

学校编码: 10384

分类号_____密级_____

学 号: 31520081153334

UDC_____

厦 门 大 学

硕 士 学 位 论 文

哼唱转 MIDI 自动记谱系统研究

Research on Humming to MIDI and Automatic Music

Transcription System

杨剑锋

指导教师姓名: 冯 寅 副教授

专 业 名 称: 智能科学与技术系

论文提交日期: 2011 年 5 月

论文答辩日期: 2011 年 月

学位授予日期: 2011 年 月

答辩委员会主席: _____

评 阅 人: _____

2011 年 5 月

厦门大学博硕士学位论文摘要库

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为()课题(组)的研究成果,获得()课题(组)经费或实验室的资助,在()实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

年 月 日

厦门大学博硕士学位论文摘要库

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

1. 经厦门大学保密委员会审查核定的保密学位论文，
于 年 月 日解密，解密后适用上述授权。

2. 不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年 月 日

厦门大学博硕士学位论文摘要库

摘要

音乐旋律的简谱或五线谱表示是人们常见的音乐旋律的标准符号化表达。而 MIDI 音乐文件则是音乐旋律的另一种符号化表达形式。哼唱转 MIDI 自动记谱 (Automatic Music Transcription) 是一种让计算机抽取人声哼唱音频信号中所描述的旋律信息(音符音高序列及其时长序列)并将其自动转化为符号化的 MIDI 乐谱的技术。这项技术的主要难点在于如何让计算机在一个哼唱得不是太准确的人声哼唱音频信号中寻找哼唱者所期望的正确旋律表达,并将其转化成 MIDI 乐谱格式。现有不少哼唱转 MIDI 自动记谱系统(如 WIDI Recognition System 等)由于无法对出现在结果中的大量不正确的短时音符进行合并、规整以及泛音过滤等方面的处理,以致其结果无法转化为利于管理、分析和应用的符号化、标准化的 MIDI 乐谱。本文提出一种基于倍音列的音高计算模型。在此基础上,针对 Krumhansl^[1, 2]调式识别模型的缺陷,提出一种基于哼唱的调式音阶参数训练方法,训练得到的参数模型能够使调式识别的第一命中率达到 90%(参见附录调式识别实验结果),最后根据人听觉的差异特性,即相对听觉误差,提出了一种浮动音高的规整方法,用于调整旋律音高识别过程中出现的半音到全音的识别误差,从而最终有效的实现旋律音符 75%左右的正确识别(其余的误识音符可以通过人机交互来调整)。

在哼唱转 MIDI 自动记谱技术应用方面,探索性地提出了一种哼唱识别检索系统数据库的构建方法,该方法具有较强的实用性。

结合以上的理论研究和算法,实现了一个哼唱转 MIDI 自动记谱系统,为自动记谱技术的研究打下了一定的基础,对进一步的深入研究具有推动和借鉴意义。

关键词: 哼唱转 MIDI; 自动记谱; 调式分析

厦门大学博硕士学位论文摘要库

Abstract

Automatic Music Transcription (AMT) is an advanced technology of transforming humming melody into symbolic and standardized MIDI score. This paper attempted to do research on the key technology of Automatic Music Transcription, including melody feature extraction, the storage and construction of feature library and the application in music retrieval.

The innovation and main work of this paper are as follow:

As to humming melody note segmentation, this paper studied the frontier technology of humming to MIDI Automatic Music Transcription and discussed about the most mature WIDI system. WIDI can approximately identify the humming melody, but can not transform the recognition result into symbolic and standardized MIDI score because of lacking the processes of short-time note mergence, pitch modulation and partials filtering. Aiming at solving these problems, a method based on time-domain feature and frequency-domain feature of humming melody was proposed.

As to pitch recognition, a floating pitch modulation method according to users' auditory error and a pitch modulation method based on music tonality theory were proposed; in the application of Automatic Music Transcription, a humming melody retrieval database was proposed.

Combining with the theory and algorithm proposed above, this paper developed a humming to MIDI Automatic Music Transcription system, which laid a solid foundation for the research of Automatic Music Transcription and had important significance for further research.

Key Words: Humming to MIDI; Automatic transcription; Tonality analysis

厦门大学博硕士学位论文摘要库

目 录

摘 要.....	I
Abstract.....	III
第一章 绪论.....	1
1.1 哼唱转 MIDI 自动记谱系统概述.....	1
1.2 研究的目的是和意义.....	1
1.2.1 研究的目的.....	2
1.2.2 研究的意义.....	2
1.3 国内外研究现状.....	3
1.4 论文的主要工作.....	4
1.5 论文结构.....	5
第二章 哼唱转 MIDI 自动记谱乐理基础.....	7
2.1 乐理基础.....	7
2.2 音频文件格式.....	11
2.3 旋律信号特征.....	14
2.3.1 时域特征.....	14
2.3.2 频域特征.....	16
2.3.3 倒谱分析.....	17
2.4 本章小结.....	18
第三章 哼唱转 MIDI 自动记谱关键技术研究.....	19
3.1 哼唱旋律预处理.....	19
3.1.1 去噪.....	19
3.1.2 预加重.....	20
3.1.3 加窗和分帧处理.....	20
3.2 音块切分技术.....	21
3.2.1 音块的定义.....	21

3.2.2 音块切分技术的定义.....	21
3.2.3 基于时频特征的音块切分.....	22
3.2.4 小结.....	28
3.3 音高识别技术.....	28
3.3.1 离散傅里叶变换(DFT).....	29
3.3.2 倍音列音高计算模型.....	30
3.3.3 基于倍频的音高计算模型.....	33
3.3.4 小结.....	34
3.4 基于调式的音高规整.....	34
3.4.1 调式的定义.....	35
3.4.2 调式音高规整技术介绍.....	35
3.5 浮动音高规整法.....	40
3.6 本章小结.....	43
第四章 自动记谱技术在音乐检索数据库构建中的应用.....	45
4.1 数据库构建.....	45
4.1.1 音高特征.....	45
4.1.2 时值特征.....	47
4.2 本章小结.....	48
第五章 系统实现与结果评估.....	49
5.1 AMT(Automatic Music Transcription)系统设计.....	49
5.1.1 哼唱旋律调式音阶参数训练模块.....	50
5.1.2 哼唱旋律音符划分和识别模块.....	51
5.1.3 人机交互模块.....	52
5.1.4 数据库录入模块.....	53
5.2 实验结果与评估.....	54
5.2.1 哼唱旋律调式识别.....	54
5.2.2 哼唱转 MIDI 自动记谱系统实验结果.....	55
5.3 本章小结.....	57
第六章 总结与展望.....	59
附录.....	61

参 考 文 献	65
攻读硕士期间发表的论文	69
致 谢.....	71

厦门大学博硕士论文摘要库

厦门大学博硕士学位论文摘要库

Contents

Abstract in Chinese	I
Abstract in English	III
Chapter 1 Introduction	1
1.1 Overview of AMT	1
1.2 Purpose and significance of the research	1
1.2.1 Purpose of the research	2
1.2.2 Significance of the research	2
1.3 Related research	3
1.4 Main work of this paper	4
1.5 Structure of this paper	5
Chapter 2 Basic theory of AMT	7
2.1 Basic music theory	7
2.2 Audio file format	11
2.3 Audio feature	14
2.3.1 Time-domain feature	14
2.3.2 Frequency-domain feature	16
2.3.3 Cepstrum analysis	17
2.4 Summary	18
Chapter 3 Research on key technology of AMT	19
3.1 Preprocessing of humming melody	19
3.1.1 Denoise	19
3.1.2 Pre-emphasis	20
3.1.3 Windowing and subframe	20
3.2 Note segmentation	21
3.2.1 Note definition	21
3.2.2 Note segmentation definition	21

3.2.3 Note segmentation based on the feature of time and frequency domain	22
3.2.4 Summary	28
3.3 Pitch recognition	28
3.3.1 DFT	29
3.3.2 Overtone series model.....	30
3.3.3 Improved overtone series model.....	33
3.3.4 Summary	34
3.4 Pitch modulation based on music tonality	34
3.4.1 Music tonality definition.....	35
3.4.2 Introduction of pitch modulation based on music tonality	35
3.5 Pitch modulation based on floating pitch interval	40
3.6 Summary	43
Chapter 4 The Application of AMT	45
4.1 Music database.....	45
4.1.1 Pitch feature	45
4.1.2 Duration feature	47
4.2 Summary	48
Chapter 5 System implementation and Result analysis.....	49
5.1 Function design of AMT	49
5.1.1 Parametric train of key profile	50
5.1.2 Note segmentation and pitch recognition	51
5.1.3 Man-machine interaction	52
5.1.4 Database application	53
5.2 Experiments and anslysis.....	54
5.2.1 Parametric train result of key profile	54
5.2.2 Experimental result of AMT	55
5.3 Summary	57
Chapter 6 Conclusion and Feature work.....	59
Appendix	61

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士论文摘要库