

Rockchip Audio Detection开发文档

文件标识: RK-KF-SF-959

发布版本: V1.0.1

日期: 2022-08-15

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

目录

Rockchip Audio Detection开发文档

1. 概述
2. 功能描述
 - 2.1 Baby Cry Detection(BCD)
 - 2.2 Abnormal Event Detection(AED)
3. 相关API介绍
 - 3.1 rkaudio_bcd_init
 - 3.2 rkaudio_bcd_destroy
 - 3.3 rkaudio_bcd_process
 - 3.4 rkaudio_aed_init
 - 3.5 rkaudio_aed_destroy
 - 3.6 rkaudio_aed_process
4. 参数介绍
 - 4.1 RKAudioBCDParam
 - 4.2 RKAudioAEDParam
 - 4.3 RKAedStat

1. 概述

本文档包括Baby Cry Detection(哭声事件检测)和Abnormal Event Detection(异常声检测)相关算法的说明和使用。两种算法相互独立，没有依赖关系。

2. 功能描述

2.1 Baby Cry Detection(BCD)

BCD实现实时检测婴儿哭声的功能。检测算法主要分为两个阶段：第一阶段先利用每帧的特征进行初步判断，如果数据符合条件，那么会进入第二阶段进行更复杂的检测算法，最后输出结果。所以输入不同的声音数据，所消耗的算力是不同的，以实际统计为准。相关接口函数以rkaudio_bcd的前缀开头。

2.2 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。相关接口以rkaudio_aed的前缀开头。

分贝（decibel），是量度两个相同单位之数量比例的计量单位，幅值为我们声音数据大小的绝对值，分贝(dB)和幅值X的关系为：

$$dB = 20 * \log_{10}(X)$$

所以幅值为倍数关系，dB为加减关系。幅值每上升一倍，dB值上升6dB。16bit的音频数据满幅值32767，此值设为0dB，所以我们这里说的dB值都在0dB以下。

信噪比(SNR)可理解为信号与噪声的比值，噪声设为0dB，信号比噪声高6dB，那么此时的信噪比为6dB。

3. 相关API介绍

该功能模块为用户提供以下API:

- [rkaudio_bcd_init](#): BCD初始化。
- [rkaudio_bcd_destroy](#): BCD销毁。
- [rkaudio_bcd_process](#): BCD执行。
- [rkaudio_aed_init](#): AED初始化。
- [rkaudio_aed_destroy](#): AED销毁。
- [rkaudio_aed_process](#): AED执行。

3.1 rkaudio_bcd_init

【描述】

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

【语法】

void* rkaudio_bcd_init(int fs, int bit, int chan, void* param_)

【参数】

参数名	描述	输入/输出
fs	采样率，当前只支持8k和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_	bcd参数，相关定义见 RKAudioBCDParam 。可通过函数rkaudio_bcd_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_bcd_destroy

【描述】

销毁BCD句柄。

【语法】

void rkaudio_bcd_destroy(void* st_ptr)

【参数】

参数名	描述	输入/输出
st_ptr	句柄。	输入

【返回值】

无

3.3 rkaudio_bcd_process

【描述】

执行哭声检测。

【语法】

```
int rkaudio_bcd_process(void* st_ptr, short* in, int in_size, int *bcd_stat)
```

【参数】

参数名	描述	输入/输出
st_ptr	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
bcd_stat	检测结果，1为检测到哭声事件，0位没有检测到。	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_aed_init

【描述】

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

【语法】

```
void* rkaudio_aed_init(int fs, int bit, int chan, void *param)
```

【参数】

参数名	描述	输入/输出
fs	采样率，当前只支持8k和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_	aed参数，相关定义见 RKAudioAEDParam 。可通过函数rkaudio_aed_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。

3.5 rkaudio_aed_destroy

【描述】

销毁AED句柄。

【语法】

void rkaudio_aed_destroy(void* st_ptr)

【参数】

参数名	描述	输入/输出
st	句柄。	输入

【返回值】

无

3.6 rkaudio_aed_process

【描述】

执行异常声检测。

【语法】

int rkaudio_aed_process(void* st_ptr, short* in, int in_size, int *aed_stat)

【参数】

参数名	描述	输入/输出
st_ptr	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
aed_stat	检测结果，两个事件结果存储在不同的bit位中。超大声检测输出为(aed_stat&LSD_STAT)，信噪比检测输出结果为(aed_stat&SNR_STAT)，具体的定义参见 RKAedStat 。	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。

4. 参数介绍

4.1 RKAudioBCDParam

【说明】

BCD算法相关参数。

【定义】

```
typedef struct RKAudioBCDParam_ {
    int      usr_mode; /* 用户模式  0:自动模式  1:用户模式 */
    int      bypass; /* 旁路功能  0:不开启  1:开启 */
    int      alarm_threshold; /* 告警阈值,越小越容易触发告警 [0 100] */
    int      time_limit; /* 告警时限,单位ms,取值大于等于一帧的时长,海思为10,我们目前为16 */
    int      time_limit_threshold_count; /* 告警次数阈值,告警时限内告警大于次数即视为检测到哭声 */
    int      interval_time; /* 两次告警之间的最小时间间隔 */
    CallFunc callback; /* 回调函数接口 */
}RKAudioBCDParam;
```

【成员】

成员名称	描述
usr_mode	使用模式，0为自动模式使用默认参数，1为用户模式则使用用户配置的参数。
bypass	bypass， 设为0则不启用BCD模块，设为1则启用BCD模块。
alarm_threshold	告警阈值, 占位用，暂未使用。
time_limit	告警时限，占位用，暂未使用。
time_limit_threshold_count	告警次数阈值，占位用，暂未使用。
interval_time	两次告警之间的最小时间间隔，暂未使用。
callback	回调函数接口，占位用，暂未使用。

4.2 RKAudioAEDParam

【说明】

AED算法相关参数。

【定义】

```
typedef struct RKAudioAEDParam_  
{  
    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; /* 信噪比检测算法灵敏度，0->2， 灵敏度等级提升。默认为1 */  
} RKAudioAEDParam;
```

【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 RKAedStat

【说明】

AED输出结果的存放。

【定义】

```
typedef enum RKAedStat_  
{  
    LSD_STAT = 1 << 0, /* 超大声状态检测结果 */  
    SNR_STAT = 1 << 1, /* 信噪比状态检测结果 */  
} RKAedStat;
```

【成员】

成员名称	描述
LSD_STAT	aed_stat & LSD_STAT，得到超大声的检测结果，1为检测到相关事件。
SNR_STAT	aed_stat & SNR_STAT，得到信噪比的检测结果，1为检测到相关事件。