

引文格式:

孔令升, 崔西宁, 郭俊广, 等. 基于时域编码结构光的高精度三维视觉引导抓取系统研究 [J]. 集成技术, 2020, 9 (2): 38-49.
Kong LS, Cui XN, Guo JG, et al. An accurate bin-picking system using temporal encoded structured light sensor [J]. Journal of Integration Technology, 2020, 9 (2): 38-49.

基于时域编码结构光的高精度三维视觉引导抓取系统研究

孔令升^{1,2} 崔西宁^{1,3} 郭俊广^{1,2} 宋展¹ 孙红雨³

¹(中国科学院深圳先进技术研究院 深圳 518055)

²(中国科学院大学 北京 100039)

³(山东科技大学 青岛 266590)

摘要 在许多自动化应用场合中, 如分拣和上、下料等过程中, 机械臂抓取是非常重要的一个环节。在有遮挡或物体杂乱放置的情况下, 对物体进行可靠、快速和精确计算位姿是机械臂抓取的难题之一。该文提出一种针对非规则目标的 3D 视觉引导抓取系统。首先, 该系统运用面结构光系统对目标进行高精度三维重建, 并建立离线 3D 点云模板库; 然后, 将标准模板与点云预处理后的场景点云进行匹配, 得到匹配参数后, 由坐标系之间的转换矩阵计算机器手抓取位姿; 最后, 引导机器人完成对目标物体的抓取。实验结果表明, 所开发的机械臂抓取系统能够对非规则目标进行可靠、快速和精确的抓取。

关键词 精确抓取; 结构光; 标定; 表面模板匹配; 位姿计算

中图分类号 TP 391 文献标志码 A doi: 10.12146/j.issn.2095-3135.20200110001

An Accurate Bin-Picking System Using Temporal Encoded Structured Light Sensor

KONG Lingsheng^{1,2} CUI Xining^{1,3} GUO Junguang^{1,2} SONG Zhan¹ SUN Hongyu³

¹(Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China)

²(University of Chinese Academy of Sciences, Beijing 100039, China)

³(Shandong University of Science and Technology, Qingdao 266590, China)

Abstract Robot bin-picking system is an important component in many automation applications, such as

收稿日期: 2020-01-10 修回日期: 2020-02-18

基金项目: 国家重点研发计划项目(2017YFB1103602); 国家自然科学基金项目(U1613213); 深圳市孔雀计划-技术创新项目(KQJSCX20170731165108047)

作者简介: 孔令升, 硕士研究生, 研究方向为计算机 3D 视觉; 崔西宁, 硕士研究生, 研究方向为计算机 3D 视觉; 郭俊广, 硕士研究生, 研究方向为计算机 3D 视觉; 宋展(通讯作者), 研究员, 博士研究生导师, 研究方向为计算机 3D 视觉, E-mail: zhan.song@siat.ac.cn; 孙红雨, 副教授, 硕士研究生导师, 研究方向为无线通信。

sorting, loading, and unloading. However, it is difficult to calculate the pose of object reliably, quickly, and accurately in the environment where the object is obscured or placed disorderly. This paper proposes a 3D vision-guided bin-picking system for free-form objects. In this system, the structured light systems are used to carry out high-precision 3D reconstruction of the targets and to establish the off-line 3D template library. Furthermore, standard template is matched with the point cloud of the scene after preprocessing. After the matching parameters are obtained, the robot's posture is calculated by the transformation matrix between the coordinate systems, and the robot is guided to complete the grasp of the target object. Experimental results show that the robot bin-picking system can grasp irregular targets reliably, quickly, and precisely.

Keywords precision bin-picking; structured light; calibration; surface model matching; pose estimation

1 引言

机械臂抓取是机器人自动化领域的一个重要研究课题,在分拣、装配、上料、焊接、喷涂等领域有着广泛的应用。现有的机械臂抓取技术包括基于 2D 图像与基于 3D 视觉抓取技术。其中,基于 2D 图像抓取技术是抓取技术中较成熟的技术,但由于缺乏 3D 姿态,机械臂只能垂直向下获取抓取位置,故在复杂的环境中,该技术已不能满足空间抓取的应用要求^[1-2]。基于 3D 视觉抓取技术由于获取物体目标空间位姿的不确定性、获取点云分辨率低、手眼标定结果误差大、场景的复杂性及遮挡等因素影响,增加了机械臂抓取物体目标的难度。随着工业技术的不断发展,机械臂抓取的需求也愈发向可靠、快速和精确方向发展,对现有机械臂抓取技术提出了严峻挑战。

现有的 3D 点云获取技术包括双目、结构光和飞行时间等^[3-5]。在不同的 3D 点云获取技术中,基于结构光的三维扫描方法以其高精度和高效率而被广泛使用。黄会明等^[6]提出了一种基于微振镜结构光的三维重建方法。该方法基于多频外差的方法对物体进行三维重建,使用了基于点对特征三维特征匹配的方式对获取的三维数据与基准模型进行匹配和定位操作,能获取任意位置放置的物体在机器人坐标系中的三维坐标和法线

信息。与 Kinect 相比,该方法无论是在精度还是在处理时间上都有了很大提升。Li 等^[7]提出了一种二维视觉和三维视觉相结合的堆垛目标拾取系统的设计方法。该系统利用结构光的方法获取三维信息,并提取出目标区域,再利用广义霍夫变换方法对目标进行匹配和位姿估计,接着将物体的位姿发送给机械手进行自动抓取,其对散落物体的捕捉非常有效,且利用二维和三维方法的结合提高了算法的效率。

手眼系统是由摄像机和机械手构成的机器人视觉系统,分为 Eye-in-Hand 式手眼系统^[8]和 Eye-to-Hand 式手眼系统^[9]。手眼标定的意义在于建立起摄像机坐标系与机械臂末端抓手坐标系之间的联系,能将在摄像机坐标系下的信息通过建立的联系转换成在机械臂坐标系下的信息。Tsai 和 Lenz^[10]早在 20 世纪 80 年代就提出了经典的方法,将手眼标定的数学模型归结为求解 $AX=XB$ 矩阵,即将手眼标定的求解具体化为矩阵的求解。Zhuang 等^[11]提出了一种同时计算世界坐标系到机械臂坐标系、机械臂末端手爪坐标系到摄像机坐标系的刚体变换方法,通过运用四元数推导 X 和 Z 的显式线性解求得 $AX=ZB$ 的齐次矩阵方程。

本文提出一种针对非规则目标的 3D 视觉引导抓取系统,系统可分为体视觉模块、点云处理

模块和抓取控制模块。在立体视觉模块中,采用格雷码与线移结合的方法作为结构光系统的编码方案,通过解码实现目标高精度的三维重建^[12]并建立 3D 模板库。点云处理模块对获取的点云去噪后,将标准模板与预处理后的场景点云进行匹配。得到匹配参数后,利用手眼标定参数计算出机器人抓取位姿。最后由抓取控制模块完成引导机器人对目标物体的抓取。

2 系统方案设计

2.1 系统硬件设计

基于立体视觉的机械臂抓取系统由工业相机、投影仪、上位机、机器人控制器、电动夹持器和机械臂等部分组成。

本文系统设计图如图 1 所示,对应的硬件功能如下:投影仪与工业相机组成立体视觉系统。具体地,首先由投影仪投出编码图案,工业相机进行拍摄并将采集到的图像传到上位机;随后上位机对图像进行处理,利用标定的参数对投影仪

投出的编码图案进行解码,得到三维点云数据后,再进行模板点云获取、点云去噪、点云分割、表面模板匹配等步骤;待获取抓取位置姿态后通过 TCP/IP (Transmission Control Protocol/Internet Protocol) 通信传送给机器人控制器,控制机械臂与电动夹持器完成抓取动作并将抓取目标物放置在指定位置。

2.2 系统软件设计

在 Windows 系统环境下,以 VS2015 作为软件开发平台,C++ 语言作为程序开发环境。系统软件结构运用类设计思想,对应的系统软件结构如图 2 所示,分为立体视觉模块、点云处理模块和抓取控制模块。

(1) 立体视觉模块。该部分先对相机、投影仪及相机投影仪之间的旋转平移矩阵参数进行标定,再对投影仪的投影图案进行编码,获得高质量的编码图像后,结合标定参数进行解码获取点云数据,并将点云数据发送至点云处理模块。

(2) 点云处理模块。该方案采用模板匹配方案对场景中物体进行抓取。首先,在工作距离下

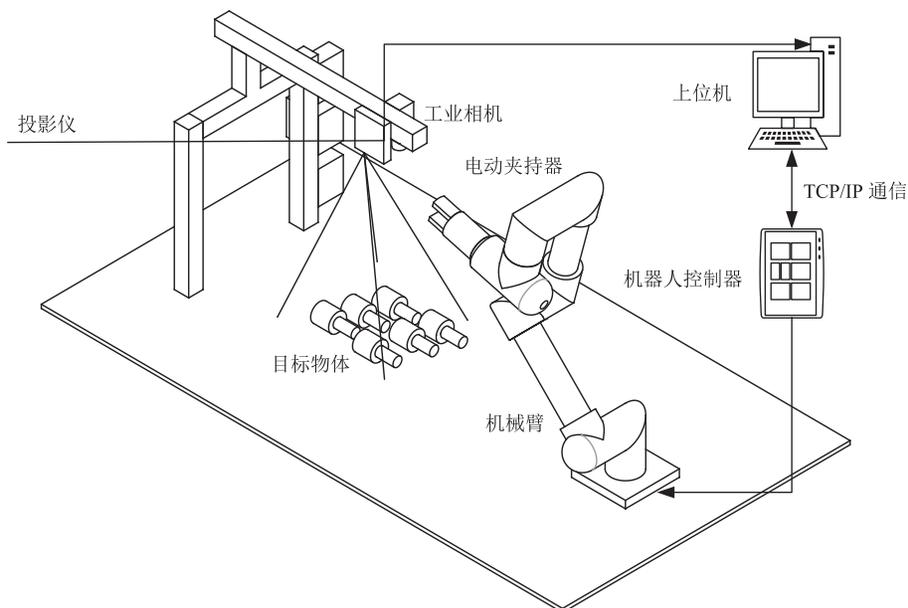


图 1 系统设计图

Fig. 1 The diagram of system design

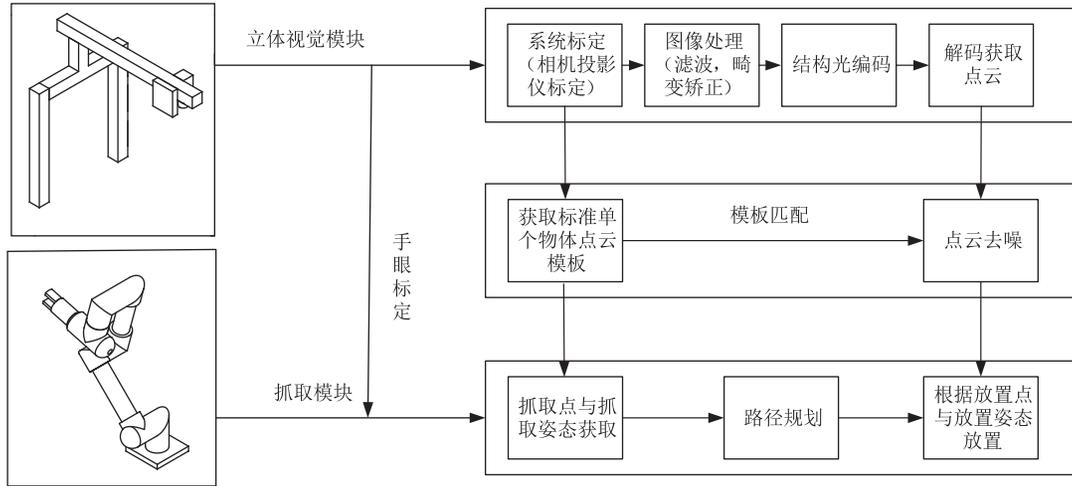


图 2 系统软件结构

Fig. 2 Structure of system software

获取标准模板, 并获取标准模板下的抓取位姿与放置位姿; 然后, 对获取场景点云进行分割, 并将分割后的点云与标准模板进行匹配, 获取抓取位姿; 最后, 通过 TCP/IP 通信将数据传送到抓取控制模块。

(3) 抓取控制模块。抓取控制模块主要包括机械臂控制模块、电动夹持器抓取控制模块和路径规划模块。机械臂控制模块通过实时获取目标位姿, 并计算机械臂控制参数来控制机械臂移动。当机械臂到达目标位姿后, 由电动夹持器抓取控制模块进行抓取, 确保抓取成功。为避免抓取后与平台发生碰撞, 根据抓取位姿与放置目标位姿合理规划路径, 确保目标物能准确放到指定位置。

3 系统实现的关键技术

本文提出的抓取系统实现的关键技术大致分 3 步: (1) 标定结构光系统内外参数与机械臂手眼标定参数, 通过解码编码图像完成点云重建, 基于此完成模板库的建立; (2) 进行场景点云与模板库中标准模板点云匹配, 找出匹配率最高的匹配参数作为最终结果, 获取转换矩阵; (3) 基于标定参数与匹配得到的转换矩阵, 获取抓取位

姿, 引导机械臂完成抓取。

3.1 系统标定与点云获取

本系统的标定主要分为结构光系统标定与机械臂标定两部分。其中, 结构光系统标定包括相机内部参数的标定、投影仪内部参数的标定、相机与投影仪畸变参数的标定, 以及投影仪与相机之间外部参数的标定。机械臂标定包括 TCP (Tool Center Point) 标定与手眼标定两部分。

结构光系统中摄像机的模型可视为理想的透视投影模型, 投影仪可视为一个逆向的相机, 故投影仪与相机具有相同的标定模型。图 3 描述了摄像机、投影仪和机械臂之间的坐标关系。根据图 3 所示几何关系及相机透视投影模型^[12]可以得到, 由图像平面上的点 $P_1(u_c, v_c)$ 和投影仪平面上的点 $P_2(u_p, v_p)$ 到世界坐标系中的点 $Q(x_w, y_w, z_w)$ 之间的变换关系如公式 (1) 和 (2) 所示。

$$s \begin{bmatrix} u_c \\ v_c \\ 1 \end{bmatrix} = \begin{bmatrix} f_x^c & f_x^c \tan \alpha & u_{c0} & 0 \\ 0 & f_y^c & v_{c0} & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} [R_C, T_C] \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (1)$$

$$= K_C [R_C, T_C] \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}$$

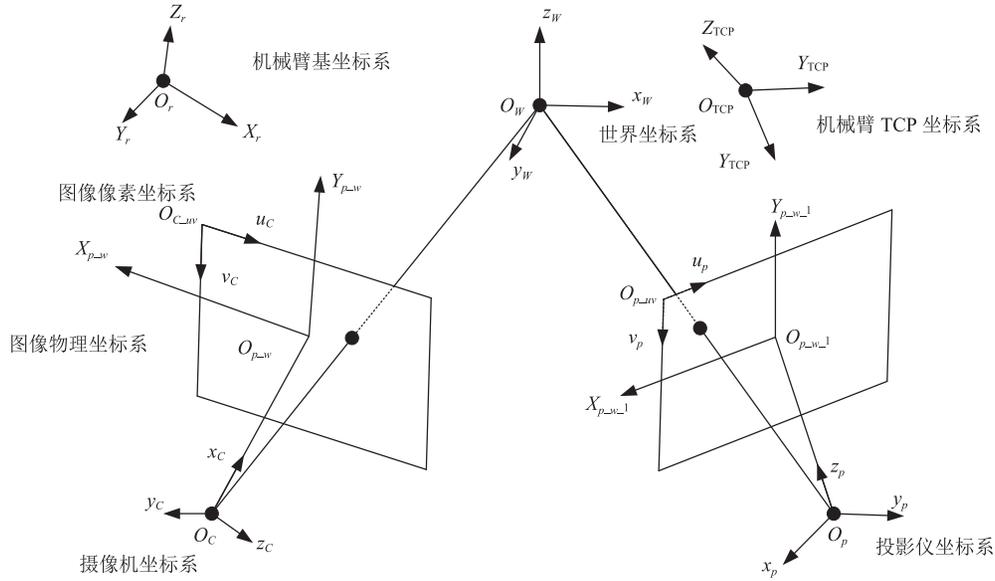


图3 系统标定原理图

Fig. 3 System calibration schematic diagram

$$s \begin{bmatrix} u_p \\ v_p \\ 1 \end{bmatrix} = \begin{bmatrix} f_x^p & f_x^p \tan \alpha & u_{p0} & 0 \\ 0 & f_y^p & v_{p0} & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \begin{bmatrix} \mathbf{R}_p \\ \mathbf{T}_p \end{bmatrix} \quad (2)$$

$$= \mathbf{K}_p \begin{bmatrix} \mathbf{R}_p \\ \mathbf{T}_p \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}$$

其中, s 为尺度因子; α 为坐标轴倾斜参数; \mathbf{K}_C 为摄像机内参; f_x^C 、 f_y^C 为摄像机在 u_C 轴和 v_C 轴方向上的尺度因子; (u_{C0}, v_{C0}) 为摄像机成像的主点; $[\mathbf{R}_C, \mathbf{T}_C]$ 为摄像机外参; \mathbf{K}_p 为投影仪内参; f_x^p 、 f_y^p 为投影仪在 u_p 轴和 v_p 轴方向上的尺度因子; (u_{p0}, v_{p0}) 为投影仪成像的主点; $[\mathbf{R}_p, \mathbf{T}_p]$ 为投影仪外参。摄像机与投影仪之间的变换关系如式(3)所示。

$$\begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = [\mathbf{R}, \mathbf{T}] \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (3)$$

$$= \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} + \begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ 1 \end{bmatrix}$$

其中, $[\mathbf{R}, \mathbf{T}]$ 为转换矩阵。

由公式(1)~(3)使用格雷码与线移结合进行编解码^[13-14], 可解出摄像机坐标系下的空间点坐标, 具体如公式(4)所示。

$$\begin{cases} z_c = \frac{(f_p t_1 - s t_3) \cdot f_c}{(s r_{13} - f_p r_{11}) \cdot x^u + (s r_{32} - f_p r_{12}) \cdot y^u + (s r_{33} - f_p r_{13}) \cdot f_c} \\ x_c = \frac{u_c \cdot z_c}{f_c} \\ y_c = \frac{v_c \cdot z_c}{f_c} \end{cases} \quad (4)$$

手眼标定的目的是获取机器人坐标系和相机坐标系之间的旋转平移矩阵, 再将视觉识别的结果转移到机器人坐标系下。TCP 标定^[15]可由四点法得到高精度的旋转平移矩阵 $[\mathbf{R}_1, \mathbf{T}_1]$ 和 $[\mathbf{R}_2, \mathbf{T}_2]$ 。现有的手眼标定算法通过移动一次机械臂建立 TCP 坐标系与摄像机坐标系之间的位置变换关系, 该关系用矩阵 \mathbf{X} 表示, 因此手眼标定问题转换为 $\mathbf{AX}=\mathbf{XB}$ 求解问题。本文提出一种新的手眼标定方案, 其原理如图4所示。设摄像机坐标系为 $O_c-x_c y_c z_c$, 机械臂坐标系为 $O_r-x_r y_r z_r$, 则机器人基坐标系和相机坐标系之间的旋转平移

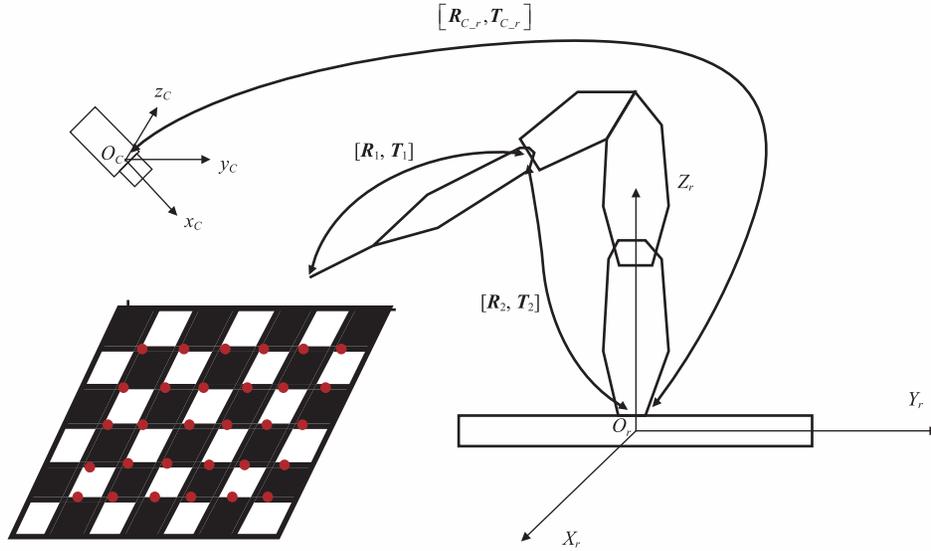


图4 手眼标定原理图

Fig. 4 System calibration schematic diagram

矩阵设为 $[R_{C_r}, T_{C_r}]$ 。首先, 当结构光光条纹图案投射在棋盘上时, 棋盘格 3D 模型可通过结构光算法来恢复, 此时可获取角点在摄像机坐标系下的坐标点 $P_i(x_{Ci}, y_{Ci}, z_{Ci})$ 。然后, 采用电动夹持器对准角点, 获取机械臂坐标系下对应的三维坐标点 $Q_i(x_{ri}, y_{ri}, z_{ri})$ 。最后, 使用最小二乘法^[16]计算 $[R_{C_r}, T_{C_r}]$, 数学模型总结如公式(5)所示。其中, γ 是使用 Zhang^[17]的方法得到的函数。

$$\min_{R_{C_r}, T_{C_r}} error = \frac{1}{2} \sum_{i=1}^n \left\| \gamma(x_{Ci}, y_{Ci}, z_{Ci}) - R_{C_r} \cdot \gamma(x_{ri}, y_{ri}, z_{ri}) + T_{C_r} \right\|_2^2 \quad (5)$$

基于上述方案获取点云后, 用三维制图软件创建匹配模板。构建 3D 模板库时, 需获取不同位姿下的物体点云构建模板库, 以解决实际场景中遮挡和堆叠情况下获取点云的多姿态问题。

3.2 三维点云匹配

本文采用点对特征的三维点云匹配算法^[18]。点对特征分为全局特征与局部特征, 其中全局特征匹配恢复三维姿态计算量巨大^[19], 而直接在点云上进行关键点、面片等局部特征提取易出现误匹配^[20]。在本文 3.1 小节中, 完成了 2D 图像到 3D

点云的转换, 通过转换矩阵的逆矩阵可将 3D 点云转换为 2D 图像, 随后在模板图像中设置关键点, 可生成 128 维关键点描述子^[21]。在局部对模型图片和场景图片进行匹配时, 选取匹配率最高的 3 对像素点映射到 3D 点云中构建点对特征。而对匹配结果进行验证时, 选取匹配误差最小的特征作为点对特征, 从而恢复模型在场景中的全局姿态。点对特征可用来描述两个点的相对位置及法线方向^[22]。假设对于任意两个点 m_1, m_2 以及它们各自的法矢量 n_1, n_2 , 令矢量差 $d = m_2 - m_1$, 并定义特征 F 为:

$$F(m_1, m_2) = [\|d\|_2, \angle(n_1, d), \angle(n_2, d), \angle(n_1, n_2)] \quad (6)$$

其中, 设分量 F_1 为 m_1 和 m_2 之间的距离; F_2 和 F_3 为法线与两点定义的向量夹角; F_4 为两法线夹角。点对特征 F 的 4 个分量如图 5 所示。

图 5 中全局模型描述表示从点对特征空间到模型的映射, 具体是把四维点对特征映射到具有相似特征向量的点对 $(m_i, m_j) \in M^2$ 。对于场景中任意一个参考点 $s_t \in S$, 若该参考点在要检测的物体目标上, 则一定会存在一个来自模板上的点 $m_t \in M$ 与之对应。因此, 只需将这两个点及法线

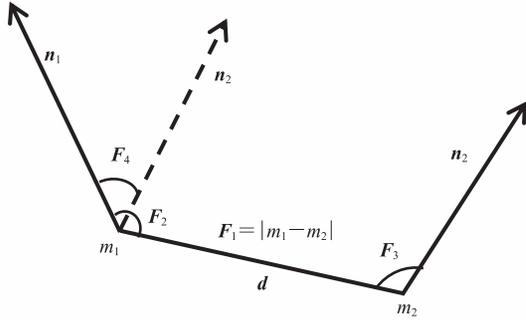


图5 特征描述

Fig. 5 Feature description

对齐,再围绕场景参考点 s_τ 的法线旋转一定的角度即可让模型和场景对齐,从而生成目标物体的姿态,完成点云匹配工作。模型到场景的矩阵变换 $[R, T]$ 可由模型上的一个参考点 m_τ 和一个旋转角 α 来表示, (m_τ, α) 即为模型相对于场景参考点 s_τ 的局部坐标,如公式(7)所示。

$$s_i = T_{s \rightarrow g}^{-1} R_x(\alpha) T_{m \rightarrow g} m_i \quad (7)$$

其中, s_i 是场景中任意点; m_i 是模型中任意点; $T_{m \rightarrow g}$ 是点 m_τ 到原点的转换矩阵; $T_{s \rightarrow g}$ 是点 s_τ 到原点的转换矩阵; $R_x(\alpha)$ 是模型到场景绕 x 轴旋转 α 角度的旋转矩阵。

点云匹配识别率由相关度大小决定匹配程度。在匹配过程中,堆叠物体会匹配多个模板,且由于不同模板在场景中遮挡程度不同,以及模板本身特征的差异,最终生成多组匹配数据。因此,取匹配率最高的一组数据作为最终识别结果。由于场景中总会有一目标处于无遮挡或接近于无遮挡状态,使用该方案在复杂的堆叠情况下也能有效识别。

3.3 机械臂抓取

模板匹配完成后,获取的仅是模板到目标场景的旋转平移矩阵。其中,模板抓取位姿在建立模板库时已确定,因此利用模板匹配求取的转换矩阵可得机械臂抓取位姿。随后使用快速扩展随机树(Rapidly Exploring Random Tree)的路径规划算法^[23],通过对状态空间中的采样点进行碰撞检测,解出合理的运动路径,并将

路径参数传送给机械臂完成抓取动作,其中位姿可由公式(7)~(11)计算得到。模板抓取位姿 (x, y, z, R_x, R_y, R_z) 可在建立模板库时获取。其中, (x, y, z) 为抓取位置; (R_x, R_y, R_z) 为以旋转向量表示的抓取姿态。对于抓取点,可由公式(8)解出。

$$\begin{bmatrix} x_{\text{grab}} \\ y_{\text{grab}} \\ z_{\text{grab}} \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} [R_{m_c}, T_{m_c}] \cdot [R_{\text{match}}, T_{\text{match}}] \cdot [R_{c_m}, T_{c_m}] \quad (8)$$

其中, $[R_{m_c}, T_{m_c}]$ 为机械臂基坐标系到摄像机坐标系下的转换矩阵; $[R_{\text{match}}, T_{\text{match}}]$ 为表面模板匹配求出的匹配矩阵; $[R_{c_m}, T_{c_m}]$ 为摄像机坐标系到机械臂基坐标系下的转换矩阵。以上矩阵皆可标定求出。

抓取姿态可由 SVD(Singular Value Decomposition)分解方法^[24]求出。在模板点云中取一组点 $P = \{p_1, p_2, \dots, p_n\}$, 由于模板抓取位姿已知,故可根据公式(8)~(10)求解出抓取姿态。

$$q_i = R_1 p_i + T_1 \quad (9)$$

$$j_i = R_{c_m} [R_{\text{match}} (R_{m_c} p_i + T_{m_c}) + T_{\text{match}}] + T_{c_m} \quad (10)$$

$$\min_{R, T} K = \min_{R, T} \left(\frac{1}{2} \|q_i - R j_i + T\|_2^2 \right) \quad (11)$$

其中, p_i 为一组点中的任意一点; q_i 为转换到机械臂坐标系下的点; j_i 为转换到抓取位姿下摄像机坐标系中的点; $[R_1, T_1]$ 为 TCP 坐标系到机械臂坐标系下的转换矩阵; R 为以旋转矩阵表示的抓取姿态。

4 实验结果与分析

4.1 实验设备

本文实验平台由工业相机、投影仪、UR5 机械臂和 RG2 夹持器组成。其中,海康工业相机(MV-CA013-20UM)分辨率为 1280×1024 , 像

元尺寸为 $4.8 \mu\text{m}$, 最高帧率为 170 fps; 镜头为 OPT(奥普特)工业定焦镜头(OPT-C1214-2M), 光圈为 1.4~16, 焦距为 12 mm, C 接口镜头; 投影仪采用德州仪器公司的投影仪(DLP4500), 分辨率为 912×1140 , 像素尺寸为 $7.64 \mu\text{m}$; 机械臂选用丹麦优傲机械臂(UR5), 有效负载为 5 kg, 有效工作半径为 850 mm, 自由度为 6 个旋转关节; 夹爪选用丹麦优傲夹持器(RG2), 有效载荷为 2 kg, 夹持力为 3~40 N, 总行程为 110 mm, 行程时间为 950 ms/200 ms。实验平台搭建如图 6 所示。

4.2 实验结果与分析

在基于结构光的机器人抓取实验中, 首先需要重建出散乱摆放系列物体的三维数据, 且重点在三维数据的获取速度和精度上。因此, 本文选用国际象棋和金属两种物体进行实验。其中, 国际象棋表面形状复杂、没有颜色特征, 且比较小, 使用其他方式很难获取高精度 3D 模型。金属物体由于高反区域及黑暗区域使重建模型容易造成缺失。本文国际象棋和金属物体三维数据获取结果如图 7 和图 8 所示。其中, 使用结构光扫描设备投射条纹, 结果如图 7(b)和图 8(b)所示。根据编解码图案和标定数据得到的物体三维

重建结果如图 7(c)和图 8(c)所示。通过象棋与金属物体进行泊松重建(如图 7(d)和图 8(d)所示), 验证了重建结果的高精度。使用如图 7(c)和图 8(c)所示模板库中的模板点云在如图 7(h)和图 8(h)的三维场景数据上进行匹配, 找到所有匹配到的物体在场景点云中的旋转平移矩阵, 匹配得到的物体用黑色显示, 如图 9 所示。完成匹配矩阵计算后, 即可通过计算获取抓取点完成抓取。图 10 为将匹配结果进行分析计算获取抓取位置姿态后完成抓取动作的过程。

本文使用标准国际象棋进行抓取实验, 其任务是在散乱放置的象棋中, 抓取匹配到的模板象棋, 并以一定姿态放置于指定位置上。其中, 象棋采用人工放置, 且随机摆放。模板匹配得到匹配参数后, 系统按照第 3.3 小节中的程序抓取象棋, 抓取实验过程如图 10 所示。

为分析抓取系统的性能, 验证精度时将机械臂移动到测量位置, 使用示教器读取机械臂在世界坐标系下的标准抓取坐标与姿态。基于 3D 视觉引导抓取系统同样可得实际抓取坐标与姿态。实验设置计算两组坐标之间的距离作为位置精度, 取 Z 轴方向的旋转角误差作为角度误差, 相机获取图片到位姿输出的时间作为处理时间(未



(a) 结构光系统



(b) 整体实验平台

图 6 实验平台搭建

Fig. 6 The experimental setup

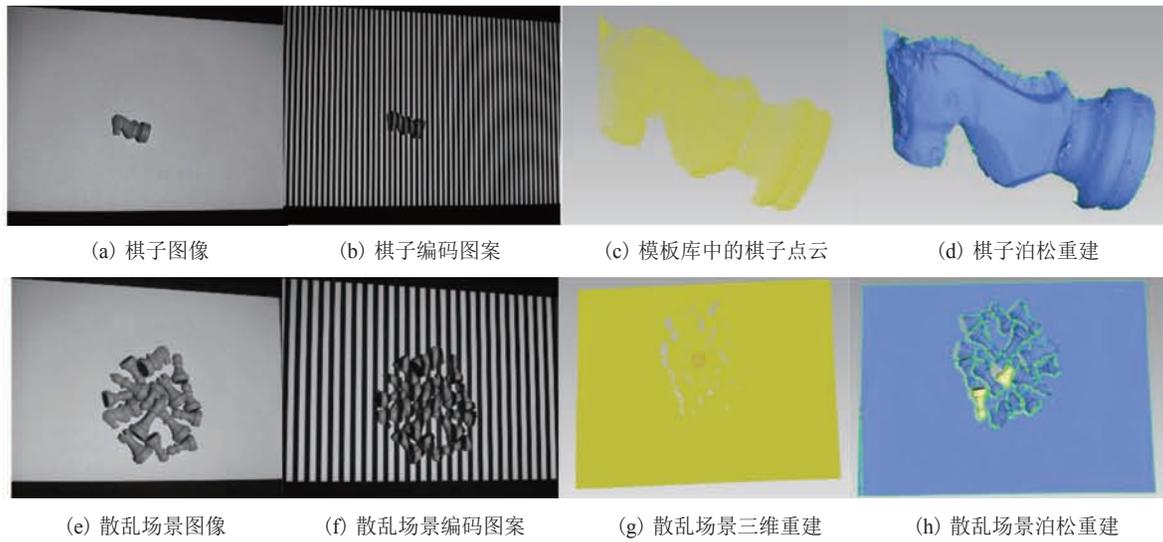


图 7 国际象棋的结构光重建

Fig. 7 3D reconstruction of chess model with structured light

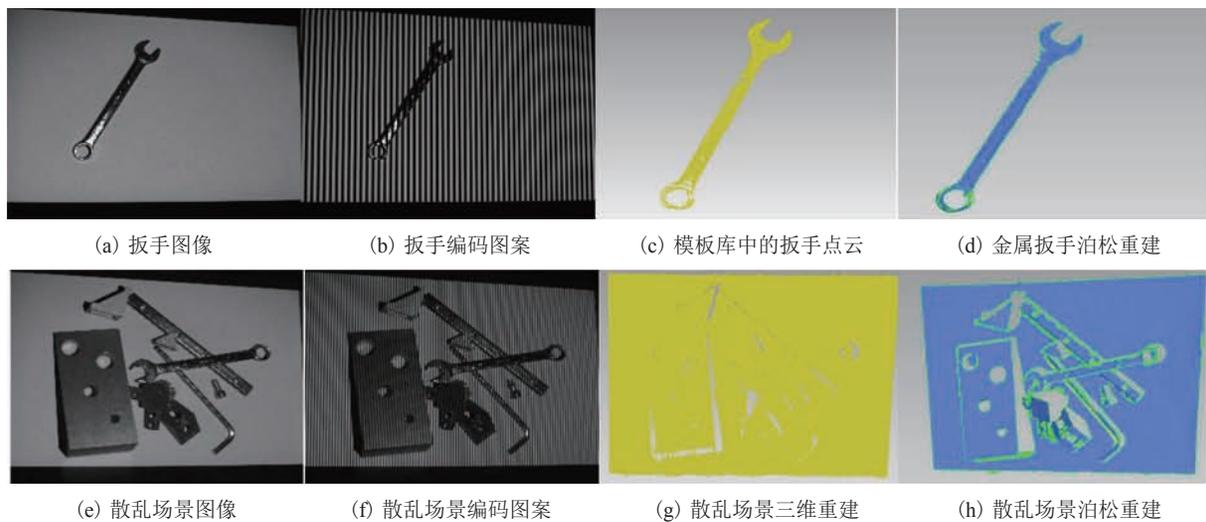


图 8 金属的结构光重建

Fig. 8 3D reconstruction of metal model with structured light

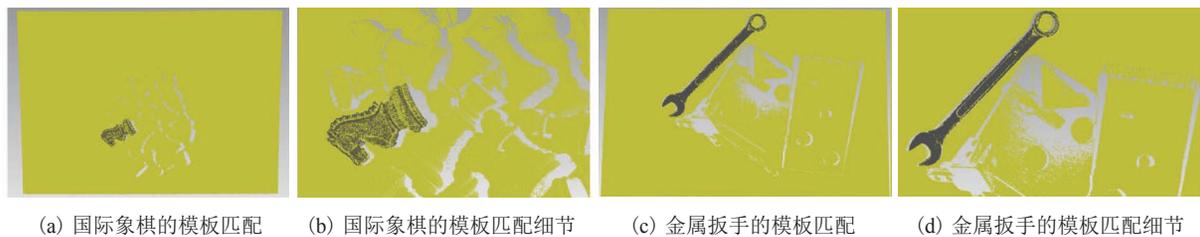


图 9 模板匹配结果

Fig. 9 Template matching results

包括机械臂运动时间)。实验共采集 100 组数据比较以确定其精度，其中 5 组数据如表 1 所示。

从实验结果来看，象棋的准确抓取验证了抓取系统的精度。此法位置精度在 0.05 mm 内，

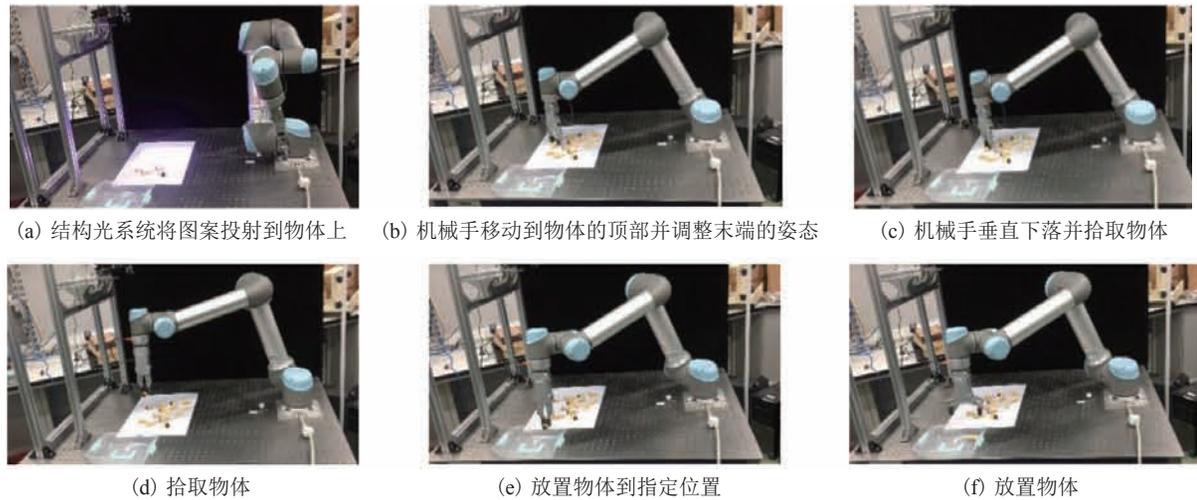


图 10 抓取的关键步骤

Fig. 10 Key steps of picking

表 1 精度与处理时间的比较

Table 1 Comparison of accuracy and processing time

序号	标准坐标 (mm)	测量坐标 (mm)	位置误差 (mm)	Z 轴测量 旋转向量($^{\circ}$)	Z 轴标准 旋转向量($^{\circ}$)	旋转向量 误差($^{\circ}$)	处理时间 (ms)
1	(327.494, 45.247, 12.539)	(327.472, 45.236, 12.519)	(0.022, 0.011, 0.020)	68.53	65.48	3.05	192
2	(297.573, 45.247, 12.284)	(297.569, 45.252, 12.282)	(0.004, 0.005, 0.002)	73.71	54.62	对称物体	132
3	(392.782, 45.247, 12.391)	(392.743, 45.261, 12.384)	(0.039, 0.015, 0.007)	65.86	64.32	1.54	214
4	(425.286, 45.247, 12.424)	(425.269, 45.282, 12.491)	(0.017, 0.035, 0.067)	72.37	76.70	4.33	178
5	(354.624, 45.247, 12.631)	(354.672, 45.273, 12.652)	(0.048, 0.026, 0.021)	75.93	75.63	0.30	195

绕 Z 轴的旋转角度误差在 5° 以内, 处理时间在 0.2 s 左右, 散乱象棋识别率为 100%, 机器人成功抓取率可达 97%。进行金属零件抓取时, 由于重建点云噪点过多, 识别率可达 94%, 机器人成功抓取率可达 86%。黄会明等^[6]提出基于微振镜结构光的方式可获得 0.1 mm 左右的重建精度, 获取位置信息处理时间约为 0.4 s, 但其并未进行姿态精度计算。Li 等^[7]提出一种二维视觉和三维视觉相结合的堆垛目标拾取系统的设计方法, 抓取精度达 1 mm。Lin 等^[25]提出一种基于深度学习的机器人视觉目标识别与姿态估计系统的设计方案, 抓取平均平移和旋转误差分别小于 5.2 mm 和 3.95° 。Le 和 Lin^[26]基于二维像素值提取三维点云数据, 抓取平均平移和旋转误差分别小于 0.23 cm 和 2.26° , 抓取一次时间约为 0.9 s,

不适合小物体抓取。Wang 等^[27]提出端到端的深度学习估计六自由度物体姿态, 机器人成功抓取率仅为 73%, 且训练模型时间长, 故不适合在工业任务中使用。Wang 等^[28]提出一种求取给定类别中不可见对象实例 6D 位姿的方法, 在散乱场景中识别率为 94.7%。Park 等^[29]提出一种新的 6D 姿态估计框架 Pix2Pose, 正确估计姿态百分比为 72.4%, 训练集所需数据量庞大, 且系统框架复杂。Park 等^[30]提出一种多任务模板匹配 (MTTM) 框架, 该框架预测分割率为 78.8%、位姿预测正确率为 52.9%, 完成抓取小型物体任务不稳定。目前, 虽然基于深度学习的识别与位姿估计方案无需提前准备 3D 模板即可完成抓取, 但在工业应用中仍有识别率过低、位姿估计鲁棒性差等问题亟需解决。与现有的方法相比, 本文

的抓取系统可作为高精度工业抓取切实可行的解决方案, 但所提方案基于 3D 模板库限制了使用场景, 仍需进一步研究。

5 结 论

针对可靠、快速和精确机械臂抓取工作, 本文搭建结构光三维重建系统, 利用结构光系统高精度、快速重建的优势, 在 10 ms 内获得了物体的高精度三维信息。为精确计算目标的姿态, 先对点云进行滤波处理, 再采用点对特征的三维点云匹配算法对场景和模型点云进行精确匹配。通过坐标系变换确定目标位姿, 以便机器人正确地抓取目标。该系统能在散乱物体中以准确位姿抓取目标, 可靠、快速和精确地完成机械臂抓取工作, 可应用于分拣、装配、上料等工业自动化领域, 尤其是小目标抓取领域, 以提高现有生产线的性能。

参 考 文 献

- [1] Yanagihara Y, Kita T. Parts picking in disordered environment [C] // International Workshop on Intelligent Robots and Systems, 1991: 517-522.
- [2] Zheng ZH, Ma Y, Zheng H, et al. Industrial part localization and grasping using a robotic arm guided by 2D monocular vision [J]. *Industrial Robot*, 2018, 45(6): 794-804.
- [3] Gu FF, Zhao H, Song Z, et al. A simple method to achieve full-field and real-scale reconstruction using a movable stereo rig [J]. *Measurement Science and Technology*, 2018, 29(6): 065402.
- [4] Huang S, Gu F, Cheng Z, et al. A joint calibration method for the 3D sensing system composed with TOF and stereo camera [C] // IEEE International Conference on Information and Automation, 2018: 905-910.
- [5] Chen Y, Song Z, Lin S, et al. Capture of hair geometry using white structured light [J]. *Computer-Aided Design*, 2018, 96: 31-41.
- [6] 黄会明, 刘桂华, 段康容. 基于微振镜结构光投射器的机器人抓取 [J]. *中国激光*, 2019, 46(2): 93-101.
- [7] Li D, Liu S, Jiang X, et al. Stacking objects picking system based on structured light [C] // International Conference on Information, Cybernetics, and Computational Social Systems, 2018: 174-179.
- [8] Liang XW, Wang HS, Liu YH, et al. A unified design method for adaptive visual tracking control of robots with eye-in-hand/fixed camera configuration [J]. *Automatica*, 2015, 59: 97-105.
- [9] Chen Z, Li S, Zhang N, et al. Eye-to-hand robotic visual tracking based on template matching on FPGAs [J]. *IEEE Access*, 2019, 7: 88870-88880.
- [10] Tsai RY, Lenz RK. A new technique for fully autonomous and efficient 3D robotics hand/eye calibration [J]. *IEEE Transactions on Robotics and Automation*, 1989, 5(3): 345-358.
- [11] Zhuang H, Roth ZS, Sudhakar R. Simultaneous robot/world and tool/flange calibration by solving homogeneous transformation equations of the form $AX=YB$ [J]. *IEEE Transactions on Robotics and Automation*, 1994, 10(4): 549-554.
- [12] Moldovan D, Wada T. A calibrated pinhole camera model for single viewpoint omnidirectional imaging systems [C] // International Conference on Image Processing, 2004: 2977-2980.
- [13] Song Z, Chung R, Zhang XT. An accurate and robust strip-edge-based structured light means for shiny surface micro-measurement in 3D [J]. *IEEE Transactions on Industrial Electronics*, 2012, 60(3): 1023-1032.
- [14] Zhang S, Yau ST. High-resolution, real-time 3D absolute coordinate measurement based on a phase-shifting method [J]. *Optics Express*, 2006, 14(7): 2644-2649.
- [15] 周星, 黄石峰, 朱志红. 六关节工业机器人 TCP 标定模型研究与算法改进 [J]. *机械工程学报*, 2019, 55(11): 186-196.
- [16] Axelsson O. A generalized conjugate gradient, least square method [J]. *Numerische Mathematik*, 1987, 51(2): 209-227.

- [17] Zhang Z. Flexible camera calibration by viewing a plane from unknown orientations [C] // Proceedings of the Seventh IEEE International Conference on Computer Vision, 1999: 666-673.
- [18] Drost B, Ulrich M, Navab N, et al. Model globally, match locally: efficient and robust 3D object recognition [C] // IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2010: 998-1005.
- [19] Chua CS, Jarvis R. Point signatures: a new representation for 3D object recognition [J]. International Journal of Computer Vision, 1997, 25(1): 63-85.
- [20] Park IK, Germann M, Breitenstein MD, et al. Fast and automatic object pose estimation for range images on the GPU [J]. Machine Vision and Applications, 2010, 21(5): 749-766.
- [21] Lowe DG. Object recognition from local scale-invariant features [C] // Proceedings of the 7th IEEE International Conference on Computer Vision, 1999: 1150-1157.
- [22] 易杰, 李蹊, 易辉成, 等. 基于点对特征的三维点云匹配算法 [J]. 传感器与微系统, 2019(9): 34.
- [23] Kang L, Zhao CX, Guo JH. Improved path planning based on rapidly-exploring random tree for mobile robot in unknown environment [J]. Pattern Recognition and Artificial Intelligence, 2009, 22(3): 337-343.
- [24] 王书民, 张爱武, 崔营营, 等. 面向全局坐标变换的点云配准实现及精度评价 [J]. 工程勘察, 2011(11): 62-65.
- [25] Lin CM, Tsai CY, Lai YC, et al. Visual object recognition and pose estimation based on a deep semantic segmentation network [J]. IEEE Sensors Journal, 2018, 18(22): 9370-9381.
- [26] Le TT, Lin CY. Bin-picking for planar objects based on a deep learning network: a case study of USB packs [J]. Sensors, 2019, 19(16): 3602.
- [27] Wang C, Xu DF, Zhu YK, et al. Densefusion: 6D object pose estimation by iterative dense fusion [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019: 3343-3352.
- [28] Wang H, Sridhar S, Huang JW, et al. Normalized object coordinate space for category-level 6D object pose and size estimation [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019: 2642-2651.
- [29] Park K, Patten T, Vincze M. Pix2pose: pixel-wise coordinate regression of objects for 6D pose estimation [C] // Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, 2019: 7668-7677.
- [30] Park K, Patten T, Prankl J, et al. Multi-task template matching for object detection, segmentation and pose estimation using depth images [C] // International Conference on Robotics and Automation, 2019: 7207-7213.