



华南理工大学

SOUTH CHINA UNIVERSITY OF TECHNOLOGY

## 学生研究计划(SRP)项目验收 项目成员个人结题报告

|           |                        |
|-----------|------------------------|
| 参加项目名称:   | 双目视觉引导机器人定位抓取技术研究      |
| 参加项目编号:   | X201910561771          |
| 参加起止时间:   | 2019 年 4 月至 2020 年 3 月 |
| 指导教师姓名:   | 邓晓燕                    |
| 指导教师所在学院: | 自动化科学与工程学院             |
| 学生姓名:     | 王贤枋                    |
| 学生学号:     | 201730646098           |
| 学生手机号:    | 15818726965            |
| 学生所在学院:   | 自动化科学与工程学院             |
| 学生所学专业:   | 自动化                    |
| 填表日期      | 2020 年 3 月             |

教务处

制

2020 年 3 月

# 目录

## 第一章 课题概述

- 1.1. 课题背景.....3
- 1.2. 文献调研.....3
- 1.3. 研究方案.....3

## 第二章 本人工作和成效

- 2.1. 项目组成员间分工 .....5
- 2.2. 本人工作及成效.....5
  - 2.2.1. 本人项目工作概述 .....5
  - 2.2.2. 视觉研究环境的搭建.....5
  - 2.2.3. 点云分割算法研究 .....6
  - 2.2.4. 物体识别算法 .....8

## 第三章 工作总结及感想

- 3.1. 研究工作总结 ..... 11
- 3.2. 工作改进思路 ..... 11
- 3.3. 参与科研工作的感想..... 11

- 参考文献 ..... 12

# 第一章 课题概述

## 1.1. 课题背景

随着科技的迅猛发展，关于机器人控制、深度视觉、机器学习领域的技术日新月异，各式各样的工业机器人和服务机器人也不断被开发和应用于不同的领域的工作中，逐步取代人类，代替人工。对于一些高危高压的工作和不良的劳动环境，人们迫切的需求推动着各种新形式的机器人的研究开发，以帮助人类进行劳动，改善我们面临的各種社会问题。机器人能协助或者代替人类从事许多危险的工作或者一些需要精确计算的工作，因此机器人的自动识别功能和抓取问题的研究尤其重要，故而本项目选取基于双目视觉的机器人灵巧手自适应抓取为研究课题。

机器人定位抓取的研究范围十分广泛，在应用方面，无论是工业机器人还是服务机器人都需要自适应抓取技术。机器人能够完成许多任务有赖于一双能够抓取多种物体的末端抓持器(机器人手)，这是机器人与外界交互的重要终端，特别是对于服务型机器人，抓取技术是其走进真实世界必不可少的功能，我们自然希望赋予机器人与手相似的抓取功能。

基于视觉的机器人可以用来自动识别物体、自动抓取。对于有视觉感知的机器人，可以极好地应用于工业、服务业，解决由于工作环境复杂引起的物品识别难、手臂运动难、抓取位置确定难等问题，实现自适应抓取物体，提高机器人的智能性。

## 1.2. 文献调研

基于视觉引导的机器人定位抓取技术主要涉及到物体的识别分割与定位、机械臂的轨迹规划、机械结构抓取位姿估计等子任务组成。在“基于深度学习的工业机械手抓取检测方法研究”<sup>[1]</sup>中，论文主要阐述基于神经网络的机械臂抓取的物体识别分割和抓取检测过程。表明神经网络在机器人抓取任务中的物体识别、图像分割及抓取检测过程中的可行性。该文在 ROS 机器人研究平台的 Gazebo 构建了仿真平台，简易物体的抓取取得一定的成果。在该类的研究基础上，本项目在进一步的实际环境中进行物体识别分割，并操控 UR5 机械臂中进行物体抓取任务。

## 1.3. 研究方案

本项目为解决机器人在复杂环境下目标精确定位问题，利用机械臂、工业相机、光源、灵巧抓手等搭建双目视觉引导机器人的抓取系统。通过棋盘标定法对机器人抓取系统进行标定，采用颜色识别和边缘分割等方法采集图像并提取工作面，编写上位机处理程序实现图像分割、目标定位和机器人运动控制等功能。

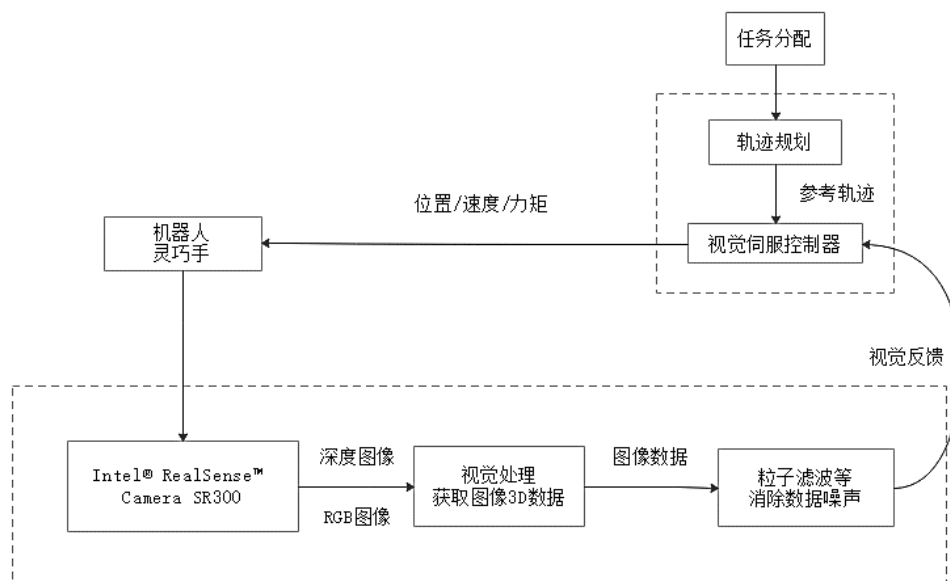


图1 系统整体框图

研究内容主要分为五个部分：

### 1. 制作上位机、搭建基于深度图像的机器人抓取平台

上位机可以实现机械臂参数的可视化调试和运动状态的分析，完善的上位机能够促进项目的进行。利用 Realsense SR300 相机、UR5 机器人和三指灵巧抓手搭建深度视觉的机器人抓取平台，并利用机器人开发系统 ROS 搭建机器人抓取仿真环境。

### 2. 基于深度摄像头的 ROS 系统的图像分割和目标定位算法研究

我们把深度图上的 2D 点转换到世界坐标的 3D 点，再把世界坐标的 3D 点再投影到彩色图像上。对图像进行特征提取。根据 RGB 图像和景深信息，容易得到图像平面内的物体颜色、形状、摆放位置。因此，我们选用了三个特征：颜色、形状、摆放位置利用分类器识别不同的物体。自适应抓取的随机物体特征提取与选择的步骤分为三步：边缘检测与图像分割、获取视野内物体的形状颜色、通过位置信息得到物品所在位置。

### 3. 机械臂轨迹规划研究

在图像中获取到目标位姿后需要控制机械臂运动到目标位姿上方，需要根据起止点进行逆运动学分析，完成机械臂的轨迹规划，从而保证轨迹的连续性，且由于实际场景中比较复杂，在抓取过程中会存在障碍物等问题，机器人运动控制需要考虑避开障碍物。

### 4. 卡尔曼滤波等动态定位算法研究

卡尔曼滤波对连续的运动控制的性能有较好的提升作用，将卡尔曼滤波算法应用到机械臂的运动过程能够改善其控制精度。

### 5. 基于强化学习对自适应抓取方案的研究

在动力学与运动学建模的基础上，为了解决对不同物体的抓取，需要研究灵巧手的抓取姿态等策略，利用强化学习以得到不同目标物体最优的抓取方案。

## 第二章 本人工作和成效

### 2.1.项目组成员间分工

本项目成员总共 7 人，都是自动化科学与工程学院学生，分别为郑安、王贤彬、邱智颖、黎恒钊、黄裕源、李耀康、刘宣彤。具体分工如下：

黄裕源主要负责强化学习等自适应算法的研究与实现。

王贤彬主要负责 3D 图像处理与目标位姿解算。

李耀康、黎恒钊、刘宣彤主要负责目标检测、图像分割算法的研究。

邱智颖主要负责机械臂运动的轨迹规划、卡尔曼滤波等算法的研究。

郑安主要负责上位机的制作与调试，视觉伺服系统的搭建。

### 2.2.本人工作及成效

#### 2.2.1. 本人项目工作概述

机械臂的抓取目标需要首先识别物体的位置甚至姿态，本人的主要工作为搭建物体视觉定位平台，对物体的定位分割算法进行学习研究。故主要工作有一部分：1. ROS 视觉识别环境平台的搭建；2. 立体视觉分割及识别算法的探索。

ROS 视觉识别环境平台的搭建主要包括 ROS 机器人研究平台搭建、深度摄像头的环境配置和深度视觉处理算法处理库的搭建。立体视觉分割及识别算法的探索主要是学习点云相关的视觉处理算法，完成物体的分割和识别方法。

立体视觉的探索建立在对点云的处理，由于机械臂抓取任务场景多为固定背景下对特定物体的抓取，故有必要对背景进行区分，和对不同的物体进行划分，二者均用到不同的点云分割算法。另外传统识别过程为对物体的特征描述和匹配的过程。本文也在探索在该场景下效果较优的描述子以及匹配方式。

#### 2.2.2. 视觉研究环境的搭建

本项目采用 Realsense SR300 相机作为视觉信息的传感器。该传感器通过 RGB 传感器获取 RGB 图像，同时使用红外 (IR) 激光投影仪系统获取到的深度数据，最终将二者信息融合后获得 RGBD 数据。Realsense SR300 配备的 RGB 传感器的分辨率可达 1920\*1080，但是深度信息的分辨率为 640\*480，故本次工作的 RGBD 数据采集分辨率将设置为 640\*480。由于任务的需求和算法计算复杂度考量该分辨率也适当。

由于 RGBD 数据的处理区别于一般图像，其包含深度信息，需要功能完善的点云专用处理库才能较好地完成 RGBD 数据处理并完成识别工作。本项目需要搭建 PCL(Point Cloud Library)作为点云(PointCloud)处理算法库，该算法库能够支持点云信息处理的众多任务。

ROS 作为机器人研究平台，有着完备的信息交互模块以及可视化、仿真功能。搭建 ROS 系统也将能够使算法参数调整和数据可视化变得便捷。故将以 ROS 系统作为交互的手段，搭建完整的深度视觉研究平台。

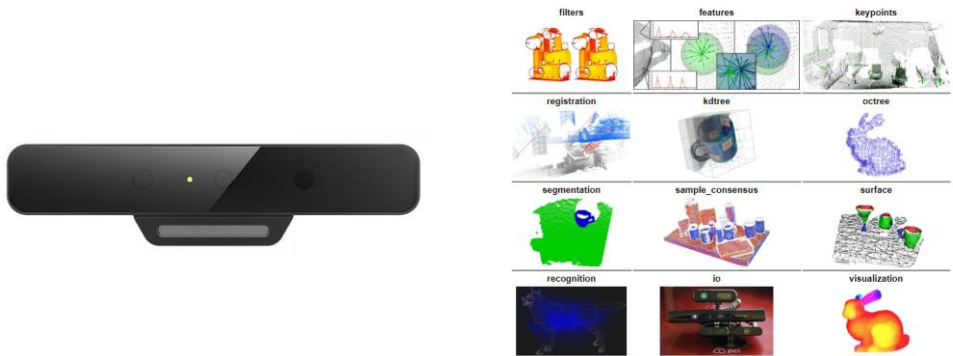


图 2 Realsense SR300 以及 PCL modules

### 2.2.3. 点云分割算法研究

将 RGBD 数据转化为点云数据后，基于传统的特征点描述及匹配算法，可以将已知物体的特征点与全局特征点进行特征描述子的匹配，即能够完成对特定已知模型物体进行识别，故在特定的任务中是可以忽略点云分割的研究过程。但是为了提高计算机的处理速率，提高识别系统的抗干扰特性，本文的思路是将全局点云先进行分割，包括对桌面背景分割和离群点去除、不同物体的分割；将符合目标物体的潜在子点云集进行特征描述，从而对物体进行识别。由于点云包含空间坐标的特殊信息，该类处理算法也主要利用了点云的位置相对信息作为主要的依据，在部分算法中辅助以 RGB 信息作为处理手段。

#### 1. 背景分割算法

在具体的抓取场景中，传统的简单背景分割算法有如下的两种应用：一是当背景为简单的平面、圆球、弧面等组成的几何体（桌面、围墙），通过识别平面和简单几何体作为背景而被去除的方法；二是当背景稍微复杂时，可以通过首先移除物体，记录背景点云，将需要背景去除的点云通过点云配准的方式，将和背景点云重合部分的点云去除，内分割的子点云即为物体点云。

##### 简单几何体的背景去除算法

算法的核心思想是将点云中简单背景（假设为平面）参数测算出来，该类算法可以类比于图像中的霍夫变换算法，在点云中也能够运用类似霍夫变换 (Hough Transform) 来对简单几何体进行识别，但由于点云数据量过大，霍夫变换算法的时间及空间开销均不理想，往往采用基于随机采样一致性算法 (RANSAC) 进行简单几何体识别。

随机采样一致性算法 (RANSAC) 算法，主要利用迭代和抽样的思想来估计选定模型的参数，当相应的模型参数估计达到一定的阈值后即可获得预估模型参数，由于该算法需要确定模型的容忍度和抽样最大次数，该模型需要设置合理的参数方能获得理想的结果。当获取到平面（背景）的模型参数后，在该模型的周围一定欧几里得距离范围阈值的点均被考虑为背景点，去除即可达到背景去除的目的，该方法需要背景足够简单。

### 固定背景去除方法

当在一个固定的场景（背景固定，同时 RGBD 传感器位姿不变）中运行物体抓取任务，该场景可以运用该方法去除背景从而获得较好的物体分割效果。该方法是将无需要抓取物体的背景作为点云 Kd-tree 的索引源，将任务场景中距离该背景点云一定欧式距离阈值内的点视作背景点去除即可。

## 2. 物体分割算法

物体之间的分割算法主要有基于常规欧式距离聚类方法（k-mean 等）、（基于颜色）区域生长、最小图割法以及点云中的基于超体素分割方法等等。每一种分割方法有各自适用的场合。

对于常用的桌面物体分割任务，本文使用基于颜色的区域生长算法 (RGB Region Growing)，该算法对于生活常用物品较为适用；该算法将颜色作为区域间区分的特征，同时利用颜色之间的相似程度作为合并的依据，同时在实现上可以根据每一个聚类之间的距离以及点的个数进行合并运算。

$$CD(P_1, P_2) = \sqrt{(R_1 - R_2)^2 + (G_1 - G_2)^2 + (B_1 - B_2)^2}$$

以上的聚类距离依据式子，同样可以类比到两点云位置均值中，得到两个点云聚类之间的欧式距离值，同阈值比较即可推得是否进行聚类合并操作。当所有迭代操作满足所有阈值条件后聚类结果即可得到。

以下是抽纸和饮料盒之间的分割结果，当两个物体相距一定距离是，分割结果较好。

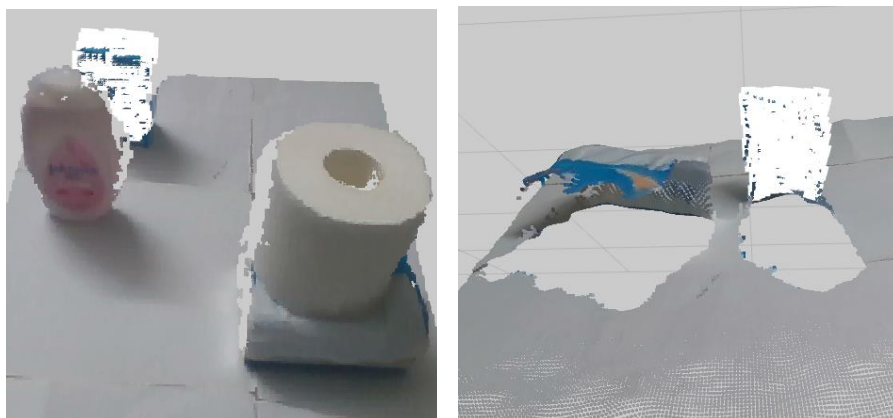


图 3 分割算法结果 1（白色部分分割的饮料盒）

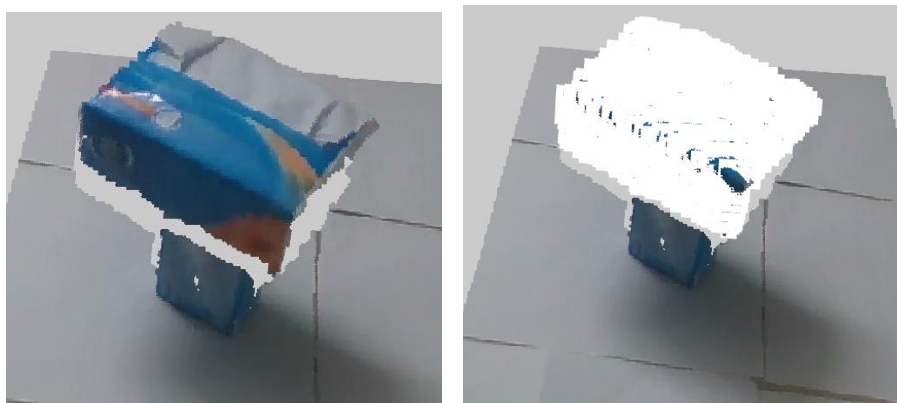


图 4 分割算法结果 2（右白部分分割的抽纸）



## 2.2.4. 物体识别算法

点云物体传统识别方法思路为将 3D 点云的特征描述提取后，再利用对应点匹配过程，将不同点云间的相同物体部分建立对应点对应，然后利用聚类的思想将相同物体的对应点进行聚类，最后通过 ICP (Iterative Closest Point 迭代最近点) 等经典的点集配准算法计算出模型与样本之间的空间变换矩阵。

由于在实验过程中，相同物体的不同点云间特征描述子对应点的计算出现问题。本文的物体识别过程改进为将分割后得到的子点云分别提取出特征描述子，然后与模型的特征描述子进行配准，将配准的得分作为物体属于哪个类别判断待测物体的类别。该改进方法适用于空间中已经存在了待检测的物体，否则仍然会返回一个错误的识别结果，而这一点可以通过设置判断相似的阈值来减少误判的几率。

### 1. 模型的提取

由于需要提前获取被识别的模型，故需要处理得到预知的模型。常用的方法为直接利用建模软件建立 3D 模型，并利用 3D 处理软件进行物体的纹理着色处理，该方法适合于物体几何较为简单的模型，而且能够获得物体的外部纹理信息。

另外一种方式为直接利用深度视觉传感器进行多方位的扫描，然后手工进行裁剪拼接得到完整的模型。该方法能够处理几何和纹理均较为复杂的模型；但是该方案局限于传感器的采集精度，特别是对于小物体，其点云模型的采集主要依赖与传感器的精度。图 5 为采集处理后的其中一个模型的点云。

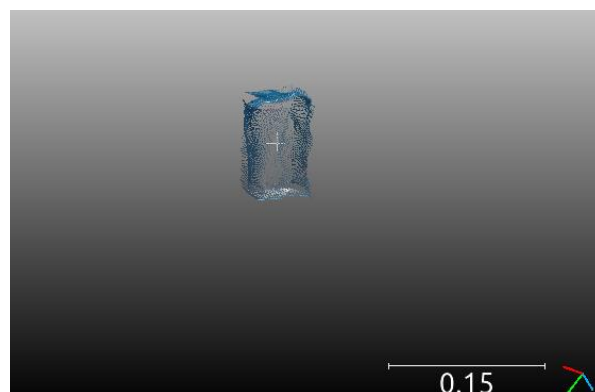


图 5 饮料盒模型扫描及手工拼接结果

### 2. 点云特征描述与提取

利用 3D 特征描述子能够完成对物体的几何拓扑和颜色特征的提取工作。效果较好的特征描述子能够对噪声敏感，辨别能力强。特征描述子概念在图像识别领域已经有广泛运用，立体视觉的特征子有类似的应用；例如图像中的 SIFT 特征子以及 SURF 特征子也被拓展到在点云的特征描述方法中。此外还有点云常用的 VFH、NARF、FPFH 特征描述子。经过多次基本实验效果，本文中使用的 FPFH 特征作为实验的特征描述子，通过分别提取已知模型的 FPFH 特征以及对应待配准的物体 FPFH 特征，进一步通过对应点的配准来达到物体定位的目标。

FPFH 特征称为快速点特征直方图 FPFH (Fast Point Feature Histograms)。通过估计  $k$  邻域或者特定半径内点的特征，通过建立点的局部坐标系进行参数的计算，结合直方图统计的思想完成对某一特征点的局部特征描述。



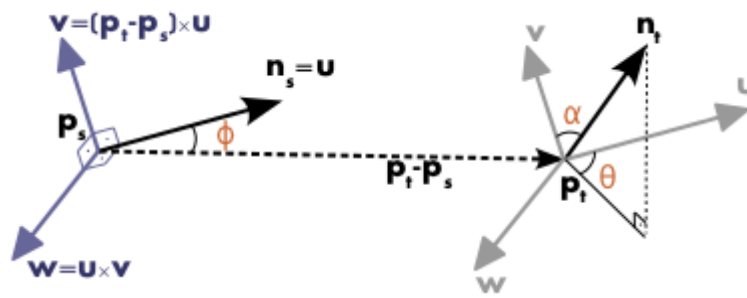


图 6 FPFH 特征计算<sup>[2]</sup>

FPFH 特征的统计如图六所示，通过统计由点和临近点相对旋转和平移  $P_t - P_s$  关系、点所在平面的法向量  $n_t, n_s$ ，统计其坐标系之间的角度关系以及距离关系的直方图结果，该特征维数为 33 维。

### 3. 基于点云特征的配准

常用的配准算法为 ICP 迭代算法，该算法仅需要通过点云之间的几何拓扑信息即可完成对已知两个相似点云进行配准。但由于 ICP 算法迭代需要的计算资源以及空间资源较高，本文采取的方式为将两个点云进行采样一致性初始配准（SAC-IA），该算法为贪婪算法，通过快速地计算拟合得到一个粗配准结果，该算法依赖与原始点云数据以及 FPFH 特征信息。通过采样来快速拟合而不必遍历所有的点云数据。算法的思路为：

- (1) 从点云 P 中选择 n 个尽量相隔一定距离的样本点；
- (2) 对于每个样本点，在点云 Q 中找到满足 FPFH 相似的点存入一个列表中。从这些点中随机选择一些代表采样点的对应关系；
- (3) 计算通过采样点定义的刚体变换和其对应变换，计算点云的度量错误来评价转换的质量；

将以上得到的粗配准变换结果进行若干次的 ICP 算法精配准，即可得到较为满意的模型匹配结果。

### 4. 物体识别实验结果

将以上算法利用 PCL 库以及 ROS 系统中可视化后，得到展示在图 7 的物体识别实验结果。该结果通过分割、特征提取、粗配准、精配准后计算的得到空间变换矩阵，通过反变换得到其空间物体姿态预估结果。

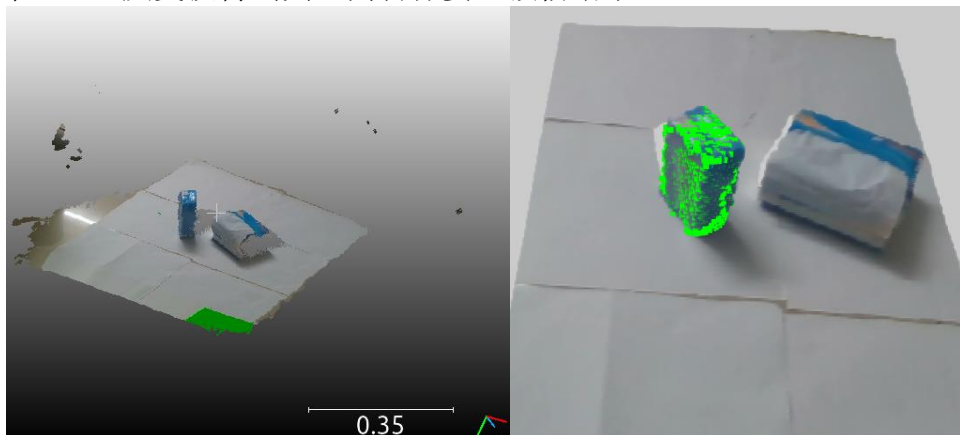


图 7 识别场景（左）及识别饮料盒姿态结果（右）

该识别结果不仅可以给出被识别物体的位置，由于已知了模型的完整点云，即同时在识别的同时已经拟合了物体的姿态。不仅给出识别结果而同时能够给出完整的物体姿态是本次工作的重点，该预估物体姿态能够在机器人柔性抓取的任务中应用。即从有限的单帧场景中，根据已知物体模型，能够获得完整的物体姿态信息。当然该先验知识局限了适用场景，接下来的工作将在此基础上进行改进。包括对特征子的描述上以及识别配准的过程加入更多的思考。

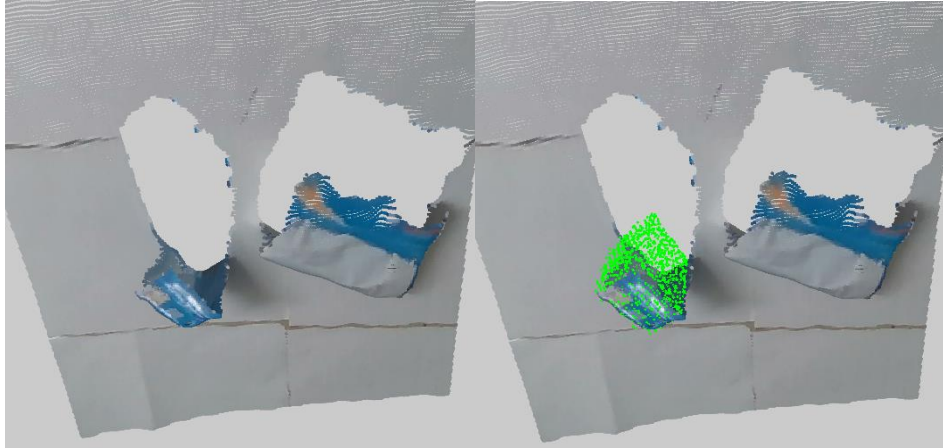


图 8 饮料盒后侧形状的估计

## 5. 基于深度学习的物体 6DoF 估计

除了使用传统方法利用特征对应点估计物体姿态外，基于深度学习的物体位姿估计也是能够得到已知物体模型在空间中的 6DoF 估计结果。该类基于深度学习的训练数据集可以是给出 pose annotation 的 2D 图像，也可以是基于深度相机（雷达）采集得到，pose annotation 后的点云数据。以下是基于 2D 图形估计物体的姿态结果示意，由于网络训练尚未达到最佳，而且拍摄的摄像头参数未知，该结果和实际位置仍存在偏差。该结果体现了物体的姿态能够由 2D 图像获得；同时基于点云数据集的神经网络也能够更好地完成该工作。目前仍在学习当中。



图 9 基于神经网络的姿态估计结果

## 第三章 工作总结及感想

### 3.1. 研究工作总结

本人在该 srp 项目实践了基于点云的物体分割识别配准工作；搭建了 3D 视觉研究环境；实践了基于传统特征匹配的物体识别工作, 对点云进行预处理、分割、特征提取、特征匹配、点云配准以及结果可视化工作。最终基本实现了物体姿态的配准识别任务。

### 3.2. 工作改进思路

本次点云识别工作完成基本的物体位姿识别工作。但仍然存在以下的局限待以改进：

1. 点云的预处理工作完成得较为粗糙。包括对原始点云的噪点处理、曲面回归处理不够恰当，该问题对后续特征点等处理存在干扰。
2. 物体标准模型的制作不够精确。可以从图 5 看出，由于采集传感器的精度问题，小物体的模型表面存在偏差。同时在边缘部分有发散趋势。应该改进方法完成模型建立工作。
3. 特征点提取及匹配算法需要改进。在识别配准的过程中，仍存在较大的概率会匹配识别以及匹配精度不佳，该原因有特征提取算法参数尚且不够好，算法处理速度不够快速等。
4. 算法计算复杂度较高。算法包含了对 3D 点云处理的多个方面，包括各种预处理、索引分割、特征计算、变换过程，处理过程较为复杂，需要改进该算法处理过程。

### 3.3. 参与科研工作的感想

本次关于机器人抓取的视觉识别研究工作周期长，从查阅文献以及前人的相关工作开始，逐渐认识到了 3D 视觉识别工作的复杂性，同时该任务经过近十年的迅猛发展，传统的视觉识别方法日益被更新，基于深度学习的物体 3D 识别方法也是研究热点。选择使用传统的方法完成该工作，是为了对点云数据有更加深层的认知，以及了解传统方法在处理 3D 视觉任务的特点以及局限性。

本次项目实践在完成前期调研工作之后，在课余时间花费约 3 个月的时间完成本次报告的实验内容，包括对环境搭建、算法效果实现、探究和识别过程完整可视化工作。在学习的过程中探索了许多点云处理过程，很多时候效果都不够理想，由此也发现虽然点云相较于图像包含有更多的信息，但其处理的难度也提升

许多。在过程中放弃了许多曾经认为不错的想法，也促使自己在这部分的学习中要更加细致。在项目结束时仍然自我感觉仅仅是完成了其中部分的入门学习，之后本部分的学习和对模型的改进工作仍任重而道远。

同时由于传统的识别工作存在以上分析的局限性，将开始探索基于深度学习的物体 6DoF 估计研究工作。该类方法已经经过了若干年的发展，在自动驾驶场景和机械臂抓取场景已经有了稳定的效果。

以上为本人在该项目的工作总结陈述，总结了过去一年在 srp 项目中实践经验，同时也为接下来的点云处理学习指明了研究方向。

## 参考文献

- [1] 皮思远. 基于深度学习的工业机械手抓取检测方法研究[D].华南理工大学, 2018.
- [2] Radu Bogdan Rusu. Semantic 3D Object Maps for Everyday Manipulation in Human Living Environments [D]. University of Munich, 2010.