

中科院声学所DNNVAD引擎

数据结构设计

目录

[1.概述 4](#_Toc91668068)

[1.1简介 4](#_Toc91668069)

[1.2 读者对象 4](#_Toc91668070)

[1.3术语定义 5](#_Toc91668071)

[2. 数据结构设计 5](#_Toc91668072)

[2.1 特征数据信息结构体 5](#_Toc91668073)

[2.2 结果回调函数数据结构 6](#_Toc91668074)

版 本 历 史

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 版本/状态 | 作者 | 参与者 | 起止日期 | 备注 |
| 1.0 | 陈向东 | 龚云波  黎塔 | 2021/04/28 |  |
|  |  |  |  |  |
|  |  |  |  |  |

# 1.概述

1

1.1简介

DNNVAD的主要功能：前端功能，对输入的原始语音进行有效音截取，使得输入到识别引擎中的语音为有效音从而减少噪音对识别引擎的干扰。本引擎主要基于DNN模型的VAD算法。对原始语音进行特征提取并输入DNN模型，根据DNN模型的输出概率值并结合维特比算法，对原始语音逐帧判别是否为有效语音，最后根据起尾点判别算法将原始语音中的有效语音截取出来。

## 1.2 读者对象

本文档是对DNNVAD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
* 有基本的音频信号处理常识

## 1.3术语定义

DNN(Deep Neural Networks）：深度神经网络

VAD（Voice Activity Detection）：有效音检测

# 2. 数据结构设计

2

声学所DNNVAD引擎主要数据结构设计如下：

## 2.1 特征数据信息结构体

struct FeatHeader

{

int nFrameNum;

int nDuration;

short nFrameSize;

short iFeatType;

};

struct SphUncertntyInfo

{

float\* Feature;

int FrameNum;

};

## 2.2 结果回调函数数据结构

typedef void (\*Dnn\_vad\_callback) (int begin\_in\_Frame, int end\_in\_Frame, void\* pObj);