

增强现实技术概论

项目导读

党的二十大报告指出“加快发展数字经济，促进数字经济和实体经济深度融合，打造具有国际竞争力的数字产业集群”。增强现实（Augmented Reality, AR）技术是新一代信息技术的重要组成部分，在数字经济中具有不可替代的作用。AR 技术在智慧教育、工业制造、医疗诊断等诸多领域应用广泛，具有重要的社会意义和巨大的商业价值。我国的 AR 技术发展迅速，近年来形成了具有自主知识产权的硬件设备、软件平台。本项目讲解了增强现实的概念、发展历程、技术分类与关键技术等，并引导读者体验几种 AR 软件。

学习目标

- 学习增强现实的基本概念、发展历程与技术分类，了解 AR 系统、设备的分类以及最新 AR 硬件设备。
- 理解增强现实的关键技术，如三维注册与几何一致性、光照一致性、显示技术和人机交互技术等。
- 了解增强现实技术在不同领域的实际应用。

职业素养目标

- 能够系统地学习增强现实的发展历程、技术分类和关键技术，产生对 AR 的兴趣。
- 应具备自主学习能力，关注 AR 技术在不同领域的实际应用。

职业能力要求

- 具备扎实的增强现实理论基础，能够在实际项目中灵活运用 AR 知识。
- 应熟悉 AR 软件常见使用方法，具备独立探索全新 AR 应用的能力。



项目重难点

项目内容	工作任务	建议学时	技能点	重难点	重要程度
增强现实技术概论	任务 1.1 体验网易洞见 APP：理解何为增强现实	1	认识增强现实	学习增强现实发展历程	★★★★☆
	任务 1.2 体验 AR 动物园：了解增强现实技术的分类	1	区分有标识 / 无标识的 AR 技术	了解无标识 AR 技术的使用方法	★★★★☆
				了解有标识 / 无标识 AR 技术的区别	★★★★☆
	任务 1.3 探索增强现实关键技术	1	增强现实的关键技术	了解增强现实的关键技术	★★★★★
任务 1.4 体验生活中的增强现实应用：AR 购物	1	增强现实的应用	了解增强现实的应用场景	★★★★☆	

任务 1.1 体验网易洞见 APP：理解何为增强现实

任务目标

知识目标：了解 AR 软件的操作方法。

能力目标：初步掌握 AR 软件的交互技巧。

建议学时

1 学时。

任务要求

学习增强现实技术的基本概念和特点，了解增强现实技术的发展历程。

知识归纳

1. 增强现实的概念

增强现实是一种把计算机生成的三维场景等数字信息叠加到物理世界，以增强用户对真实环境感知和理解的技术，是在虚拟现实的基础上发展起来的，计算机图形学、计算机视觉、人机交互和显示技术等是其关键的支撑技术。虚拟现实和增强现实的典型区别在于，虚拟现实技术表达的内容是纯虚拟的，而增强现实技术突出的是虚实融合。1994 年多伦多大学的 Milgram 等首次提出了混合现实连续统（Mixed Reality Continuum），认为虚拟世界到真实环境是一个连续渐变的过程，如图 1-1 所示。将虚拟环境和真实环境的融合技术称为混合现实（Mixed Reality，MR），其中用虚拟信息来增强用户对真实环境的感知称为增强现实。



图 1-1 混合现实连续统

2. 增强现实的发展历程

1968年，计算机图形学和增强现实之父 Sutherland 开发出第一套增强现实系统“达摩克利斯之剑（The Sword of Damocles）”，能够将简单线框图转换成三维效果。如图 1-2 所示，这套系统使用光学透视式头戴显示器，配有六自由度的机械跟踪器和超声波跟踪器。

1992年，波音公司 Caudell 和 Mizell 在辅助布线系统中提出了增强现实的概念，以可视化的方式指导工程师完成布线工作。1993年，哥伦比亚大学 Feiner 等提出了基于知识的 AR 维修支援（Knowledge-based Augmented Reality for Maintenance Assistance, KARMA），指导用户安装和维修打印机。1997年，哥伦比亚大学 Feiner 等人开发了第一个户外 AR 游览系统，由背包计算机和传感器组成，可在移动过程中输出三维图形。



图 1-2 “达摩克利斯之剑”的头戴式显示器系统

1999年，Kato 和 Billinghurst 开发了第一个 AR 开源框架 ARToolKit，使用模板匹配方法进行目标识别与跟踪。2005年，Henrysson 等人在塞班手机上使用 ARToolKit 进行目标跟踪，实现了多人协同的 AR 网球游戏。

2007年，Klein 和 Murray 提出了一种实时同步定位和建图方法，在重建三维稀疏场景的同时，计算相机空间位置和朝向可为 AR 应用提供实时位置与姿态（简称位姿）估计。

2008年，Wikitude 公司在第一款 Android 手机 G1 上开发了旅游辅助软件 Wikitude AR Travel Guide，为智能手机用户带来了全新的旅游体验。

2012年，谷歌公司推出了可运行 AR 应用的智能眼镜 Google Glass，具有 GPS 定位、拍照等功能。2015年，微软公司发布了 HoloLens 1，能够呈现虚实融合的影像，成为当时的 AR 眼镜行业标杆。2017年，苹果公司和谷歌公司分别推出了 iOS 和 Android 系统上的移动 AR 开发平台 ARKit 和 ARCore，促进了移动 AR 技术体验的应用。2019年，微软公司更新发布了 HoloLens 2，与第一代相比，在计算能力和人机交互等方面得到进一步提升。

近年来，随着移动计算、5G 通信、智能芯片和显示技术迅速发展，体积小、功耗低、性能强的传感器和终端设备不断涌现，极大地推动了 AR 技术的研究和应用推广。我国出现了一批具有自主知识产权的 AR 硬件企业，如广东虚拟现实科技有限公司（Ximmerse）、优奈柯恩（北京）科技有限公司（XREAL）、亮风台（上海）信息科技有限公司（HiAR）等。同时，也出现了一批国产 AR 软件开发平台，如网易洞见 AR、百度 AR、商汤科技



SenseAR、视辰 EasyAR 等,AR 产业生态日趋完善。2022 年“天宫课堂”在中国空间站开讲,太空教师陈冬、刘洋、蔡旭哲借助 MR 智能眼镜为广大青少年带来了一场精彩的太空科普课(图 1-3)。



图 1-3 问天实验舱授课现场

任务实施

步骤 1 初体验网易洞见 AR 软件。

下载并安装网易洞见 APP。启动网易洞见 APP,使用手机扫描现实环境,与虚拟角色互动。

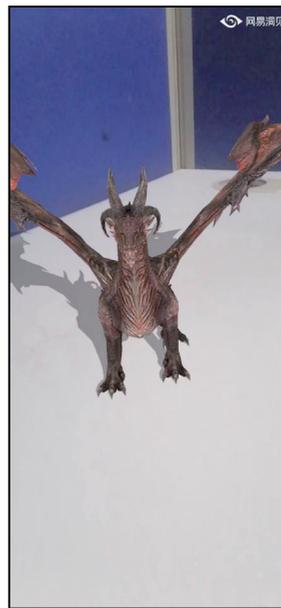
(1) 下载并安装网易洞见 APP。

(2) 打开网易洞见 APP,进入首页,如图 1-4(a)所示。单击首页相机图标,可以看到一个虚拟角色出现在相机前方平面上,如图 1-4(b)所示。

(3) 用户可触摸屏幕使虚拟角色放大、缩小、旋转或移动,从各个角度观察虚拟角色。



(a) 首页



(b) 虚拟角色

图 1-4 网易洞见 APP

步骤2 体验过 AR 应用后，思考以下问题。

- (1) 以前体验过 AR 应用吗？上述任务与以前体验的 AR 应用给你留下了什么印象？
- (2) 分组讨论网易洞见 APP 可以如何改进，可以与虚拟角色做哪些交互。

任务 1.2 体验 AR 动物园：了解增强现实技术的分类

任务目标

知识目标：区分有标识与无标识的 AR 技术。

能力目标：对比网易洞见 APP 与 AR 动物园的使用体验和操作过程，了解有标识与无标识 AR 技术的区别，以及无标识 AR 技术的使用方法。

建议学时

1 学时。

任务要求

深入了解 AR 技术的基本分类，包括有标识和无标识 AR 技术。同时，掌握 AR 技术的常见设备，学习设备分类以及国内外最新 AR 设备。按任务步骤要求进行操作，探讨有标识 / 无标识 AR 技术的体验差别。

知识归纳

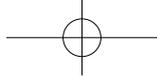
1. AR 系统分类

AR 系统可分为有标识 AR 系统和无标识 AR 系统。

(1) 有标识 (Marker-based) AR 系统。有标识 AR 系统在现实世界中放置特殊标识 (如 Logo、海报、二维码等)，用来进行目标识别与跟踪。用 AR 设备 (如手机) 扫描标识时，AR 系统可以识别标识的位置和方向，并叠加数字内容。如图 1-5 所示，亮风台 (上海) 信息科技有限公司研发的西柏坡纪念馆 AR 红色旅游项目丰富了景区情景传达方式，AR 应用叠加的数字内容让游客身临其境，以更生动的方式学习历史知识。



图 1-5 西柏坡纪念馆 AR 红色旅游项目



(2) 无标识 (Markerless) AR 系统。无标识 AR 系统无需用户事先准备标识模板, 使用算法 (如 ORB、SURF 等) 对场景周围进行对象识别。无标识 AR 技术使得用户体验 AR 系统更为自然, 但通常算法复杂, 对硬件计算能力要求高。成都弥知科技有限公司研发的 Kivicube 平台可在网页、微信小程序中运行 AR 应用。该公司开发的无标识 AR 试衣系统, 可自动检测身体为用户带来了全新的购物体验, 如图 1-6 所示。

2. AR 设备分类

近年来, AR 应用不断普及, 各种新型 AR 设备不断涌现, 为用户提供丰富的视觉、听觉和触觉体验。这些 AR 设备可以分为移动式、头戴式和体感交互式三种类型, 如图 1-7 所示。



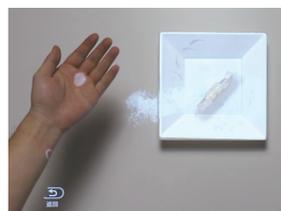
图 1-6 Kivicube AR 试衣系统



(a) 移动式AR设备



(b) 头戴式AR设备



(c) 体感交互式AR设备

图 1-7 常见 AR 设备

(1) 移动式 AR 设备: 通常是智能手机或平板电脑, 使用触控交互。优点是硬件成本较低、携带方便, 是使用最为广泛的 AR 应用载体; 缺点是屏幕较小、沉浸感不足, 如图 1-7 (a) 所示。

(2) 头戴式 AR 设备: 使用头部姿态、手势和手柄等方式进行交互, 用于工业仿真、医疗诊断等场景。优点是沉浸感强, 缺点是硬件成本高, 如图 1-7 (b) 所示。

(3) 体感交互式 AR 设备: 可感知用户肢体动作和姿态, 用身体姿势或动作进行内容交互, 广泛应用于医疗康复、体育锻炼等场景。优点是交互自然、趣味性强; 缺点是没有移动式 AR 方便, 如图 1-7 (c) 所示。

得益于移动计算和显示技术的进步, 头戴式 AR 眼镜产品不断朝着轻量化、多模态交互和大视场角等方向发展, 如 Google Glass 和国产的 XREAL Air、Rhino X Pro 等 AR 设备如图 1-8 所示。



(a) Google Glass



(b) XREAL Air



(c) Rhino X Pro

图 1-8 三种头戴式 AR 设备

任务实施

步骤1 体验有标识的 AR 项目——AR 动物园。

(1) 进入微信小程序平台,搜索“AR+ 动物园”,找到“AR 口袋动物园”并打开运行,如图 1-9 (a) 所示。

(2) 使用小程序内的摄像头,对准特定的图片,如图 1-9 (b) 所示,即可弹出对应的三维动物模型,如图 1-9 (c) 所示。读者可以尝试不同的动物标识图片。

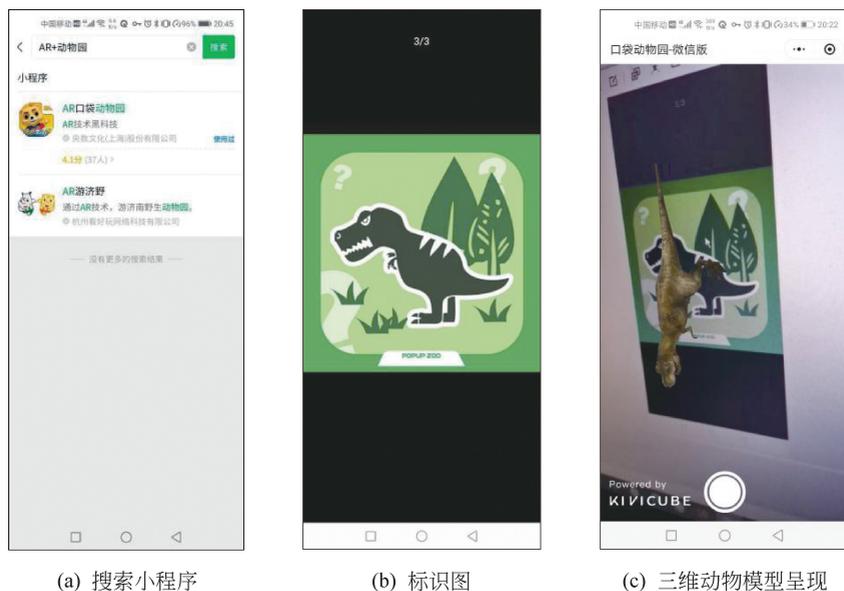


图 1-9 有标识 AR 识别与三维模型展示

(3) 在体验该 AR 游戏时,读者可以关注以下问题:该动物模型是否会运动;运动是否有规律。读者还可以捧着智能手机围绕动物模型位置转上几圈,仔细观察该动物模型及其与标识图片平面之间的位置关系。

步骤2 体验过有标识的 AR 项目后,思考以下问题。

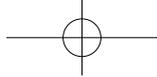
(1) 与任务 1.1 中的网易洞见 APP 相比,有标识 AR 技术的优点和缺点是什么?在使用体验上有什么差别?

(2) 有标识的 AR 系统需要事先在物体上贴上 AR 标签,这对于一些应用场景是不可行的。分组讨论,哪些场景适合使用有标识的 AR 应用?

任务 1.3 探索增强现实关键技术

任务目标

知识目标:了解增强现实的关键技术:三维注册与几何一致性、光照一致性、人机交互技术和显示技术。



能力目标：对比“AR 摆摆看”小程序和网易洞见 APP 使用体验和操作过程，了解 AR 关键技术在真实应用中是如何发挥作用的。

建议学时

1 学时。

任务要求

透彻理解本任务中的知识点，熟悉增强现实的关键技术。按任务步骤要求进行操作，探讨 AR 应用中的增强现实关键技术如何影响交互体验。

知识归纳

1. 三维注册与几何一致性

增强现实将虚拟场景融合到现实场景，再呈现给用户。三维注册的目的是重建三维现实场景和计算用户相机的实时位姿信息，从而保证虚拟场景与真实场景的几何一致性。几何一致性是指虚拟场景与现实场景共享同一空间。如果不能保证几何一致性，会严重影响用户交互体验。增强现实的三维注册技术分为三类：基于传感器的跟踪注册、基于计算机视觉的三维注册和混合三维注册技术。

(1) 基于传感器的跟踪注册技术的方法是依靠传感器设备实现三维注册。其技术原理是利用收发信号装置感知物体位置信息，进而求解相机在世界坐标系下的位姿信息完成三维注册。其缺点是易受干扰、灵活性差。

(2) 基于计算机视觉的三维注册技术可分为基于人工标识的方法和基于自然特征的方法。

基于人工标识的三维注册技术所使用的人工标识物通常具有规则的形状，而且与环境差别明显。计算机通过特定的视觉算法将标识物从环境中提取出来。因为标识物形状规则，所以处理算法复杂度较低，实时性较高。人工标识方法已经较为成熟，对硬件处理能力要求不高，且鲁棒性很高，被 AR 引擎广泛使用。图 1-10 展示了几种典型的 AR 标识物。



图 1-10 AR 标识物

基于自然特征的三维注册技术不依赖标识物，缺点是计算量大。这类注册技术先提取场景内物体的自然特征，再求解出相机位姿，进而完成三维注册。特征点的选取和匹配是

重要的影响因素；常用的特征点检测与特征匹配方法主要有 SIFT、SURF、BRIEF 等。

(3) 混合三维注册技术是指综合应用上述注册技术。与单一注册技术相比，混合三维注册方法能够兼顾高精度和鲁棒性，其难点在于不同类型的位姿数据的同步和融合。

2. 光照一致性

AR 系统的最大特点是虚拟物体和现实世界共存，不仅需要实时稳定的三维注册技术，还需要虚实物体间具有光照一致性。光照一致性指虚实物体所受到的光照情况相同。实现光照一致性需要先复原现实世界的光照模型，然后在虚拟世界中构造相似的光照模型并考虑虚拟物体由此所受的影响。虚拟物体的呈现效果与其形状、位姿、材质、纹理、光源以及周围环境密切相关。如图 1-11 所示，人物模型及阴影很好地反映了光照一致性，与现实场景融合恰当、效果逼真。



图 1-11 光照一致性

3. 人机交互

AR 中的常见交互技术有触控交互、手势交互、控制器交互和实物交互。

(1) 触控交互：以触觉作为主要感知通道的交互技术。系统获取用户手指触摸的位置信息，触发相应命令，完成对虚拟物体的选择、移动和旋转等操作，如图 1-12 所示。触控交互产品按工作原理不同可分为四种：电阻式触摸屏、电容式触摸屏、红外触摸屏和表面声波式触摸屏。触控交互的优点是操作精度高，缺点是双手被占用。



图 1-12 触控交互

(2) 手势交互：属于自然人机交互，符合日常行为习惯。手势交互依赖手势识别，常见的方法有基于传感器的手势识别和基于计算机视觉算法的手势识别。如图 1-13 所示，双手在 AR 眼镜前拖曳，可与虚拟场景中的太阳系交互。手势交互的优点是交互自然，缺点是需要专用设备（如 Leap Motion）支持。

(3) 控制器交互：使用光投射器、空间位置跟踪器、交互笔或手柄等控制设备进行交互。可将控制器的六自由度位姿、按钮状态和加速度等信息绑定在 AR 系统中的虚拟对象上，进行选择与交互。控制器交互的优点是操控精准，缺点是需要专用设备。

(4) 实物交互：将虚拟对象与真实物体联系起来，构建知觉共存的交互体验。实物既可以是输入，也可以是输出，反馈信息与环境融为一体。如图 1-14 所示，实物交互充分利用了现实环境和实物，因此学习成本较低、交互自然，缺点是需要实物定制设备。



图 1-13 手势交互



图 1-14 实物交互

4. 显示技术

增强现实系统显示技术可分为头戴式、移动手持式和投影式显示技术。头戴式显示技术一般是指借助 AR 眼镜进行虚实融合内容的显示。按照实现原理，头戴式显示器分为视频透视式显示器和光学透视式显示器。视频透视式显示器将摄像头获取的真实环境实时视频，与虚拟场景相互融合，最终虚实内容融合显示出来，如图 1-15 (a) 所示。显示技术的优点是虚实融合效果较好，缺点是视频方式显示不自然、真实感较差。

光学透视式显示器利用光的反射原理，在用户的眼前放置一块半透明的光学融合器，使用户能够看到虚拟三维物体与真实场景的融合画面。光学融合器是部分透明的，用户可看见真实环境；光学融合器又是部分反射的，用户可看见监视器反射的虚拟场景，如图 1-15 (b) 所示，其优点是结构简单、性价比高，缺点是易受环境光影响。

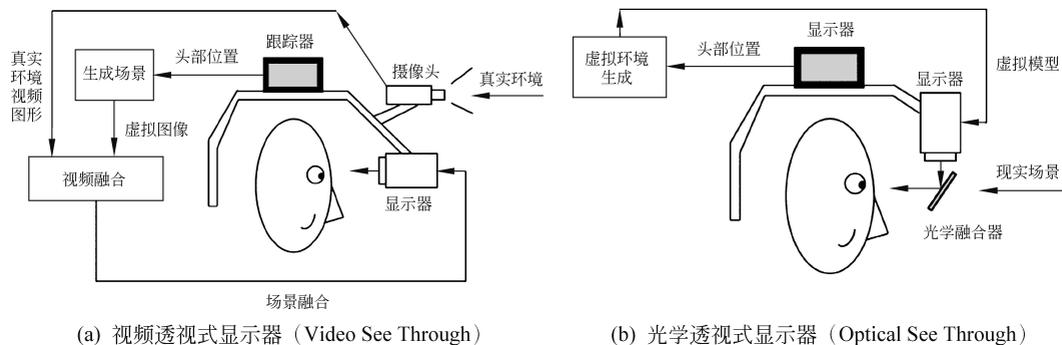


图 1-15 头戴式显示技术

移动手持式显示技术是用智能手机或平板电脑的显示屏作为 AR 显示设备。手持式显示器是以摄像头作为输入设备，进行触控交互。如图 1-16 所示，上海亮风台开发的 AR 汽车 APP 使用摄像头对准车内部件，可在手机上显示对应部件的相关信息。该技术的优点是智能手机终端广泛使用，缺点是用户双手被占用，人机交互受限。

投影式显示技术利用投影仪将虚拟场景直接投影到真实环境中，达到对现实世界增强的效果，为用户提供更加直观的展现方式。如图 1-17 所示，2019 年 CCTV 元宵戏曲晚会将孔明灯场景投影到真实舞台上，虚实融合、美轮美奂。该技术的优点是用户摆脱了佩戴头戴式显示器的束缚，缺点是成本高、场地受限。