

基于增强型点对特征的三维目标识别方法

鲁荣荣^{1,2,3,4,5}, 朱枫^{1,2,4,5*}, 吴清潇^{1,2,4,5}, 陈佛计^{1,2,3,4,5}, 崔芸阁^{1,2,3,4,5}, 孔研自^{1,2,3,4,5}

¹中国科学院沈阳自动化研究所, 辽宁 沈阳 110016;

²中国科学院机器人与智能制造创新研究院, 辽宁 沈阳 110016;

³中国科学院大学, 北京 100049;

⁴中国科学院光电信息处理重点实验室, 辽宁 沈阳 110016;

⁵辽宁省图像理解与视觉计算重点实验室, 辽宁 沈阳 110016

摘要 针对基于原始点对特征的三维目标识别算法中存在的内存浪费、效率不高的问题, 提出了一种基于增强型点对特征的三维目标识别算法。通过在原始点对特征的第四个分量上乘以一个符号函数, 得到了一种区分性更强的点对的特征, 消除了原始点对特征存在的二义性。考虑到待识别目标三维模型存在的自遮挡, 利用点对之间的视点可见性约束, 剔除了目标三维模型哈希表中存在的大量冗余点对, 减小了内存开销并提高了三维目标识别算法的准确率和效率。在开放数据集以及实际采集的数据集上的实验结果表明, 与基于原始点对特征的算法相比, 本文提出的三维目标识别方法在识别准确率以及识别效率上都有一定程度的提升。

关键词 机器视觉; 点对特征; 三维目标识别; 可见性约束

中图分类号 TP391 **文献标识码** A

3D object recognition based on enhanced point pair features

Lu Rongrong^{1,2,3,4,5**}, Zhu Feng^{1,2,4,5*}, Wu Qingxiao^{1,2,4,5}, Chen Foji^{1,2,3,4,5},

Cui Yunge^{1,2,3,4,5}, Kong Yanzi^{1,2,3,4,5}

¹ *Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, Liaoning 110016, China;*

² *Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang, Liaoning 110016, China;*

³ *University of Chinese Academy of Sciences, Beijing 100049, China;*

⁴ *Key Laboratory of Opto-Electronic Information Process, Shenyang, Liaoning 110016, China;*

⁵ *The Key Laboratory of Image Understanding and Computer Vision, Shenyang, Liaoning 110016, China*

Abstract Aiming at the problem of memory waste and low efficiency in three-dimensional (3D) object recognition algorithm based on original point pair feature (PPF), a 3D object recognition algorithm based on enhanced point pair feature (EPPF) is proposed. By multiplying the fourth component of the original PPF with a sign function, a more

收稿日期: 年-月-日; **修回日期:** 年-月-日; **录用日期:** 年-月-日

基金项目: 国家自然科学基金(U1713216)、机器人学国家重点实验室自主课题(2017-Z21)

*E-mail: fzhu@sia.cn; **E-mail: lurongrong@sia.cn

distinguishing PPF is obtained, which eliminates the ambiguity of the original PPF. Considering the self-occlusion of the 3D model of the target to be identified, the large numbers of redundant point pairs existing in the target 3D model hash table are eliminated by means of the viewpoint visibility constraint between the point pair, which reduces the memory overhead and improves the accuracy and efficiency of the 3D object recognition algorithm. The experimental results on the open dataset and the actual collected dataset show that the proposed 3D object recognition algorithm in this paper has a certain degree of improvement in recognition accuracy and recognition efficiency.

Key words machine vision; point pair feature; 3D object recognition; visible constraint

OCIS codes 150.1135; 100.5010; 150.6910

1 引言

三维目标识别主要包括目标检测与位姿估计两个任务,即在采集的三维场景数据中检测是否存在目标模型,如果存在需要返回其在场景中的位置和姿态(位姿)。在工业背景下,待处理的金属零件一般表现为弱纹理、结构单一且表面易产生高光,这些因素都会给基于图像的二维目标识别算法带来挑战。其次,遮挡、场景杂乱等因素的干扰使得大部分的二维目标识别算法一般不易获取目标精确的三维位姿。与之相比,三维目标识别算法处理的对象一般为三维点云数据,其抗光照、视角变化的能力更强,鲁棒性和适应性更好。此外,随着近些年消费级三维传感器的兴起,大量学者从2D视觉研究逐渐转向3D视觉研究,大大加快了3D视觉领域的研究进展。针对三维目标识别这一视觉任务,目前主要有三类研究方法:基于视角投影图匹配的方法、基于点特征对应的方法以及基于投票的方法。下面分别对这三类方法进行简单概述。

基于视角投影图匹配的方法:其基本思路是利用一个虚拟相机离线渲染3D模型在不同视角下的2D视图,从而建立相机位姿与2D视图的对应关系。在线识别阶段,首先将这些2D模板逐一在场景图像中滑动匹配计算目标分布响应图。然后根据设定的阈值剔除弱响应位置,并采用极大值抑制的策略获取目标最终的位置。最后根据当前采用的模板得出目标的三维位姿。这类方法实现简单,扩展性好。对于新增的模型,只需添加新的2D模板即可。但是,为了提高位姿估计的精度,通常需要为每种模型渲染上千幅2D视图模板,这无疑会增加计算开销,降低识别效率。Ulrich等人提出基于尺度空间的搜索策略^[1],加快了识别的速度。Hinterstoisser等人采用梯度图建立2D模板,并利用线性化内存的方式加快模板匹配的速度^[2],但是随着模型数量增长,算法的计算成本仍会线性递增。潘旺等人利用多传感器融合技术,克服了模板视图退化所引起的误匹配问题,提高了识别定位算法的鲁棒性和精

度^[3]，但是系统的硬件成本也随之升高。

基于点特征对应的方法：这类方法通过建立模型与场景的点对应关系进行目标识别与定位。主要包含离线模型特征库建立与在线特征匹配识别两个阶段。离线阶段，提取每个三维模型表面的特征点并对其进行局部特征描述。在线阶段，利用同样的策略提取场景的特征点并计算相应的局部特征。然后通过场景特征与模型特征的近邻匹配建立若干点对应关系。最后通过对应性分组以及位姿假设验证等手段完成目标的识别与定位。该类方法依赖于建立的点对应关系的好坏，适用于几何特征丰富的模型的识别。由于采用局部特征匹配的策略，使其对遮挡以及场景杂乱具有较好的抵抗力。研究的热点主要集中在三维特征点的提取，三维特征描述以及对应性剔除等环节上。读者可从相关研究的综述性文章^[4-6]中获取更多细节信息。

基于投票的方法：主要包含基于随机采样一致性^[7]以及点对特征投票^[8-10]的三维目标识别方法。基于随机采样一致性的方法用于从场景中提取具有显示几何模型的物体，如平面结构、圆柱以及球体等，本质上是一种模型拟合的方法。通过预先设定的几何模型，利用场景点对该模型进行投票，从而提取相应的几何构型。而基于点对特征投票的方法利用哈希索引的方式建立一系列弱的点对对应关系，并基于这些点对对应生成众多的候选位姿变换。然后根据位姿的得票数高低选取若干可靠的候选位姿。最后对候选位姿集进行聚类获取目标的识别结果。点对特征计算简单，灵活性好，不仅适用于几何特征丰富的物体，对特征较少、结构单一的工业零部件的识别也具有很好的性能。本文针对原始点对特征存在的二义性以及构建模型点对特征哈希表中出现点对冗余的问题，提出了一种在视点可见性约束的前提下，基于增强型点对特征的三维目标识别方法。提高了点对特征的区别性并节省了内存开销，改善了原始算法的识别性能。

2 基本原理

2.1 增强型点对特征

点对特征是由一对有方向的点之间所包含的四种简单几何属性构成的四元数组。以图 1 中所示的点对 (p, q) 为例，其点对特征为

$$F(p, q) = (\|d\|_2, \alpha, \beta, \theta) = (\|d\|_2, \angle(n_1, n_3), \angle(n_2, n_3), \angle(n_1, n_2)), \quad (1)$$

其中 n_1, n_2 分别是 p, q 对应的法向量， $d = q - p$ ， $n_3 = d / \|d\|_2$ 。由此定义可知，点对特征不是

对称的，即通常 $F(p, q) \neq F(q, p)$ 。如果给定 (p, q) 的点对特征 $F(p, q)$ 以及点 p 的三维坐标和法向量 n_1 ，根据点对特征第一和第二分量的约束，可以确定点 q 位于空间中的某个圆周上，该圆周上的点与点 p 的距离为 $\|d\|_2$ ，且点 p 与该圆周上任意一点的连线与法向量 n_1 的夹角为 α 。在该圆周上任取一点 q ，根据点对特征第三和第四分量的约束，可以得到两条满足约束条件的法向量 n_2, n'_2 ，即 $\angle(n_2, n_3) = \angle(n'_2, n_3) = \beta, \angle(n_1, n_2) = \angle(n_1, n'_2) = \theta$ ，它们关于由 n_1, n_3 确定的平面 Π 对称。从这一分析可以看出，原始的点对特征无法区分点 q 的法向量究竟是 n_2 还是 n'_2 ，将其称为原始点对特征的二义性问题。事实上，三维点的法向量刻画了该点邻域曲面的某种几何属性。该属性对于区分不同位置的点具有一定的参考意义。鉴于此，提出一种增强型的点对特征，试图保留这一区分性。同样以点对 (p, q) 为例，其增强型点对特征为

$$F_e(p, q) = (\|d\|_2, \angle(n_1, n_3), \angle(n_2, n_3), \delta(n \cdot n_2) \angle(n_1, n_2)), \quad (2)$$

其中 $n = n_3 \times n_1$ ，表示平面 Π 的法向量， $\delta(x) = 1$ ，当且仅当 $x \geq 0$ ，否则 $\delta(x) = -1$ 。基于该设定，点 q 的法向量取 n_2 或者 n'_2 时将对应不同的点对特征，由此便可消除上述二义性问题，提高点对特征的区分性。

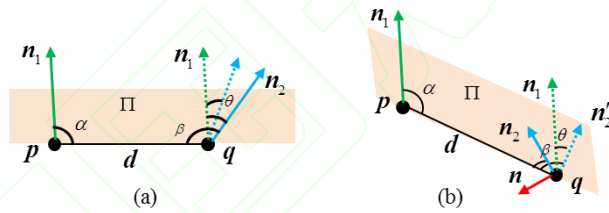


图1 原始点对特征二义性示意图。(a) 正视图；(b) 侧视图

Fig. 1 An illustration of the ambiguity of the original point pair feature. (a) front view; (b) side view

2.2 视点可见性约束

所谓视点可见性约束是指一对点能否在某个视角下被同时观测到。如图 2 (a) 所示，假设长方体不透明，上表面的红色点和绿色点总能被同时观测到，而下表面的蓝色点则无法与红色点被同时观测到，就称红色点和绿色点满足视点可见性约束，红色点与蓝色点不满足可见性约束。本文所提及的视点可见性问题主要是由物体的自遮挡产生的。

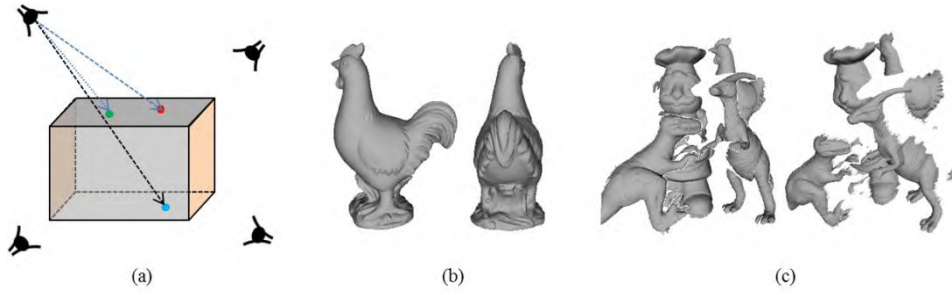


图 2 (a) 视点可见性示意图; (b) 三维模型; (c) 2.5 维场景

Fig. 2 (a) An illustration of the visible constraint between two points; (b) 3D model; (c) 2.5D scene

在基于点对特征的三维目标识别算法中，需要离线建立三维模型的点对特征哈希表。原始的目标识别算法并没有考虑视点可见性约束，而是直接计算所有可能的点对特征，并根据其对应的哈希值将相应的点对映射到模型哈希表中。由于扫描得到的场景点云通常是单一视角下的 2.5D 数据（如图 2 (c) 所示），即场景中的任意点对之间自动满足视点可见性约束。而待识别的目标往往具有完整的 3D 模型（如图 2 (b) 所示），因此模型哈希表中会出现许多不满足视点可见性约束的点对，这些点对对于三维目标识别算法并没有实质上的贡献，且会增加系统的内存开销，降低算法的运行效率。

针对这一问题，提出视点可见性强弱值的概念。给定一对有向点 (p, n_1) 与 (q, n_2) ，二者之间的可见性强弱值 $\rho(p, q) = n_1 \cdot n_2$ ，即二者法向量一致性越好，可见性概率越高（同时被观测的几率越大）。相反，当二者法向量恰好反向时，则完全不可见。通过预先设定的最小可见性强弱值 τ ，可以从模型中找出与每个点互为可见的点，即 $\rho(p, q) \geq \tau$ ，则称点对 (p, q) 互为可见，否则不可见。以图 2 (b) 中的公鸡模型为例，选取模型上的一点 p （图 3 (a) 中红色标记点），统计 τ 取不同值时，与 p 互为可见的点的数量。为了方便观察，令 $\tau = \cos(\theta)$ 。图 3 (b) 显示了与 p 互为可见的点的数量随着角度 θ 变化的曲线图，并给出了 θ 取 30 的倍数时与 p 互为可见点的分布情况（绿色代表可见，灰色代表不可见）。从图中可以看出，当 $\theta = 90^\circ$ ，即 $\tau = 0$ 时，与其互为可见的点约为总数的 1/2。这说明通过设定合理的可见性阈值，可以有效剔除模型哈希表中存在的大量冗余点对，节省内存开销。

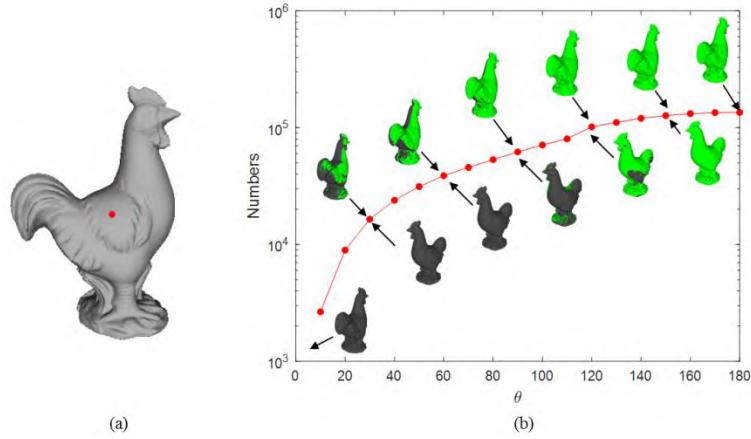


图3 视点可见性约束。(a) 点 p 的位置；(b) 与 p 互为可见点的数量及分布随着 θ 变化的示意图
 Fig. 3 The viewpoint visibility constraint. (a) Location of the point p ; (b) The changes of the distribution and number of the points satisfying the visible constraint with point p with the increase of angle θ

2.3 三维目标识别定位算法

基于增强型点对特征的三维目标识别算法共分为两个阶段：离线训练阶段与在线识别阶段。其中，离线阶段通过将目标三维模型中所有满足视点可见性约束的点对根据其增强型点对特征所生成的哈希值存于哈希表中从而完成对目标模型的训练。在线识别阶段主要分为四个步骤：a) 种子点的选取；b) 位姿投票；c) 假设生成以及 d) 位姿聚类与验证。整个识别定位算法的流程如图4所示，下面对算法的具体实现细节进行一一介绍。

离线训练阶段：由于增强型点对特征的计算利用了模型点的法向量信息，因此需要预先估计目标模型中每个点的法向量。对于目标模型中的任意点对 (p, q) ，如果其视点可见性强弱值 $\rho(p, q)$ 小于 τ （所提方法中将 τ 设置为 0），即不满足视点可见性约束，则不考率该点对；否则，根据公式 (2) 计算其增强型点对特征 $F_e(p, q)$ 。所提方法所采取的哈希策略是将 $F_e(p, q)$ 的 4 个分量分别以固定步长离散化为 4 个整数，不妨记为 N_1, N_2, N_3 以及 N_4 ，然后根据这 4 个整数值计算其相应的哈希值 $K(p, q) = N_1 N_{\text{ang}}^3 + N_2 N_{\text{ang}}^2 + N_3 N_{\text{ang}} + N_4$ ，其中 N_{ang} 为三个角度余弦分量的最大离散数，所提方法中设置为 15。当视点可见性阈值 τ 设置为 0 时，增强型点对特征的第四个分量的取值范围为 $[-\pi/2, \pi/2]$ ，为了保证离散化得到的整数值非负，通过增加一个偏置值 $\pi/2$ 将其变换到 $[0, \pi]$ ，故后三个分量的范围取值范围皆为 $[0, \pi]$ ，则 $N_i = \lfloor N_{\text{ang}} f_i / \pi \rfloor, i=2,3,4$ ，其中 $f_i, i=1,2,3,4$ 表示增强型点对特征 $F_e(p, q)$ 经过规范化后的 4

个分量的数值。由于增强型点对特征的第 1 个分量是点对之间的欧式距离，为了使得该分量的离散化能够适应不同模型之间的尺寸变化，故将该分量的离散化步长 δ_d 设置为 $0.05 D_M$ ，其中， D_M 是目标三维模型的直径（模型中距离最远的两个点之间的距离），则 $N_1 = \lfloor f_1 / \delta_d \rfloor$ 。最后，将所有满足视点可见性约束的点对 (p, q) 根据其哈希值 $K(p, q)$ 存储到该模型的哈希表中从而完成离线训练。对于所有的目标模型，其哈希表的构建过程只进行一次，在线识别之前根据需要预加载即可。

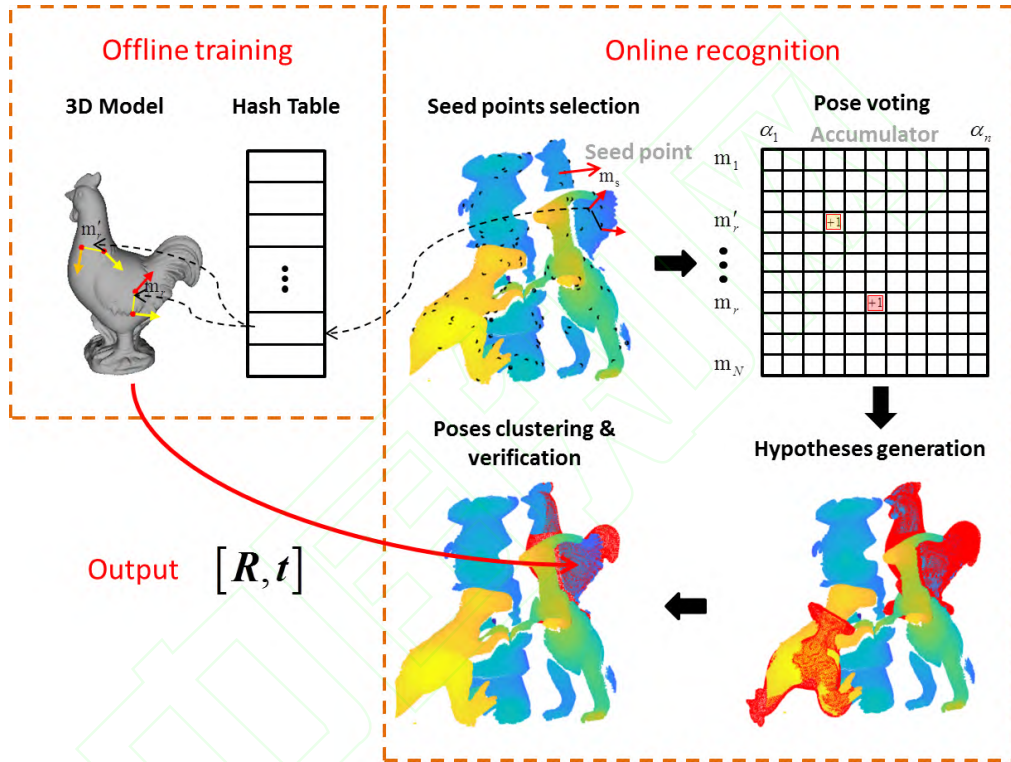


图 4 基于增强型点对特征的三维目标识别算法流程图

Fig. 4 The flow chart of the 3D object recognition based on the enhanced point pair feature

在线识别阶段：对于一个可能包含待识别目标的 2.5D 场景，首先加载目标模型的哈希表并计算场景点云中每个点的法向量，以便后续增强型点对特征的计算。然后在场景三维点云中随机选取 N_{seed} 个点作为种子点 p_s ，以该点为中心，在场景中搜索与之距离不超过 r 的点集 $Q(p_s) = \{q_1, q_2, \dots, q_k\}$ 。依次遍历 $Q(p_s)$ 中的每个点 $q_i, i = 1, 2, \dots, k$ ，计算点对 (p_s, q_i) 的增强型点对特征 $F_e(p_s, q_i)$ 以及它关于待识别目标模型而言的哈希值 $K(p_s, q_i)$ ，然后根据该哈希值到目标模型的哈希表中查找与之匹配的候选点对集合

$M(K(\mathbf{p}_s, \mathbf{q}_i)) = \{(\mathbf{m}_{i1}, \mathbf{m}_{k1}), (\mathbf{m}_{i2}, \mathbf{m}_{k2}), \dots, (\mathbf{m}_{in}, \mathbf{m}_{kn})\}$ 。场景点对 $(\mathbf{p}_s, \mathbf{q}_i)$ 与 $M(K(\mathbf{p}_s, \mathbf{q}_i))$ 中的每个模型点对 $(\mathbf{m}_{ij}, \mathbf{m}_{kj}), j=1, 2, \dots, n$ 匹配可以得到一个旋转角 α_{ij} ，然后在预先初始化为零的累加器 A_s 的坐标 $(\mathbf{m}_{ij}, \alpha_{ij})$ 处加 1，其中每个种子点 \mathbf{p}_s 都有一个与之对应的累加器 A_s ， A_s 的行对应目标模型中每个点，列对应一个将 $[0, 2\pi]$ 等间距离散为 30 份后的夹角值。当种子点 \mathbf{p}_s 与其邻域点集中的每个点都完成上述计算后， A_s 中值最大的位置对应的行即为目标模型中与之最匹配的点，列对应的角度即为二者之间的旋转角，由此可以得到一个候选位姿变换 $[\mathbf{R}_s, \mathbf{t}_s]$ 。当遍历完所有的种子点后，可以得到 N_{seed} 个候选位姿，将这些候选位姿按照各自累加器中最高得票数的高低降序重排，并将得票数低于全局最高得票数一半的候选位姿直接剔除。最终得到的候选位姿集合记为 $C = \{(\mathbf{R}_1, \mathbf{t}_1), (\mathbf{R}_2, \mathbf{t}_2), \dots, (\mathbf{R}_k, \mathbf{t}_k)\}$ 。

得到候选位姿集 C 后，需要从中找出真正的位姿变换（可能没有）。由上述分析可知，只有当种子点落在场景中的目标上，才可能生成正确的位姿变换。由于场景中可能有其他物体的干扰，随机选取的 N_{seed} 个种点中通常只有一部分落在目标上。为了从候选位姿集中提取正确的位姿变换，本文采用聚类的策略进行提取。位姿聚类的算法如下：

Step1: 以 C 中的第一个位姿建立一个新的类别，中心为 $T_1 = [\mathbf{R}_1, \mathbf{t}_1]$ ，然后从剩余的位姿集中找到属于该类别的位姿。所谓属于该类别，是指位姿变换 $[\mathbf{R}_i, \mathbf{t}_i]$ 与类别中心 $[\mathbf{R}_1, \mathbf{t}_1]$ 之间满足以下距离约束，即

$$\arccos\left(\frac{\text{trace}(\mathbf{R}_1^{-1}\mathbf{R}_i) - 1}{2}\right) \leq \varepsilon \wedge \|\mathbf{t}_i - \mathbf{t}_1\|_2 \leq \sigma, \quad (3)$$

其中 ε 与 σ 分别是旋转矩阵角度阈值以及平移向量距离阈值，所提方法中将 ε 与 σ 分别设置为 $\pi/6$ ， $0.2 D_M$ 。保存该类别，然后从 C 中剔除所有属于该类别的位姿。

Step2: 令 C 等于剩余的位姿集合，重复 step1，直至 C 为空。统计每个类别包含的位姿个数，并按照包含位姿个数的多少对所有类别降序重排，记第一个类别中包含的位姿个数为 N_{max} 。剔除元素个数小于 $N_{\text{max}}/2$ 的类别，将剩余的类别返回，算法结束。

假设位姿聚类算法最终返回 N_{valid} 个类别，对于每个类别，求出它的位姿中心。具体的做法是将平移向量的平均值作为位姿中心的平移向量，将所有旋转矩阵转化为四元数，计算四元数的平均值再将其转化为旋转矩阵作为位姿中心的旋转矩阵。最终会得到 N_{valid} 个粗匹配位姿。利用每个粗匹配位姿，将模型变换到场景中，并通过迭代最近邻（Iterative closest

points, ICP) 算法进行位姿优化, 如果拟合后的残差大于设定的阈值 τ_{rms} , 则剔除该位姿, 否则, 计算模型与场景点云的重叠率 (即与模型点距离小于等于 d_0 的场景点的数量与模型点总数的比值), 如果重叠率低于设定的阈值 τ_0 , 则剔除该位姿, 所提方法中分别将 $\tau_{\text{rms}}, \tau_0$ 与 d_0 设置为 3.5mm, 0.15 与 $2pr$ (pr 表示目标模型的平均分辨率)。最后返回经过 ICP 优化并且满足重叠率的位姿作为三维目标识别算法的结果。

3 实验结果与分析

本文提出的算法在联想笔记本 Y430P 上基于 Matlab2017b 平台实现。笔记本的配置为 Intel i7-4710MQ 处理器, 8GB RAM, 64 位操作系统, 算法实现没有依赖 GPU 等并行加速优化工具。所提算法与基于原始点对特征的三维目标识别算法分别在网上公开数据集和实验室实际采集的数据集上进行了效率与识别准确率的对比。下面对实验的具体情况介绍分析。

3.1 数据集介绍

为了测试算法的有效性以及实用性, 分别在实际采集的数据集 (Real dataset) 以及西澳大利亚大学 (University of Western Australia, UWA) 三维目标识别数据集^[10]上对本文提出的算法进行验证。下面对这两个数据集进行简要介绍。

针对工业自动化中常见的机械臂抓取分拣任务, 利用实验室的线结构光扫描平台采集了三种类型的数据集: a) 单目标混叠装箱场景 (Real dataset one, R1), 目标的平均遮挡率约为 50%; b) 多目标混叠装箱场景 (Real dataset two, R2) 以及 c) 多目标随机放置场景 (Real dataset three, R3), 分别如图 5 (b) 到 (d) 所示, 目标的平均遮挡率约为 45%。这三种类型的数据集分别包含 8, 4 以及 5 个不同的场景。其中待识别的目标模型主要包括眼镜盒和鼠标, 它们的 3D 模型由采集的不同视角的 2.5D 数据借助 Meshlab 软件手动拼接而成, 如图 5 (a) 所示。该数据集每个场景均匀采样 50 个点作为种子点。

UWA 数据集包含 5 个完整 3D 模型以及 50 个 2.5D 场景 (5 个模型与部分场景如图 6 所示), 其中, 犀牛模型主要用于干扰。每个 2.5D 场景中包含 4 到 5 个模型, 模型受遮挡的程度范围在 65%-95% 之间。所谓模型受遮挡的程度是指其在场景中未出现部分的顶点数与整个模型顶点数的比值。遮挡程度为 100% 即表示场景中没有出现该物体。由于该数据集相对复杂, 故在每个场景均匀采样 100 个点作为识别算法的种子点。

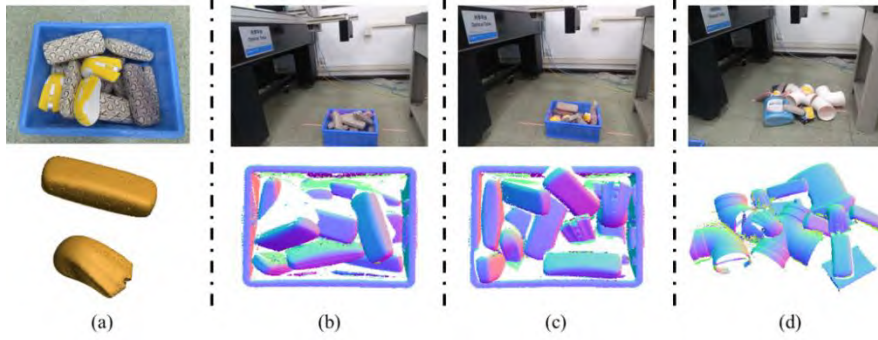


图5 实际采集的数据集。(a) 眼睛盒与鼠标模型；(b-d) 三种类型的数据集
Fig. 5 Real dataset. (a) Glass box and mouse model; (b) R1; (c) R2; (d) R3

3.2 模型哈希表建立

由于模型的点对特征哈希表需要离线建立,表1统计了上述数据集中总共6个目标模型的基本情况,包括模型的顶点数、不考率视点可见性约束的原始点对的个数、视点可见性约束阈值 $\tau=0$ 时的增强型点对的个数以及构建模型哈希表花费的时间成本。可以看到,本文提出的视点可见性约束可以缩减大约一半的冗余点对,且大幅加快了构建模型哈希表的速度。

表1 六个目标模型的基本情况介绍
Table 1 Summary of the six models

Models	#points	Original PPF		Enhanced PPF	
		#PPFs	Time cost	#EPPFs	Time cost
Chef	3351	11225850	85.1s	5557422	34.7s
Chicken	2643	6982806	44.0s	3434958	18.9s
Para	2507	6282542	43.2s	3114468	17.5s
T-rex	2337	5459232	35.9s	2713202	14.8s
glass Box	1134	1284822	3.94s	638622	1.87s
Mouse	1358	1842806	6.16s	916592	2.96s

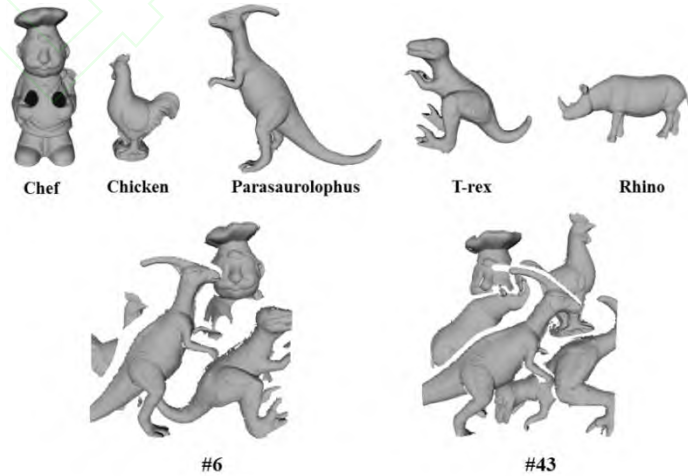


图6 UWA 数据集中的5个模型以及两个场景
Fig. 6 Five models and two sample scenes of UWA 3D object recognition dataset

3.3 UWA 数据集识别结果分析

本文提出的三维目标识别算法在 UWA 数据集上的识别结果如表 2 所示, 分别列出了每个模型的识别情况, 包括正确的识别数、失败的场景编号以及对应模型在该场景中被遮挡的程度。这里需要特别指出的是, 每个模型的识别结果都是以它作为目标, 分别与每个场景进行匹配识别得到的, 即场景中除了它以外的其他物体都是干扰。因为有些识别算法^[11-12]在该数据集上给出的结果是联合识别得到的, 即以 4 个模型为目标, 分别与每个场景进行匹配。区别在于后者在识别的过程中会将识别到的物体从场景中不断分割出去, 从而使得后面的目标更容易被识别。从表 2 中得知, Chef 模型在 43 号场景中 (如图 6 所示) 识别失败了, 主要原因是该模型在该场景中被遮挡的程度太高, 其次是由于其它物体对它的干扰。可以看到, 另外三个模型在该场景中都识别正确了。如果按照联合识别的策略, 将容易识别的目标成功识别后分离出场景, 那么当场景中只剩下 Chef 模型时, 再去识别就没有干扰了。事实上, 对于一个三维目标识别算法而言, 直接无分割式识别更能体现其性能的好坏。因此, 本文所有的识别实验结果都是在原始场景上直接识别得到的。由表 2 可知, 识别失败的案例大多在原始场景中被遮挡的程度过高, 导致其有效信息不足, 加之场景中其他物体的干扰, 使得点对投票阶段目标模型获得的票数不够, 导致算法最终识别失败。由于文献 8 中只统计了遮挡率低于 84% 的识别结果, 为了与之进行客观比对, 表 3 给出本文算法在遮挡率低于 84% 的情形下的识别结果。可以看到, 所提方法较原始点对特征方法识别率略有提升, 但是识别速度较后者有明显提升。说明利用视点可见性约束缩减模型哈希表并不会降低识别准确率, 且基于增强型点对特征匹配能够加快识别速度。

表 2 本文方法在整个 UWA 数据集上的识别结果

Table 2 Recognition results on the whole UWA dataset of our proposed algorithm

Models	#correct/total	Failed scene	Occlusion of the targets
Chef	49/50	43	91.3%
Chicken	45/48	6,26,32	89.7%,86.5%,89.5%
Parasaurolophus	40/45	7,10,38,41,50	86.4%,91.4%,89%,87%,83.9%
T-rex	41/45	4,10,34,48	84%,80.2%,83.8%,77.3%
average	175/188 (93.1%)	-	-

表 3 四种方法在 UWA 数据集上的识别结果对比 (目标模型的遮挡率低于 84%)

Table 3 Comparison of four algorithms on the UWA dataset (occlusion of the targets less than 84%) in terms of recognition rate

algorithms	Recognition rate	Time cost for one object
Proposed	164/168 (97.6%)	10s
PPF ^[8]	97%	85s

3.4 实际采集的数据集的认识结果分析

本节给出了本文方法与原始点对特征方法在三种实际采集的数据集上的识别结果。其中原始点对特征方法指的是没有使用增强型点对特征且不考虑视点可见性约束的方法。除此之外，两种算法的其他流程及参数完全一致，每个场景选取的种子点也相同。下面针对这两种算法在这三种数据集上的识别结果进行性能对比分析。

单目标混叠装箱数据集：图 7 给出了本文算法在 8 个场景上的识别结果，图中每个格子从左到右依次是：原始场景，识别结果，识别结果中遮挡率最低的物体。由于原始点对特征与本文算法的识别结果一致，故只显示了本文算法的结果。图 8 给出了两种算法在每个场景中的时间开销（不包括建立模型哈希表花费的时间）。由图 8 可知，在识别率一致的情形下，本文算法计算效率更高。主要原因是增强型点对特征区分性更强，使得在候选位姿集生成阶段中与每个场景点对匹配的模型点对更可靠且数量更少，从而加速了投票过程，提升了算法的整体效率。从图 7 中还可以看到，本文算法能够同时识别出多个目标实例，这对于提高整个分拣系统的实时性具有重要意义。

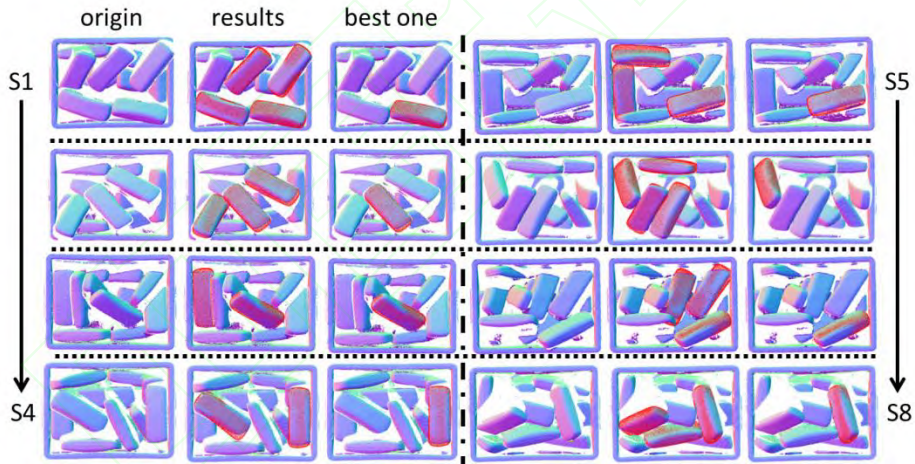


图 7 八个单目标混叠装箱场景的识别结果

Fig. 7 Recognition results on the eight scenes of Real dataset one (R1)

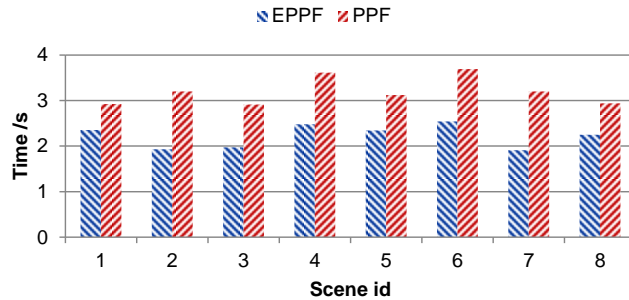


图 8 本文方法与原始点对特征方法在 R1 数据集上的计算效率对比

Fig. 8 Time cost comparison between the proposed method and the original PPF based method on the R1 dataset

多目标混叠装箱数据集：图 9 展示了本文算法在 4 个场景上的识别结果，同样，由于原始点对特征方法也取得了相同的识别结果，故没有另外单独显示。图 9 中每一列从上到下依次是：原始场景，眼镜盒识别结果，鼠标识别结果。图 10 展示了两种算法的时间开销对比。可以看到，本文算法效率仍然优于原始点对特征方法。

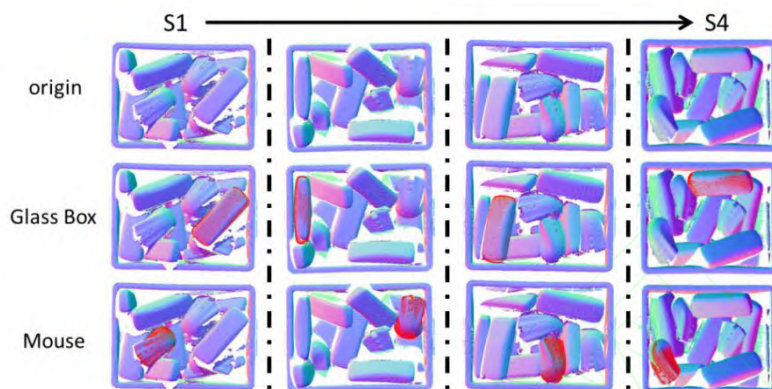


图 9 四个多目标混叠装箱场景的识别结果

Fig. 9 Recognition results on the four scenes of Real dataset two (R2)

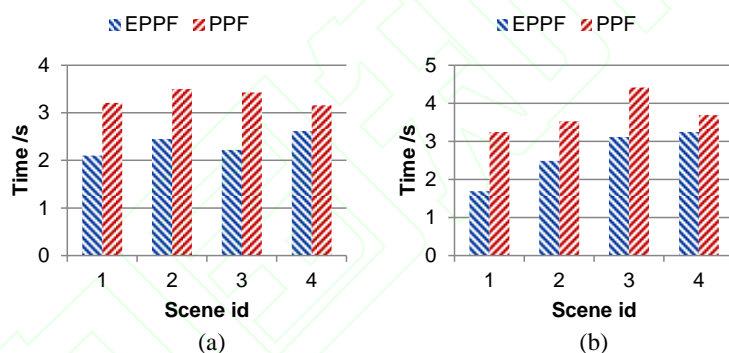


图 10 本文方法与原始点对特征方法在 R2 数据集上的计算效率对比

Fig. 10 Time cost comparison between the proposed method and the original PPF based method on the R2 dataset

多目标随机放置场景数据集：图 11 显示了本文算法与原始点对特征方法在 5 个场景上的识别结果，图 12 给出了它们的时间开销。与前两种数据集相比，由于该数据集干扰物相对较多，所以两种算法整体的时间成本都有所提高，不过所提方法时间开销依然优于后者。此外，从识别准确率上来看，所提方法在这 5 个场景中都准确识别到目标，识别率为 100%，而原始点对特征方法没有正确识别场景 1 和场景 4 中的鼠标模型，识别率为 60%。可见所提方法与原始点对特征方法相比识别可靠性也有所提升。主要原因仍然是增强型点对特征的区别性要强于原始点对特征，从而排除了一些错误的点对匹配，使得投票阶段能够保留正确的位姿，才有可能在位姿聚类的时候从候选位姿集中将其筛选出来。总之，本文提出的基于增强型点对特征的方法对于提升了原始点对特征方法的性能。

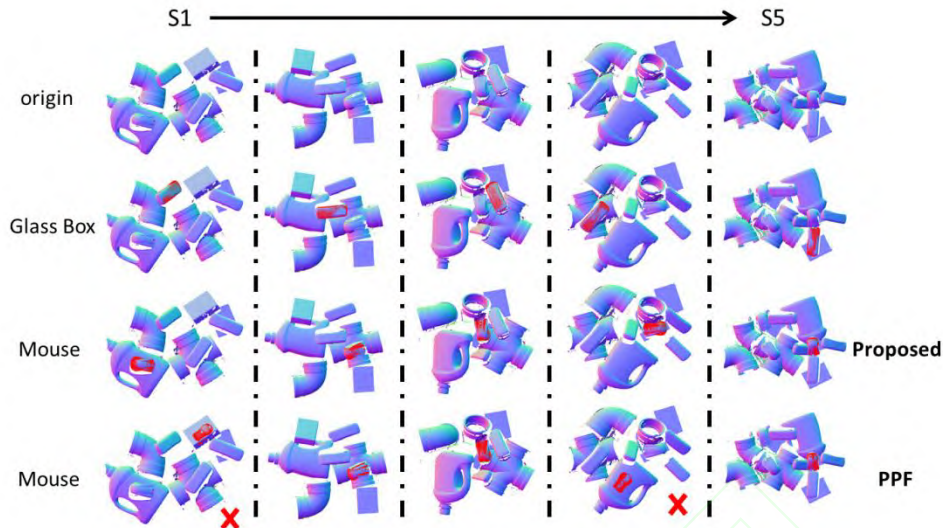


图 11 五个多目标随机放置场景的识别结果

Fig. 11 Recognition results on the five scenes of Real dataset three (R3)

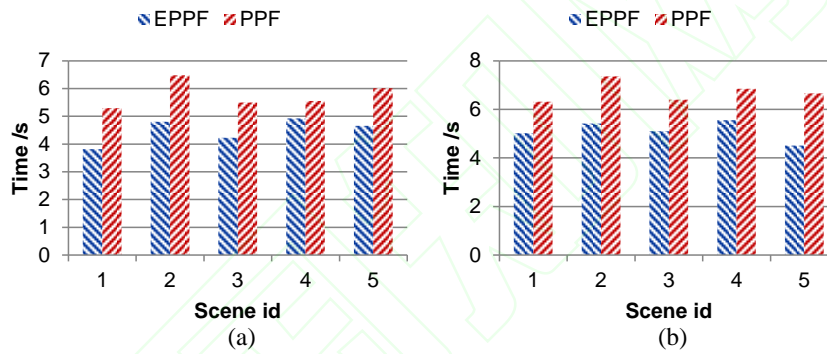


图 12 本文方法与原始点对特征方法在 R3 数据集上的计算效率对比

Fig. 12 Time cost comparison between the proposed method and the original PPF based method on the R3 dataset

4 结论

本文提出了一种新型点对征，增强了原始点对特征的区别性。通过建立模型点对之间的可见性约束，剔除了大量冗余点对，加快了构建模型点对特征哈希表的速度并节省了存储开销。在 UWA 数据集和实际采集的数据集上得到的识别实验结果表明，本文方法较原始点对特征方法在识别准确率与识别效率上皆有所提升，证明了本文方法的有效性与实用性。事实上，对于单目标混叠场景，可以通过设置少量的种子点进一步加快识别的效率。因为单目标混叠场景中的大部分的点都是由目标实例组成的，有效种子点被选中的概率很大。此外，对于较为复杂的场景，如果目标在场景中占比较小，可适当多选一些种子点，以提高识别率。

参考文献

- [1] Ulrich M, Wiedemann C, Steger C. Combining scale-space and similarity-based aspect

- graphs for fast 3D object recognition[J]. IEEE transactions on pattern analysis and machine intelligence, 2012, 34(10): 1902-1914.
- [2] Hinterstoisser S, Cagniart C, Ilic S, *et al.* Gradient response maps for real-time detection of textureless objects[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(5): 876-888.
- [3] Pan W, Zhu F, Hao Y M, *et al.* Pose Measurement Method of Three-Dimensional Object Based on a Multi-Sensor System[J]. Acta Optica Sinica, 2019, 39(02): 0212007.
潘旺, 朱枫, 郝颖明, 等. 基于多传感器的三维目标位姿测量方法[J]. 光学学报, 2019, 39(02): 0212007.
- [4] Tombari F, Salti S, Di Stefano L. Performance evaluation of 3D keypoint detectors[J]. International Journal of Computer Vision, 2013, 102(1-3): 198-220.
- [5] Guo Y, Bennamoun M, Sohel F, *et al.* 3D object recognition in cluttered scenes with local surface features: a survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36(11): 2270-2287.
- [6] Yang J, Xian K, Xiao Y, *et al.* Performance evaluation of 3D correspondence grouping algorithms[C]//2017 International Conference on 3D Vision (3DV). IEEE, 2017: 467-476.
- [7] Schnabel R, Wahl R, Klein R. Efficient RANSAC for point - cloud shape detection[C]//Computer graphics forum. Oxford, UK: Blackwell Publishing Ltd, 2007, 26(2): 214-226.
- [8] Drost B, Ulrich M, Navab N, *et al.* Model globally, match locally: Efficient and robust 3D object recognition[C]//2010 IEEE computer society conference on computer vision and pattern recognition. Ieee, 2010: 998-1005.
- [9] Drost B, Ilic S. 3d object detection and localization using multimodal point pair features[C]//2012 Second International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission. IEEE, 2012: 9-16.
- [10] Choi C, Christensen H I. 3D pose estimation of daily objects using an RGB-D camera[C]//2012 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2012: 3342-3349.
- [11] Guo Y, Sohel F, Bennamoun M, *et al.* Rotational projection statistics for 3D local surface description and object recognition[J]. International journal of computer vision, 2013, 105(1): 63-86.
- [12] Guo Y, Sohel F, Bennamoun M, *et al.* A novel local surface feature for 3D object recognition under clutter and occlusion[J]. Information Sciences, 2015, 293: 196-213.

网络首发:

标题: 基于增强型点对特征的三维目标识别方法

作者: 鲁荣荣, 朱枫, 吴清潇, 陈佛计, 崔芸阁, 孔研自

收稿日期: 2019-03-05

录用日期: 2019-05-05

DOI: 10.3788/aos201939.0815006

引用格式:

鲁荣荣, 朱枫, 吴清潇, 陈佛计, 崔芸阁, 孔研自. 基于增强型点对特征的三维目标识别方法[J]. 光学学报, 2019, 39(08):0815006.

网络首发文章内容与正式出版的有细微差别, 请以正式出版文件为准!

您感兴趣的其他相关论文:

基于深度信息的大豆株高计算方法

冯佳睿 马晓丹 关海鸥 朱可心 于菘

黑龙江八一农垦大学电气与信息学院, 黑龙江 大庆 163319

光学学报, 2019, 39(5):0515003

基于通道可靠性的多尺度背景感知相关滤波跟踪算法

尹明锋 薄煜明 朱建良 吴盘龙

南京理工大学自动化学院, 江苏 南京 210094

光学学报, 2019, 39(5):0515002

基于归一化算法的一维标定物多相机标定

全燕鸣 覃镇波 李维诗 张瑞

华南理工大学机械与汽车工程学院, 广东 广州 510640

光学学报, 2019, 39(4):0415001

基于形变多样相似性的空间正则化相关滤波跟踪

毛宁 杨德东 李勇 韩亚君

河北工业大学人工智能与数据科学学院, 天津 300130

光学学报, 2019, 39(4):0415002

解耦光流运动场模型的车载平台仿真

乌萌 郝金明 付浩 高扬

信息工程大学地理空间信息学院, 河南 郑州 450052

光学学报, 2019, 39(4):0415005