

基于 ARCore 的移动增强现实应用系统设计

宋燕燕

(南京传媒学院,江苏 南京 211172)

摘要:伴随着计算机软硬件计算能力的不断提升,增强现实技术的核心算法日益成熟,给人机交互技术及交互方式带来前所未有的突破。通过对增强现实技术的运动跟踪、环境理解及光线估计等技术进行研究,并采用 ARCore 进行架构和核心方法设计,提出一种移动增强现实应用系统设计方案。通过对现实场景中的图像或者物体进行运动追踪和定位,计算出虚拟世界与现实世界坐标系的对应关系,采用 ARCore 的网格化技术自动识别表面,捕捉每帧画面的特征点云、锚点、二维平面集合和跟踪状态绘制平面。利用可识别特征以及设备的朝向和加速度传感器保持实时更新,将虚拟物体按照正确的空间透视关系叠加到现实场景并确定位置,实现可视化仿真及互动教学等功能,系统最终在移动平台进行发布。增强现实技术以各种形式应用于教学中,具有良好的操作性和交互性,降低了教学实验成本,增加了教学的趣味性。

关键词:增强现实;ARCore;运动跟踪;环境理解;网格化

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2022)07-0196-05

doi:10.3969/j.issn.1673-629X.2022.07.034

Design of Mobile Augmented Reality System Based on ARCore

SONG Yan-yan

(Communication University of China, Nanjing 211172, China)

Abstract: With the continuous improvement of the computing power of computer software and hardware, the core algorithm of augmented reality technology is becoming more and more mature, which has brought an unprecedented breakthrough to human-computer interaction technology and interaction mode. The motion tracking, environment understanding and light estimation of augmented reality technology are studied. ARCore is used to design the architecture and core method. A design scheme of mobile augmented reality application system is proposed. As a result, the corresponding relationship between the coordinate systems of the virtual world and the real one can be calculated by tracking and positioning the motion of the images or objects on the real scene. The surface can be automatically identified using meshing technology of ARCore. The plane is drawn by capturing the feature point cloud, anchor, two-dimensional plane set and tracking state of each frame. Real time updates are maintained using recognizable features and the orientation and acceleration sensors of the device. And then the virtual objects could be superimposed to the real scene according to the correct spatial perspective relationship before determining the location. All this leads to the functions of visual simulation and interactive teaching. The system is finally released on the mobile platform. Augmented reality technology has been applied to teaching in various forms, with excellent operability and interaction, which reduces the cost of teaching experiment and excites the students' interest.

Key words: augmented reality; ARCore; motion tracking; environment understanding; meshing

0 引言

增强现实(augmented reality, AR)是一种结合虚拟化技术观察世界的方式,它利用计算机技术将虚拟的物体实时地叠加到一个真实画面或空间,形成具有实时交互的三维图像画面,给用户带来更真实的体验与感受^[1]。在发展信息时代教育的过程中,现代教育技术以其先进的观念、手段和方法发挥着重要的作用。从木制黑板到电动黑板、投影仪、平板电脑、网课教学

等,教育水平和技术在不断改善和进步。增强现实这项新兴技术与教育相结合,可以更加促进教育水平的提升,在教育的进程上实现一个飞跃。

谷歌 ARCore 的工作原理是集成了安卓设备的加速传感器、陀螺仪、上下文识别信息等功能^[2]。首先 ARCore 从相机环境中甄别出一些可见的特征,并通过手势追踪、传感器传来的坐标位移等,实现现实世界与虚拟事物的一个映射功能。AR 系统具有三个突出的

收稿日期:2021-08-06

修回日期:2021-12-07

基金项目:2020 年度江苏省高校哲学社会科学一般项目(2020SJA2313);全国高等院校计算机基础教育研究会计算机基础教育教学研究项目(2020-AFCEC-275);2019 年度南京传媒学院校级“金课”建设项目(JK201906)

作者简介:宋燕燕(1978-),女,硕士,副教授,CCF 会员(F9498M),研究方向为数字媒体技术。

特点:(1)真实世界和虚拟世界的信息集成;(2)具有实时交互性;(3)在三维空间中添加和定位虚拟物体。ARCore 可以在移动设备移动时追踪它们的位置,并且会逐渐利用自己的方式对现实世界进行理解^[3-4]。

1 关键技术

将 AR 技术应用于移动终端,意味着系统能够提供以下主要功能模块:视频获取设备、实现高精度定位功能的跟踪装置、计算机视觉运算功能、3D 渲染功能、人机交互功能^[5]。需要综合考虑如何实现快速高效的跟踪定位算法、大场景范围下的多目标标识、移动设备与用户可见的自然交互等多方面的技术问题。

1.1 运动跟踪

英国牛津大学机器视觉组的 Davison 在 2003 年发表文献^[6],首次提出了基于主动重建的跟踪定位方法。该方法的思想源自机器人导航领域的同步定位和地图创建算法(simultaneous localization and mapping, SLAM)。近年来,随着 SLAM 算法的不断更新,已经成为一个研究热点问题。Davison 提出系统运行过程中,在估计当前摄像机位置的同时,对新观察到的特征点进行三维重建,以恢复其坐标位置,供后续位姿估计过程使用。算法采用扩展卡尔曼滤波技术,将摄像机位姿和特征点三维位置相结合作为状态变量 $\hat{x}$ 。

$$\hat{x} = \begin{pmatrix} \hat{x}_e \\ \hat{y}_1 \\ \hat{y}_2 \\ \vdots \end{pmatrix} \quad P = \begin{pmatrix} P_{xx} & P_{xy1} & P_{xy2} & \cdots \\ P_{y1x} & P_{y1y1} & P_{y1y2} & \cdots \\ P_{y2x} & P_{y2y1} & P_{y2y2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

其中, $\hat{x}_e$ 表示摄像机的运动状态, $\hat{y}_1, \hat{y}_2 \cdots$ 表示空间环境中待构建特征点的世界坐标, P 表示协方差矩阵。

Klein 等在 Davison 研究的基础上提出并行的跟踪定位与地图创建算法^[7],并应用于增强现实系统中。将 SLAM 系统分为跟踪和制图两个流程,跟踪定位的目的是在地图中估计摄像机的位置,地图创建的目的是利用跟踪获得的摄像机位置将图像上的特征点三角化得到空间坐标,对地图进行优化,最终得到空间点的最优估计。其地图创建的流程如图 1 所示。

目前使用的大部分视觉 SLAM 系统都是基于 Klein 提出的并行的跟踪定位与地图创建算法框架改进而来。Mur - Artal 等^[8]于 2015 年提出 ORB - SLAM,是现代 SLAM 系统中做的非常完善的系统之一。跟踪线程负责对每幅图像提取 ORB 特征点,并与最近的关键帧进行比较,计算特征点的位置并估计相机位姿。

目前通过光学摄像机可实现平面矩形图案、二维编码、自然图像以及立体物体的实时识别追踪。基于光学或深度摄像机的图像实时识别追踪会确保增强现

实应用的稳定性,因此它对识别追踪的速度、准确性、环境光的适应能力,对识别多标识同时追踪的容错能力有极高的要求。

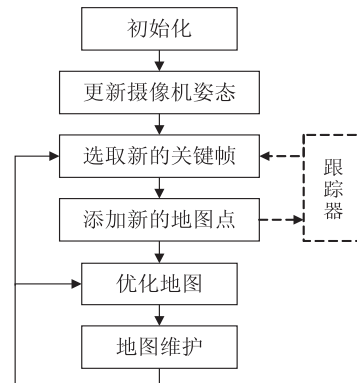


图 1 地图创建流程

提高基于图像识别追踪增强现实应用性能的方法有^[9-10]:利用图像分割与光流法相结合,高效与准确的运动捕捉,快速运动模糊图像中识别标识;通过平面自然图像特征点的离线与在线,进行实时训练来提高自然图像识别的速度与适应性;使用位移与旋转运动平滑过滤器,来减少图像识别误差带来的抖动影响;通过实时检测现实环境亮度并对图像亮度阈值进行调整来实现在不同光线条件下的自适应能力等。

1.2 环境理解

当移动 AR 技术被用在大范围环境或者需要对场景中的物体进行信息增强时,系统只需识别当前场景中的目标,并将相应的虚拟信息叠加显示在图像中。

1.2.1 环境理解的关键点

(1) 帧(Frame)。

即相机获取的一帧图像。在 ARCore 中,帧包含更丰富的内容,提供了某一个时刻 AR 的状态^[11]。这些状态包括:当前帧的环境光线,在绘制内容的时候根据光线控制物体的颜色;当前帧中检测到的特征点云和它的位置用来绘制点云;当前帧中包含的锚点和二维平面集合用于绘制内容和平面;设备当前的位置、帧获取的时间点、AR 跟踪状态等^[12-13]。

(2) 特征点云(PointCloud)。

系统在检测平面的时候,显示的白色小点就是特征点云。特征点云包含了被观察到的三维点和中心值的集合以及被 ARCore 检测到的时间。

(3) 二维平面(Plane)。

ARCore 中所有的内容,都要依靠平面类进行渲染。系统中的虚拟物体只有在检测到网格的地方才能放置。

ARCore 中平面可分为水平朝上、朝下和非水平平面类型。其描述了对一个真实世界二维平面的认知、平面的中心点、x 和 z 轴方向长度以及组成平面多边

形的顶点。ARCore 检测到的平面分为三种状态:正在跟踪、可恢复跟踪和永不恢复跟踪。如果是没有正在跟踪的平面,包含的平面信息可能不准确。如果发生两个或者多个平面被自动合并成一个父平面,可以通过子平面找到它的父平面。

(4) 平面交集 (PlaneHitResult)。

点击平面的时候,从设备点击开始手机朝向方向发出一条射线,看下和被检测平面是否有交集。HitResult 就是射线和真实世界几何体的交集。如果是平面交集就是 PlaneHitResult,如果是点云就是 PointCloudHitResult。PlaneHitResult 中可以判断交集点是否在被检测的集合范围内,是否在平面的正面。

(5) 锚点 (Anchor)。

每个物体的位置就是 Anchor。Anchor 位置会随着手机或者摄像头的位置变化而变化,这样才能保持“位置不变”。Anchor 描述了在现实世界中的固定位置和方向。

这个位置的数值描述将会随着 ARCore 对空间理解的改进而更新,使其保持在物理空间的固定位置。使用 `getPose()` 获取当前数值位置,这个位置每次被调用的时候都可能会改变,但不会自发地改变。

(6) 空间位置 (Pose)。

Pose 表示从一个坐标系,到另一个坐标系的精确的转换。在所有的 ARCore APIs 中,Pose 总是描述从物体的局部坐标系到世界坐标系的变换,来自 ARCore API 的 Pose 被认为等同于 OpenGL 模型矩阵。

1.2.2 网格化与环境

网格化与环境是能够识别对象的特征或是理解用户环境的基础,利用这些特征来帮助识别平面。ARCore 可通过网格化技术来自动识别表面,在其跟踪表面的过程中,需要多次执行网格化。

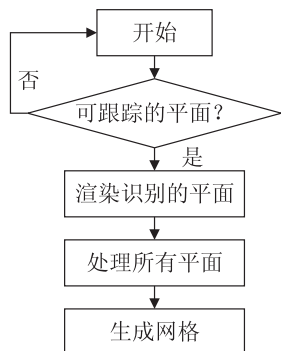


图 2 网格化算法过程

网格是一个显示多边形和构成该多边形的顶点有序集合的嵌入图形^[14]。顶点的顺序是沿逆时针方向连接,这些点的连接顺序会导致表面朝向的网格亮暗不同。渲染一个场景时只会看到朝向相机的表面,而远离相机的表面就会被去除或剔除,连接这些点的顺

序称为缠绕。

网格化是指获取特征点集合并从中构建网格的过程,然后生成的网格通常设置阴影并渲染到场景中,由 ARCore 生成并放置表面或平面的网格。网格化算法过程如图 2 所示。

1.3 光线估计

光线估计是一种复制现实世界中的光线条件并将其应用到三维虚拟对象上的技术^[15-16]。ARCore 利用图像分析技术从摄像机采集的图像中读取光线强度,并将其转换成全局光强度或颜色值,以及着色器上设置的光线的值。程序处理包括漫游方式、交互功能等。以往的瞬时移动可能会让体验者在频繁切换的场景中迷失方向,因此需要改善移动方式。交互功能一般情况下采用触碰的方式。而昼夜切换是在程序中预设昼夜变换系统,通过光照来感知。

2 ARCore 原理

利用移动设备摄像机来实现设备的移动跟踪,包括陀螺仪、电子罗盘、加速度计、相机视觉信息、SLAM 算法等^[17],采用 PL-VIO^[18]等算法利用点和线特征的紧密耦合的单目视觉惯性里程计系统,利用视觉惯性测距环境理解来识别水平和垂直平面,通过光线估计估算光源的角度和强度。

2.1 架构原理

ARCore 在运行时两路通道会同时打开,运用 SLAM 算法来支持分析原因。在 SLAM 算法中,视野越大效果越好,用于 SLAM 运算的通道是用其中一路来进行渲染。

Session 类是 SDK 对外接口的核心,用来封装相机预览到的现实世界和手机惯性传感器信息。通过 Session 的核心算法,来获取特征点、平面、光照估计等。

Ancho 类描述了建模的物体在现实世界中的固定位置和方向。这个位置的数字描述信息会改变,它会随着自身对空间更新的理解而不断地改进。

Pose 类表示从一个坐标空间到另一个坐标空间位置不变的转换。

LightEstimate 类保存真实场景中光照的估计信息。

Trackable 接口类使 ARCore 与锚点绑定在一起并且可以进行跟踪。

Camera 类用于捕获图像的摄像头信息。Camera 是一个长期存活的对象,每次调用 `Session.update()` 都会更新 Camera 的属性。

2.2 ARCore 核心方法

ARCore 核心方法如下:

(1) public Session (Context context): 通过 Session 获取帧和平面。

(2) public void resume: 打开相机和惯性传感器设备获取信息和执行核心算法。

(3) public Frame update: 从 Session 中得到帧, 包含特征点、姿态、光照、图像等信息。

(4) public void setCameraTextureName (int textureId): 获取现实世界的纹理 ID。

(5) public Collection<Anchor> getAllAnchors: 获取所有锚点。

(6) getAllTrackables (Class<T> trackable): 获取所有继承的平面和点。

(7) public Anchor createAnchor (Pose var1): 根据空间位置创建锚点。

(8) public List<CameraConfig> getSupportedCameraConfigs: 获取相机所有的支持项。

(9) frame. getLightEstimate (). getColorCorrection (colorCorrectionRgba, 0): 获取光线。

(10) PointCloud pointCloud = frame. AcquirePointCloud: 获取点云。

(11) backgroundRenderer. draw (frame): 根据模型矩阵、光照估计、投影矩阵绘制虚拟物体。

ARCore 的核心算法使虚拟环境在现实中更加真实。

3 系统设计

采用 ARCore 进行核心方法设计, 实现移动平台上的场景可视化与互动教学。系统设计方案如图 3 所示。

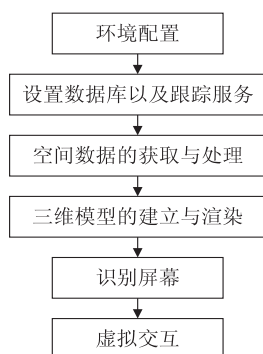


图3 系统设计方案

利用 ARCore 构建 Android Unity 和 Wech 平台下的增强现实, 基于 Firebase 的跟踪服务能够跟踪用户的运动, 实现在多个应用程序跟踪用户, 或同时跟踪多个用户, 编写服务器程序、建立数据库、制定一种架构模式。在移动增强现实应用系统建模的制作过程中, 前期要收集整理项目资料以及需要项目的架构图, 并且对真实环境进行数据采集。

3.1 ARCore 的可视化运动跟踪

ARCore 通过实时跟踪用户来具体使用跟踪服务, 并在 AR 应用程序中建立可视化跟踪数据。设置一个重复计时器, 使之每秒调用一个匿名函数。这样是为了跟踪每秒的运动, 也可以像在多人游戏中那样跟踪每一帧运动。最后在设备的浏览器中运行, 可以让设备四处移动。切换到数据库终端, 可观察到数据流入数据库。任意扩展数据点, 并查看捕获的数据值, 就可以看到采集的数据了。把这些数据在二维或三维中进行可视化, 利用发送到数据库的相同数据来绘制用户经过的三维路径。核心代码如图 4 所示。

```

var material = new THREE.LineBasicMaterial({ color: 0x0000ff});
var geometry = new THREE.Geometry();
geometry.vertices.push(new THREE.Vector3(pos.x, pos.y, pos.z),
new THREE.Vector3(lastPos.x, lastPos.y, lastPos.z));
var line = new THREE.Line(geometry, material);
scene.add(line);
lastPos = { x: pos.x, y: pos.y, z: POs.z};
  
```

图4 可视化核心代码

调用 THREE.LineBasicMaterial 创建一条基准线 material, 并赋予十六进制的颜色值。接下来, 根据当前的 pos 和 lastPos 变量以及 material 创建 geometry, 即 line。利用每个位置上 x、y 和 z 的值构建 Vector3 对象来实现。最后, 利用 scene.add(line) 函数在场景中添加 line。当利用设备四处运动时就会有一条可视化运动轨迹的蓝色的线。ARCore 中的运动跟踪是通过识别和跟踪用户周围的可识别特征来完成的, 然后利用这些点以及设备的朝向和加速度传感器保持跟踪更新。如果没有这些信息, 就无法实现准确跟踪。

3.2 环境理解与交互

ARCore 可以提供在用户周围识别到的特征点和平面, 并且根据这些识别点或平面, 生成虚拟对象。由于 ARCore 可以跟踪这些点和平面, 因此在用户移动对象时, 附着在平面上的虚拟对象应保持固定不变。

通过 USB 或远程将设备与开发机器相连, 在设备上构建和运行应用系统。绘制点云数据需要花费一段时间, 并且随着点的个数的增加, 在一个方向保持设备的时间就越长。这些点表示用于跟踪和解释环境的可识别特征点。放置对象用光线投射技术, 光线投射技术可在两个维度上获取触摸点, 并向场景投射光线, 然后该射线测试场景中产生碰撞的其他对象。

3.3 光线方向估计

ARCore 根据来自设备的当前图像, 通过图像分析算法确定光线强度, 并作为全局光线用于三维对象。通过 Shader.SetGlobalFloat (“_GlobalLight - Estimation”, 1.0f) 语句, 将着色器变量设为 1, 也就是将光线强度设为 1.0。调用 Frame.LightEstimate.Pix-

elIntensity,从相机读取图像确定当前像素强度,并归一化处理。最终通过线性模型进行光线估计。光线方向的估计也是利用 Frame.LightEstimate. PixelIntensity 读取相机方向的光强度,检查该值是否小于之前的所有值,如果小于,则在相机的相反方向上旋转光线。

在互动教学中,可以将摄像头对准任何平面,生成网格化平面的效果如图 5 所示,利用数据显示运动轨迹,点击平面后叠加虚拟场景,通过触摸移动事件和表面跟踪,在场景放置物体(如椅子)并进行位置调节等完成室内空间的布局,如图 6 所示。

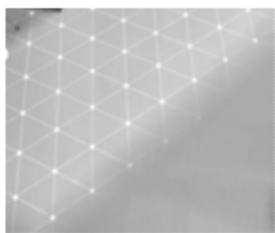


图 5 生成网格化平面



图 6 虚拟场景互动效果

4 结束语

现在人工智能和增强现实技术在社会中的地位稳步上升,这些领域的核心关键技术正在逐步实现突破。SLAM 跟踪定位给增强现实在构建完整地图提供了良好的开端,采用 ARCore 进行架构和核心方法设计,通过运动追踪和定位,环境理解及光线估计计算出虚拟世界与现实世界坐标系的对应关系,进行移动增强现实应用系统设计,实现交互体验。在这些应用系统设计中,不希望跟踪定位等占用所有计算资源,所以对小型化和轻量化提出了强烈的需求。另一方面,利用高性能计算设备实现精密的三维重建和场景理解成为必然。未来将和深度学习技术结合,为构建准确的语义地图打下基础,将神经网络方法引入到物体识别和分割,甚至位姿估计和检测中,使增强现实系统能够在各

种干扰的条件下稳定运行。

参考文献:

- [1] 宋燕燕,秦 军,邢艳芳,等. 基于多智能体的图优化 SLAM 构建方法[J]. 计算机技术与发展,2020,30(11): 205-209.
- [2] GÜNGR C. Augmented reality development tools and google ARCore[C]//ISMSIT 2017 1st international symposium on multidisciplinary studies and innovative technologies. Muffler. Tokat, Turkey: [s. n.], 2017.
- [3] 周 玲,陈 强. 城市地下管网增强现实系统研究[J]. 计算机工程与应用,2020,56(1): 251-256.
- [4] 王 崑,彭勃宇,周 诚,等. 面向增强现实的多传感器组合位姿注册算法[J]. 传感器与微系统,2015,34(7): 132-136.
- [5] 王涌天,陈 靖,程德文. 增强现实技术导论[M]. 北京:科学出版社,2015.
- [6] DAVISON A J. Real-time simultaneous localisation and mapping with a single camera[C]//IEEE international conference on computer vision. Nice: IEEE, 2003: 1403.
- [7] KLEIN G, MURRAY D. Parallel tracking and mapping for small AR workspaces[C]//IEEE & ACM international symposium on mixed & augmented reality. Nara: IEEE, 2008: 1-10.
- [8] MUR-ARTAL R, MONTIEL J M M, TARDOS J D. ORB-SLAM: a versatile and accurate monocular SLAM system [J]. IEEE Transactions on Robotics, 2015, 31(5): 1147-1163.
- [9] 于敬延. 移动增强现实组网与应用研究[D]. 北京:北京邮电大学,2017.
- [10] 陈和恩. 基于单目摄像机的增强现实场景感知技术研究[D]. 广州:广东工业大学,2016.
- [11] 刘 宇. 深度学习在图像处理中的应用[D]. 成都:电子科技大学,2020.
- [12] 邢志伟,朱晓蕊,何 超. 基于行人行为学习的机器人同时定位与可通行区域制图[J]. 机械工程学报,2019,55(11): 36-45.
- [13] 张臻炜,赵 勇,布树辉. 一种基于计算机视觉的无人机实时三维重建方法[J]. 机械与电子,2015(12): 19-22.
- [14] 于瑞云,苏 展,谢 青,等. 基于空间分割的人体模型骨骼提取算法[J]. 计算机学报,2019,42(9): 2049-2061.
- [15] 陈宝权,秦学英. 混合现实中的虚实融合与人机智能交融[J]. 中国科学:信息科学,2016,46(12): 1737-1747.
- [16] 宁瑞忻,朱尊杰,邵碧尧,等. 基于视觉的虚拟现实与增强现实融合技术[J]. 科技导报,2018,36(9): 25-31.
- [17] 王 巍,王志强,赵继军,等. 基于移动平台的增强现实研究[J]. 计算机科学,2015,42(S2): 510-519.
- [18] HE Yijia, ZHAO Ji, GUO Yue, et al. PL-VIO: tightly-coupled monocular visual-inertial odometry using point and line features[J]. Sensors, 2018, 18(4): 1159.