

Rockchip Sound Event Detection开发文档

文件标识: RK-KF-SF-959

发布版本: V1.0.2

日期: 2022-08-20

文件密级: ☐绝密 ☐秘密 ☐内部资料 ☒公开

免责声明

本文档按“现状”提供, 瑞芯微电子股份有限公司(“本公司”, 下同)不对本文档的任何陈述、信息和内容的准确性、可靠性、完整性、适销性、特定目的性和非侵权性提供任何明示或暗示的声明或保证。本文档仅作为使用指导的参考。

由于产品版本升级或其他原因, 本文档将可能在未经任何通知的情况下, 不定期进行更新或修改。

商标声明

“Rockchip”、“瑞芯微”、“瑞芯”均为本公司的注册商标, 归本公司所有。

本文档可能提及的其他所有注册商标或商标, 由其各自拥有者所有。

版权所有 © 2022 瑞芯微电子股份有限公司

超越合理使用范畴, 非经本公司书面许可, 任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部, 并不得以任何形式传播。

瑞芯微电子股份有限公司

Rockchip Electronics Co., Ltd.

地址: 福建省福州市铜盘路软件园A区18号

网址: www.rock-chips.com

客户服务电话: +86-4007-700-590

客户服务传真: +86-591-83951833

客户服务邮箱: fae@rock-chips.com

产品版本

芯片名称	内核版本
全系列	通用

读者对象

本文档（本指南）主要适用于以下工程师：

技术支持工程师

软件开发工程师

修订记录

版本号	作者	修改日期	修改说明
V1.0.0	廖华平、江迪	2022-07-23	初始版本
V1.0.1	廖华平、郑兴	2022-08-15	整理文档格式
V1.0.2	廖华平	2022-08-20	更新接口，加入蜂鸣器检测

目录

Rockchip Sound Event Detection开发文档

1. 概述
2. 功能描述
 - 2.1 Abnormal Event Detection(AED)
 - 2.2 Baby Cry Detection(BCD)
 - 2.3 Buzz Detection(BUZ)
3. 相关API介绍
 - 3.1 rkaudio_sed_init
 - 3.2 rkaudio_sed_destroy
 - 3.3 rkaudio_sed_process
 - 3.4 rkaudio_sed_param_init
 - 3.5 rkaudio_sed_param_destroy
 - 3.6 rkaudio_sed_param_aed
 - 3.7 rkaudio_sed_param_bcd
 - 3.8 rkaudio_sed_param_buz
4. 参数介绍
 - 4.1 RKAudioSedParam
 - 4.2 SedAedParam
 - 4.3 SedBcdParam
 - 4.4 SedBuzParam
 - 4.5 RKAudioSedRes
 - 4.6 RKAudioSedEnable

1. 概述

本文档主要描述声音事件检测(Sound Event Detection)功能。当前包含婴儿哭声检测(Baby Cry Detection)、异常声检测(Abnormal Event Detection)和蜂鸣器检测(Buzz Detection)。三个模块通过统一接口调用，但功能相互独立，可通过开关使能关闭其中任何模块。

2. 功能描述

2.1 Abnormal Event Detection(AED)

AED实现实时异常声检测功能，包括超大声检测和信噪比检测。超大声检测实现对dB的检测，超过设定的dB值输出1，否则输出0。信噪比检测主要实现对噪声和信号进行检测，这里说的噪声主要是环境中的平稳噪声和录音的底噪，当信噪比大于设定阈值后输出1，否则为0。

信噪比（SNR）可理解为信号与噪声的比值：

$$SNR = 20 * \log_{10}(S/N)$$

假设噪声平稳不变，信号幅度每增大一倍，SNR就增加6dB，并以此来作为判断的阈值。

2.2 Baby Cry Detection(BCD)

BCD实现实时检测婴儿哭声的功能。通过深度学习的方式进行婴儿哭声检测，信噪比高的时候效果较好。

2.3 Buzz Detection(BUZ)

BUZ实现实时检测蜂鸣器检测的功能。主要检测常见的警报声，包括烟雾报警、防空报警、防盗报警等。通过深度学习的方式进行警报声检测，信噪比高的时候效果较好。

3. 相关API介绍

该功能模块为用户提供以下API:

- [rkaudio_sed_init](#): SED初始化。
- [rkaudio_sed_destroy](#): SED销毁。
- [rkaudio_sed_process](#): SED执行。

3.1 rkaudio_sed_init

【描述】

初始化并返回SED的操作句柄，此句柄用于[rkaudio_sed_process](#)。使用结束后，执行[rkaudio_sed_destroy](#)销毁。

【语法】

void *rkaudio_sed_init(int fs, int bit, int chan, [RKAudioSedParam](#) *param)

【参数】

参数名	描述	输入/输出
fs	采样率，AED支持8k和16k，BCD和BUZ只支持16k。	输入
bit	每个数据的bit数，一般使用的都是16bit数据。	输入
chan	通道数，如果输入多通道数据，使用的是第一个通道的数据。	输入
param	SED参数，相关定义见 RKAudioSedParam 。可通过函数 rkaudio_sed_param_init 构建，也可自行构建相关函数和初始化系数。	输入

【返回值】

返回值	描述
NULL	失败。
非NULL	成功。

3.2 rkaudio_sed_destroy

【描述】

销毁SED句柄。

【语法】

void rkaudio_sed_destroy(void *st_)

【参数】

参数名	描述	输入/输出
st	句柄。	输入

【返回值】

无

3.3 rkaudio_sed_process

【描述】

进行声音事件检测，返回结果存于res中。

【语法】

int rkaudio_sed_process(void *st_, short *in, int in_size, [RKAudioSedRes](#) *res)

【参数】

参数名	描述	输入/输出
st_	句柄。	输入
in	输入数据的指针。	输入
in_size	输入数据的长度，8k数据size应为128的倍数，16k数据size应为256的倍数。	输入
res	检测结果结构体指针，此结构体需在外部申请，定义参见 RKAudioSedRes 。	输出

【返回值】

返回值	描述
大于等于0	执行成功，此返回值为执行数据的长度。
小于0	执行失败。

3.4 rkaudio_sed_param_init

【描述】

初始化SED模块参数，进行子模块使能，并调用各个子模块参数初始化函数。此函数源码对外开放，并且各参数默认值已设置，也可根据实际数据进行适当调整。使用完后，调用[rkaudio_sed_param_destroy](#)销毁。如果在SED调用过程中要对参数或者模块使能进行调整，需要将SED模块销毁后，重新初始化才能生效。

【语法】

RKAudioSedParam *rkaudio_sed_param_init()

【返回值】

SED参数指针，定义参见[RKAudioSedParam](#)。

3.5 rkaudio_sed_param_destroy

【描述】

销毁SED模块参数。

【语法】

void rkaudio_sed_param_destroy([RKAudioSedParam](#) *param)

【参数】

SED参数指针。

3.6 rkaudio_sed_param_aed

【描述】

初始化AED模块参数，在rkaudio_sed_param_init函数中调用，在rkaudio_sed_param_destroy函数中销毁。

【语法】

[SedAedParam](#) *rkaudio_sed_param_aed()

【返回值】

AED模块参数指针。

3.7 rkaudio_sed_param_bcd

【描述】

初始化BCD模块参数，在rkaudio_sed_param_init函数中调用，在rkaudio_sed_param_destroy函数中销毁。

【语法】

[SedBcdParam](#) *rkaudio_sed_param_bcd()

【返回值】

BCD模块参数指针。

3.8 rkaudio_sed_param_buz

【描述】

初始化BUZ模块参数，在rkaudio_sed_param_init函数中调用，在rkaudio_sed_param_destroy函数中销毁。

【语法】

[SedBuzParam](#) *rkaudio_sed_param_buz()

【返回值】

BUZ模块参数指针。

4. 参数介绍

4.1 RKAudioSedParam

【说明】

SED算法相关参数。

【定义】

```
typedef struct RKAudioSedParam_  
{  
    int            model_en;  
    SedAedParam    *aed_param;  
    SedBcdParam    *bcd_param;  
    SedBuzParam    *buz_param;  
} RKAudioSedParam;
```

【成员】

成员名称	描述
model_en	通过设置bit位开启子模块，各bit定义参见 RKSedEnable ，如要开启AED和BCD，则应设为EN_AED EN_BCD。
SedAedParam	AED模块参数，定义参见 SedAedParam 。
SedBcdParam	BCD模块参数，定义参见 SedBcdParam 。
SedBuzParam	BUZ模块参数，定义参见 SedBuzParam 。

4.2 SedAedParam

【说明】

AED算法相关参数。

【定义】

```
typedef struct SedAedParam_  
{  
    float    snr_db;           // 信噪比大于snr输出1，单位为db  
    float    lsd_db;           // 响度大于db值输出1，最高为0db  
    int      policy;           // vad灵敏度，0->2，灵敏度等级提升。默认为1。  
} SedAedParam;
```

【成员】

成员名称	描述
snr_db	语音信噪比阈值，大于则输出1。
lsd_db	超大声阈值，大于则输出1。最大为0dB。
policy	信噪比检测算法灵敏度，取指范围为[0， 2]， 值越大越灵敏，越容易满足检测阈值。默认取1。

4.3 SedBcdParam

【说明】

BCD算法相关参数。

【定义】

```
typedef struct SedBcdParam_  
{  
    int      frm_len;           // 符合条件的帧长，越长越难唤醒  
    int      max_blank_frm;    // 重置帧长，超过此帧长重新统计  
    float    cry_energy;       // 哭声能量阈值  
    float    jug_energy;       // 能量判断阈值  
    float    cry_thres1;       // 哭声阈值1  
    float    cry_thres2;       // 哭声阈值2  
} SedBcdParam;
```

【成员】

成员名称	描述
frm_len	统计值大于此阈值则为唤醒， 越长越难唤醒， 此为主要调试参数。
max_blank_frm	重置帧长， 越小越容易重置， 则越难唤醒。
cry_energy	能量阈值0， 小于此能量则统计值-1。
jug_energy	能量阈值1， 小于此能量重置帧长+1。
cry_thres1	哭声概率大于此阈值则统计值+1。
cry_thres2	哭声概率小于此阈值则统计值-2。

4.4 SedBuzParam

【说明】

BUZ算法相关参数。

【定义】

```
typedef struct SedBuzParam_
{
    int      frm_len;           // 统计帧长，越长越难唤醒
    int      max_blank_frm;     // 重置帧长，超过此帧长重新统计
    float    mean_energy;      // 平均能量
    float    max_energy;       // 最大能量
    float    buz_thres1;       // 数据阈值1
    float    buz_thres2;       // 数据阈值2
} SedBuzParam;
```

【成员】

成员名称	描述
frm_len	统计值大于此阈值则为唤醒，越长越难唤醒，此为主要调试参数。
max_blank_frm	重置帧长，越小越容易重置，则越难唤醒。
mean_energy	平均能量阈值，小于此能量则重置帧长+1。
max_energy	最大能量阈值，小于此能量则重置帧长+1。
buz_thres1	警报声概率大于此阈值则统计值+1。
buz_thres2	警报声概率小于此阈值则统计值-5。

4.5 RKAudioSedRes

【说明】

SED模块返回结果。

【定义】

```
typedef struct RKAudioSedRes_ {
    int snr_res;
    int lsd_res;
    int bcd_res;
    int buz_res;
} RKAudioSedRes;
```

【成员】

成员名称	描述
snr_res	SNR返回结果，1为满足信噪比阈值，0不满足。
lsd_res	LSD返回结果，1为满足超大声阈值，0不满足。
bcd_res	BCD返回结果，1为检测到哭声，0没检测到。
buz_res	BUZ返回结果，1为检测到警报声，0没检测到。

4.6 RKAudioSedEnable

【说明】

使能各模块，将此值赋给model_en， 则可使能对应模块。如要使能多个模块，则使用EN_AED| EN_BCD方式。

【定义】

```
typedef enum RKAudioSedEnable_  
{  
    EN_AED = 1 << 0,  
    EN_BCD = 1 << 1,  
    EN_BUZ = 1 << 2,  
} RKAudioSedEnable;
```

【成员】

成员名称	描述
EN_AED	使能AED模块。
EN_BCD	使能BCD模块。
EN_BUZ	使能BUZ模块。