

<Audio Augmented Reality: Real-Time Sound Replacement>

<Zhenyuan Shen>

School of Computing Science

Sir Alwyn Williams Building

University of Glasgow

G12 8RZ

A dissertation presented in part fulfillment of the requirements of the Degree of Master of Science at the University of Glasgow

<Date of submission placed here>

**Abstract**

Audio augmented reality (AAR) enriches the user experience by overlaying virtual sounds in the real-world auditory environment. This study explores the innovative use of real-time sound replacement in AAR, where everyday sounds (like human voices) are transformed into other interesting sounds (like a dog barking or thunder) instantly. Leveraging advanced audio processing technology and deep learning algorithms, this research achieves efficient and real-time sound replacement, demonstrating the potential application value of AAR in entertainment, education, virtual training, games, and artistic creation. The system's ability to swiftly and accurately replace sounds while preserving audio quality ensures an immersive user experience. Furthermore, this study uncovers the potential of sound replacement technology in privacy protection and assistive hearing devices, highlighting the broad impact of AAR. By presenting these findings, we aim to inspire new ideas and application avenues for AAR, fostering its innovation and application in diverse fields and instilling optimism about its future.

Education Use Consent

I hereby give my permission for this project to be shown to other University of Glasgow students and to be distributed in an electronic form.

<**Please note that you are under no obligation to sign this declaration, but doing so would help future students.>**

Name: Signature:

Acknowledgements

<Acknowledgements go here>

Contents

<Update the table of contents by right-clicking on it and selecting Update Field… and then select page numbers only.>

Chapter 1 Introduction <This is Heading 1> 1

1.1 A section <This is style Heading 2> 1

1.1.1 A subsection <This is style Heading 3> 1

Chapter 2 Survey 2

Chapter 3 Further Chapters 3

Chapter 4 Conclusion 4

Appendix A <Name of appendix> 1

Appendix B <Another appendix> 2

Bibliography 3

# Introduction <This is Heading 1>

Introduce the project.

<This is style Normal. We recommend you make use of styles to simplify creating a well-formatted document. We have used “space before” and “space after” in defining these styles, in order to space the headings and paragraphs appropriately. You should never need to enter a blank line.>

## A section <This is style Heading 2>

Please note your dissertation need not follow the included section headings – this is only a suggested structure. Also add subsections etc. as required.

### A subsection <This is style Heading 3>

Try to avoid this too much, but it’s here if you need it.

在当今的数字化时代，技术的迅速发展极大地丰富了人类的感官体验并且随着VR, AR的发展，音频增强现实（AAR）作为一种创新的技术，通过在现实世界中增加虚拟声音层，为用户提供了一种全新的交互方式。与传统耳机相比，AAR技术的一个显著优势在于声学透明性，即用户可以自然感知到周围环境因此一方面对于马路上的使用者来说安全性毋庸置疑但是人们的隐私又是一点需要解决的难题。然而，尽管技术上取得了一定的进展，实时声音替换技术在应用中仍面临许多挑战，如实时性、准确性和用户体验的优化等。

本研究旨在探索音频增强现实中的实时声音替换技术，通过深度学习（YAMNet model）和音频处理技术（双声道麦克风判断左右方向），判断出目标声音的类别的同时实现对于声源的定位将日常的某些声音即时转换为其他某种有趣的声音（如人声转换为狗叫声）。

本文将详细介绍使用YAMNet模型进行声音识别和音频处理转换的方法，并通过现场实验验证该Unity场景的有效性。此外，本研究还将探讨实时声音替换技术在提高用户交互体验和实际应用潜力方面的影响以及关于如何优化等一系列的问题。通过这项研究，我们希望不仅能够推动AAR技术的发展，还能为未来声音增强设备的设计和应用提供理论和实践指导。

# Survey

<Each new chapter should appear on a new page.>

Background concepts (if required) and overview of relevant previous work (critically evaluate strengths and weaknesses).

2.1 **AAR技术的定义与起源**

* 介绍AAR技术的基本概念和定义。
* 增强现实（Augmented Reality, 简称AR）是在一个特定计算机感知的虚拟信息与真实世界的动态叠加，将虚拟世界与现实世界结合起来的技术。这技术通过对人的视觉、听觉、嗅觉、触觉等感官进行模拟和增强，允许感官信息能够在空间中交互并动态共融。目前广泛接受的一种定义又是Azuma在1997年提出的，他认为AR应该具有三个特征：
* • 结合真实与虚拟
* • 实时交互
* • 支持三维注册
* 而我们所要讨论的增强现实，主要是音频增强现实其核心在于
* 音频增强现实（AAR）是一种在用户的现实世界音频体验中添加虚拟音频元素的技术。与视觉增强现实类似，AAR旨在通过增强自然声音或添加非现场声音来增强用户的听觉感知。AAR的技术起源可以追溯到2000年代初期，当时主要用于军事和航空领域的高级通信系统。随着智能手机和便携式设备的普及，AAR技术开始被商业化应用，尤其是在娱乐和消费电子产品中。

2.2 **AAR技术的关键组件**

* 描述构成AAR系统的核心技术，如音频捕捉、处理和输出设备。
* 详细讨论声源定位技术和环境声音处理技术的最新进展。

2.3 **当前AAR应用的实例分析**

* 分析AAR在不同领域（如娱乐、教育、辅助设备等）的应用案例。
* 评估这些应用中AAR技术的具体优势和挑战。

2.4 **技术挑战与研究方向**

* 讨论当前AAR技术面临的主要技术挑战，如实时处理能力、系统集成和用户接口优化等。
* 展望未来的研究方向，包括潜在的技术革新和市场需求。

# Further Chapters

The content of these chapters depends on the project and should be agreed with your supervisor (e.g. description of the solution, evaluation results, etc).

<Figure below is in style “figure” which continues to style “figure caption” when you press Enter and then back to “Normal” when you press Enter again.>

Figure 1: Some important shapes.

<If you wanted to show any code fragments, you could use the following style called code, which could then be followed by figure caption..>

*# This is a little bit of Python*

**for** i in range( 10 ):

**for** j in range( 10 ):

**print** i\*j,

**print**

Figure 2: A crucial algorithm for the project.

# Conclusion

Main conclusions of your project. Here you should also include suggestions for future work.

###### <Name of appendix>

<Use Heading 6 for the Appendix heading>

###### <Another appendix>

# Bibliography

[1] C. Baier and J.-P. Katoen. *Principles of Model Checking*. MIT Press, 2008.