

基于增强现实的装配引导技术综述

刘宇，徐锋，张丹，阮桥，左敦稳

南京航空航天大学 江苏南京 210016

摘要：阐述了增强现实的内涵，总结了基于增强现实的装配引导技术的国内外现状。并对其中关键技术进行了分析总结，指出基于深度学习的增强现实装配引导技术是未来的发展趋势。增强现实技术通过将虚拟信息叠加到真实环境中与用户进行交互，显著提升装配效率，为制造业向精密化、复杂化发展提供技术支持。

关键词：增强现实；智能制造；装配引导；三维跟踪注册

1 序言

在工业4.0时代，质量和效率是制造业产品增值的关键因素。增强现实技术（AR）凭借其能够帮助操作人员建立物理世界与数字信息环境连接的特性，在制造业中已经被视为提高效率的强大技术工具。增强现实区别于虚拟现实（VR），它以交互的方式，实时地在真实环境中注册虚拟对象，是对真实物理环境的补充，这显著减少了其对硬件的依赖，提高了工业应用能力。工业增强现实适用性分布研究报告中指出，在机器人操作、维护、制造等10个AR工业应用类别中，增强现实技术在手动装配领域的应用最广，占比达到32%。基于增强现实的装配引导技术已经成为了智能制造研究的热点，其重要性得到了普遍认可。

2 研究进展

基于增强现实的装配引导技术一般以系统平台的形式进行开发部署，将增强现实技术应用于产品装配引导，以实现产品、真实环境和虚拟信息与操作人员实时交互。图1所示为典型的视觉增强现实装配引导系统运行流程图，主要分为视频流采集、图像处理、跟踪注册和信息交互4个步骤。其中图像处理对采集的视频帧进行分析，为跟踪注册或信息交互提供响应信号。

在国外，基于增强现实的装配引导技术正趋于集成化与智能化，并已应用于真实的装配环境中。Zhl等研发出了面向机械装配的智能增强现实装配系统，该系统的视觉模块由两个单目相机组成，其中AR相机基于人工标识进行跟踪注册，识别相机

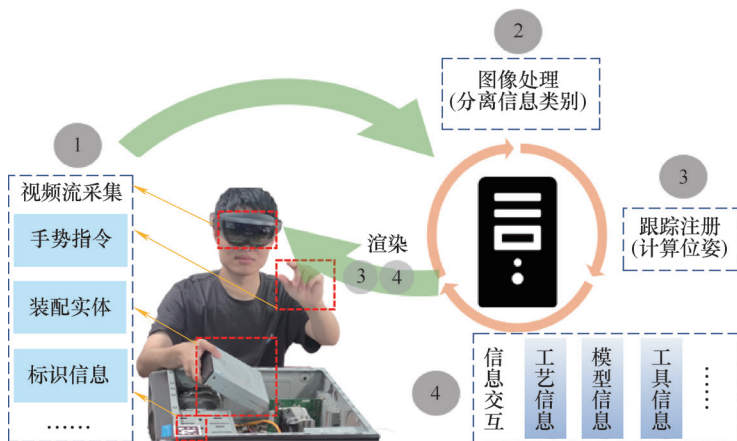


图1 典型视觉增强现实装配引导系统运行流程

基于R-CNN神经网络对装配工具或装配部件进行智能识别，并标记其在视场的位置。此系统被应用在数控雕刻机的装配中，通过增强现实的指令引导，操作人员装配时间及错误率显著降低。Mura的研究团队开发了支持车身面板对准装配的增强现实原型系统，它通过传感器实时测量汽车所需装配面板之间的间隙及其他公差信息，并与矩阵模板库进行对比，最后转换成增强现实指令以校准工人的装配误差。系统对装配误差的及时检测与回馈，提高了装配的精确度，极大减少了装配过程对操作人员经验的依赖。然而，此系统仍然存在部署繁琐、矩阵模板测量困难等问题，还存在很大的改善空间。

在国内，此项技术的研究主要集中在高校，距离工业应用还存在一定差距。南京航空航天大学的杨康康等针对跟踪注册流程提出一种基于RGB-D数据的配准方法，基于此方法开发了增强现实装配指导系统，同时以发动机模型的装配为例，验证了配准方法的稳定性与准确性。上海交通大学的刘然等以车门驱动电动机为装配对象，使用基于自然特征和LINEMOD算法分别对装配基体和安装的零部件进行识别、跟踪注册，最后根据增强指令进行零部件装配。此方法对大基体、小零部件的工业产品具有良好的适用性。

目前，国内外对增强现实装配引导技术的研究正处于如火如荼的关键时期，其整体的发展正向着集成化、智能化、普适化和精确化的目标前进，但这仍然需要研究人员的艰苦攻关。

3 关键技术

基于增强现实的装配引导技术的关键是三维跟踪注册技术，它直接影响虚拟信息的可读性以及信息呈现在真实环境中的位姿。而信息显示与交互技术是基于增强现实的装配引导技术的重要组成部分，是联系人机的纽带。

3.1 三维跟踪注册技术

三维跟踪注册是将虚拟信息实时地与真实环境进行配准融合，其中应用最广泛的是基于计算机视觉的跟踪注册技术，根据环境中有无标志物可再细分为基于人工标识和无标识的三维跟踪注册方法。

(1) 基于人工标识的三维跟踪注册方法 基于人工标识的三维跟踪注册方法应用最为广泛，并且已开发出成熟的增强现实软件开发工具包

(SDK)，如AR TOOL Kit。该方法（见图2）预先在装配场景中放置预定义的人工标识，相机对其捕捉识别后进行配准，而这些标识大多由黑白两色组成，易被识别。基于人工识别的三维跟踪注册方法的工作流程为特征点提取、位姿矩阵计算、虚实融合3个步骤。该方法鲁棒性、实时性好，受环境因素影响小，但还存在着如下问题：进行装配操作时标识易被遮挡，使跟踪注册失效；对零部件造成二次污染；操作视野小。这些都制约着此方法在增强现实装配引导技术中的发展和应用。

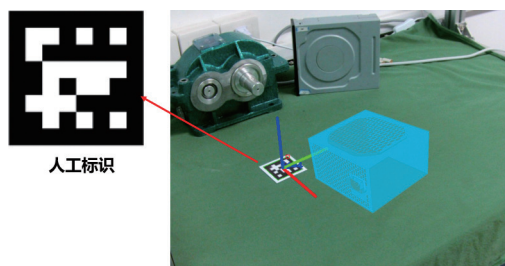


图2 基于人工标识的三维跟踪注册方法

(2) 无标识的三维跟踪注册方法 有以下几种。

1) 基于自然特征的跟踪注册方法。基于自然特征与基于人工标识的方法，其工作流程类似，区别在于基于自然特征的方法以具有丰富特征的自然图像作为计算位姿变换矩阵的依据。对自然场景图像进行特征点提取与匹配的相关算法已经非常成熟，其中具有代表性的有SIFT (Scale-Invariant Feature Transform)、ORB (Oriented FAST and Rotated BRIEF) 等，利用ORB算子进行特征点匹配的效果如图3所示。各个算法的侧重点不同，其选取和改进的方向依赖于自然场景的具体特征，否则将会影响增强现实系统的实时性与鲁棒性。此方法避免了标识对装配环境的“污染”，但受环境影响较大且对弱纹理的工业零件识别效果差，具有一定的局限性。

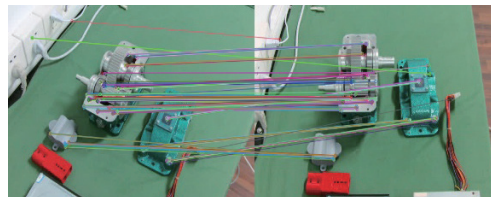


图3 利用ORB进行特征点匹配

2) 基于模型的跟踪注册方法。基于模型的跟踪注册方法是解决弱纹理零件跟踪注册的主要方法，

其中边跟踪与基于点云配准的方法发展最快。此类方法通过三维重建或Solidworks等绘图软件建立零件模型，并以此获得零件的轮廓特征或点云数据作为先验知识，最后在捕捉的视频帧或RGB-D数据中寻找最优关系，进行配准和注册。目前，高通公司的Vuforia发布了model target插件，在Unity平台实现了基于模型的跟踪注册技术（见图4）的商业应用，加快了增强现实装配引导系统的开发流程。基于模型进行跟踪注册仍有很大的局限性，它依赖于零件本身：特征单一、轮廓简单的零件误识别率高；特征繁多，轮廓复杂的零件计算量大，实时性识别难以保证。

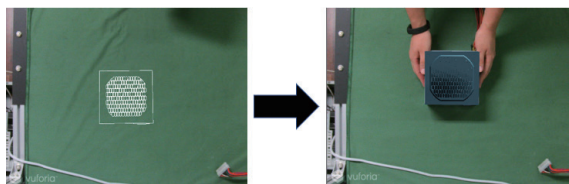


图4 Vuforia中model target的跟踪注册

3) 基于深度学习的跟踪注册方法。近几年，深度学习快速发展，在增强现实领域中，各种深度学习的方法及体系结构凭借其强大的计算能力可以更有效、更鲁棒地执行和解决跟踪注册的任务，其中直接以RGB图像为输入，以目标物体六自由度位姿为输出的深度学习网络模型发展最为迅猛。此类方法主要以李代数、四元数或控制点来表示目标物体的旋转和位移，将位姿估计问题转化为回归问题，主要流程如图5所示。

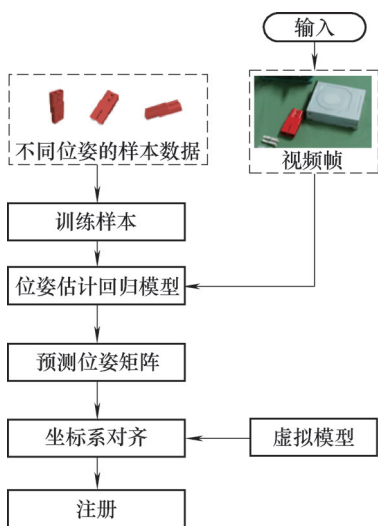


图5 基于深度学习的位姿估计方法主要流程

Tekin等基于YOLOv2网络提出了YOLO-6D网络模型。以光驱为识别对象的YOLO-6D位姿预测通过输入的RGB图像来实时预测目标物体3D边界框的8个角点及中心点的2D投影坐标，最后使用PnP（Perspective-n-Point）算法来计算目标对象在三维空间的旋转和平移矩阵（见图6）。Kstner等基于YOLO-6D网络，在HoloLens上实时检测和标注移动机器人的位姿，在实时性与预测准确性方面都取得了不错的效果，张德等提出通过YOLO和关键点检测器网络（KPD），分别实现目标检测和标记被训练关键点的2D投影位置的功能，然后根据关键点2D-3D关系，利用概率神经网络算法计算位姿，此方法无需再进行位姿的二次优化就能达到较高的精度。

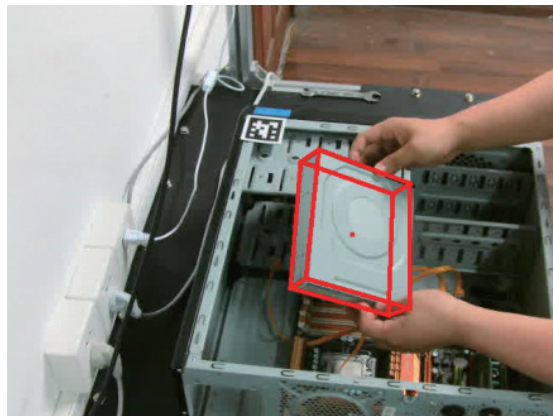


图6 基于YOLO-6D的位姿预测

基于深度学习的跟踪注册方法较于传统方法具有较高的精度和鲁棒性，但还存在所需数据庞大、数据集制作困难以及模型训练周期长等问题，限制了它的快速发展。

3.2 信息显示与交互技术

信息显示与操作人员的装配过程直接相关，是引导装配的重要一环。信息显示取决于硬件设备，比较经济的方案是利用多目摄像头和显示器，来组成系统的场景信息收集和信息显示模块，但是此方案存在视野固定，调试部署周期长，设备零散等问题。目前较流行的还是Microsoft公司发布的HoloLens2，它集成了深度相机、磁力仪等多种传感器，使操作人员可以更灵活地进行装配操作，具有较好的沉浸式体验。虽然HoloLens2因为存在成本较高、易致人眩晕等问题，而没有在工业上广泛应用，但信息显示硬件载体将会继续沿着集成化的

方向发展。

信息管理是根据对装配作业的认知，对装配工艺等信息进行获取、维护和管理。常见的方法是信息建模或运用知识图与语义网技术。华中科技大学的蓝珊将基本的装配工艺信息与装配过程的多媒体信息进行IDEF1X建模，将它们分类组织、转换、关联，从而转换成在增强现实装配引导的数据信息库，并在装配过程中进行调用。语义网技术主要是将各种文本格式的文件及其他非结构化的信息编码转化成计算机能理解的结构化信息和推理规则集，而知识图更侧重于数据和语义的图形化结构。知识图与语义网技术相辅相成，通过上下文信息，极大地丰富了数据的内容。Walczak等提出了一种分布式增强现实服务的体系结构，利用语义网技术有效地搜索上下文描述的分布式资源，构成交互式增强现实演示，以提供搜索和反馈服务。

随着传感器技术的发展，人机交互的方式越来越多样化。传统的键鼠交互方式虽然稳定，但是操作效率低，无法实现虚实场景的无缝衔接，已经渐渐被摒弃。而基于深度学习和多传感器的语音、手势、眼控等交互方式具有简单、自然、效率高等特点，是目前人机交互研究的重点和热点。董琼等就在利用增强现实进行装配的场景中，通过提炼的关键语音命令搭建了指导装配的语音交互系统，提高了人机交互的效率和用户体验。

4 结束语

基于增强现实的装配引导技术为用户提供了新的交流和获取信息的方式，能够有效提高产品装配的效率。目前，此项技术正向着集成化和智能化方向发展，而深度学习是智能化的重要技术工具，它能够提高整个系统的效率和有效性，并将“智能”灌输到系统中，在跟踪注册、人工交互等技术领域相比于传统方法，在效率和精确度方面取得了不错的应用效果。同时，科研人员提出使用虚拟合成数据来代替部分真实数据的方法来弥补训练深度学习模型时所需数据量大、数据集制作困难等问题，从而显著降低了深度学习应用的门槛。未来，基于增强现实的装配引导技术的发展会与深度学习进行更深层的结合，将进一步强化此项技术的功能、应用和服务，这还需要科研人员不断的研究和创新。

参考文献：

- [1] Azuma R, et al. Recent advances in augmented reality [J]. IEEE Comput Graphics Appl, 2001, 21 (6) : 34-47.
- [2] Luís Fernando de Souza Cardoso a b, Fcmqm A, Erza C. A survey of industrial augmented reality [J]. Computers & Industrial Engineering, 2020, 139 (8) : 1-12.
- [3] Zhi A, et al. Smart augmented reality instructional system for mechanical assembly towards worker-centered intelligent manufacturing [J]. Journal of Manufacturing Systems, 2020 (55) : 69-81.
- [4] MD Mura, Dini G. An augmented reality approach for supporting panel alignment in car body assembly [J]. Journal of Manufacturing Systems, 2021 (59) : 251-260.
- [5] Yang K, Guo Y, Tang P, et al. Object registration using an RGB-D camera for complex product augmented assembly guidance [J]. Virtual Reality & Intelligent Hardware, 2020, 2 (6) : 501-517.
- [6] 刘然, 范秀敏, 尹旭悦, 等. AR辅助装配中基体及零件位姿组合估计方法 [J]. 机械设计与研究, 2018, 34 (6) : 119-125, 137.
- [7] 韩玉仁, 等. 增强现实中三维跟踪注册技术概述 [J]. 计算机工程与应用, 2019, 55 (21) : 25-34.
- [8] Lowe D G. Distinctive Image Features from Scale-Invariant Keypoints [J]. International Journal of Computer Vision, 2004, 60 (2) : 91-110.
- [9] 张德, 李国璋, 王怀光, 等. 基于深度学习的目标位姿估计方法综述 [J]. 飞航导弹, 2019, 4 (9) : 72-76.
- [10] 蓝珊. 人工装配过程增强现实辅助技术研究 [D]. 武汉: 华中科技大学, 2018.
- [11] D Rumiński, Walczak K. Large-Scale Distributed Semantic Augmented Reality Services – A Performance Evaluation [J]. Graphical Models, 2020, 107 (1) : 1-15.
- [12] 马金盾, 等. 增强现实装备维修引导系统研究综述 [J]. 兵器装备工程学报, 2020, 41 (7) : 169-176.
- [13] 董琼, 李斌, 董剑, 等. 面向头戴式眼镜的增强现实装配语音交互实现 [J]. 制造业自动化, 2020, 42 (10) : 77-80.

MW 20210801