

中科院声学所叠音检测引擎

数据结构设计

目录

[1.概述 5](#_Toc90286993)

[1.1简介 5](#_Toc90286994)

[1.2 读者对象 5](#_Toc90286995)

[1.3术语定义 6](#_Toc90286996)

[2. 数据结构设计 7](#_Toc90286997)

版 本 历 史

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 版本/状态 | 作者 | 参与者 | 起止日期 | 备注 |
| 1.0 | 张英楠 |  | 2021/04/28 |  |
|  |  |  |  |  |
|  |  |  |  |  |

# 1.概述

1

1.1简介

叠音自动检测在语音技术领域中具有十分重要的地位。该技术是指从声音信号中寻找出多个说话人语音的重叠部分。进行有效的叠音自动检测能够对声音信号中多个说话人互相打断对话的情况进行检测。

## 1.2 读者对象

本文档是对识别引擎的内部代码流程和相关技术点的整体说明，因此读者最好满足如下几点条件后再读此文档：

* 有一定的C和C++语言基础
* 了解语音信号处理过程和原理

## 1.3术语定义

DNN(Deep Neural Networks）：深度神经网络

# 2. 数据结构设计

2

本节主要讲解叠音检测类主要内容，叠音检测主要的工作就是检测出原始语音中的叠音部分的起尾时间点。

* 实现源文件：DetectOverlap.h DetectOverlap.cpp
* 类名：cDetectOverlap
* 模块接口函数说明：

（1）cDetectOverlap();

功能：

构造函数，初始化引擎

形参：

返回值：

（2）bool InitDetectOverlap(const char \*systemDir, const char \*cfg);

功能：

初始化函数

形参：

const char \*systemDir：系统目录地址

const char \*cfg: 配置文件名

返回值：成功返回true，失败返回false

（3）bool Process(float \*newSet, int nFrameNum, int startFrm, int nDuration, short nFrameSize);

功能：

叠音检测接口函数

形参：

float \*newSet：输入语音特征

int nFrameNum：输入语音帧数

int startFrm：输入语音起点绝对时间（在整通语音中）

int nDuration, short nFrameSize：特征系数

返回值：成功返回true，失败返回false

（4）bool WriteOutSmoth(int \*curSegId, int beginTime, int iFrameNum, int \*Dnnstate, float Ftime, float \*postp);

功能：

叠音段输出函数，判断连续的叠音帧是否大于阈值

形参：

int \*curSegId：叠音ID

int beginTime：输入语音起点绝对时间（在整通语音中）

int iFrameNum：输入语音帧数

int \*Dnnstate：DNN判断出的每帧状态数组

float Ftime：帧长

float \*postp：每帧叠音概率数组

返回值：成功返回true，失败返回false