

# AI 虚拟导游在高职虚拟仿真实训基地的应用研究

## ——以广州番禺职业技术学院为例

黄国荣, 刘炜

(广州番禺职业技术学院, 广东广州 511483)

**摘要:** 以广州番禺职业技术学院为例, 研究 AI 虚拟导游在高职虚拟仿真实训基地的应用。文章首先介绍 AI 虚拟导游技术, 其次分析 AI 虚拟导游的模型开发过程, 分别从虚拟导游建模与骨骼绑定、ASR 语音识别与 TTS 语音合成、Audio2Face 基于面部表情训练和识别, 以及虚拟导游 UE 引擎实时驱动四个方面进行阐述, 最后对 AI 虚拟导游在高职虚拟仿真实训中的应用进行探讨和分析。

**关键词:** 人工智能; 虚拟导游; 虚拟仿真实训

中图分类号: TP18; G434

文献标识码: A

文章编号: 2096-4706 (2024) 02-0092-05

## Research on the Application of AI Virtual Tour Guide in Vocational Virtual Simulation Training Base

—A Case of Guangzhou Panyu Polytechnic

HUANG Guorong, LIU Wei

(Guangzhou Panyu Polytechnic, Guangzhou 511483, China)

**Abstract:** Taking Guangzhou Panyu Polytechnic as an example, this paper investigates the application of AI virtual tour guides in vocational virtual simulation training bases. It first introduces AI virtual tour guide technology, and then analyzes the model development process of AI virtual tour guide. It elaborates on four aspects: virtual tour guide modeling and bone binding, ASR speech recognition and TTS speech synthesis, Audio2Face facial expression-based training and recognition, and real-time driving of virtual tour guide UE engine. Finally, the application of AI virtual tour guide in vocational virtual simulation training is discussed and analyzed.

**Keywords:** Artificial Intelligence; virtual tour guide; virtual simulation training

## 0 引言

根据《教育信息化“十四五”发展规划》的指导, 虚拟仿真基地在学生实践能力和职业素养培养方面扮演着重要的角色。它为学生提供了真实的实践场景和模拟环境, 促进他们实际操作能力的提升和职业技能的培养。AI 虚拟导游结合人工智能技术, 具备导览、讲解和互动等功能<sup>[1-3]</sup>。广州番禺职业技术学院的虚拟仿真基地作为示范性项目, 在高职教育领域有广泛的知名度和影响力。然而, 随着参观人数的增加, 基地面临着接待任务繁重和人力不足的问题。本研究旨在深入探索 AI 虚拟导游在高职虚拟仿真基地中的潜力, 并提出创新的接待解决方案, 以提高效率和质量。通过研究问题和方法的设计, 我们致力于深入研究 AI 虚拟导游的融合应用。研究的目的是为高职虚拟仿真

基地提供创新的接待解决方案, 并为推动 AI 虚拟导游在教育领域的发展提供实证研究和有益的参考。

## 1 AI 虚拟导游技术简介

AI 虚拟导游综合应用人工智能、虚拟现实和人机交互等技术进行设计开发, 包括三个核心组件: 前端用户界面、后端算法模块和输出模块。AI 虚拟导游技术架构如图 1 所示。

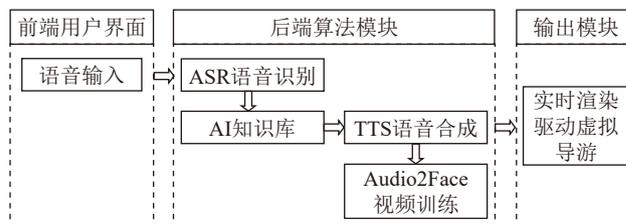


图 1 AI 虚拟导游技术架构

首先是前端用户界面, 它提供了用户与虚拟导游进行交互的界面。前端用户界面主要采用语音识别技术, 用户可以通过语音输入与虚拟导游进行对话和互

收稿日期: 2023-05-31

基金项目: 广东省教育科学规划课题 (2021GXJK617)

动。其次是后端算法模块, 它涵盖了自然语言处理、语音识别等关键技术, 能够实现对用户语义的理解、语音识别等功能。使用 ASR 语音识别将语音转换为文本数据, 然后调用人工智能知识库返回问题的答案。最后, 文本再被转换成语音 (TTS), 经过深度训练的虚拟导游由 UE4 引擎接收到语音数据实现实时驱动。目前人工智能智库识使用了科大讯飞的 SDK, 这个知识库包含了丰富的知识, 虚拟导游可以通过调用知识库中的数据来提供准确和有用的导览服务。通过数据库管理, 我们还可以不断增加知识库的内容, 使虚拟导游能够提前学习和了解广州番禺职业技术学院虚拟仿真基地的知识, 提供更加专业和全面的导览。

## 2 AI 虚拟导游模型开发

图 2 展示了虚拟导游制作的核心技术与流程。下面将分四个部分详细说明虚拟导游模型开发的相关技术。



图 2 虚拟导游模型开发流程

### 2.1 虚拟导游建模与骨骼绑定

为了制作逼真的虚拟导游, 我们采用了 3D 建模和渲染技术<sup>[4, 5]</sup>, 通过专业的建模软件 (Maya 或 Blender) 进行人物建模和细节塑造, 包括身体结构、面部特征和服装等。随后, 用 Maya 软件进行骨骼绑定, 骨骼绑定是将虚拟人物的 3D 模型与骨骼系统相连接的过程。骨骼系统是由一系列关节组成的层次结构, 模拟了真实人体的骨骼结构。在骨骼绑定过程中, 模型的顶点被分配到相应的骨骼上, 形成与骨骼之间的关联关系, 效果如图 3 所示。这样, 当骨骼运动时, 模型的顶点也会相应地进行变换和变形, 从而实现虚拟人物的动画效果。

### 2.2 ASR 语音识别与 TTS 语音合成

ASR (Automatic Speech Recognition) 语音识别和 TTS (Text-to-Speech) 语音合成技术在虚拟导游系统开发中的语音处理领域扮演着重要角色。它们是实现语音交互和语音应用的关键技术组成部分。ASR 语音识别技术将人类的语音输入转换为文本形式, 使计算

机能够理解和处理语音信息。该技术利用机器学习和深度学习算法, 如循环神经网络 (RNN) 和注意力机制, 对语音信号进行分析和建模。它通过将语音信号与训练数据中的语音模型进行匹配, 识别出语音中所包含的语音单位, 例如音素、单词或短语。TTS 语音合成技术则是将文本转化为自然流畅的语音输出。它通过模拟人类的语音产生过程, 将文字信息转换为可听的语音。TTS 技术也利用了机器学习和深度学习算法, 如循环神经网络和变分自编码器 (Variational Autoencoder), 对文本进行语音合成模型的训练。合成的语音可以具有多种音色和表达风格, 使得听者可以以自然的方式理解和接受信息。我们的 ASR 和 TTS 是调用科大讯飞的, 下面是基于 Python 调用科大讯飞的语音识别 (ASR) 以及调用科大讯飞知识库返回答案, 再进行语音合成 (TTS) 的技术流程, 如图 4 所示。

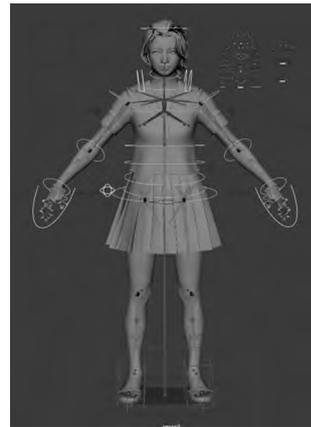


图 3 虚拟导游骨骼绑定截图

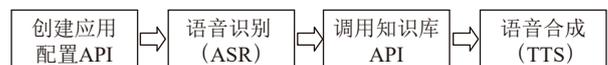


图 4 语音识别 (ASR) 和语音合成 (TTS) 技术流程

具体步骤如下:

1) 在科大讯飞开放平台注册一个账号, 在控制台中创建一个应用, 将会获得一个唯一的 AppID 和相应的 SecretKey。在 Python 代码中导入科大讯飞的相关模块, 配置 API 相关信息, 包括 APPID、密钥和 API 地址。

2) 语音识别 (ASR): generate\_signa 函数用于生成请求的签名, 然后定义 recognize\_speech 函数, 该函数发送 POST 请求给语音识别 API, 并解析响应结果为 JSON 格式。最后调用 recognize\_speech 函数进行语音识别, 并根据识别结果打印相应的信息。

3) 调用知识库 API: 构建请求参数, 包括用户的问题和 API 密钥。发送这些参数到知识库 API。通常, API 的响应是一个 JSON 格式的数据, 其中包含

与用户问题相关的答案。根据 API 的响应结果，可以解析返回的数据，并根据需要进行相应的处理。图 5 是针对番职院虚拟仿真实训基地创建的虚拟导游知识库管理平台，我们可以通过这个平台添加虚拟导游的问题与答案，管理知识库。



图 5 虚拟导游知识库管理平台截图

4) 语音合成 (TTS): 准备待合成的文本, 构建请求参数, 发送 POST 请求给语音合成 API, 并解析响应结果, 获取合成后的音频数据。

### 2.3 Audio2Face 基于面部表情训练和识别

此外, 虚拟导游还具备脸部表情识别和情感分析的能力。通过构建大规模的面部表情数据集并采用深度学习算法 (如卷积神经网络和循环神经网络), 我们训练了人脸识别和情感分类模型。这些模型能够识别用户的表情状态, 并相应地做出回应和互动, 提升用户体验和情感交互效果。

神经网络模型的训练通常涉及三个关键阶段: 数据采集制作、数据预处理和数据模型训练, Audio2Face 语音生成视频过程模型如图 6 所示。

在数据采集制作的第一阶段, 我们收集两类数据: 声音数据和相应的动画数据。声音数据主要涵盖录制的中文字母表的发音以及一些特殊的爆破音, 同时也包括尽可能多样的发音文本。动画数据则是

根据录制的声音数据, 在 Maya 软件中根据模型的面部特征创建对应的发音动画。如图 7 所示, 我们利用 FACEGOOD 的 P1 头盔通过 AVATARY 来进行数据采集。第二阶段是数据预处理阶段, 主要通过 LPC (线性预测编码) 对声音数据进行处理, 将其分割成与动画对应的帧数据, 并导出 Maya 动画帧数据。第三阶段是数据模型训练阶段, 其中将经过预处理的数据作为神经网络的输入, 进行训练直至损失函数收敛。通过这三个关键阶段的处理, 我们能够神经网络模型提供高质量的训练数据, 使其能够学习和模拟声音与动画之间的对应关系。这个过程是基于数据的驱动, 通过充分利用声音和动画数据之间的关联, 实现了声音到动画的准确映射。

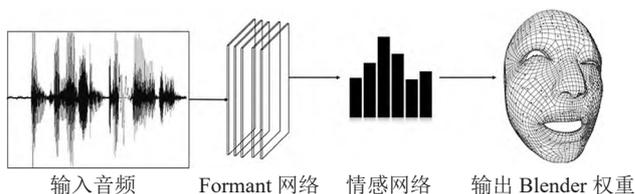


图 6 Audio2Face 语音生成视频过程模型



图 7 虚拟导游面部表情训练场景

### 2.4 虚拟导游 UE 引擎与实时驱动

如图 8 所示, 将虚拟导游模型放入 UE 引擎, 利用 UE 引擎调整效果, 包括模型贴图、骨骼网格和防碰撞体, 以及增加光影效果和真实感, 使虚拟导游在虚拟仿真环境中呈现出逼真的外观。先在 UE 引擎中安装 LiveLink 插件, 用于接收训练视频数据。然后运行项目, 让 UE 应用连接 AI 服务器, 接收 AI 模型产生的驱动数据, 继而驱动虚拟导游。



图 8 UE 引擎接实时驱动虚拟导游

通过以上的设计和开发过程, AI 虚拟导游系统实现了高度逼真的虚拟数字人的制作和表现, 并融合了语音合成、自然语言处理、语音识别和脸部表情识别等关键技术。该系统为用户提供沉浸式的导览服务和互动体验, 推动了虚拟导游与人工智能技术在教育领域的进一步发展。

### 3 AI 虚拟导游在高职虚拟仿真实训的应用

虚拟导游作为人工智能驱动的技术, 担任导览应用、实训课程的引导与辅助以及评估与优化的重要角色<sup>[6-11]</sup>, 具体介绍如下。

#### 3.1 在高职虚拟仿真基地的导览应用

如图 9 所示, 虚拟导游在高职虚拟仿真基地的导览应用中, 扮演着向参观者提供详细解说的角色。通过虚拟导游, 参观者可以了解虚拟仿真实训基地的建设情况, 并获得对各种 VR 设备的详细介绍。虚拟导游通过语音合成技术, 以自然流畅的语音与访客进行对话和交流。向参观者介绍不同的实训教室以及学校各专业的 VR 课程资源。通过虚拟导游的引导, 参观者可以深入了解虚拟仿真基地的设施和资源, 增强他们的参观体验。



图 9 虚拟导游在虚拟仿真基地的应用

#### 3.2 在实训课程中的引导与辅助

虚拟导游在高职虚拟仿真基地的实训课程中发挥着重要的引导和辅助作用。他们通过与学生的互动, 提供实训课程的引导和指导。虚拟导游可以解释实训操作的步骤和注意事项, 并回答学生提出的问题。通过与虚拟导游的交流, 学生可以更好地理解和掌握实训内容, 提高实践能力和技能水平。虚拟导游还能根据学生的学习进度和需求, 提供个性化的实训辅助, 监测学生的操作并及时纠正错误, 提供指导和建议, 帮助学生顺利完成实训任务。

#### 3.3 在高职虚拟仿真基地的评估与优化

为了不断提升虚拟导游的性能和用户体验, 对其进行评估与优化是必要的。评估工作可以通过参观者的反馈和评价, 以及虚拟导游的性能指标来进行。参观者的反馈可以通过问卷调查、用户访谈等方式收集, 以了解他们对虚拟导游的满意度和体验感受。同时, 虚拟导游的性能指标可以通过技术指标和系统性能的

评估来衡量, 如语音识别准确率、对话交互的流畅性等。基于评估结果, 可以对虚拟导游进行优化。优化工作可以包括改进语音识别和自然语言处理技术, 提升虚拟导游的对话能力和交互体验。此外, 还可以增加虚拟导游的知识库和学习能力, 使其能够更全面地解答参观者的问题。通过评估与优化, 可以不断改进虚拟导游的性能和功能, 提供更好的用户体验和参观效果。

## 4 结论

本文探讨了 AI 虚拟导游的开发技术, 以及在高职虚拟仿真基地中的应用实践。AI 虚拟导游在导览应用方面能够提供个性化的导览服务, 通过语音合成和自然语言处理实现与参观者的流畅交流, 提升交互体验。在实训课程中, 虚拟导游能够引导和辅助学生, 提供实训指导和问题解答, 帮助提高实践能力和技能水平。虚拟导游在评估与优化方面发挥重要作用, 通过参观者的反馈和情感分析改进自身性能, 提供个性化和情感化的互动体验。虽然 AI 虚拟导游在应用实践中展现了优势和潜力, 但仍存在以下局限性: 语音识别和自然语言处理技术的准确性和稳定性需要进一步提高, 虚拟导游系统的开发和运营成本较高, 需要寻求更有效的开发和运营模式。

随着持续的研究和创新, AI 虚拟导游将朝着以下方向迭代发展: 外貌的精致度将进一步提高, 形象更加逼真, 动作更加自然。AI 虚拟导游将推动高职教育的发展, 为人们带来更精彩和有价值的体验。它将不仅仅是一项技术的突破, 更是人工智能与教育、文化、娱乐等领域融合的杰出典范, 为我们构建一个更加智能、互动和丰富多彩的未来。

### 参考文献:

- [1] 腾讯研究院, 腾讯云智能, 创业黑马. 好看的皮囊到有趣的灵魂——数字人产业发展趋势报告 [R/OL].[2023-04-30]. <https://www.tisi.org/24879>.
- [2] 清华大学新闻与传播学院, 元宇宙文化实验室. 虚拟数字人研究报告 2.0 版 [J/OL].[2023-04-30]. <http://www.100ec.cn/index.php/detail--6619418.html>.
- [3] 陈芳. AI 虚拟数字人技术在融合媒体生产中的运用 [J]. 数字通信世界, 2021 (1): 21-23.
- [4] REN Y, HU C X, QIN T, et al. FastSpeech 2: Fast and High-Quality End-to-End Text-to-Speech [J/OL]. arXiv: 2006.04558 [eess.AS]. (2022-08-08) .[2023-04-26]. <https://arxiv.org/abs/2006.04558v1>.
- [5] KARRAS T, AILA T, LAINE S, et al. Audio-driven facial animation by joint end-to-end learning of pose and emotion [J/OL]. ACM Transactions on Graphics, 36 (4): 1-12[2023-04-26]. <https://doi.org/10.1145/3072959.3073658>.
- [6] 郭全中. 虚拟数字人发展的现状、关键与未来 [J]. 新闻与写作, 2022 (7): 56-64.



图 9 集群中安装成功的 MariaDB 服务

角色类型	状态	主机
MARIADB_SERVER NODE	已启动	cdh2
MARIADB_SERVER NODE	已启动	cdh3

图 10 CM UI 显示的 MariaDB 在各节点的分布情况

### 3.2.3 MariaDB 服务连接测试

MariaDB 服务部署到 cdh2 和 cdh3 个节点, 如图 11 所示, 通过 DBeaver<sup>[11]</sup> 的 MariaDB 服务, 显示成功连接, 说明 CM 成功实现 MariaDB 的部署, 验证 CM 能够部署和管控 MariaDB 关系数据库的设想。



图 11 MariaDB 连接 cdh2 和 cdh3 服务成功示意图

### 3.2.4 停用测试

执行停止服务命令后, 如图 12 所示, cdh2 和 cdh3 主机状态处于停止状态, 与图 10 启动状态呈现鲜明对比, MariaDB 服务运行端口号和进程均已退出, 这表明通过 CM 成功实现停止 MariaDB 服务, 测试通过。

角色类型	状态	主机
MARIADB_SERVER NODE	已停止	cdh2
MARIADB_SERVER NODE	已停止	cdh3

图 12 停用 MariaDB 服务成功示意图

通过 CMUI 进行信息检查、软件部署和功能操作以及使用 DBeaver 数据库客户端连接测试, 并且均以无异常显示, 证实了利用 CM 能够扩展 MariaDB 关系数据库。

## 4 结论

本文利用 CM 扩展方法研发了 MariaDB 的 Parcel 应用包和 CSD 管理包, 并提供了这两个扩展包的使用方法, 进而实现 CM 对关系数据库 MariaDB 的部署和管控。通过实验安装和服务测试, 验证了此方法有效性, 为 CM 用户使用关系数据库功能开拓渠道, 同时为 CM 扩展更多应用软件提供了方法。

### 参考文献:

- [1] CLOUDERA.CDH Overview [DB/OL]. (2021-09-29). [https://docs.cloudera.com/documentation/enterprise/6/6.3/topics/cdh\\_intro.html](https://docs.cloudera.com/documentation/enterprise/6/6.3/topics/cdh_intro.html).
- [2] 张延松, 王珊, 周烜. 内存数据仓库集群技术研究 [J]. 华东师范大学学报: 自然科学版, 2014 (5): 117-132.
- [3] 田燕军, 王玥. hbase 与 hive 整合研究 [J]. 山西电子技术, 2022 (1): 81-83.
- [4] 李翀, 张彤彤, 杜伟静, 等. 基于 Hive 的高可用双引擎数据仓库 [J]. 计算机系统应用, 2019, 28 (9): 65-71.
- [5] POL U R. Big data and hadoop technology solutions with cloudera manager [J]. International Journal, 2014, 4 (11).
- [6] CHAI R, GAO S, LAN J, et al. Efficient Computing Resource Metric Method in Computing-First Network [J]. Journal of Computer Research and Development, 2023, 60 (4): 763-771.
- [7] TURLAND M, SLADE P. Farmers' willingness to participate in a big data platform [J]. Agribusiness, 2019, 36 (1): 20-36.
- [8] MARIADB. MariaDB Server [DB/OL]. (2023-01-24). <https://mariadb.com/docs/server>.
- [9] GITHUB. Cloudera Manager Extensions [DB/OL]. (2020-12-12). [https://github.com/cloudera/cm\\_ext/wiki](https://github.com/cloudera/cm_ext/wiki).
- [10] GITHUB.CM 的 MariaDB 扩展 [DB/OL]. (2023-04-20). <https://github.com/huangjie155/Store>.
- [11] DBeaver. Universal Database Tool [DB/OL]. (2023-05-07). <https://dbeaver.io>.

**作者简介:** 黄杰 (1992—), 男, 汉族, 重庆人, 硕士研究生, 讲师, 研究方向: 大数据部署监控、模式匹配。

(上接 95 页)

[7] 朱永琼, 宋章通, 方浩. “文旅元宇宙”中虚拟数字人的应用 [J]. 传媒, 2023 (3): 55-57.

[8] 徐琦. 主流视听媒体虚拟数字人应用创新与优化策略 [J]. 中国电视, 2023 (1): 102-107.

[9] 覃祖军, 杨静. 元宇宙中国教育范式研究视角下虚拟数字人辅助的双师课堂教学范式实践探索 [J]. 中国现代教育装备, 2023 (2): 1-5.

[10] 李晓明, 张宾, 魏晓鹏, 等. 基于多模态智能交互的虚拟数字人 [J]. 电力大数据, 2022, 25 (12): 36-43.

[11] 夏钦. 虚拟数字人: 好看的皮囊和有趣的灵魂缺一不可 [J]. 工会博览, 2022 (29): 38.

**作者简介:** 黄国荣 (1975.10—), 男, 汉族, 广东广州人, 工程师, 本科, 主要研究方向: 信息技术、虚拟仿真; 刘炜 (1981.08—), 男, 汉族, 甘肃平凉人, 副研究员, 本科, 主要研究方向: 信息技术、虚拟仿真。