

卡内基梅隆大学研发新型算法 准确预测日常家具操纵方法

在日常生活中，我们接触到的家具多为“铰接物体”（articulated objects）。何谓铰接物体呢？

比如，抽屉有一条可以抽拉的轨道，门有一个垂直的旋转轴，烤箱有一个水平的旋转轴。这种由关节连接起来的物品，我们就叫它“铰接物体”。由于特定关节的存在，铰接物体的零件会受到关节的运动学约束，让这些零件只有一个自由度（DoF）。

铰接物体在我们生活中无处不在，构成了我们日常生活重要的一部分。而作为人类的我们，无论是看到什么样的铰接结构的家具，我们都能快速知道如何去操纵以及启动它。好似我们知道这些物体的每个关节是如何运动的。

那么机器人能不能也像人类一样，拥有这种可以预测家具如何启动的能力呢？如果可以的话，对居家机器人来说将是一种很大的提升。

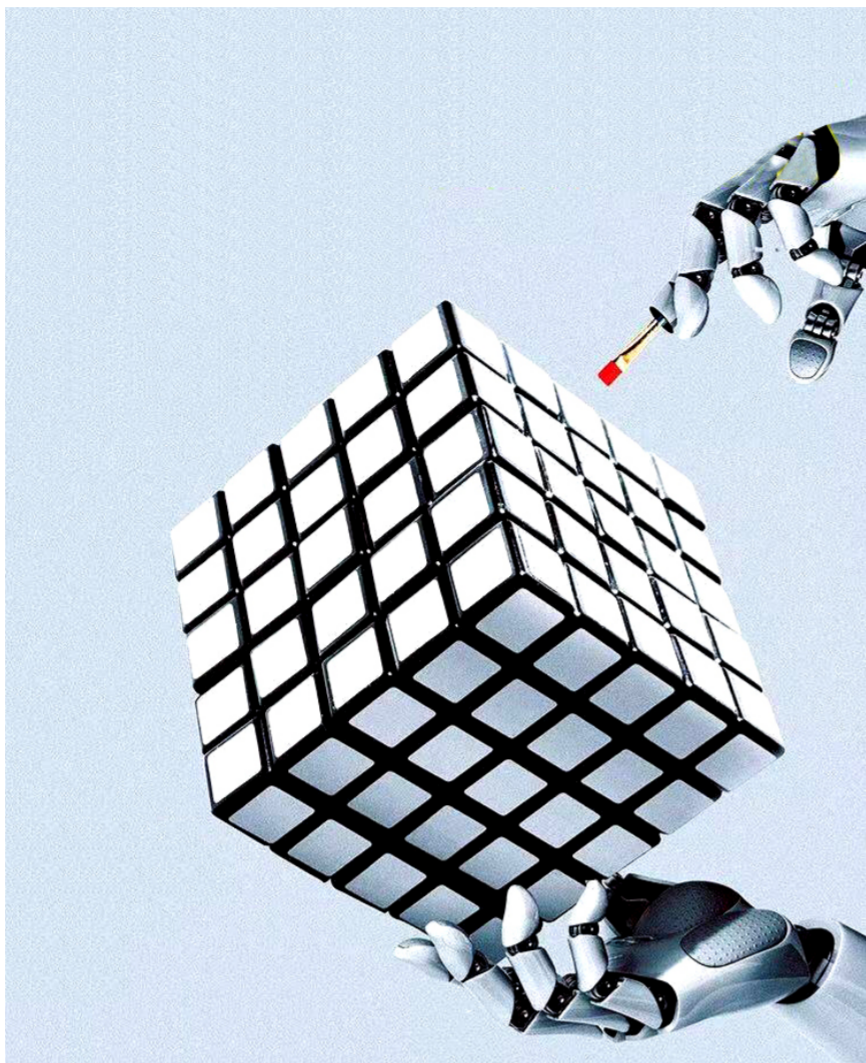
卡内基梅隆大学研发新算法

近日，来自卡内基梅隆大学（CMU）机器人学院 David Held 教授的 Robots Doing and Perceiving (R-PAD) 实验室的两名学生 Ben Eisner 和 Harry Zhang 在有效操控复杂铰接物体（articulated objects）方面取得了重大突破，推出了一种基于 3D 神经网络，能有效表达，预测日常家具等铰接物体的零件运动轨迹的算法 FlowBot 3D。

该算法包含两个主要子策略，其一是利用 3D 深度神经网络（PointNet++）去预测被操纵的物体点云数据（pointcloud）的瞬时运动轨迹（3D Articulated Flow/3DAF），其二是利用预测出来的运动轨迹去选择机器人下一步的动作。二者完全在模拟器中学习，并可以直接部署到真实世界中，不需要重新训练或者微调。在 FlowBot 3D 算法的帮助下，机器人可以像人一样随意操纵日常家具等铰接物体。

以往的日常家具操控要么是已知被操纵物体的几何特征（比如连接轴的位置及方向），要么通过模仿人类去学习如何操控一个给定的物体。二者都没有较好的可泛化性并且需要大量人类数据去训练。

与上述这些不同，FlowBot 3D 是第一个基于模拟器学习，通过学习



用物体的运动轨迹在模拟器中可以准确计算出来。

在训练中，机器人观测到一个被操纵物体的点云（pointcloud）数据，然后在这个点云数据上，机器人的视觉模块利用 PointNet++ 来预测每一个点（per-point）在受外力下的下一步的位置。此运动轨迹的真实数据可以通过正向运动学（forward kinematics）准确计算出来。将计算出来的下一步坐标减去目前的坐标就能得到被操纵物体零件的运动轨迹（3D articulated flow/3DAF）。因此，训练的时候只需要去最小化预测出的 3DAF 和真实数据的 3DAF 的 L2 loss。

FlowBot 3D 通过在模拟器中学习多种铰接物体在运动学约束下运动的轨迹，来预测新物体的操纵方

法的遮挡有一定的鲁棒性。另外，由于 FlowBot 3D 算法是闭环算法，机器人可以在下一步对自己可能出现的错误进行修正。

在真实世界中部署

FlowBot 通过使用两个子系统（视觉和操纵系统）克服了泛化性的挑战。在真实世界中，FlowBot 3D 可以准确预测出 3DAF 物体运动轨迹。只要这个轨迹可以准确预测出，那么操纵物体就只需要跟随这个轨迹这么简单。在真实世界中，FlowBot 3D 只需要使用模拟器中训练出来的一个模型就可以操控真实物体。

即使真实世界中的物体和模拟器中的外貌有很大不同，只要两者的运动学约束相似（抽拉式或者旋转式），那么 FlowBot 3D 就可以准确

需要几百万的训练数据，有时还需要人工指导才会学习新型物体的操控方式，使得这些机器人在现实世界中，尤其是家用机器人场景中不现实。

实验表明，支持 FlowBot 3D 算法的机器人能成功地操控较有难度的新型铰接物体，表现优于基于模仿学习部署的机器人。研究者使用相同的策略执行所有现实世界的部署，而不需要任何模拟校准或现实世界的微调。

实验结果显示，FlowBot 3D 在操作多数物体时的开闭误差，都能达到全开距离 10% 以下。然而其他基于模仿学习或者强化学习的方法差得很远。

FlowBot 3D 是机器人技术一项激动人心的进步，它可以无需微调在现实世界中部署高效且准确性强的家用机器人。这项工作还表明，计算机视觉的进步可以改变机器人领域，增强机器人的能力，同时使这些改进更易于扩展到新的条件。纯粹依靠模拟学习的方法有可能直接部署到真实世界里，并且有较强的泛化性，这将大大降低未来家用机器人训练和学习的成本。

实验室和课题组

这项研究是在卡内基梅隆大学机器人学院的 Robots Doing and Perceiving (R-PAD) 实验室进行及完成的。R-PAD 实验室在可变形物体操纵（如学习理解布料）课题上处于世界领先地位。实验室的主旨是利用计算机视觉方法（perceiving）帮助机器人去完成复杂的任务（doing）。

David Held 教授相信，即使对端学习方式很简单，但是真正部署到机器人上的时候会遇到许多问题，所以研究者应该将视觉（perception）与策略（policy）分开考虑，但是同时思考二者如何互相增加能力。在 FlowBot 3D 这项工作中，这个主旨深度体现了出来：3D Articulated Flow 是一个可以极大简化策略和规划的视觉表示方式，在正确 3D Articulated Flow 预测的基础上，策略会被简化成追踪 flow 向量。

研究报告的第一作者是 R-PAD 实验室二年级学生 Ben Eisner，他对 3D 视觉学习有极大兴趣，本科毕业于普林斯顿大学，加入 CMU 之前，Ben Eisner 曾在谷歌和三星北美研究所工作，从事机器人学习（robot learning）的研究。报告的共同第一作者是 R-PAD 实验室一年级学生 Harry Zhang，他对视觉和控制学都有极大兴趣，本科毕业于加州大学伯克利分校，加入 CMU 之前，Harry Zhang 在伯克利 BAIR 实验室负责可变形物体动态操纵（learning dynamic deformable objects manipulation）项目。

下一步，研究在尝试将 flow 这种视觉理解方式应用到铰接物体以外的物体上面，比如如何用 flow 预测 6 自由度的物体轨迹。同时，作者在尝试将 flow 应用到强化学习里，以增加强化学习的学习效率。

强化学习与灵巧机器人

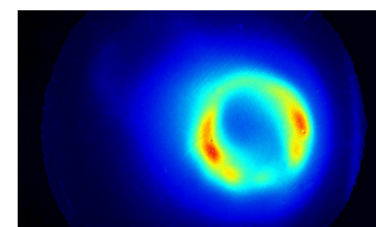
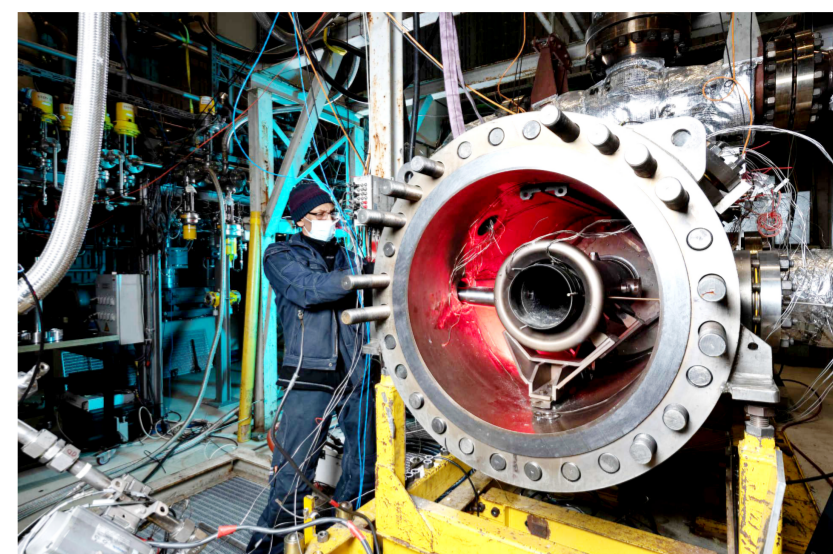
得益于 DeepMind 开发的计算机 AlphaGo 在围棋领域取得的惊人成就，强化学习这项技术在近几年备受关注。而随着设备计算速度的提升，以及深度学习架构的兴起，强化学习也得到了快速的成长。

强化学习的优势在于可以不依赖标注，而是由奖励函数来主导学习过程，这一点与人类在大多数情况下的学习方式是如出一辙的。

强化学习的出现，不仅改善了自动驾驶汽车的技术，还让机器人可以在无明确指导的情况下，掌握对一些物体的简单操作。就像前文提到的那样，机器人也能做到和人一样，在看到物体后就立刻明白如何将其拿起或打开。

相信在不久的将来，各种能够模仿人类行为的灵巧机器人，将以不同的方式走进千家万户。（麻省）

德国宇航中心开发 氢燃烧特种测量技术



测量技术方面拥有多年的专业经验。

为了能在燃烧室的恶劣环境中完成测量，该光学探针还配置了有保护作用的散热片，该散热片是在 DLR 技术市场营销部支持的“Keyhole”项目框架下开发而成。该项目负责人 Guido Stockhausen 表示，通过紫外内窥镜，可以观察空间氢火焰结构，从而为工业合作伙伴提供更多信息，并加快了未来燃烧室部件设计的进程。

为了在设定的时间框架内实现拟定的气候目标，科学界正在努力为其提供必要的技术，包括为能源和航空运输领域开发适用的燃烧室设备。到目前为止，还未开发出合适的光学测量技术来分析燃烧过程中的火焰特性，尤其是应用于全尺寸涡轮的极其昂贵的高压燃烧室测试。与科学实验室设置相比，在测试台上的观察受限。因此，必须找到内窥解决方案。

德国宇航中心（DLR）表示，目前已成功开发出一种用于专门研究氢燃烧的测量技术，并已完成氢和氨的燃烧室测试，这种新的测量技术可以“穿透锁孔”进入燃烧室内部。

新的测量技术所采用的内窥探头由 DLR 推进技术研究所发动机测量系统部开发。该部门在开发适用于高压和高温等不利条件的高精度

与天然气或煤油的火焰不同，氢火焰在可见光谱中是无色的，因此必须在紫外线（UV）光谱范围内观察其反应区域。这对内窥镜的光学元件来说是一个特别的挑战。经过两年的努力，DLR 与工业合作伙伴 Thomassen 能源公司最终合作开发出紫外内窥镜，并在科隆 HBC-2 高压燃烧室测试台的真实测试环境中实现首次应用。测试结果将用于优化固定式燃气轮机中的氢气燃烧。

在下一阶段的开发中，研究人员将向 NH3 储罐氢气中添加氨气，以观察其对燃烧的影响。氨由氮和氢分子组成，比纯氢更易储存和运输。未来，在向氢经济过渡期间，氨可以作为通过可再生能源实现发电的临时储存介质。（彩林）

美国国家制造科学中心 展示增强无损检测能力的 创新热成像技术



美国国家制造科学中心（NCMS）一份新的报告介绍了大间距 / 大面积热成像（LASLAT）的优势，这是一种先进的热成像工具，扩展了无损检测（NDI）的能力。该技术简报强调了 NCMS 如何与热像仪创新者和美国空军合作，使军用飞机维护人员能够以更短的时间和更低的成本提高飞机的战备状态。

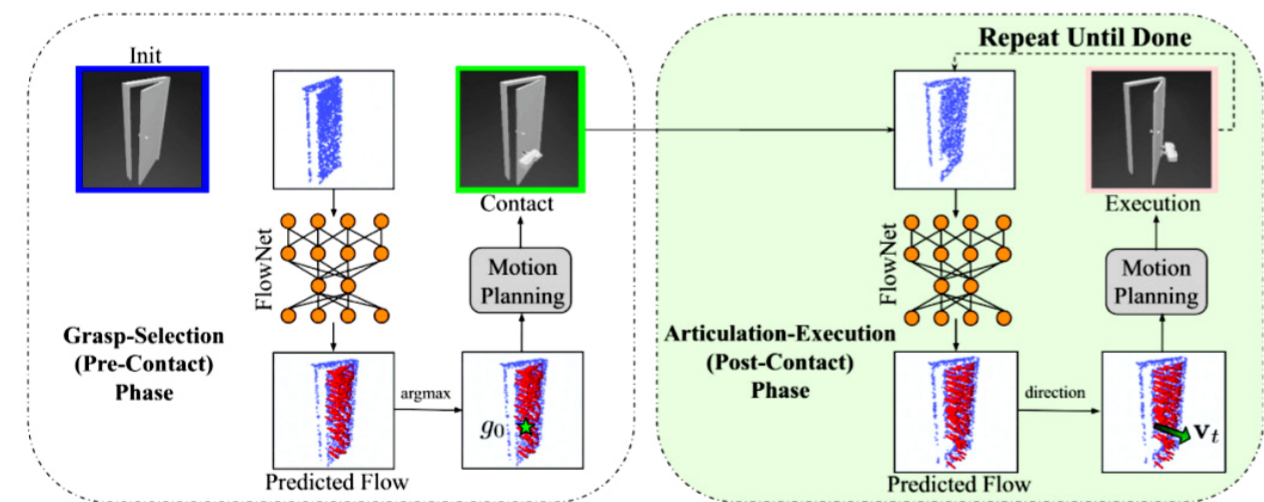
NCMS 表示，在过去的 20 年里，热成像已经成为一种被广泛接受的无损检测方法。热成像检查可以使用组件的继承热流（被动热像）或感应热流（主动热像）来揭示组件异常。被动热像法用于远距离检测物体，而闪光热像法（一种主动热像技术）主要用于近距离。NCMS 行业合作伙伴美国热波成像公司已经开发了一种称为热成像信号重建（TSR）的方法，该方法利用了近距闪光热成像的成功经验和被动热成像的大距离目标的原理。据报道，该技术具有前所未有的灵敏度，并提升探测的深度范围和“分辨率”。

LASLAT 系统的一个关键优势是它的便携性：技术人员可以直接将系统移动到飞机上，而无需将飞机移动到另一个大型系统中。红外摄像机监测表面温度变化，以检测表面以下的缺陷或异常，包括分层、

分离、流体侵入、冲击损伤和其他缺陷。自动化检测程序消除了手动定位，系统可以覆盖 17 英尺 × 15 英尺（约合 5.18 米 × 4.57 米）的检测区域，速度为每分钟 7 平方英尺（约合 0.65 平方米）。使用软件程序中的数据进行分析工具，技术人员可以确定缺陷的准确位置和测量值，然后生成一份飞机使用寿命期间的检查结果档案。

通过 NCMS 项目，美国空军发现，以前需要大约 10~14 小时才能扫描的大型飞机螺旋桨桨叶的检查现在需要 3~4 小时，从而增加了检查的文件记录和可追溯性。在该项目完成之后，该系统被海军用于检查 V-22 “鱼鹰”螺旋桨桨叶。这种检查方法的每年平均可节省 625 个工时。

虽然 NCMS 项目侧重于 LASLAT 系统的军事应用，但在其他行业也有许多潜在的应用，特别是在航空和航天部门。它的灵敏度和深度范围将使先进的复合材料以及各种金属材料的检测成为可能。此外，它还可以应用于其他具有高价值资产的行业，如发电、汽车、船舶、电子等。（阴鸣艳）



即使真实世界中的物体和模拟器中的外貌有很大不同，只要两者的运动学约束相似（抽拉式或者旋转式），那么 FlowBot 3D 就可以准确预测出操纵物体的策略。

每个零件的瞬时运动轨迹，使机器人可以计算出最优的物体操纵路径。这个特性使 FlowBot 3D 可以泛化到训练没有见到的物体上，并且可以直接部署到真实世界。

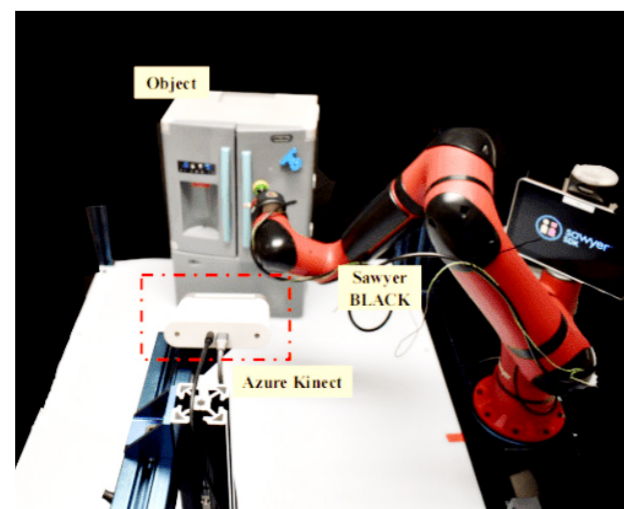
具体来说，FlowBot 3D 完全在模拟器中进行监督学习，从而学习出日常家具等铰接物体零件的瞬时运动轨迹（3D Articulated Flow/3DAF）。3DAF 是一种 3D 视觉表达方式，它能够极大简化策略的复杂度，从而提高操纵效率。机器人只需要紧密跟随这条瞬时轨迹，并闭环地重新预测，就能成功操纵一个铰接物体。

FlowBot 3D 赋能的 Sawyer 机器人不仅可以迅速识别哪个家具零件可以被操控，还可以准确预测出开启动件的运动轨迹。

机器人是如何做到的？

手动编程（hardcoding）的策略可以提高机器人在受控环境中的性能，但是想要机器人真正地理解家用物品的操控方式，唯一的方法是教机器人像人一样去预测这些物体的运动轨迹以及运动学约束（kinematic constraint）。

想要赋予机器人预测物体运动轨迹能力，研究者通过监督学习在模拟器中训练机器人的视觉模块，因为家



FlowBot 3D 在真实世界中部署的设置。

