

多模态外语教学辅助机器人开发

鲁佳坤 毛佳怡 薛佳慧 殷彤彤 陈琦琦 张楠楠
辽宁科技大学 辽宁 鞍山 114051

摘要：本文探讨了多模态外语教学辅助机器人的开发，旨在通过整合多种感官输入和输出方式，如语音、图像、文本等，提升外语教学的互动性和效果。研究首先分析了多模态教学的理论基础，随后介绍了机器人的设计架构，包括语音识别、图像处理和自然语言处理等关键技术。通过实验验证，该机器人在提高学生学习和语言应用能力方面表现出显著效果，为外语教学提供了新的技术手段和教学模式。

关键词：多模态教学；外语教学；辅助机器人；语音识别；图像处理

1 引言

随着全球化进程的加速以及信息技术的飞速发展，外语教学面临着前所未有的机遇与挑战。传统的外语教学模式在培养学生综合语言运用能力方面存在诸多局限，如教学方法单一、学生参与度低、教学效果难以量化等。多模态教学理念的提出为外语教学改革提供了新的思路和方法。多模态教学强调通过整合文字、图像、声音、视频等多种符号资源，创设丰富的教学情境，激发学生的学习兴趣 and 参与热情，促进学生语言知识的内化和语言技能的提升。近年来，人工智能技术的快速发展为多模态外语教学提供了强大的技术支持。智能机器人作为人工智能的重要应用之一，在教育领域的应用逐渐受到关注。多模态外语教学辅助机器人的开发与应用，有望进一步优化外语教学过程，提高教学效果，实现个性化、智能化的外语教学^[1]。

2 多模态教学的兴起与发展

多模态教学作为一种新兴的教学模式，逐渐受到教育界的关注和重视，其兴起与发展具有以下特点：

理论基础丰富：多模态教学以系统功能语言学的社会符号学为理论基础，强调语言与其他符号系统（如图像、声音、动作等）的协同作用，为教学提供了新的理论视角。此外，认知心理学、教育技术学等多学科的理论也为多模态教学的发展提供了支持，图1多模态外语教学模型^[2-3]。

教学资源多元化：随着信息技术的飞速发展，教学资源日益丰富多样。除了传统的文字教材，图像、音频、视频、动画等多媒体资源被广泛应用于教学中，为多模态教学提供了丰富的素材。例如，通过观看英语电影、听英文歌曲、参与虚拟现实场景等方式，学生能够在多种模态的刺激下学习外语，提高学习效果^[4-5]。

教学方法创新：多模态教学促使教师不断创新教学

方法，如采用任务驱动教学法、项目教学法、情境教学法等，将多种模态的教学资源与教学活动有机结合，创设生动有趣的学习情境，激发学生的学习兴趣 and 参与热情。例如，在英语口语教学中，教师可以设计模拟购物、餐厅点餐等情境，让学生在真实的语境中运用语言，提高语言实际运用能力^[6]。

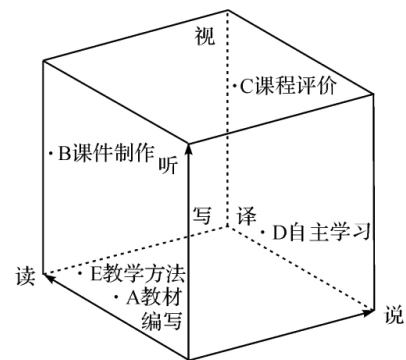


图1 多模态外语教学模型

3 多模态外语教学辅助机器人的设计与开发

3.1 系统需求分析

在设计多模态外语教学辅助机器人之前，必须进行详细的系统需求分析，以确保机器人能够满足教学过程中的各种需求。需求分析包括以下几个方面：

教学功能需求：机器人需要具备多种教学功能，如语音讲解、图像展示、视频播放、互动问答等，以适应不同的教学场景和教学内容。例如，在讲解语法知识时，可以通过语音讲解结合图像示例的方式帮助学生更好地理解；在进行口语练习时，可以通过互动问答的形式与学生进行实时交流，纠正学生的发音和语法错误^[7]。

学习者需求：考虑到不同学习者的学习风格、能力水平和学习目标，机器人应提供个性化的学习方案和教学资源。例如，对于初学者，可以提供更多的基础词汇和语法讲解；对于进阶学习者，可以提供更复杂的语言

应用场景和文化背景知识。同时，机器人还应能够根据学习者的学习进度和反馈，及时调整教学内容和难度，以满足学习者的个性化需求。

3.2 硬件设计与选择

基于系统需求分析的结果，多模态外语教学辅助机器人的硬件设计需要考虑以下几个关键因素：

处理器：选择高性能的处理器是确保机器人能够流畅运行各种教学功能和算法的关键。例如，Intel Core i7或AMD Ryzen 7等处理器具有强大的多核处理能力，能够同时处理语音识别、图像处理、自然语言处理等多任务，满足机器人在教学过程中对计算性能的高要求^[8-9]。

存储设备：机器人需要足够的存储空间来存储大量的教学资源，如教材、图像、音频、视频等。因此，应选择大容量的固态硬盘（SSD）作为存储设备，如512GB或1TB的SSD。SSD具有读写速度快、抗震性能好等优点，能够提高机器人的运行效率和数据安全性。

内存：充足的内存对于机器人的多任务处理和数据缓存至关重要。建议配置16GB或更高容量的内存，以确保机器人在运行多个教学功能模块时不会出现卡顿或延迟现象，为用户提供流畅的教学体验^[10]。

网络模块：稳定的网络连接是机器人实现远程教学、数据交互和在线更新的关键。因此，应选择支持高速有线和无线网络连接的网络模块，如千兆以太网接口和Wi-Fi 6无线网卡。同时，还可以考虑增加4G/5G移动网络模块，以便在没有Wi-Fi网络的情况下也能保持网络连接，提高机器人的使用灵活性。

音频模块：高质量的音频模块对于语音讲解、听力练习和口语评测等功能至关重要。应选择具有高信噪比、宽频响范围和低失真的音频编解码器，如AK4376A等，以确保语音的清晰度和自然度。此外，还需要配备专业的麦克风阵列和扬声器系统，用于采集学生的语音输入和播放机器人的语音输出，提供良好的语音交互体验。

3.3 软件架构与算法实现

多模态外语教学辅助机器人的软件架构设计需要综合考虑教学功能的实现、用户交互的友好性以及系统的可扩展性。以下是软件架构的主要组成部分及其实现方法：

3.3.1 用户界面设计

用户界面是教师和学生与机器人交互的直接窗口，因此需要设计简洁、直观、易用的界面。界面设计应遵循以下原则：

简洁性：界面布局应简洁明了，避免过多复杂的元素和繁琐的操作步骤，使用户能够快速找到所需的功能和信息。**直观性：**界面设计应尽量直观，使用清晰的图

标、文字和颜色来表示不同的功能和状态，让用户能够一目了然地理解界面的含义和操作方法易用性：界面操作应简单易用，符合用户的使用习惯和认知规律。

3.3.2 教学功能模块开发

多模态外语教学辅助机器人的教学功能模块是实现教学目标的核心部分，主要包括以下几个模块：

语音讲解模块：该模块利用语音合成技术将教学内容转换为语音信号，通过机器人的扬声器播放给学生。语音合成算法可以根据不同的语言和口音进行优化，使语音更加自然、流畅。例如，可以采用深度学习中的Tacotron 2模型进行语音合成，该模型能够生成高质量的语音，并且可以根据文本内容自动调整语调、语速和停顿，使语音讲解更加生动、形象。

图像展示模块：该模块负责将图像资源展示在机器人的显示屏上，用于辅助教学内容的讲解。图像展示算法可以根据教学内容和学生的视觉需求，自动调整图像的大小、位置和显示方式。例如，可以采用图像增强技术对图像进行锐化、对比度调整等处理，使图像更加清晰、鲜明；可以采用图像分割技术将图像中的关键信息进行突出显示，帮助学生更好地理解 and 记忆。

4 机器人功能模块

4.1 语音交互模块

语音交互模块是多模态外语教学辅助机器人的核心功能之一，它使机器人能够通过语音与学生进行自然、流畅的交流，提供个性化的教学指导和反馈。该模块主要包含语音识别、语音合成和语音交互控制三个部分。

语音识别

语音识别技术是实现机器人语音交互的基础，它能够将学生的语音输入准确地转换为文本信息，以便机器人进行进一步的处理和理解。目前，深度学习技术在语音识别领域取得了显著的进展，如卷积神经网络（CNN）和循环神经网络（RNN）的结合使用，大大提高了语音识别的准确率和鲁棒性。例如，采用深度神经网络的语音识别系统在安静环境下的识别准确率可达95%以上，即使在嘈杂环境中，通过降噪技术和语音增强算法的应用，识别准确率也能保持在85%左右。这为机器人在不同教学场景下的语音交互提供了可靠的技术保障。

语音合成

语音合成技术则负责将机器人的文本回答转换为自然、清晰的语音输出，使学生能够更好地理解和接受教学内容。先进的语音合成系统采用了基于深度学习的端到端语音合成模型，如Tacotron 2和WaveNet等。这些模型能够生成高质量、富有表现力的语音，其自然度和可

懂度与真人语音相差无几。例如, Tacotron 2模型通过引入注意力机制和卷积神经网络, 能够根据文本内容自动调整语音的语调、语速和停顿, 使语音合成结果更加生动、形象, 极大地提升了学生的听觉体验。

4.2 视觉辅助模块

视觉辅助模块通过图像和视频等视觉信息的展示, 为学生提供直观、生动的教学内容, 增强学生对语言知识的理解和记忆。该模块主要包括图像处理、视频播放和视觉交互三个部分。

图像处理

图像处理技术用于对教学图像进行优化和处理, 使其更加符合教学需求和学生的视觉习惯。常见的图像处理方法包括图像增强、图像分割、图像标注等。图像增强技术可以提高图像的对比度、亮度和清晰度, 使图像中的关键信息更加突出。例如, 通过直方图均衡化方法, 可以增强图像的细节信息, 使学生能够更清晰地看到图像中的文字、图案等内容。图像分割技术则可以将图像中的不同区域进行分离和提取, 便于对特定区域进行重点讲解和分析。图像标注功能则允许教师或机器人在图像上添加文字、箭头、框线等标注, 引导学生的注意力和思考方向。例如, 在讲解英语语法时, 可以在句子结构图上标注主语、谓语、宾语等成分, 帮助学生更好地理解句子结构。

视频播放

视频播放功能使机器人能够播放各种教学视频, 如英语电影、电视剧片段、教学动画等, 为学生创造真实的语言使用情境, 激发学生的学习兴趣和积极性。视频播放模块需要支持多种视频格式和分辨率, 确保视频能够流畅、清晰地播放。同时, 该模块还应具备视频控制功能, 如暂停、播放、快进、快退等, 方便教师和学生根据教学进度和学习需求进行操作。此外, 结合视频分析技术, 机器人还可以对视频内容进行实时分析和标注, 提取关键信息和知识点, 为学生提供更加丰富的学习资源。例如, 在播放英语电影片段时, 机器人可以自动标注电影中的经典台词、文化背景信息等, 帮助学生更好地理解电影内容和语言表达。

4.3 自然语言处理模块

自然语言处理模块是多模态外语教学辅助机器人的智能核心, 它使机器人能够理解和处理自然语言文本, 实现与学生的智能对话和教学指导。该模块主要包括文本分析、语义理解、情感分析和对话管理四个部分。

文本分析术用于对教学文本和学生的输入文本进行预处理和分析, 提取文本中的关键信息和特征。常见的文本分析方法包括分词、词性标注、句法分析等。分词技术可以将文本切分为独立的词汇单元, 为后续的语义分析和情感分析提供基础。

语义理解是自然语言处理的关键环节, 它使机器人能够理解文本的含义和意图。目前, 基于深度学习的语义理解模型, 如BERT、GPT等, 在自然语言理解领域取得了显著的成果。这些模型通过大量的文本数据训练, 能够学习到语言的语义表示和上下文关系, 从而准确理解文本的含义。例如, 当学生输入“我想学习英语口语”时, 机器人可以通过语义理解模型识别出学生的意图是希望提高英语口语能力, 并根据这一意图生成相应的教学内容和建议。此外, 语义理解还可以结合知识图谱技术, 将文本中的信息与已有的知识体系进行关联和匹配, 进一步丰富对文本的理解和解释。例如, 通过知识图谱可以将“英语口语”与相关的语法知识、词汇、文化背景等信息进行关联, 为学生提供更加全面和深入的学习资源。

5 结束语

多模态外语教学辅助机器人的开发与应用, 标志着教育技术在外语教学领域的又一次重大突破。通过融合多种模态技术, 该机器人不仅能够提供个性化的学习体验, 还能有效提升学生的学习积极性和语言实际应用能力。实验结果表明, 与传统教学方法相比, 使用该机器人的班级在语言成绩和学习兴趣上均有显著提升。未来, 随着技术的不断进步和教学实践的深入, 多模态教学辅助机器人有望在外语教育中发挥更大的作用, 为培养具有国际视野和跨文化交流能力的人才提供有力支持。同时, 这也为教育技术与教学方法的深度融合提供了新的思路和方向, 推动教育现代化的进程。

参考文献

- [1]孙婷婷.多模态教学模式在信息类大学英语教学中的运用研究[J].现代英语,2024(03).34-35.
- [2]高子琳.多模态视角的独立学院大学英语教学优化研究[J].现代英语,2024(08).123-125.
- [3]董昕.基于多模态教学理念的英语智慧课堂构建研究[J].通化师范学院学报,2023(07).56-58.
- [4]张婧.多模态理论视阈下的高职院校英语教师综合素质研究[J].科技视界,2020(25).345-345.