Cyberon VAD
Programming Guide

Version: 1.0
Date of issue: 2020/12/22

Cyberon Corporation

Software solution provider for embedded system

http://www.cyberon.com.tw/

Contents

History：

Ver.	Data	Update
1.0	2020-12-22	1. First release.

1. Voice Activity Detection

1.1. Introduction

Cyberon voice activity detection(VAD) is used to detect human voice. There are four state of speech:

1. Silence, or say no-speech precisely.
2. Voice start.
3. Voice active.
4. Voice end.

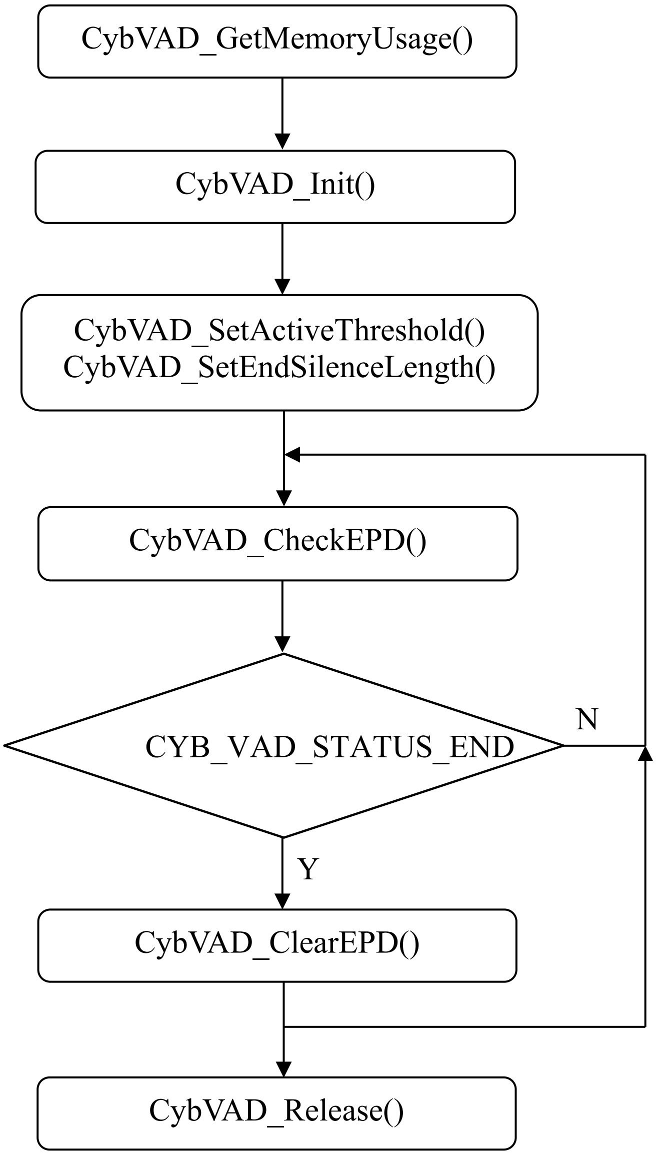
At current time, Cyberon VAD only support 16 KHz, mono, 16 bits PCM input format. The frame time of VAD is 10 ms, means it detect VAD for every 10 ms.

1.2. Specification

1. Code & table: 20.4 K
2. RAM: 5.5 KB.
3. CPU usage: 4.1 MCPS, test on CM4 with -O3 compile option.

2. Cyberon Voice Activity Detection SDK

2.1. Calling Flow Chart



2.2. API Function

CybVAD_GetMemoryUsage

Get the memory usage.

Prototype

INT CybVAD_GetMemoryUsage(INT nSampleRate)

Parameters

nSampleRate (IN):

It must be 16000.

Return value

CYBVAD_SUCCESS(0) if success, negative value if error.

CybVAD_Init

Initialize a VAD object.

Prototype

HANDLE CybVAD_Init(INT nSampleRate, BYTE *lpbyMemPool, INT nMemSize, const BYTE *lpbyLicense, INT *pnErr)

Parameters

nSampleRate (IN):

It shall be 16000.

lpbyMemPool (IN):

Memory buffer for the VAD.

nMemSize (IN):

The size in bytes of lpbyMemPool. It shall be great or equal than the return value of CybVAD_GetMemoryUsage().

lpbyLicense (IN):

The license data..

pnErr (OUT):

The error code.

Return value

The handle of VAD object or NULL if error.

CybVAD_Release

Release a VAD object.

Prototype

INT CybVAD_Release(HANDLE hCybVAD)

Parameters

hCybVAD (IN):

The handle of VAD object.

Return value

CYBVAD_SUCCESS(0) if success, negative value if error.

CybVAD_SetActiveThreshold

Set the active threshold of VAD.

Prototype

INT CybVAD_SetActiveThreshold(HANDLE hCybVAD, INT nActiveThreshold)

Parameters

hCybVAD (IN):

The handle of VAD object.

nActiveThreshold (IN):

The value range is 1 ~ 6, default value is 1, lower threshold means the VAD is more easy to become active.

Return value

CYBVAD_SUCCESS(0) if success, negative value if error.

CybVAD_SetEndSilenceLength

Set the ending silence of VAD.

Prototype

INT CybVAD_SetEndSilenceLength(HANDLE hCybVAD, INT nSilenceIn10ms)

Parameters

hCybVAD (IN):

The handle of VAD object.

nSilenceIn10ms (IN):

The length of ending silence, unit is 10 ms. The valid range is 10 ~ 200. The default value is 80.

Return value

CYBVAD_SUCCESS(0) if success, negative value if error.

CybVAD_CheckEPD

Detect the endpoint status(EPD) of current frame.

Prototype

INT CybVAD_CheckEPD(HANDLE hCybVAD, const SHORT *lpsSample, INT nNumSample)

Parameters

hCybVAD (IN):

The handle of VAD object.

lpsSample (IN):

The 16 bits, mono PCM audio input data.

nNumSample (IN):

The sample count of lpsSample.The VAD frame time is 10 ms, 160 samples for 16 KHz, so nNumSample must be the multiple of 160.

Return value

CYB_VAD_STATUS_INIT:

After CybVAD_Init(), VAD need 80 ms data to initial its internal state. So, CybVAD_CheckEPD will return this value during this period.

CYB_VAD_STATUS_NO_SPEECH:

VAD can't find human speech.

CYB_VAD_STATUS_START:

VAD find the start point of human speech.

CYB_VAD_STATUS_ACTIVE:

VAD find the human speech is still active.

CYB_VAD_STATUS_END:

VAD find the end of human speech.

Other negative value if error.

CybVAD_ClearEPD

After CybVAD_CheckEPD() finish the speech detection(return CYB_VAD_STATUS_END), we need to call this API to clear VAD state for prepare the next detection.

Prototype

INT CybVAD_ClearEPD(HANDLE hCybVAD)

Parameters

hCybVAD (IN):

The handle of VAD object.

Return value

CYBVAD_SUCCESS(0) if success, negative value if error.