

中科院声学所DNNVAD模块

工程使用说明

目录

[1 概述 4](#_Toc84942980)

[2 编译环境 4](#_Toc84942981)

[3 整体框架 4](#_Toc84942982)

[4 主要源代码介绍 6](#_Toc84942983)

[5 接口调用流程 7](#_Toc84942984)

[5.1“单线程VAD”函数调用流程 7](#_Toc84942985)

版 本 历 史

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 版本/状态 | 作者 | 参与者 | 起止日期 | 备注 |
| 1.0 | 张英楠 |  | 2021/09/30 |  |
|  |  |  |  |  |
|  |  |  |  |  |

## 1 概述

VAD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。

DNNVAD是基于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。

## 2 编译环境

（1）编译器要求：

安装4.7以上版本的gcc和g++

（2）机器环境配置：

要求64位linux系统，Centos或Redhat 6.3以上企业开发版

（3）编译步骤：

进入DNNVAD目录执行： make

即可生成DNNVAD静态库：libDnnVad.a

（4）依赖库

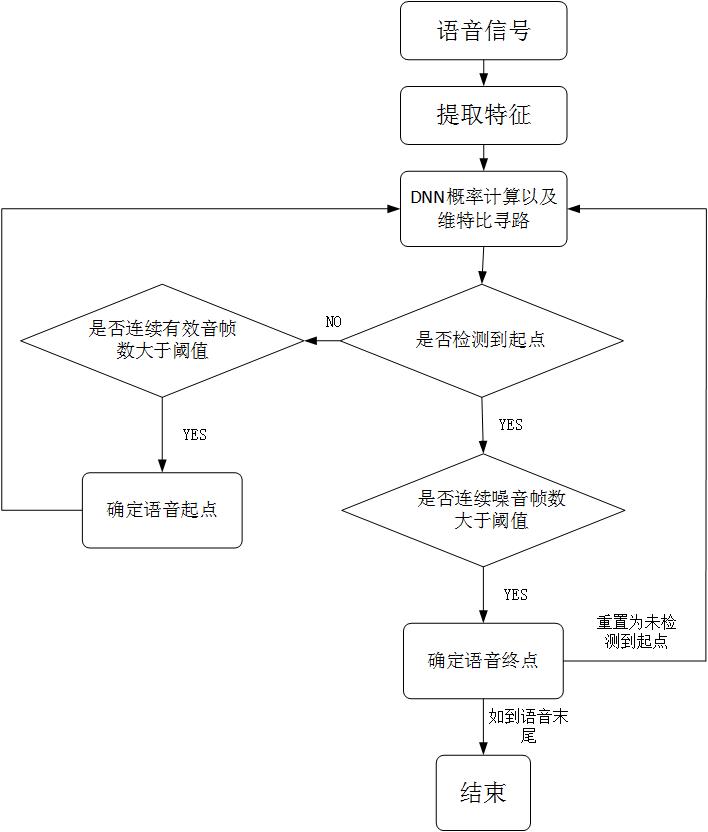
无

（5）第三方依赖库

无

## 3 整体框架

DNNVAD主要模块主要流程包括DNN模型推理、维特比寻址、起尾点判断，大致流程图如下：



## 4 主要源代码介绍

DNNVAD目录：用于编译生成libDnnVad.a

├──DNNVAD目录：编译目录

├──Makefile：Makefile文件

├──FenZng.cpp: 实现DNNVAD各功能模块的组合，对外接口API。

├──DSpSingle.cpp： 通过类DSpSingle来控制数据流程，进行VAD工作。

├──config\_DNNVAD.cpp： 实现读取配置

├──ChunkInfo.cpp： 实现将数据打包为chunk进行神经网络推理

├──BasicMath.cpp： 实现一些基本的数学运算功能

├──MatMath.cpp： 实现一些基本的矩阵运算功能

├──NeuralNet.cpp： 通过类FPNeuralNet类来整理构建神经网络

├──NetComponent.cpp： 实现神经网络各层Component的功能

## 5 接口调用流程

### 5.1“单线程VAD”函数调用流程

