

基于大语言模型的移动应用可访问性增强方法

唐洪斌

中译语通科技股份有限公司 北京 100000

摘要：移动应用可访问性增强需关注用户多样性，包括不同障碍类型用户的需求。功能需求上，应提供适应各类用户的特性，如字体调整、语音控制等。界面设计需简洁直观，支持高对比度、大字体，并兼容辅助技术。基于大语言模型，可增强移动应用可访问性：通过智能语音交互实现无手操作；图像与界面元素解读帮助视觉障碍用户理解内容；交互流程辅助简化操作步骤；无障碍测试与反馈确保应用对所有用户友好。这些方法共同提升移动应用的可访问性和用户体验。

关键词：大语言模型；移动应用；可访问性增强

引言：在数字化时代，移动应用已成为人们生活中不可或缺的一部分。然而，对于存在视觉、听觉或运动障碍的用户来说，许多应用仍存在使用障碍。大语言模型的出现，为移动应用可访问性的增强带来了新机遇。通过利用大语言模型的强大能力，我们可以实现更智能的语音交互、更准确的图像解读、更流畅的交互流程以及更全面的无障碍测试。这将有助于打造更加包容、易用的移动应用，让科技的力量惠及每一位用户，共创无障碍的数字未来。

1 大语言模型在移动应用可访问性增强中的作用

大语言模型在移动应用可访问性增强中扮演着至关重要的角色，随着人工智能技术的飞速发展，大语言模型以其强大的自然语言处理能力和逻辑推理能力，为移动应用的可访问性提升带来了革命性的变化。在移动应用可访问性领域，大语言模型能够深入理解用户的语言输入，无论是文本还是语音，都能进行准确的理解和响应。这对于视觉障碍、听觉障碍等特殊群体来说尤为重要，因为他们可能无法直接通过视觉或听觉与移动应用进行交互。大语言模型通过文字转语音、语音转文字等功能，使得这些用户能够轻松获取应用中的信息，并进行有效的交互^[1]。此外，大语言模型还能根据用户的历史行为和偏好，提供个性化的服务。例如，通过分析用户的搜索记录、浏览历史等数据，大语言模型可以预测用户的需求，并主动推荐相关的内容或服务。这种个性化的服务不仅提高了用户的使用体验，还增强了移动应用的可访问性。在移动应用的可访问性增强过程中，大语言模型还可以与其他辅助技术相结合，形成更加完善的无障碍解决方案。例如，与眼动追踪技术、手势控制技术相结合，可以为用户提供更加便捷、自然的交互方式。

2 移动应用可访问性增强需求分析

2.1 用户多样性识别

2.1.1 残障用户群体的需求识别

残障用户群体在移动应用使用中面临着诸多挑战。对于视觉障碍用户，他们无法像正常人一样通过视觉获取界面信息，因此移动应用需要提供文字转语音功能，将界面文字转化为语音输出，同时采用高对比度色彩和可调节的字体大小，确保信息清晰可见。听觉障碍用户则依赖字幕和振动反馈来接收信息，应用应确保所有音频内容都配有准确字幕，并提供振动提醒功能。此外，对于肢体障碍用户，移动应用应支持语音控制、外部辅助设备接入等特性，简化操作流程，降低使用难度，让残障用户也能轻松享受移动应用的便利。

2.1.2 老年用户群体的需求识别

老年用户作为移动应用的重要群体，其使用需求尤为特殊。由于视力、听力等生理机能的逐渐下降，他们在使用移动应用时往往遭遇诸多不便。因此，移动应用的设计应充分考虑老年用户的实际需求，提供大字体、高对比度的界面设计，确保信息清晰可读；同时，界面应简洁明了，避免信息过载。在操作流程上，应用应给予明确的操作提示，简化操作步骤，剔除冗余功能，让老年用户能够轻松上手。此外，针对老年用户记忆力减退的特点，移动应用还应提供个性化设置和记忆辅助功能，如设置常用功能快捷入口、保存历史记录等，以更加贴心的设计提升老年用户的使用体验，让他们也能享受数字生活的便利。

2.1.3 其他特殊用户群体的需求识别

除了残障用户和老年用户，移动应用还需关注其他特殊用户群体的可访问性需求。在特定工作环境中，如嘈杂的工厂或繁忙的医院，用户可能难以听清应用的

声音提示。为此,移动应用应提供振动或视觉提醒功能,确保用户能及时接收到重要信息。对于色盲或色弱用户,应用设计时应采用高对比度或易于区分的颜色搭配,甚至使用图案标识来辅助信息传达,确保他们能够准确理解界面内容。随着全球化的不断深入,移动应用应支持多语言界面,并提供本地化内容,以满足来自不同语言和文化背景用户的需求,让科技真正惠及每一个人,构建更加包容的移动应用环境。

2.2 功能需求

在移动应用可访问性增强的功能需求方面,我们主要关注如何使用应用更加适应不同用户群体的特殊需求,提升用户体验。第一,对于视觉障碍用户,移动应用应提供文字转语音功能,确保用户能够听到界面上的文字内容。应用界面应采用高对比度设计,使文字和背景色彩形成鲜明对比,便于用户识别。还应支持字体大小调节功能,让用户根据自己的视力情况调整字体大小,确保阅读舒适。第二,对于听觉障碍用户,应用应提供全面的字幕服务,确保所有音频内容都配有准确、清晰的字幕。同时,为了增强用户的交互体验,应用还应提供振动反馈功能,让用户在无法听到声音提示的情况下,也能通过振动感知到操作结果或系统通知。第三,针对肢体障碍用户和其他特殊需求用户,移动应用应支持多种交互方式,如语音控制、手势识别等,以降低操作难度^[2]。同时,应用还应提供个性化设置选项,让用户根据自己的使用习惯和需求,调整应用的功能和界面布局。

2.3 界面设计需求

在移动应用可访问性增强的界面设计中,需关注以下五个关键方面:(1)直观性与简洁性:界面设计应直观易懂,避免复杂繁琐的布局。元素排列应有序,功能区域划分明确,使用户能够快速找到所需信息。简洁的设计不仅能提高普通用户的使用效率,对认知障碍用户也尤为重要。(2)可调整性:界面应支持字体大小、颜色对比度等个性化调整。视觉障碍用户可以根据自身需要,调整界面元素以适应其视力状况,确保信息清晰可读。(3)色彩与对比度:色彩选择应考虑色盲和色弱用户的需求。使用高对比度色彩搭配,确保重要信息和按钮在视觉上突出,易于识别。(4)交互多样性:提供多种交互方式,如触摸、语音、手势等,以满足不同用户的需求。对于肢体障碍用户,语音控制和手势识别可以提供更加便捷的操作途径。(5)辅助技术兼容性:界面设计应与屏幕阅读器、放大镜等辅助技术兼容,确保所有用户都能无障碍地使用应用。这要求开发者在设计阶段就考虑辅助技术的接入和使用,确保应用的可访问性。

3 基于大语言模型的移动应用可访问性增强方法分析

3.1 智能语音交互

3.1.1 语音识别技术的集成

智能语音交互的精髓,在于语音识别技术的深度集成。大语言模型作为这一技术的核心支撑,能够精准捕捉并识别用户的语音输入,迅速将其转化为文本信息,从而架起用户与移动应用之间的沟通桥梁。对于视觉障碍者或肢体不便的用户来说,这项技术无疑是一大福音。它彻底颠覆了传统的操作方式,用户无需再依赖触摸屏幕或费力寻找按键,只需简单发出语音指令,就能轻松完成各种操作。而且,大语言模型具备强大的持续学习和优化能力,不断提升语音识别准确率,使得交互过程更加流畅自然。这一技术的不断进步,正为特殊用户群体带来前所未有的便捷体验。

3.1.2 语音合成的个性化定制

智能语音交互中,语音合成技术同样扮演着举足轻重的角色。大语言模型凭借其强大的处理能力,能够根据用户的独特需求和偏好,生成富有个性化特色的语音输出。对于听觉障碍用户而言,清晰、准确的语音合成无疑是沟通的桥梁,帮助他们更好地理解 and 接收应用反馈的各种信息。更值得一提的是,语音合成的语速、音调等关键参数,也可以根据用户的实际听力状况进行灵活定制。这样的设计,不仅确保了信息传达的准确性,还极大提升了用户的使用舒适性,让智能语音交互技术更加贴心、更加人性化,真正服务于每一个需要的用户。

3.1.3 语音交互场景的拓展

随着大语言模型的不断精进与完善,语音交互技术在移动应用领域的应用愈发广泛而深入。从日常的搜索查询、路线导航,到购物支付、社交聊天、娱乐休闲等多元化场景,语音交互都以其独特的便捷性和高效性脱颖而出。用户只需轻轻一声语音指令,便能轻松实现商品购买、消息即时发送、音乐随心播放等多种操作,真正享受到了无手操作的极致便捷。这种创新的交互模式,不仅极大地提升了用户的使用体验,更使得移动应用的可访问性和包容性得到了前所未有的增强。它为更多用户打开了数字世界的大门,带来了前所未有的便捷与乐趣。

3.2 图像与界面元素解读

在移动应用中,图像与界面元素的解读对于提升可访问性至关重要,大语言模型在此方面发挥着重要作用。(1)图像识别与描述:大语言模型能够识别应用中的图像内容,并生成准确的描述性文本。对于视觉障碍用户,这一功能使他们能够通过听觉或触觉设备了解图

像所展示的信息,从而更全面地理解应用界面。(2)界面元素标签化:为了增强可访问性,大语言模型可以为界面元素添加语义标签。这些标签不仅帮助屏幕阅读器准确读取元素信息,还使用户能够通过语音指令快速定位和操作界面元素,提高使用效率。(3)自动化测试与优化:大语言模型可以参与自动化测试过程,检测界面元素的可访问性问题。通过模拟不同用户的交互场景,模型能够识别出潜在的障碍点,并提供优化建议,确保应用对所有用户都友好。(4)个性化辅助设置:基于大语言模型,应用可以提供个性化的辅助设置。用户可以根据自己的需求调整图像描述的详细程度、界面元素的朗读方式等,获得更加定制化的使用体验。(5)持续学习与更新:大语言模型具备持续学习能力,能够随着应用界面的更新和用户需求的变化,不断优化图像与界面元素的解读能力,保持可访问性的持续提升。

3.3 交互流程辅助

在移动应用中,交互流程的顺畅与否直接影响用户的使用体验,特别是对于存在认知或操作障碍的用户来说,一个复杂或混乱的交互流程可能成为他们使用应用的巨大障碍。大语言模型可以通过分析用户的行为数据和反馈,识别出交互流程中的瓶颈和难点。例如,模型可以检测用户在完成某项任务时所需的步骤数、操作时间以及错误率,从而评估流程的复杂度和用户友好性。基于这些分析结果,大语言模型可以提出改进建议,如简化步骤、优化界面布局、提供明确的操作引导等^[3]。这些建议可以帮助开发者设计出更加直观、易懂的交互流程,降低用户的学习成本和操作难度。在用户进行操作时,模型可以根据用户的输入和上下文信息,提供智能化的提示和建议,引导用户顺利完成任务。这种实时的辅助不仅可以提高用户的操作效率,还可以增强用户的自信心和满意度。

3.4 无障碍测试与反馈

在移动应用开发中,无障碍测试是确保应用对所有用户,包括那些有视觉、听觉或运动障碍的用户,都能友好使用的重要环节。大语言模型可以模拟不同障碍用户的交互行为,对移动应用进行全面的无障碍测试。通过模拟视觉障碍用户的屏幕阅读器使用、听觉障碍用户的字幕需求,以及运动障碍用户的语音控制等操作,模型能够揭示出应用中可能存在的无障碍问题。在测试过程中,大语言模型不仅能够识别出具体的障碍点,如颜色对比度不足、文字大小不可调、缺乏语音导航等,还能提供详细的测试报告和反馈。这些反馈包括问题的具体位置、影响程度以及改进建议,为开发者提供了明确的优化方向。基于大语言模型的测试反馈,开发者可以针对性地进行应用改进,提升应用的可访问性。同时,模型还可以持续监控应用的无障碍性能,确保在后续的更新和迭代中,无障碍特性得到持续保持和优化。

结语

未来,我们将不遗余力地继续探索大语言模型在移动应用可访问性领域的无限可能。随着技术的不断进步,我们将深入挖掘大语言模型的潜力,将其更广泛地应用于移动应用的各个层面,从语音交互到图像识别,从界面设计到交互流程,全面提升应用的可访问性。我们相信,科技的力量应当惠及每一个人,无论他们是否存在障碍。我们将秉持这一理念,不断前行,让科技之光照亮每一个角落,与所有人共同筑造一个无障碍、包容、和谐的数字世界。

参考文献

- [1]马琦珉,李向民,周雅倩.基于大语言模型的移动应用可访问性增强方法[J].计算机科学,2024,51(12):223-233.
- [2]任元凯,谢振平.大语言模型领域意图的精准性增强方法[J].计算机应用研究,2024,41(10):2893-2899.
- [3]李向民,沈立炜,董震.基于录制回放的移动应用可访问性增强方法[J].计算机科学,2023,50(12):32-48.