

浙江大学

硕士研究生读书报告



题目 增强现实浏览器

作者姓名 朱坤坤

作者学号 21551012

指导教师 李启雷

学科专业 软件工程

所在学院 软件学院

提交日期 二〇一五年 十二 月

The Effect Of The Requirements Analysis On The System Design

A Dissertation Submitted to
Zhejiang University
in partial fulfillment of the requirements for
the degree of
Master of Engineering

Major Subject: Software Engineering

Advisor: Qilei Li

By

Kunkun Zhu

Zhejiang University, P.R. China

2015

摘要

尽管低水平的硬件设备已经能允许我们在真实的世界里显现虚拟的物体,但是以更为结构化的方式去管理它仍然有待于加强。虚拟现实浏览器技术可以类比为一种路由,用于连通结构化的软件平台与任意位置的增强现实用户。在现实世界中增强现实浏览器是 Web 浏览器的替代品,这种浏览器能够将相互作用的多媒体内容覆盖在现实世界的物体上。如今的增强现实浏览器仅能差强人意地允许我们看到虚拟物品简单悬浮于现实世界之上,与现实自然无缝地连接仍处有待探索阶段。本篇报告将介绍如何创造下一代更为富有、连续、灵活的增强现实浏览器。并讨论在这一领域的主要挑战,和未来的研究方向。

关键词: 增强现实浏览器、移动设备

Abstract

As low-level hardware will soon allow us to visualize virtual content anywhere in the real world, managing it in a more structured manner still needs to be addressed. Augmented reality (AR) browser technology is the gateway to such structured software platform and an anywhere AR user experience. AR browsers are the substitute of Web browsers in the real world, permitting overlay of interactive multimedia content on the physical world or objects they refer to. As the current generation allows us to barely see floating virtual items in the physical world, a tighter coupling with our reality has not yet been explored. This paper presents our recent effort to create rich, seamless, and adaptive AR browsers. We discuss major challenges in the area and present an agenda on future research directions for an everyday augmented world.

Keywords: AR browser; augmented reality (AR); mobile devices

1 引言

自从 1990 年迈出了第一步之后，增强现实技术一直在不断地大力发展。当年的增强现实体验需要携带笨重的定制化硬件设备，现今，一台智能手机就足矣。这一便利性大大促进了增强现实的大众化使用，现今增强现实已经大量用于市场（增强一些杂志的纸质体验）、游戏（在桌面或街道上显现出虚拟的角色），或是家居购物（在你现实的家中放置虚拟的家居）。

一个最为成功的增强现实应用是这样一款东西，它是电脑或手机上用于现实真实世界的 Web 浏览器，通常被认为是增强现实浏览器。从历史上看，这一术语是有 SPRX mobile 提出，先与学院派和工业界。这一类型的应用现今已经下载到或内置于 5000 万的智能手机上。并且伴随着低消耗的头部可穿戴设备的出现（如谷歌眼镜），我们可以预想这一类的应用会大量整合在我们的日常生活中。

AR 浏览器可以直接将现实世界以视频流的形式展现出来，比如一个餐馆的图标，这一方式可以直接减少认知的工作，并且能够提供传统基于地理位置接口的产品（地图、lists）的优势。为了便于浏览，登记的图标可以超链接到更多的内容，如果有需要的话。

尽管 AR 浏览器已经经历了在硬件和软件方面的大量基础性整合，但商业的 AR 浏览器仍没有达到它最大的潜能。他们的内容方面仍然十分地贫瘠，不够流畅，并且是静态的。

在这个报告中，我们讨论一些有争论的话题关于现今这一代的 AR 浏览器和他们的历史发展（有学术性的也有商业性的）。并且还会陈述，如何让我们的下一代 AR 浏览器的内容更为充足，画面衔接更为真实，并且与我们真实世界更能灵活适配。

我们倡导下一代增强现实浏览器的内容应该更为富足，通过以更为有效和高质量的方式去整合大量的不同类型的数字媒体，需要有建筑学方面的知识，和大量不同类型的多媒体接口支持。我们同样也强调社会化媒体。进一步地，我们希望 AR 浏览器能够更融合地方式将这些多媒体内容整合到物理世界中。要求我们一方面能准确地追踪，一方面也要提高管理技术。最后，我们强调 AR 浏览器对当前环境的适配，让它能够充分理解用户现在所处的环境。我们从我们现在的一些研究成果中选择一部分内容来阐述这些问题，并且讨论一些开放的问题和研究

的方向。

这一论文通过以下方式组织。第二部分我们对现今的 AR 浏览器的结构和做出下一代 AR 浏览器存在的挑战做一下鸟瞰。第三-五部分展现一些研究的背景和我们在应对这些挑战上所作的工作。第六部分 从我们的工作引出的一些讨论。第七部分是该论文的结论。

总的来说,这篇论文的贡献是有组织地分析了关于 AR 浏览器和研究成果(AR 社区和我们自己的研究) 方面的主要研究话题。

2 AR 浏览器鸟瞰

增强现实应用是复杂的多媒体系统,定位于将虚拟混合的数字信息整合入现实的物理世界中。下面的部分我们将给出 AR 浏览器在数据结构、软件组件、限制方面的鸟瞰。

A.数据结构

AR 浏览器的主要数据项是基于地理坐标点或兴趣目标参考点。基于地理坐标的信息有多种组成形式,有文字的、插图的、少部分是以视频、语音、3D 媒体的形式。信息通常是以地理位置为导向,用到经纬度的信息,或者通过虚拟的指纹与特定的物体联系在一起。

存储的结构通常是类似于 XML 的数据库,最近 HTML 也有少量地运用。与新闻推送相似,内容频繁地被组织在各种主题频道内。

现在大量的增强现实浏览器复用了传统的 Web 应用里的基于 GPS 标签的信息来提高可获取信息的数量。典型的例子是, Twitter 的反馈, Flickr 的图片, Wikipedia 的文章。它们也支持专门为 AR 浏览器创作的信息,通常通过由浏览器厂商提供的基于地图的桌面工具来完成。

B.软件组件

与大多数移动信息系统相似, AR 浏览器大致的结构也遵循 C/S 的模型该服务器负责内容存储和检索的查询客户端(例如,提供在我的位置有 10 米半径的 3-D 模型)。客户端是负责渲染的内容,让用户交互(浏览)。虽然移动 AR 浏览器软件的基本架构类似于 3D 游戏或模拟, AR 需要一些额外的功能。

首先,移动设备需要确定其姿态(位置和方向)在物理世界。这是需要识别环境信息(是什么在我的位置?),同时也将任何本地信息到当前用户的观点(我

可以从我目前的位置看？)。这些过程通常由注册组件完成。商业 AR 浏览器通常采用内置一些传感器，GPS 用于定位，加速度计和陀螺仪用于确定方向。

然后，使用由内容组件此空间信息来启动从服务器到客户机的流相关的内容。所述用户界面组件是负责数据呈现所述内容。可以光学来实现的真实和虚拟内容结合视觉演示文稿，但今天最为方便的选择是计算机图形添加到的数码相机馈送内置在移动装置的照相机和呈现的结果移动设备的屏幕。

C.限制

尽管 AR 浏览器，其主要架构已经有很长的历史的研究，可以追溯到原始的旅游机在世界 - 板真正的万维网浏览器中，当前一代的 AR 浏览器只能被看作是对履行已经表示在这些早期作品的最终愿望的方式的一个中间步骤。还有而言显著的缺点：1) 注册准确性。 2) 质量不足和内容的数量;3) 缺乏灵活性和专用软件的 AR 浏览器架构;并且，的信息呈现终于，4) 可用性较差。

因此，未来的 AR 浏览器需要处理的后续研究的挑战。

•准确和全球注册：

虽然许多技术用于计算移动设备的姿态存在，大多数解决方案不支持大和无法控制的环境。这会影响虚拟内容的无缝的集成，用户都背负着注册问题。

•无缝登记：

今天的 AR 浏览器依靠的登记方法时，例如用于户外环境（GPS）或基于对象的登记（特征匹配，基准标记）的组合。这些方法之间的切换破坏用户体验，应在未来的 AR 的浏览器（从室外移动到室内时，例如，）制成无缝。

•内容密度：

AR 浏览器依靠位于媒体，但位于介质的分布显著变化。虽然城市中心可以与兴趣点进行杂乱，农村地区可能缺乏任何有趣的内容。理想情况下，我们希望有密集的 CON-10 吨报道随处可见，适当的自动过滤和选择工具来管理混乱，当浏览 AR 浏览器。

•丰富的内容：

虽然富媒体丰富的 Web 应用程序，AR 浏览器的数据库严重依赖文本内容，缺乏其他媒体类型。

•无缝内容整合：

一个视觉上增强场景的用户的理解是虚拟信息的融合在物理世界中的质量依赖，在空间和感知。不幸的是，AR 浏览器不提供在这些方面令人满意的体验呢。

•适应性：

AR 浏览器可以在物理场所的大型弗吉尼亚州 riety 使用。由于我们的现实世界是一直亮着我们身边的变化，AR 浏览器必须能够适应。例如，计算机生成的项目的对比度可能需要变化显着从晴天到阴天或室内环境。

而其中的一些挑战，必须通过改进现有的组件，如注册或用户界面来解决，其他挑战需要纳入新的部件，如自适应视图管理。在本文的其余部分，我们提出我们对这些挑战的工作。

3.环境录入

每个 AR 应用的基础部分是将虚拟内容在现实世界坐标系中的精准对位。通常，登记是由两部分构成，定位步骤和随后的跟踪步骤。定位将提供一张精确的绝对位置，而跟踪提供了一个连续的相对位置。照片应该由 6 个自由的角度并且以 30HZ 以上的速度进行追踪，由此产生方向和距离上的精度。在实践中满足这些要求仍然具有挑战性，尤其是在室外环境中。

A.背景与挑战

在 AR 浏览原始研究中使用的昂贵的检测设备，如实时动态 GPS 获得可接受的准确度[9]。现今在智能手机上的 AR 浏览器只会使用消费级传感器用于他们的处置。即使有其它传感器如 WiFi 或用于全球移动通信系统（GSM）的协助下，中值位置误差被报告为 5- 8.5 米并且工作频率也只有 1-2 赫兹。从磁力取向的错误只有在几度范围内，但很大程度上受外部因素影响，如附近的磁力环境。这样的表现显然是不够细粒度标注的要求。Feiner 已在 1997 年说明，由于不准确的传感器，它仅仅可能做到分配标注到建筑物上，但无法具体到细节（例如，窗口，广告）。

相比之下，计算机视觉算法在定位和具有较高的精度和性能跟踪上已经足够成熟。与非视觉传感器相比，相机硬件和移动设备的计算能力正在迅速提高。在下文中，我们给出了新的计算机视觉为基础的技术用于在下一代的 AR 浏览器的室外登记。我们组织这部分的内容是根据先验知识在环境测量上是否可用来的，

我们也考虑将非视觉传感器与计算机视觉技术的结合。

B. 无先验知识下的录入

在同时定位和地图测算 (**SLAM**) 最近的结果里表明, 在没有任何事先场景知识的条件下, 以 6 个自由角度来确定相机的角度是可能的 (例如, 放置在环境或三维激光扫描场景中的基准标记)。因为这样的先验知识通常是不适用, **6DOF SLAM** 可以大大延长 **AR** 可使用的范围。然而, **6DOF SLAM** 有缺点: 除了高计算要求外, 它还需要测量人走一段长距离来达到对大型实体 (建筑) 的可靠测量, 这很容易产生错误的积累。因为 **6DOF SLAM** 可以在移动设备上, 它被限制于小规模的环境中, 因此大多在室内使用。

为了使室外录入变得可行, 我们可以通过组合的高分辨率取向跟踪与低分辨率位置跟踪 (通过 **GPS**) 的方式来重构在未知场景精确跟踪的问题。这被事实所证明, 在室外环境, 旋转误差比位置误差更容易被感知。

4. 结论

我们认为, 新一代的 **AR** 浏览器需要无缝注册, 丰富的内容, 以及自适应用户和控制台, 才能取得成功。我们讨论了这些方面是如何影响 **AR** 的关键组成部分, 即登记, 内容和用户界面。

无缝登记仍然是一个挑战性的课题, 这需要进一步的研究, 特别是与其他组件强烈的配准的质量, 以实现所希望的结果依赖。新的追踪技术来实现这一目标, 如全景跟踪和映射方法结合基于服务器的图像查询或多个传感器融合。

丰富的内容将在下一代 **AR** 浏览器的成功显著的作用。因此, 各种各样的内容形式如文本, 视频, 音频, 和 **3-D** 内容的无缝集成是一个重要的发展。对于这个问题, 我们了解到经由 **AR** 全景跟踪的精确放置的文本注释, 音频注释结合视觉方式, 原位制作, 和视频内容无缝例子以及创作和三维内容使用。

丰富的内容也将导致更多的挑战。因此, 改善视图管理, 统一的用户界面是新一代的 **AR** 浏览器的重要组成部分。