EasyVAD

Technical Document
Version 2.0－ Revision 2006-1-11

Table of Contents

Introduction

EasyVAD is an implementation of VAD（Voice Active Detection）. EasyVAD support multiple channels concurrent. There is no limit on concurrent channels and it can up to thousands channels.
The binary library of EasyVAD only support 8kHz samples. If you want to support samples other than 8kHz, you need buy the source code of EasyVAD.

EasyVAD specifications	
Speech sampling rate(Hz)	8000
Samples in one Frame	Any

EasyVAD has a binary release version on Windows and Linux. The source code of EasyVAD is written by C/C++, so you can easily port it to UNIX, PPC,DSP, Vxworks or other operation system that support C/C++.

PACKAGE CONTENTS

EasyVAD.pdf	This document
EasyVAD.lib	Win32 statically linkable library of VAD for Pentium and compatible processors.
libVAD.a	Linux statically linkable library of VAD for Pentium and compatible processors.
EasyVAD.h	API prototypes and constants declarations required by the sample programs.
test_vad directory	Microsoft VC6.0 sample application and Linux GCC sample application. Demonstrating how to use VAD API calls.

The input requires raw 16-bit mono PCM speech data sampled at 8000 Hz as input, i.e., without any header information. For every speech frame, consisting of any number samples

ALGORITHM COMPLEXITY

The complexity of VAD is represented as percentage of CPU usage, and is as follows when tested on an Intel 800 MHz Celeron-MMX:
VAD less than 1% CPU time

ABOUT THE SAMPLE PROGRAMS

The sample program under test_vad directory is used to simulate the VAD and demonstrate how to initialize and call the VAD process. The program is run as follows (where **infile** raw 16 bit PCM files sampled at 8 kHz):

```
    test_vad infile
```

To build the VAD sample programs on Windows, you can open test_vad.dsw with VC6.0 or later version. After compiler and link, it will create the execute program of test_vad.exe, you can test it with following command.

```
    test_vad test.pcm
```

To build the VAD sample programs on Linux, you only need rum **make** command. After you successfully finished make command, you can run **make run** to test VAD.

EasyVAD API FUNCTIONS

EasyVAD_init

Description Initializes the memory needed by the VAD process. This function must be called prior to opening or re-opening a channel.

Syntax

```
#include “EasyVAD.h”

VAD_HANDLE EasyVAD_init();
```

Arguments none

Returned value Return a handle that represent an VAD channel, this value will used at EasyVAD_IsSilence and EasyVAD_release

EasyVAD_IsSilence

Description Detect the input voice samples and decide if it is a silence voice or not..

Syntax

```
#include “EasyVAD.h”

bool   EasyVAD_IsSilence(VAD_HANDLE hVAD, short *input, int numSamples);
```

Arguments

hVAD	The handle returned by EasyVAD_init
input	Input speech buffer containing 16-bit PCM speech data.
numSamples	The samples number in input speech.

Returned value Return true if the input samples is silence voice, return false if the input samples is not silence voice.

EasyVAD_release

Description release the memory allocated by the VAD process. This function must be called before you quit your program. If not, it will cause the memory leak.

Syntax

```
#include “EasyVAD.h”

bool   EasyVAD_release(VAD_HANDLE hVAD);
```

Arguments

hVAD	The VAD handle returned by EasyVAD_init

Returned value Return true if successful, return false if failed.

FAQs

Here are some frequently asked questions about the EasyVAD.

Q — What type of speech input format is required?
A — Raw 16-bit mono PCM sampled at 8000Hz. Do not use .WAV files. They contain a header that will produce distortion at the start of a decoded audio sample because the encoder interprets the header as speech data.

Q — How can I convert my .WAV files to raw 16 bit mono PCM sampled at 8000 Hz?
A — Use an audio editing tool such as SoX - Sound eXchange. See home.sprynet.com/~cbagwell/sox.html for more information

Q — Can I get link on platforms other than Pentium or compatible?
A — The object code provided in this package is Microsoft Win32 and Linux x86 compatible. It is compiled for the Pentium family of processors. If you want to use EasyVAD on other platforms, you should buy the source code of EasyVAD. Then you can compile and link.

Q — Is the EasyVAD able to handle multiple channels?
A — Yes, It can handle multiple channels. There is no limited.

Q — Is the EasyVAD codec free to use?
A — No, The version you get freely is a version only for test. If you want to use it in commercial, you must buy it from www.imtelephone.com. This version has the same function with the formal release version, but It can only run less than 6 hours continuously.

Q — How much does the EasyVAD codec cost?
A — The object code of Windows or Linux is $2000/year. The source code is $20000/year. You can buy it from www.imtelephone.com.