

基于Kinect的服务机器人任务操控学习方法的研究

陈楠^{1,2} 胡颖^{1,2} 张俊^{1,2} 夏泽洋^{1,2}

¹ (中国科学院深圳先进技术研究院 深圳 518055)

² (香港中文大学 香港 999077)

摘要 本文围绕中国科学院深圳先进技术研究院认知技术研究中心自行研发的搭载有Kinect传感器的服务机器人操控平台,从Kinect传感器带来的彩色图像、深度图像和真三维点云信息中提取基于图像的2D和基于点云的3D特征,并将它们进行融合,作为待识别物体几何模型归类的依据,为手爪选择合适抓取姿态提供判断准则。同时结合人体示范学习框架(Learning from demonstration, LFD),研究了一种通过提高机器人的认知学习能力来完成人类生活环境中室内日常物品操控任务的方法,如:自行识别门把手的位置并完成开门动作,从橱柜中识别出目标物,抓取目标物体并送到指定目标地点等。最后,我们通过实验验证,该方法能够保证服务机器人成功抓取一些类似圆柱状、长方体等几何形态的物体并能在抓取之后顺利完成与周围环境进行交互过程中的轨迹规划这一复杂任务。

关键词 Kinect传感器; 物品操控; 服务机器人; 认知学习

Researches on Methods about Robot Manipulation Using Kinect

CHEN Nan^{1,2} HU Ying^{1,2} ZHANG Jun^{1,2} XIA Ze-yang^{1,2}

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

²(The Chinese University of Hong Kong, Hong Kong 999077, China)

Abstract This paper presents the integrated service robot control platform developed at the Cognitive Technology Research Center in Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences. The platform is equipped with a sensor named Kinect. We can extract 2D and 3D features from the Kinect's color images, depth images and point clouds, then classify the objects' geometric models and choose the right grab attitude for the robot's hand after those two types of features have been fused. Following the theory of Learning from demonstration (LfD), we have developed an approach to improve robot's cognitive learning ability for indoor object manipulation tasks in everyday household environments, such as identifying the door handle's location and opening the door, recognizing the targets in the cupboard, grabbing the object and sending it to the designated location, etc. Finally, we design an experiment to prove our method, which can teach the robot to grasp some cylindrical, rectangular or other geometrical objects and then finish the complicated task of planning the motion trajectory while interacting with the surrounding environment.

Keywords kinect; object manipulation; service robot; cognitive learning

1 引言

智能服务机器人是在非结构环境下为人类提供必要服务的多种高技术集成的智能化装备^[1]。其中,应

用在半结构化环境中的室内服务机器人是近年来智能服务机器人研究的一个热点。不同于应用在结构化环境中的工业机器人,室内服务机器人面对的是以满足人类居住舒适性而设计的室内环境。此类环境存在复杂多变的因素,如:(1)不确定的游离的“障

基金项目:深圳市基础研究重点项目(JC201005270375A)。

作者简介:陈楠,硕士研究生,研究方向为计算机视觉、机器学习、机器人学习, E-mail:nan.chen@siat.ac.cn; 胡颖,研究员,博士生导师,研究方向为并联机器人、医疗机器人; 张俊,博士研究生,研究方向为机器人控制、服务机器人; 夏泽洋,副研究员,研究方向为机器人与生物力学。

碍物”的存在，如：人，宠物；（2）变化的工作环境，如：橱柜，书桌，冰箱，门等家具摆设的位置和姿态；（3）种类繁多的任务操控对象，如：装水的水壶，喝咖啡的马克杯，炒菜的锅铲等有不同外形和材质的物品。上述因素给服务机器人在处理日常物品操控的任务时带来了诸多具有挑战性的难题。与此同时，近年来计算机视觉技术得到了快速发展，特别是各类视觉传感器的应用，已经受到了机器人认知学习研究领域的高度关注，越来越多地被各国研究人员运用到解决机器人任务操控的问题中。

为解决服务机器人任务操控规划系列问题，完整的机器人操控平台通常都包含在众多完全集成设计的移动的或非移动的机器人系统中。近年来出现了一些典型的机器人操控平台。MIT的人形机器人小组发明了一款防人型机器人Domo^[2]。Domo的手臂上安装了串联弹性力传感制动器（Series Elastic Actuators, SEA），同时采用视觉感知，力反馈协同机械手控制，它可以在不依赖先验的感知模型和完善的规划预处理的前提下安全地完成一些精细操作任务。UMass Amherst的研究者设计了好几款机器人操作平台，如单臂的UMan^[3]，双臂的Dexter^[4]，他们都使用了巴雷特技术（Barrett Technologies）的7自由度的WAM（Whole Arm Manipulators）机械臂和三指手爪这一组合来作为操作器。其中，UMan的腕部还设置了一个Sony XCD710CR彩色相机，可以监控手臂运动的位置和方向；Dexter的双臂之间严格地通过准确的力感应来协同工作，它的视觉监控系统由搭载在TRC BiSight头部位置上能提供4个自由度的两台相机组成，多传感器的协同工作帮助它完成更加灵巧的操作任务。Carnegie Mellon University的机器人研究所利用Bumblebee2立体相机，SR4000 TOF相机，加上Barrett WAM和3指Barrett Hand组合成的机器人操作平台实现了在抓取和操控日常物体时主动力感应和运动之间的协同调控机制^[5]。其中，该学校的个人机器人实验室设计的室内移动服务型机器人HERB^[6]，则是搭载了包含视觉，激光，触觉传感器，机械臂在内的多传感集成系统，它通过融合多传感器信息、开环地执行规划子动作等功能模块，能完成诸如“去某地方取回某种物品”的指令。不同于前者，Willow Garage公司的明星产品PR2^[7]，它将2010年微软公司推出的3D Kinect相机作为视觉输入设备（图1）^[8]，利用其便捷一体化的设计为机器人操控任务提供丰富有效的信息源数据，是完全架构在机器人操作系统（Robot Operation System, ROS）之上的一款提供科学研究或开发应用的综合性平台。

加拿大英属哥伦比亚大学的Marius Muja和Willow Garage的Radu Bogdan Rusu^[9]等人就基于装载有Kinect传感器的PR2平台，设计出多种新型的机器人感知分类器，实时性的解决物体识别和姿态辨识等问题，辅助PR2完成多种复杂的日常室内操控任务。

上述各系统在解决特定的操控问题上已经取得较好的结果。但是有两个问题仍然有待进一步研究：

（1）机器人操控规划问题通常都包括一系列子问题（诸如：针对特定的操控对象选择合适的抓取位置，选择合适的抓取姿势，如何依据机械臂的已有模块结构确定它的演变轨迹等），已有的工作大多倾向于有针对性地解决孤立的子问题；（2）已有的机器人平台所使用的关键器件，如上述系统中所使用的3D激光扫描仪和3D ToF型相机等多传感器搭载的视觉系统，对比于仅含Kinect传感器的视觉系统，通常价格成本高而且后期多传感器之间的信息融合算法较复杂，不利于服务器人未来的广泛应用。

因此，本文从计算机视觉的角度出发，使用Kinect这一新型的传感器装置，结合传统的视觉识别、定位以及追踪技术，将提升服务机器人任务操控的学习能力，完成一系列室内日常物品操作任务这一难题作为一个整体来进行研究的首次尝试。



图1 Kinect

本文后续章节安排如下：第2节是结合2D的SURF特征和3D的VFH特征完成对静态非联动物体实现识别和抓取操作；第3节则是研究对联动物体的运动轨迹进行动态规划和操控；第4节即设计实验来验证本文中构建的机器人学习框架；最后给出全文的总结和展望。

2 静态非联动物体的识别和抓取

研究室内服务机器人操控的对象时，我们发现在人类居家和工作等室内环境下，日常生活中处理的对象大致可以分为静态的联动物体和动态的非联动物体两大类。前者包括水杯，书籍，纸盒等，这一类物体的共同点即是，当机器人完成抓取工作之后，物体即

可以和机械臂末关节视为一个整体, 其后续的运动轨迹即和机器臂末端的运动轨迹保持一致; 而后者包括各种门上的门把手, 抽屉上面的拉手, 甚至电器件上面的开关, 这一类物体在被机械手爪成功抓取之后, 其后续的操控运动轨迹不仅仅与机械臂的工作区间及自由度有关, 还要受限于与目标物联系紧密的主控物体的活动区间及活动轴的位置与方向等客观约束因素, 如: 安置门把手的门, 固定拉手的抽屉, 及配备开关的主电器件。本小节主要研究基于 Kinect 的静态非联动物体识别, 并用它来辅助平行式二指手爪完成抓取操作的相关技术。

2.1 特征的选择及提取

对于 2 维的图像信息, 图像特征的选择对于快速、高效的图像识别算法的实施具有重要的影响, 不合适的特征选择不能滤除无效的噪声干扰, 甚至因为不能提供充分有效的目标信息而直接导致视觉识别失败。常用于图像匹配的局部特征子有 SIFT^[10], SURF^[11], ORB^[12] 等角点特征。同时, 基于真三维点云, 提取 3D 特征即是指从周围空间的点中获取信息, 从而表征点云中某个确定的三维点或三维位置的几何模式。全局的 3 维特征有 VFP^[13], CVFH^[14] 算子等。为满足室内服务机器人操控任务实时性的需求, 本文选取 2 维 SURF 特征与 3 维 VFH 特征相互映射进行定位, 相关的流程框架如下图 2:

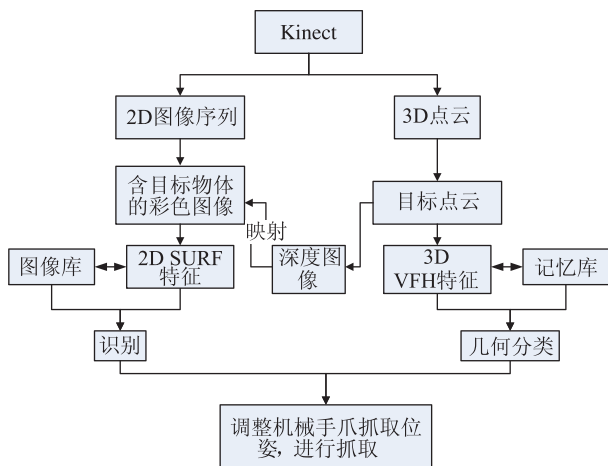


图2 基于 Kinect 的抓取操作流程

2.2 机械手抓取位置和抓取姿态的确定

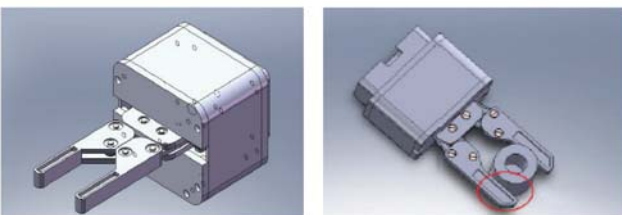


图3 (a) 平行式二指手爪 (b) 机械手爪抓取圆柱状物体

我们所研发的服务机器人操控平台上搭载的是平行式二指手爪(图3), 当机械臂运动到预抓取的范围后, 不同于多指灵巧手可以通过在线手指接触力的变化分析研究目标物体的几何特征, 从而自行找到满足形封闭条件的抓取点, 二指手爪与目标物体之间的抓取姿态和受力状态直接关系到抓取的稳定与否。在考虑到日常环境中, 服务机器人操控的对象大多为刚性目标物(如: 杯子, 纸盒, 书籍, 勺子等), 我们对模型进行简化处理, 采取开环操作。利用 Kinect 相机先在识别出目标物体之后, 对目标物进行粗建模归类, 对于规则几何形态的物体采取几何中心位置(或质心)作为抓取实验中的施力点。同时, 在手指上安装 FlexiForc 触觉传感器, 采集手指上的压力分布状况, 通过辅助性检测抓取过程中触觉反馈的力来对手爪夹紧位移量进行控制。这种将视觉和触觉传感结合起来分析抓取过程状态的方法, 能很好地提高抓取的准确率和稳定性。

2.3 具有学习能力的智能识别和抓取框架建立

本文中将有学习能力的识别框架的范畴进行如下定位——过去对于物体 A, 机器人已经采取甲种位姿来进行操作; 对于物体 B, 机器人采取乙种位姿来进行操作。这部分记忆即形成, 可以当做机器人的操作经验之一存放在机器人的记忆库中。当机器人面对新的操作对象 C 时, 机器人会先在记忆库中搜寻与 C 几何形态最相近的物体, 不妨设作物体 A, 则机器人可以在免除人为设定先验模型的前提下采取甲种位姿来完成对目标物 C 的操控, 同时新的操作经历将会被当作新的经验值存入机器人记忆库中, 为未来完成新一轮新的操控任务时提供参考借鉴。

这里, 我们选择使用开源的 FLANN (Fast Library for Approximate Nearest Neighbors) 框架^[15], 它通过构建 octrees 或 KD-trees 结构为我们在解决从日渐庞大的机器人记忆库中搜索与待操控的目标物相似度最大的记忆片段时, 提供一个在有效时间内快速得到相对的最优结果的解决方案。

3 联动物体运动轨迹的规划和操控

通过观察总结我们人类日常的生活经验, 我们发现人类在面对许多日常物品的复杂操控运动时, 不论任务操控者有什么差别(包括年龄, 身高, 性别), 也不论是否是在同样的室内环境中, 这些复杂任务的操控都可以分解成一系列有先后执行顺序的旋转和平

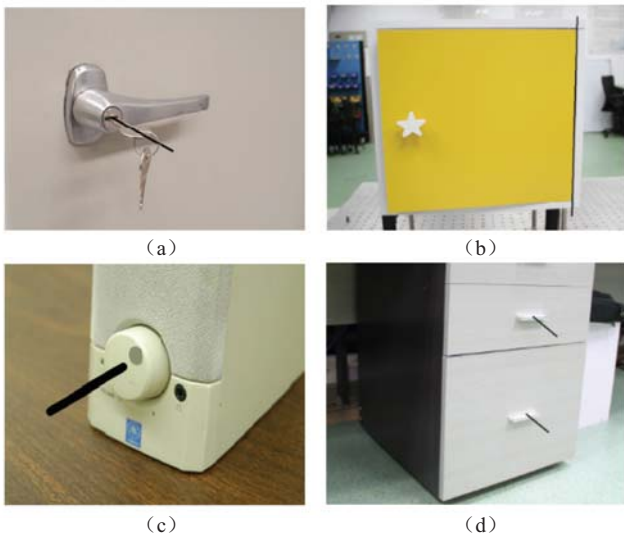


图4 室内日常联动物体

移子运动，这些原子型运动即可以构建成我们的操控基元。如图4(a)，旋转门把手即可抽象成一个按顺时针旋转四分之一圆弧的旋转运动；(b)开橱柜的门即可视为一个在平行于地板的平面上旋转90°的旋转运动；(c)开音响的开关则可抽象为一个垂直于按钮平面向里的一个平移动作；(d)拉抽屉即可视为一个垂直于抽屉平面向外做的一个平移动作。而对于较复杂操控任务，我们则可以将其描述成一个由多个有先后顺序的旋转或平移型操控基元组织成的向量。这种将原子任务操作抽象成一个类别的基元，充分考虑到人类手臂与机械臂无论是在感知未知物体时的建模，运动控制时的力反馈，还是执行操作时的自由度及正常工作范围等客观因素上的差异，前者可离线实施，发生失误时也不会给后者带来严重后果的设计思想，在描述服务机器人日常处理不同类型的操作任务时具有一定的普适性。一些有关操控基元的相关工作，如：参考文献[16]从分析刚体之间相关运动的角度来分析任务操控基元类别；而参考文献[17]则从抽象的任务描述给出了操控基元的定义，同时基于它结合待操控物体的几何模型和机器人手爪给出了任务描述子的定义。

本文中则基于人体示范学习框架(Learning from demonstration, LFD)^[18]，将机器人通过学习来完成对复杂任务操控的过程分解成人类示范和机器人操控两个子过程。其中，前者通过Kinect在一些热点捕获人手的3维位置，构成由一系列点组成的示范基元；后者在综合考虑了机器人与人类在自身组成结构和对外界的感知方式之间的不同因素的同时，基于前面获取的示范基元，计算出不同操控运动轨迹中的平移或旋转基元轴的位置和方向，得到适宜机器人操控的操

作基元，来指导机器人完成计划中的操控任务。大致的总流程框架如下图5：

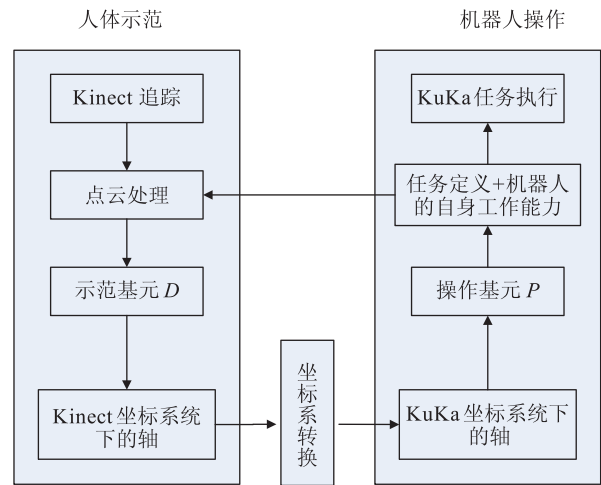


图5 基于人体示范学习框架的机器人开门操作

3.1 示范基元和操控基元的设计

(1) 示范基元

PrimeSense公司提供的NITE^[19]中间件，配合Kinect的官方驱动OpenNI^[20]使用，借助于Kinect实时生成的深度图像，可以从3D场景中提取人体的15个骨骼点，其中包括左右两个手部关节点。当手指夹住某一操控对象(如：门把手，抽屉拉环等)，手背在与该对象间隔8厘米左右的位置保持静止一段时间，Kinect即可以记录下该位置对应的一团点云。结合Mean-shift算法^[21]，针对每一团点云(可视为一个类别)，可以提取一个类表征点(即对应点云中点分布密度最大的位置)。由此，一个示范基元可被设计为：

$$D = \langle p_1, p_2, p_3, \dots, p_n \rangle \quad (1)$$

其中， $p_i = [p_{x_i}, p_{y_i}, p_{z_i}]$ 即对应示范过程中Kinect记录下的某一点3D位置。特别的，即为 p_1 完成该旋转或平移基元动作时的初始位置， p_n 为完成该基元动作的终点位置。

(2) 操控基元

对于上面设计的旋转或平移型的示范基元，我们都可以提取与其对应的示范子动作轨迹中隐含的旋转或平移轴。例如：开门动作中的门轴、拉开抽屉过程中垂直于抽屉表面的无形的轴。显然，对于旋转操作，轴垂直于示范基元中包含所有点所在的平面，轴与平面的交点对应着旋转圆心；对于平移操作，轴平行于示范基元中包含所有点所在的直线，轴心即对应着示范基元中的初始点位置(也可以是直线上任何一点的位置)。由此，一个操控基元可被设计为：

$$M = \langle pos, dir, m_s, m_e \rangle \quad (2)$$

其中, $pos = [p_{x_{axis}}, p_{y_{axis}}, p_{z_{axis}}]$ 是指轴心的位置; $dir = [d_{x_{axis}}, d_{y_{axis}}, d_{z_{axis}}]$ 是执行基元操作时轴的方向; $m_s = [m_{x_s}, m_{y_s}, m_{z_s}]$ 是机械臂执行单个基元操作时的开始位置, 对应 D 中的 p_1 ; $m_e = [m_{x_e}, m_{y_e}, m_{z_e}]$ 是机械臂执行单个基元操作时的终止位置, 对应 D 中的 p_n 。

(3) 任务描述子

当服务机器人需要处理一个复杂任务链时, 我们可以将它的运动轨迹规划为几个子操控基元的组合, 例如: 移动型服务机器人从一间屋子进入另一间屋子时, 需要在准确抓取了门把手 (假设是如图4 (a) 中的门把手类型) 以后先顺时针旋转门把手, 打开门锁, 然后绕着门轴推 (或拉) 门把手, 旋转小段距离后门即被打开, 这一系列子操作完成后, 机器人就可以完成操控任务, 进入到目标房间里。由此, 我们将复杂任务描述子定义为:

$$T = \langle M_1, M_2, M_3 \cdots M_n \rangle \quad (3)$$

其中, $M_i (i=1, 2, 3 \cdots n)$ 是指复杂任务 T 轨迹中所包含的某一个操控基元, 同时 M_i 中的终止位置 m_e 对应着 M_{i+1} 中的开始位置 m_s 。

3.2 基于人体示范学习框架的运动轨迹生成

(1) 示范基元的生成

考虑到利用 NITE 中间件从 Kinect 采集的深度图中提取手的位置时, 来自周围环境结构的变化, 操作者的身体结构变动, 目标物上反射的光照变化等参数带来不可避免的噪声影响, 使得使用 Kinect 采集即使在人眼中看来相对静止的 3D 手部位置时, 也不会只是包含一个点的准确坐标, 而是以一团点云的形式被记录下来。为获得每一团点云中密度分布最大处的位置, 我们使用自适应带宽的 Mean-shift 算法, 它能使用无参数模型的密度梯度估计的迭代方式来获得最大密度估计, 即示范基元 D 中的 p_i 。

具体步骤如下:

①通过 PCA 算法分析点云中各点的坐标, 将 3D 空间中的点投影到坐标成分变化最大的平面上, 这样对应每一个示范位置, 我们将获得一系列的 2D 点集 $X = \{x_1, x_2 \cdots x_i \cdots x_n\}$;

②用一个固定带宽 h_0 初始化带宽 h_i 和由高斯密度核构成的密度估计函数 \hat{f} , 其中 $\hat{f}(x) = \frac{1}{nh_i^d} \sum_{i=1}^n K(\|\frac{x-x_i}{h_i}\|^2)$;

③计算中间变量 λ , 其中 $\log \lambda = n^{-1} \sum_{i=1}^n \log \hat{f}(x_i)$;

④对于每一个 x_i , 计算 $h_i = h(x_i) = h_0[\lambda / \hat{f}(x_i)]^{1/2}$;

⑤先使用 x_1 来初始化最大密度估计位置 x_j , 然后使用公式 (4) 计算 x_{j+1} 直到结果收敛, 即 $\|x_{j+1} - x_j\| < \varepsilon$, 否则则返回步骤 2; x_i 所对应的即为所求 p_i 。

$$x_{j+1} = \frac{\sum_{i=1}^n \frac{x_i}{h_i^{d+2}} g(\|\frac{x_j - x_i}{h_i}\|^2)}{\sum_{i=1}^n \frac{1}{h_i^{d+2}} g(\|\frac{x_j - x_i}{h_i}\|^2)}, j = 1, 2, \cdots \quad (4)$$

(2) 操控基元的生成

分析操控基元的组成成分, 我们发现其中最为关键的即是关于基元轴的计算。通过基元轴的转换, 我们可以排除人手和机械手爪在抓取过程中姿态和抓取点位置的不一致, 机械臂和人手臂在运动过程中各关节轴之间的耦合调度不相同等因素带来的干扰。类似于 [17], 我们使用 PCA 算法来区别子动作是旋转型操控基元还是平移型操控基元。假设示范基元中点集的特征值是 $\lambda_1 \geq \lambda_2 \geq \lambda_3$, 定义一个阈值 $T_0 \in [0, 1]$, 则有判断准则:

$$Motion = \begin{cases} rot & \text{if } \lambda_3 < T_0 S \ \& \ \lambda_2 > T_0 S \\ transl & \text{if } \lambda_1 > (1 - T_0) S \end{cases}$$

其中, $S = \sum_{i=1}^3 \lambda_i$ 。

定义完操控基元的类型后, 基元轴就比较容易得到——旋转轴的轴心即为 X 系列投影点所对应的圆心, 平移轴的轴心即为示范基元 D 中离散点对应的直线上的任一点, 同时, 结合操控基元的起始点和终止点位置, 轴的方向也可以实时地获取。

考虑到人手操控真实的物体, 可能会因为手势的差别产生一个固定的距离偏离值, 不妨设作 $\nabla l = [\nabla l_x, \nabla l_y, \nabla l_z]$, 所以:

$$m_s = p_1 - \nabla l;$$

$$m_e = p_n - \nabla l;$$

其中, 当人手采取如图 7 中所示的两种姿势示教时——手背近似平行于 Kinect 镜头所在的平面, 有 $\nabla l_x = 0, \nabla l_y = 0, \nabla l_z \in [5cm, 10cm]$ 。

3.3 具有学习能力的机械臂运动轨迹规划和操控框架的建立

通过人体示范学习过程, 机器人完成对操控基元的获取, 进而基于任务描述子的定义, 规划出处理复杂操控任务的策略并执行操作。实验证明, 经过学习过程, 机器人能够自主地完成相似类别的门 (如: 具有一个旋转轴的微波炉门, 冰箱门, 橱柜门等)。参考文献 [17] 中使用将复杂轨迹通过分类算法分割成多

个基元轴控制的一系列基元操作，将抽象的思想扩展到一些复杂精细的操作任务上，机器人通过对门旋转系列运动的学习，完成一些更精细的操控任务（如：依据一个旋转轴拧开水瓶盖）。

4 实验

为了讨论我们的方法能提升服务机器人的任务操控的综合学习能力，本文中我们设计了一个复杂任务操控的实验来讨论这个框架的普适性。整个实验过程可以划分为以下几个阶段：第一阶段，Kinect识别门把手的位置和几何形态，确定手爪抓取位置和姿态；第二阶段，机械臂通过人体示范学习框架规划出运动轨迹，并通过点到点的运动方式打开橱柜门；第三阶段，Kinect识别橱柜中的目标物（如：可乐罐）——不存在机器人之前的记忆库中，并在记忆库中搜索与此目标物几何形态相似的操控记忆片段，从中选择合适手爪位姿来完成操控任务。设计本实验的特点在于，借助于本论文中阐述的思想，它可以将机器人操控任务进行模块化分离，使之包含基于Kinect目标识别和位置测量模块，手爪抓取状态检测模块和机械臂运动控制模块三大部分，降低了操控复杂任务时多传感器信息的耦合分析的复杂度。

我们的实验平台主要包括三个单元器件——一个Kinect 3D摄像机，一个含有6自由度的kuka机械臂，一个贴有FlexiForced触觉传感器平行二指手爪。软件框架上，一方面，我们使用开源项目OpenCV^[22]，PCL^[23]来处理Kinect采集的局部图像和点云数据，同时帧地映射两种特征的几何位置，辅助手爪抓取状态的选取；另一方面，我们借助于NITE^[19]中人体骨骼建模技术，使用Kinect追踪人手的移动变化位置，再经过Kinect坐标系与kuka坐标系之间的转换，指导kuka机械臂完成开门操作。

4.1 门把手和橱柜中目标物的抓取

这里主要涉及的是实验的第一、三阶段。普通物如门把手的识别，在计算机视觉中已经有多种方法，我们采用的是简单直接的点云建模。鉴于门把手的形状类型有限，使用平行式二指手爪完成此类任务时我们假定机器人只需要完成最简单的位移靠近，然后手爪闭合等子操作。因此我们只需从场景中分割出目标点云，然后对其几何造型进行建模归类（如：是属圆柱状还是长方体状），最后计算出模型的质心位置提供给手爪，即可完成操作。

另一种抓取操作，如：让机器人取出橱柜中的饮料（形状未知）。假定机器人以前完成过从橱柜中取过维生素饮料的任务，并将完成此任务时的维生素饮料瓶的几何形状以及抓取此饮料瓶的姿态——可能涉及复杂的位移和角度调整，当作一个记忆片段存储在记忆库中。因此，当打开橱柜门，通过SURF算子识别后发现是可乐时，机器人先要在记忆库中搜寻与可乐罐几何形态（表现为VFH特征）最相似的物体（假设是维生素饮料瓶），然后从记忆库中提取与之匹配的抓取姿态，实时地完成对可乐罐的抓取任务。

如图6（a）所示，物体下面的数字越小，表明两物体的几何形态差异越小——即维生素瓶与保温杯相比较，前者与可乐罐子的形态更接近；（b）表明从记忆库中提取经验值后，机械手爪完成即定的抓取任务。

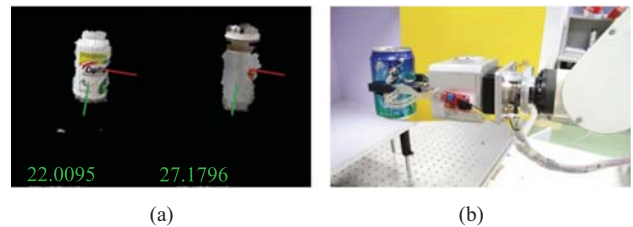


图6 (a) 基于VFH描述子的物体几何形态描述和 (b) 平行式二指手爪对可乐罐的抓取

4.2 机器人学习开门实验

本小节主要阐述的是实验的第二阶段，这里我们选择构建旋转型基元作为实验的代表。它分作人体示范和机械臂执行两个子模块，二者通过基元轴的计算相关联。

在人体示范阶段，当人手以图7中两种姿势近距离操控物体时， ∇I 即可被确定。这样，针对每一个热点位置，就可以收集一团点云，最后利用Mean-shift算法分析出该点云中的密度分布最大的位置，如图8（a）

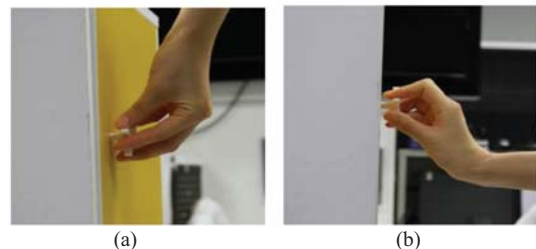


图7 两种抓取门把手的姿势

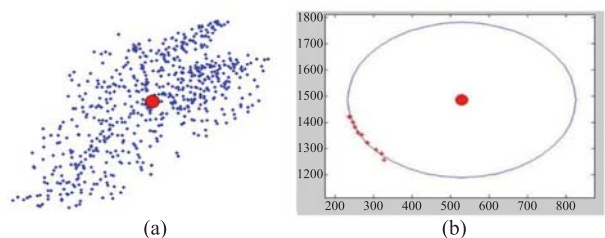


图8 Mean-shift计算轴心位置

图中红点所示; 多个离散位置的中心点即可得出基元轴的轴心位置, 如右边 (b) 图中中间的圆心点。

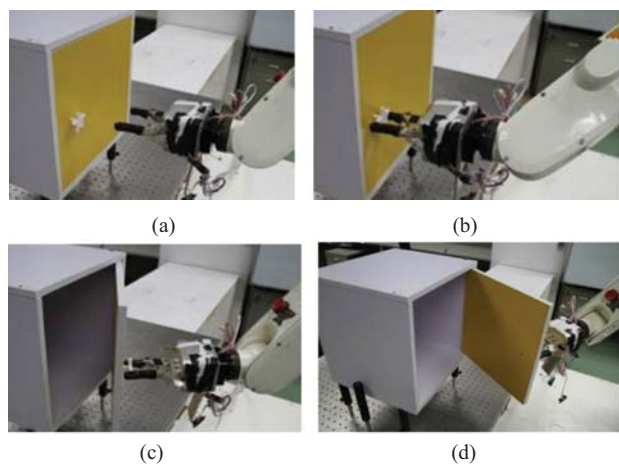


图9 (a)-(d) 机械臂执行旋转型的开门动作

图9 (a) - (d) 表示通过人体示范学习框架, 机器人提升了自身的认知操控能力, 完成对橱柜门的打开操作——即在旋转基元的指导下完成对应的旋转操作。

5 总结和展望

本文依靠自行研发的服务机器人操控平台, 充分利用 Kinect 传感器带来丰富信息, 结合计算机视觉中定位识别和动态追踪技术, 研究了一种能够让机器人从自身或外部获取对物品操控的经验, 转化为完成人类生活环境中新一轮日常多类别物品操控任务的技能, 进一步地提升机器人对未知物品操控的认知学习能力。

随着科技的进步, 包括 Kinect 在内的新型传感器层出不穷, 特别是将来自触觉传感器^[24], 力传感器, 惯性测量单元与视觉传感器等多源信息的融合, 将极大促进机器人任务操控这一综合性课题的研究。同时, 从本文中仅单独地分析机械臂末端的运动状态调整, 而拓展到研究多关节轴机械臂的运动调度、姿态规划等领域的研究, 也是提升服务机器人物品操控能力的一个重要内容。另外, 结合人体示教的框架, 在新型传感器的帮助下, 采取新的示教方法, 传授更多的操控技能给服务机器人, 也是本文后期展开更高层次研究工作的一个重要方向。

参考文献

[1] 国务院. 国家中长期科学和技术发展规划纲要(2006~2020年) [EB/OL]. [2006-02-09]. http://www.gov.cn/j=s/2006-02/09/content_183787.htm.

[2] Edsinger G, Weber J. Domo: a force sensing humanoid robot for manipulation research [C] // IEEE-RAS International

Conference on Humanoid Robots, 2004: 273-291.

[3] Katz D, Horrell E, Yang O, et al. The umass mobile manipulator uman: an experimental platform for autonomous mobile manipulation [C] // Workshop on Manipulation in Human Environments at Robotics: Science and Systems, 2006.

[4] Hart S, Ou S, Sweeney J, et al. A framework for learning declarative structure [C] // Proceedings of Robotics Science and Systems Workshop: Manipulation for Human Environments, 2006.

[5] Bagnell J A, Cavalcanti F, Cui L, et al. An integrated system for autonomous robotics manipulation [C] // IEEE/RSJ International Conference on Intelligent Robots and Systems, 2012.

[6] Srinivasa S, Ferguson D, Helfrich C, et al. Herb: A home exploring robotic butler [C] // Autonomous Robots, 2009.

[7] Garage W. Pr2 platform [EB/OL]. <http://www.willowgarage.com/pages/pr2/overview>.

[8] Microsoft kinect for X-BOX 360 [EB/OL]. <http://www.xbox.com/en-US/kinect>.

[9] Muja M, Rusu R B, Bradski G, et al. REIN-A fast, robust, scalable recognition Infrastructure [C] // in Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), Shanghai, China, 2011.

[10] Lowe D. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60: 91-110.

[11] Bay H, Tuytelaars T, Gool L V. Surf: speeded up robust features [C] // In the ninth European Conference on Computer Vision, 2006.

[12] Rublee E, Rabaud V, Konolige K, et al. ORB: an efficient alternative to SIFT or SURF [C] // In Proceedings of the IEEE International Conference on Computer Vision, 2011, 13.

[13] Rusu R B, Bradski G, Thibaux R, et al. Fast 3d recognition and pose using the viewpoint feature histogram [C] // IEEE/RSJ International Conference on Intelligent Robots and Systems, 2010.

[14] Aldoma A, Vincze M, Blodow N, et al. CAD-model recognition and 6DOF pose estimation using 3D cues [C] // IEEE International Conference on Computer Vision Workshops, 2011.

[15] Marius Muja. FLANN fast library for approximate nearest neighbors [EB/OL]. <http://www.cs.ubc.ca/~mariusm/index.php/FLANN/FLANN>.

[16] Morrow J D, Khosla P K. Manipulation task primitives for composing robot skills [C] // Proceedings of the IEEE International Conference on Robotics and Automation, 1997 I, 4.

[17] Hao D, Peter K A. Robot learning of everyday object manipulations via human demonstration [C] // IEEE/RSJ International Conference on Intelligent Robots and Systems, 2010, 1284-1289.

[18] Argall B, Chernova S, Velosoand M, et al. A survey of robot learning from demonstration [J]. Robotics and Autonomous Systems, 2009, 57: 469-483.

[19] Prime Sense. Prime sensor NITE 1.3 algorithms notes, 1.0 edition [Z]. 2010.

[20] Openni homepage [EB/OL]. <http://www.openni.org/>.

[21] Comaniciu D, Ramesh V, Meer P. The variable bandwidth mean shift and data-driven scale selection [C] // Proceedings of the Eighth IEEE International Conference on Computer Vision, 2001, 1: 438-445.

[22] Opencv homepage [EB/OL]. <http://opencv.willowgarage.com/wiki/>.

[23] PCL Point Cloud Library [EB/OL]. <http://www.pointclouds.org/>.

[24] Tegin J, Wikander J. Tactile sensing in intelligent robotic manipulation—a review [J]. Industrial Robot: An International Journal, 2005, 32(1): 64-70.