

分类号: 按中国图书分类法, 学位办网上可查

单位代码: 10335

密 级: 注明密级与保密期限

学 号: _____

浙 江 大 学

硕士学位论文



中文论文题目: 流行歌曲自动翻译与歌声合成研究

英文论文题目: Automatic Song Translation and Singing Voi

申请人姓名：_____

指导教师：_____

合作导师：_____

学科(专业)：_____

研究方向：_____

所在学院：_____

论文递交日期

递交日期

流行歌曲自动翻译与歌声合成研究



论文作者签名: _____

指导教师签名: _____

论文评阅人 1: _____

评阅人 2: _____

评阅人 3: _____

评阅人 4: _____

评阅人 5: _____

答辩委员会主席: _____

委员 1: _____

委员 2: _____

委员 3: _____

委员 4: _____

委员 5: _____

答辩日期 _____

Automatic Song Translation and Singing Voice Synthesis on Pop Song



Author's signature: _____

Supervisor's signature: _____

External reviewers: _____

Examining Committee Chairperson:

Examining Committee Members:

Date of oral defence:

浙江大学研究生学位论文独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得 浙江大学 或其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

学位论文作者签名:

签字日期:

年

月

日

学位论文版权使用授权书

本学位论文作者完全了解 浙江大学 有权保留并向国家有关部门或机构送交本论文的复印件和磁盘，允许论文被查阅和借阅。本人授权 浙江大学 可以将学位论文的全部或部分内容编入有关数据库进行检索和传播，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

(保密的学位论文在解密后适用本授权书)

学位论文作者签名:

导师签名:

签字日期:

年

月

日

签字日期:

年

月

日

勘误表

序言

摘要

Abstract

缩略词表

英文缩写	英文全称	中文全称
NLP	Natural Language Processing	自然语言处理
AI	Artificial Intelligence	人工智能
SVS	Singing Voice Synthesis	歌声合成
NMT	Neural Machine Translation	神经机器翻译

目录

勘误表.....	I
序言	III
摘要	V
Abstract	VII
缩略词表	IX
目录	XI
图目录.....	XIII
表目录.....	XV
1 绪论	1
1.1 引言	1
1.2 国内外研究现状	1
1.3 研究意义及内容	1
1.4 章节安排	1
1.5 章节小结	2
2 自动歌曲翻译	3
2.1 歌曲翻译数据集	3
2.2 歌曲翻译数据的收集和预处理	3
2.3 Transformer Encoder-Decoder 模型的结构	3
2.4 Adaptive Computation Time 算法原理	3
2.5 歌曲翻译的评价指标	3
2.6 本章小结	3
3 歌声合成	5
3.1 歌声合成数据集	5
3.2 歌声合成数据的收集和预处理	5
3.3 扩散模型	5
3.4 歌声合成的评价指标	5
3.5 本章小结	5
	XI

4 总结和展望	7
4.1 总结.....	7
4.2 挑战与未来展望	7
附录	9
A 一个附录	9
B 另一个附录.....	9

图目录

图 A.1 附录中的图片.....	9
-------------------	---

表目录

‘

1 绪论

1.1 引言

文字和语音都是自然语言处理领域的研究对象。无论是比较早期的句法、词法分析，音素研究，还是诞生稍晚的针对语音内容的识别、拼接合成，对于文字和语音内容的研究可称得上互为表里。语音是文字可感知的信号载体，文字是语音内容的本质表示。机器翻译一直以来都是自然语言处理领域针对不同语言文字的一项重要技术。对于机器翻译的研究最早始于 20 世纪 30 年代，它的发展一直与计算机技术、语言学和信息论等学科密切相关。从早期的字典匹配、结合专家知识的规则翻译等朴素方法，到结合概率统计学和语言学的统计机器翻译，再到近年来随着运算硬件性能提高和深度神经网络发展而兴起的神经机器翻译，机器翻译技术取得了世人瞩目的成功，这项技术本身也逐渐从学界中的理论发展走向工业界的落地实践。

自动歌曲翻译是神经机器翻译在这一基础上针对非常规语体的拓展性研究，这一任务旨在。

歌声合成则是由语音合成任务衍生而来。语音合成是仅以文本为输入、以梅尔特征图或声波波形为输出的生成任务，歌声合成和语音合成任务不同的是，其文本发声时所对应的音高和时长都为设定好的乐谱所限制。歌声合成使得歌曲翻译结果的直观评测成为可能。同时，由于歌声合成可以承接歌曲翻译的输出结果而连接成为完整的、级联式的歌曲到歌声的翻译系统，这为歌曲翻译的实际应用打下了坚实的基础。

1.2 国内外研究现状

1.3 研究意义及内容

1.4 章节安排

本文各章节组织如下：

第一章：绪论。第一章主要介绍了歌曲翻译技术和歌声合成技术的定义、应用和

发展情况、自动歌曲翻译和歌声合成的国内外研究现状、以及自动歌曲翻译和歌声合成的研究背景及意义。另外，第一章节还阐述了本文将基于 xxx 来研究自动歌曲翻译任务、基于扩散模型来研究歌声合成任务。

第二章：

第三章：

第四章：

1.5 章节小结

2 自动歌曲翻译

2.1 歌曲翻译数据集

2.2 歌曲翻译数据的收集和预处理

2.3 Transformer Encoder-Decoder 模型的结构

2.4 Adaptive Computation Time 算法原理

2.5 歌曲翻译的评价指标

2.6 本章小结

3 歌声合成

3.1 歌声合成数据集

3.2 歌声合成数据的收集和预处理

3.3 扩散模型

3.4 歌声合成的评价指标

质量 MOS, 速度 RTF

3.5 本章小结

4 总结和展望

4.1 总结

4.2 挑战与未来展望

附录

A 一个附录

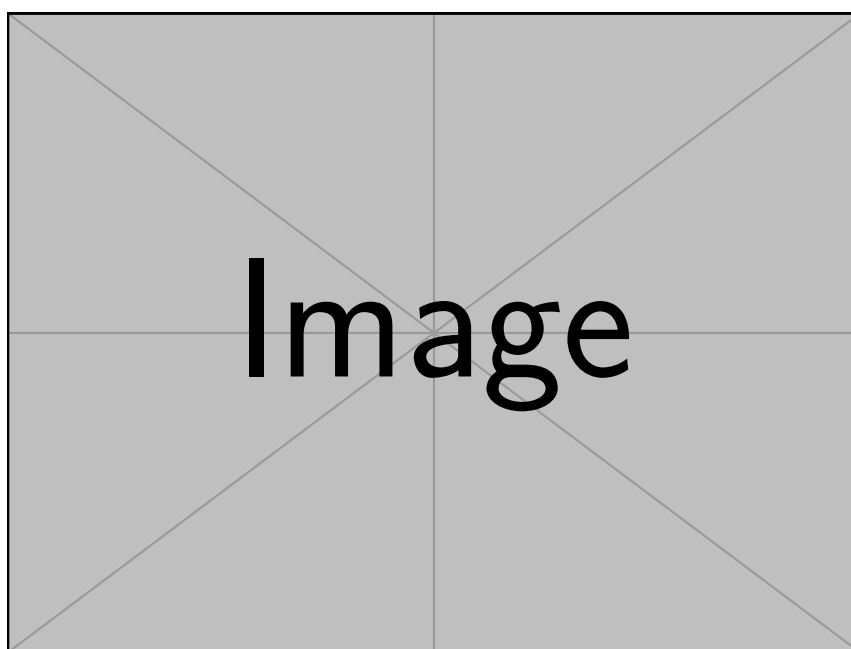


图 A.1 附录中的图片

B 另一个附录