

VR头盔的虚拟形象动画生成

张有为*, 周航宇, 张志华

北京信息科技大学计算机学院, 北京

收稿日期: 2022年11月17日; 录用日期: 2022年12月17日; 发布日期: 2022年12月26日

摘要

现如今VR技术迅速发展, 相关的游戏应用层出不穷, VR头盔手柄等硬件也有十足进展。本论文意图从人机交互入手, 提出采用动作识别的方式, 当用户做出不同动作时, 虚拟人物可以依据用户的动作进行相应的行动, 实现用户和动画形象之间的高度互动, 在虚拟形象的设计方面, 我们利用VRoidStudio进行模型设计, 并导出模型。我们创新地使用动作捕捉为3D模型添加骨骼数据, 实现3D动画生成, 使用动作识别技术丰富VR设备使用者的体验感。

关键词

VR技术, 交互虚拟, 动作捕捉, 动作识别

Virtual Image Animation Generation of VR Headset

Youwei Zhang*, Hangyu Zhou, Zhihua Zhang

Computer School, Beijing Information Science & Technology University, Beijing

Received: Nov. 17th, 2022; accepted: Dec. 17th, 2022; published: Dec. 26th, 2022

Abstract

With the rapid development of VR technology, there is an endless stream of relevant game applications, and hardware such as VR headset gamepad has also made full progress. The purpose of this paper is to start with human-computer interaction and propose the method of action recognition. When the user makes different actions, the virtual character can take corresponding actions according to the user's actions, so as to achieve a high degree of interaction between the user and

*第一作者。

the animated image. In terms of the design of the virtual image, we use VRoidStudio to design the model and export the model. We innovatively use motion capture to add bone data to 3D models, realize 3D animation generation, and use motion recognition technology to enrich the experience of VR device users.

Keywords

VR Technology, Interactive Virtual, Motion Capture, Action Recognition

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

虚拟数字人的应用场景众多,可广泛与各行业领域相结合,包括电商直播、虚拟偶像以及品牌营销等方面。目前,网络中已经有了大量的虚拟动画形象,在3D虚拟动画形象的设计过程中有采用“关键帧法”的方法进行设计动画,通过Proser库中的各种素材对虚拟人进行外形设计[1],这种方法的优点是使设计出的动作具有逼真的效果,便于建立起模型动作库,但是缺点是制作周期时间长,设计出的动画不能根据变化的需求而改变。或者有的需要动作提供者佩戴感应器进行动作捕捉,这样的方法被广泛应用到电影、动漫的制作中[2],优点是能够做出高难度的动作,动作与佩戴者做到一致,但是需要的硬件设备条件较为严格。

在传统的VR用户交互方面,用户需借助模拟手在虚拟现实运动的手套或者手柄和携有显示屏的头盔,进入计算机生成的虚拟世界,与虚拟世界里的人与物等“现实”进行交互[3]。但是这种方式受设备型号影响,不具备通用性。

本项目提出采用动作捕捉技术,通过获取动作演示者的骨骼数据,将骨骼数据与3D模型结合绑定,实现动画的生成,基于动作识别技术,仅利用设备自带的摄像头,当用户戴上VR头盔后即可与虚拟动画形象进行互动,丰富了VR设备的使用体验感。

2. 项目工作概述

2.1. Oculus Quest2 头盔开发

Oculus Quest2是Facebook在2020年9月发布的VR一体机,特点是与PC端连接功能强以及较为强大的动作追踪技术,可以支持手部追踪,脱离传统的手柄控制模式,这也是选择这款头盔进行开发的原因。同时相较于HTC等其他品牌的VR头盔而言,无需接线无需定位器的设计更为便捷。

2.1.1. 激活环境

Oculus系列的VR头盔因为是Facebook开发并发售的原因,对于中国大陆的用户并不怎么友好,繁杂的Facebook注册过程不说,同时还需要搭载专用的vpn网络,但好在随着时间网站上各类教程都有且不算困难,跟着教程搭建好无线热点即可使用Oculus Quest2联网。

2.1.2. 开发者模式

想要在Oculus上进行VR软件开发需要下载Oculus软件并使用Facebook注册Oculus账户,在Oculus

Quest2 头盔联网后，软件会查找头盔并将头盔与账号绑定，届时可以在设置里开启开发者模式，将开发软件通过 Type-C 数据线与 VR 头盔进行连接，即可游玩 PC 端的 VR 软件或者在 VR 中运行自己开发的 VR 软件。

2.2. Unity 开发

Unity 是现如今 VR 软件开发比较常用的几款软件之一，支持 3D、2D 画面开发的同时，也预制了许多与 VR 头盔配套的开发者插件，依靠这些插件，开发者能够较为快捷方便的进行 VR 软件开发，而不用全部都从零开始[4] [5]。

2.2.1. 环境配置与连接头盔

选用 Oculus 系列头盔的另一个原因则是 Unity 本身就有为 Oculus 开发设计的软件包。在包管理器中找到“XR Interaction Toolkit”、“XR Plugin Management”、“Oculus XR Plugin”三个软件包并下载，前两个是虚拟现实开发插件，最后一个则是用于连接头盔的专用插件。

要让 Unity 与 Oculus Quest2 进行连接，需要先配置好适合头盔的虚拟环境。因为 Oculus 系列的头盔从技术上讲属于安卓系统，所以需要将 Unity 项目的文件 - 生成设置中将平台改为 Android，同时将纹理压缩改为 ASTC (见图 1)。



Figure 1. Android configuration

图 1. 安卓系统配置

之后打开项目设置，在 XR Plugin Management 这一栏中选中安卓系统，再勾选 Oculus 选项即可(见图 2)。

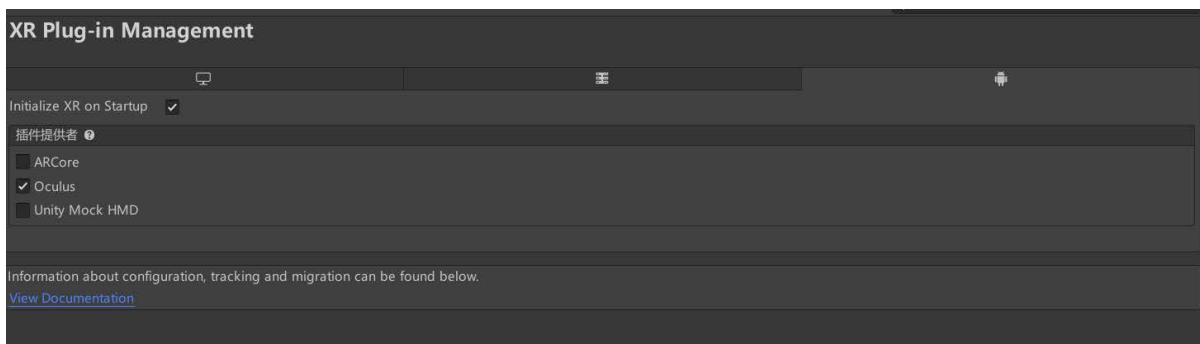


Figure 2. XR Plugin Management page

图 2. XR Plugin Management 页面

在这之后需要调节识别安卓系统连接的最低 API 等级，在项目设置 - 玩家 - 识别这一栏中找到最低 API 级别，将其调整至 Android 6.0 ‘Marshmallow’ (API level 23) (见图 3)。

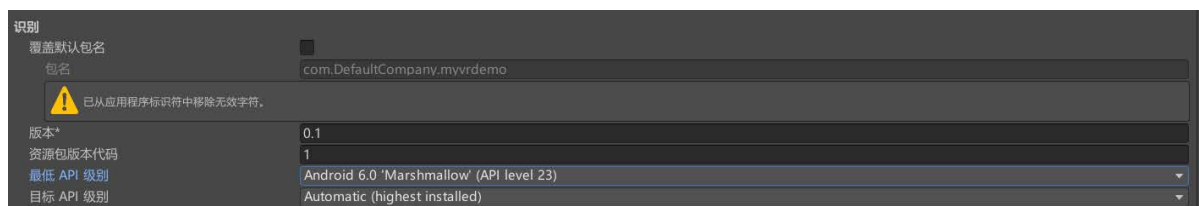


Figure 3. Android system connection identification

图 3. 安卓系统连接识别

通过以上操作即可完成 Unity 项目对于 Oculus 头盔的环境配置。在包管理器中还有很多种可以用于 Unity 游戏、软件开发的资源包，可供开发者使用，在此不再多进行赘述。

将头盔和电脑用数据线连接，在 Oculus 软件上绑定头盔并开启开发者模式，这时头盔内会显示是否需要 USB 调试，点击确认。在 unity 中打开生成设置，在运行设备一栏选中“Oculus Quest2”。全部设置完成后点击生成即可在 VR 头盔中运行我们在 Unity 上开发的项目。

2.2.2. 动作模型

在 Unity 中运行的 glb 格式文件，也就是模型文件一般分为两个部分，一是模型本身的形态和材质，二是动画 Animator 插件，有了 Animator 模型才会动起来。开发者可以在动画编辑器中对 glb 文件的动作进行剪辑和编辑，而 glb 模型文件的生成在之后会详细解释。

2.2.3. 手柄交互

为了实现 VR 手柄的交互功能，我们可以在 Unity 上用其他组件代替两个手柄。创建两个空组件并为它们添加 XR-Controller (Device-based)这一插件，并设置左右手，连接 VR 头盔后这两个组件就会作为手柄用来与模型进行交互。图 4 是在 Oculus Quest2 头盔中用手柄与方块模型交互的测试。

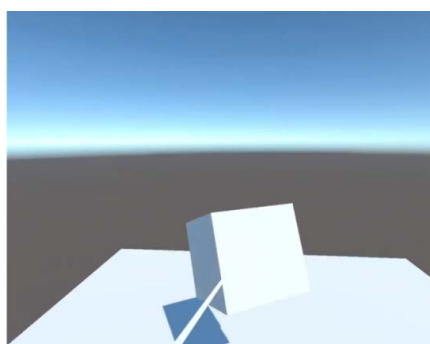


Figure 4. Gamepad interaction test

图 4. 手柄交互测试

2.3. 生成虚拟形象模型以及利用动作捕捉为虚拟形象添加动作

目前 3D 虚拟形象建模主要借助精准建模和后期细节处理，应用人工智能、大数据等先进技术，对人体动作做半自动化的数据标注，产生出大量的 3D 运动数据，结合语音数据、图像数据，共同做多模态建模，最后生成一个可驱动的 3D 数字人。这样固然能生成一个很好的 3D 虚拟形象，但是这种方法需要实验者佩戴动作传感设备，对采集的环境也有相当的要求，同时制作虚拟形象的周期也较长。本项目使用 VRoidStudio 进行快速创建模型。VRoidStudio 能让使用者通过类似绘画的方式构建虚拟形象模型，在 VRoidStudio 中我们可对想要的虚拟形象进行快速配置，比如 Face、Hairstyle、Body、Outfit 等(如图 5 所示)配置完成后导出 vrm 文件即可。目前项目拥有的虚拟形象有共用 3 个，如图 6 所示：

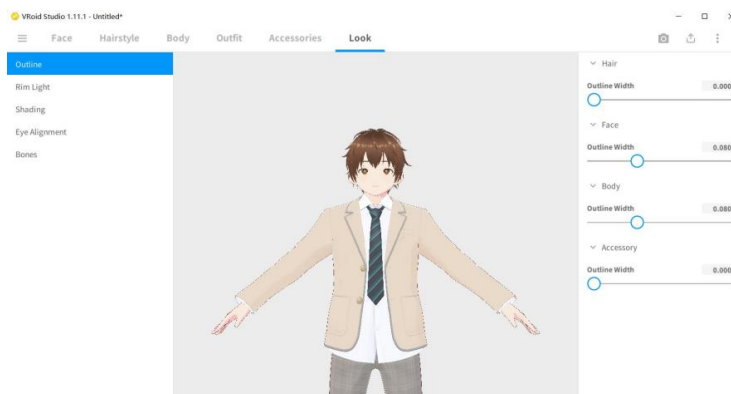


Figure 5. Configuration image
图 5. 配置形象



Figure 6. Avatar model
图 6. 虚拟形象模型

本项目为实现动作捕捉本项目采用了简便的工具 ThreeDPoseTracker。ThreeDPoseTracker 是 Windows 10 的应用程序，允许仅使用 USB 摄像机或含有具体动作的视频执行动作捕捉。该工具支持捕捉视频中人物的动作，也支持通过摄像机对人物的动作进行捕捉，但是仅支持一人。本项目选择先将人物的动作录制成视频，通过该工具获取骨骼数据后将虚拟动画的骨骼同骨骼数据绑定，实现虚拟形象的动作设计，如图 7，图 8 所示。

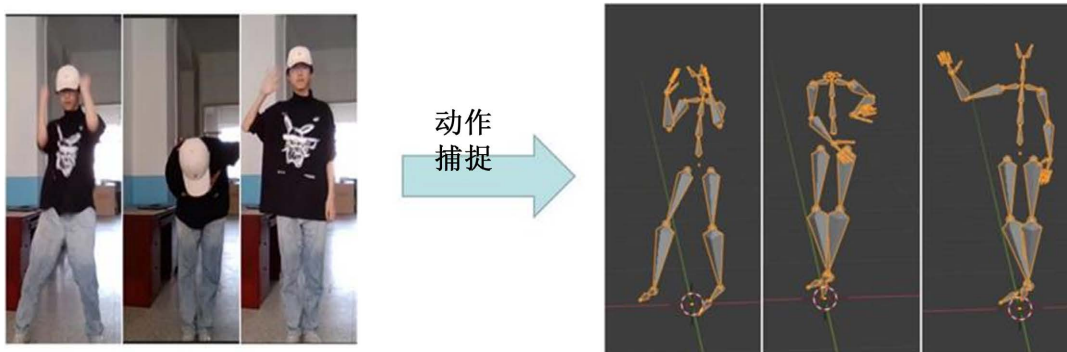


Figure 7. Motion capture-1
图 7. 动作捕捉-1

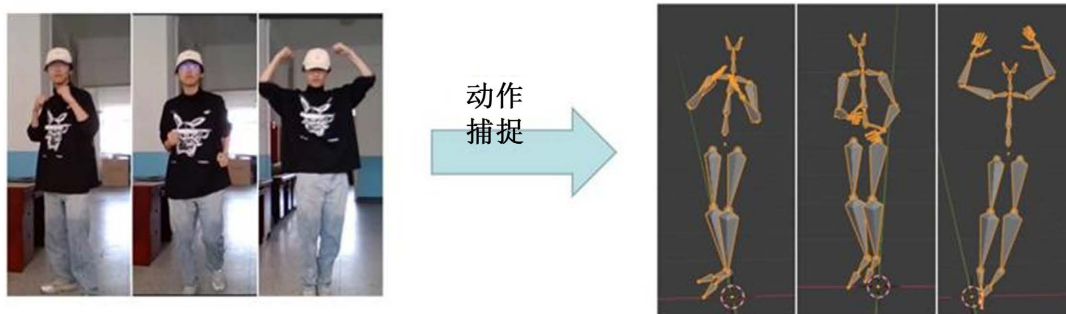


Figure 8. Motion capture-2
图 8. 动作捕捉-2

将生成的模型导入到 blender 中，在 blender 对生成的模型进行加工，修改部分可能出错的地方。利用 Rokoko 在 blender 中的插件实现骨骼数据与虚拟动画的骨骼数据绑定，至此模型便能动起来了。Rokoko 是一个运动库插件，能在 Maya、unity 等软件中都可使用，本项目主要使用它的绑定骨骼数据功能。以挥手动作为例，3D 模型将根据骨骼数据进行运动，如图 9 所示。



Figure 9. Waving
图 9. 挥手动作

2.4. 进行动作识别

本项目为实现对使用者的动作进行识别采用了 MMAction2，MMAction2 是一款基于 PyTorch 的视频理解开源工具箱。其包含了视频理解领域的常见任务，比如动作识别，时序动作检测，时空动作检测等。我们使用该工具对参与者进行了动作识别，识别结果如图 10 所示，使用多个标签将人物的动作进行识别标注。

2.5. 调用 Unity 实现的接口

在 Unity 中设计好 python 接口，让动作识别结果与 3D 虚拟形象的动作形成对应关系，通过 Unity 的接口控制 VR 头盔中的形象生成，同时 3D 虚拟形象的动作还会随使用者的动作的变化而变化。在 Unity 中看到的虚拟形象如图 11 所示：



Figure 10. Display of action recognition results
图 10. 动作识别结果显示



Figure 11. The model acts accordingly
图 11. 模型做出相应动作

3. 相关算法

Temporal segment network

在动作识别中项目使用到了 MMAction2 中 TSN—temporal segment network，这是一种基于视频的动作识别的新框架，它结合了稀疏时间采样策略和视频级监督，使整个视频得到高效和有效的学习[6]。TSN 网络结构图如图 12 所示：

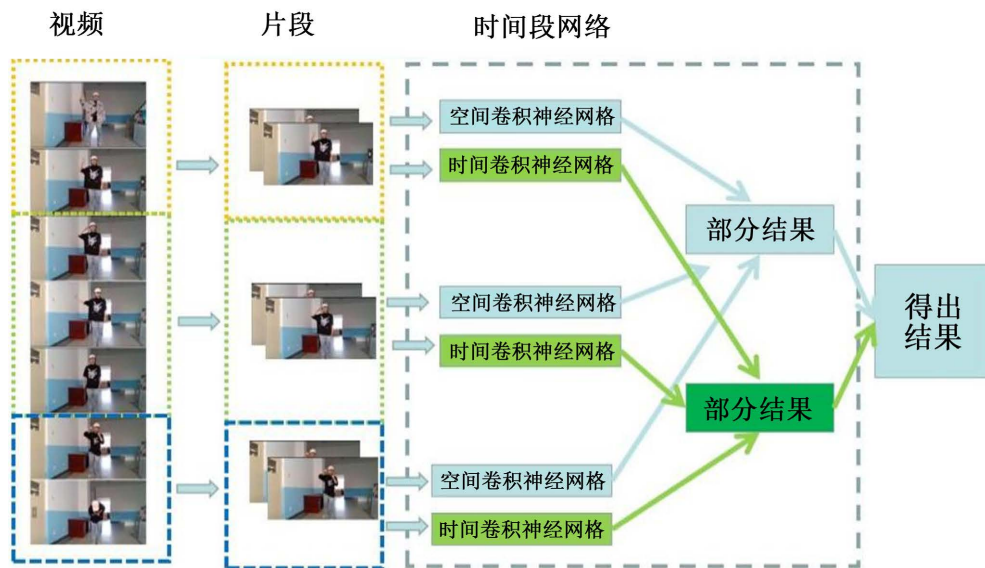


Figure 12. Time segment network structure diagram

图 12. 时间片网络结构图

首先将输入的视频分成 n 个片段，随机取出一个片段进行取样分析，再通过基于空间的卷积神经网络和时间卷积神经网络将输出结果进行类别分布，对部分结果进行均值处理，最后得出结果。

稀疏时间采样策略如下式所示：

$$TSN(T_1, T_2, T_3, \dots, T_n) = H(G(F(T_1; W), F(T_2; W), F(T_3; W), \dots, F(T_n; W))) \quad (1)$$

SlowFast Stream

SlowFast Stream 是一种用于视频识别的用于视频识别的双流模型(如图 13)。一个分支用来捕获图像或稀疏帧所提供的语义信息，它以较低的帧率和较慢的刷新速度运行。另一个分支负责快速捕获运动的变化，它以较快的更新速度和高时间分辨率运行[7] [8]。快速流具有轻量级的特性，不需要花费时间池化，而慢速流的时间速率比较低，能更专注于空间的信息。在慢速流中包含了更多的空间信息，而在快速流中包含更多的时序信息。

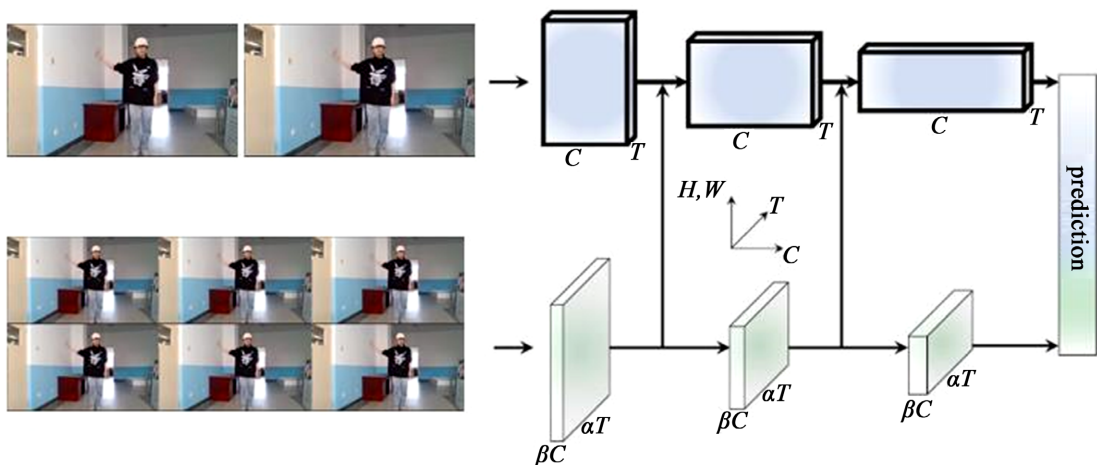


Figure 13. SlowFast two-flow model

图 13. SlowFast 双流模型

4. 实验分析

本项目提出的模型制作方法能够大幅度缩短制作周期，约为使用“关键帧法”制作时间的十分之一，制作的模型动画质量大部分取决于动作捕捉时的精确度。动作捕捉在单人、背景单一，人物四肢动作清晰缓慢的情况下捕捉效果较好，骨骼数据与形象模型绑定稳定。

利用动作识别，对使用者的动作进行捕捉，让虚拟形象同 VR 头盔使用者进行互动添加趣味性。本项目系统架构图如图 14 所示。

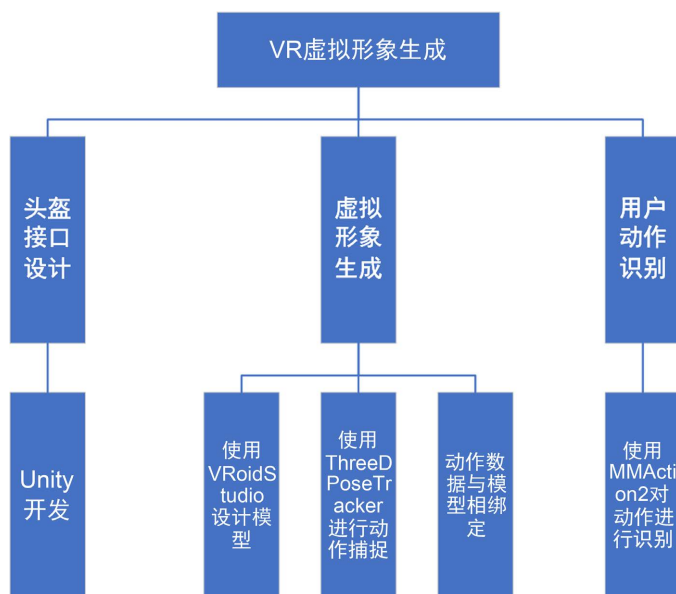


Figure 14. System structure diagram
图 14. 系统结构图

5. 结语

虚拟动画形象应用类型可以分为服务型 and 身份型。服务型的顾名思义，就是可以替代服务行业的某些职业，身份型的可以用于文娱领域，譬如虚拟偶像。本项目主要实现在 VR 头盔中设计模型形象，并实现与使用者进行互动，丰富 VR 设备的使用体验感。在模型形象的设计过程中，我们创新地采用了动作捕捉的方式为模型添加动作，在与用户的交互方面，我们除了能通过 VR 头盔的手柄同设计的虚拟形象交互外，我们还利用动作识别技术将 VR 头盔使用者的动作数据线传递给模型，让模型形象能同 VR 头盔使用者进行交互，这将丰富 VR 头盔的体验感，也验证了通过动作识别技术形成交互这一做法的可行性。

基金项目

由北京信息科技大学大学生创新创业训练计划项目——计算机学院(5112210832)支持。

参考文献

- [1] 张华忠. 基于 Proser 的虚拟人外形设计和动作控制[J]. 自动化应用, 2018(2): 65-66.
- [2] 刘宇洋, 周子怡. 浅析动作捕捉技术在影视动画作品中的应用[J]. 今古文创, 2022(13): 99-101.
- [3] 解放. 渐入佳境的 VR 虚拟现实[J]. 家庭影院技术, 2019(5): 86-91.
- [4] 杨芳圆, 蔡青青. 基于 Unity3D 平台的 VR 黄酒博物馆漫游交互系统开发[J]. 电视技术, 2022, 46(8): 26-28+33. <https://doi.org/10.16280/j.videoe.2022.08.006>

-
- [5] 杜汝涛. Unity3D 射击类游戏的开发[J]. 电脑编程技巧与维护, 2022(10): 141-144.
<https://doi.org/10.16184/j.cnki.comprg.2022.10.049>
- [6] 杜书升. 基于 android 的设备盘点软件系统的设计与实现[D]: [硕士学位论文]. 天津: 天津大学, 2016.
- [7] 智洪欣. 基于深度学习的视频分类技术研究[D]: [硕士学位论文]. 郑州: 战略支援部队信息工程大学, 2018.
- [8] Feichtenhofer, C., Fan, H.Q., Malik, J. and He, K.M. (2019) SlowFast Networks for Video Recognition. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*. Seoul, 27 October 2019-2 November 2019, 6201-6210.
<https://doi.org/10.1109/ICCV.2019.00630>