针对 AR 辅助维修信息增强对实际维修场景中,目标物识别正确性、位姿估计准确性以及识别过程实时性的需

表1 场景帧的描述符平均计算匹配耗时和目标平均位姿

场景目标物	场景描述符平均计算匹配耗时/ms	位置坐标/m	姿态矩阵
减速器底座		$\begin{bmatrix} -0.001 & -0.119 & 0.844 \end{bmatrix}$	$\begin{bmatrix} -0.997 & -0.078 & -0.029 \\ -0.082 & 0.980 & 0.183 \\ 0.014 & 0.185 & -0.983 \end{bmatrix}$
减速器上盖		$\begin{bmatrix} -0.012 & 0.141 & 0.774 \end{bmatrix}$	$\begin{bmatrix} -0.999 & -0.026 & -0.029 \\ -0.028 & 0.997 & 0.078 \\ 0.027 & 0.078 & -0.997 \end{bmatrix}$
蜗轮	33.6	$\begin{bmatrix} -0.361 & -0.057 & 0.940 \end{bmatrix}$	$\begin{bmatrix} -0.910 & -0.414 & 0.020 \\ -0.362 & 0.771 & 0.524 \\ 0.201 & -0.485 & -0.851 \end{bmatrix}$
蜗轮轴		$\begin{bmatrix} 0.304 & 0.119 & 0.815 \end{bmatrix}$	$\begin{bmatrix} -0.197 & -0.920 & -0.339 \\ -0.861 & 0.003 & 0.509 \\ -0.469 & 0.392 & 0.791 \end{bmatrix}$
轴承端盖		$\begin{bmatrix} 0.331 & -0.069 & 0.950 \end{bmatrix}$	$\begin{bmatrix} -0.947 & -0.320 & -0.020 \\ -0.263 & 0.814 & -0.518 \\ 0.184 & -0.484 & -0.855 \end{bmatrix}$
套筒	失败,无数据		失败,无数据
扳手	失败,无数据		失败,无数据

求,本文深入分析研究了目前最先进的点云全局特征 3D 物体描述符、全局正交对象描述符,在此基础上改进并提出了附带纹理信息的新颖点云全局特征描述符及其 3D 物体识别算法。经过实验分析和蜗轮蜗杆减速器拆装维修场景的实例识别验证,证明本文所提算法在 AR 辅助维修信息增强的目标识别应用中具备实际可行性、真实有效性,进一步助推了 AR 辅助维修技术的发展与应用。

参考文献:

- [1] NOVAK Marcincin J, BARNA J, JANAK M, et al. Augmented reality aided manufacturing [J]. Procedia Computer Science, 2013, 25:23-31.
- [2] 赵敏,刘秉琦,武东升. 增强现实装配和维修系统的技术研究[J]. 光学仪器, 2012, 34(2):16-20.
- [3] 崔波,王崑,瞿珏,等. 基于增强现实的装备诱导维修系统设计与实现[J]. 火力与指挥控制, 2016, 41(11):176-181.
- [4] 兰佳,叶文华. 基于立体视觉的工件识别与定位系统研究[J]. 机械制造与自动化,2017,46(5):188-193.
- [5] WOHLKINGER W, VINCZE M. Ensemble of shape functions for 3D object classification [C]. IEEE International Conference on Robotics and Biomimetics, IEEE, 2012:2987-2992.
- [6] LIMA J P S D M, TEICHRIEB V. An efficient global point cloud descriptor for object recognition and pose estimation [C]. Graphics, Patterns and Images, IEEE, 2017:56-63.
- [7] KASAEI S H, TOMÉ A M, LOPES L S, et al. GOOD: a global orthographic object descriptor for 3D object recognition and manipulation [J]. Pattern Recognition Letters, 2016, 83:312-320.
- [8] BESL P J, MCKAY H D. A method for registration of 3D shapes. IEEE Trans Pattern Anal Mach Intell [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1992, 14(2):239-256.

收稿日期:2019-03-08