

RK3308

VAD 寄存器配置

发布版本:**1.1**

日期:**2018.09**

免责声明

本文档按“现状”提供，福州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 © 2018 福州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。

福州瑞芯微电子股份有限公司

Fuzhou 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：+86-591-83991906

客户服务传真：+86-591-83951833

客户服务邮箱：www.rock-chips.com

前言

概述

本文档主要介绍 RK3308 VAD 相关寄存器配置及调参说明。

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

日期	版本	作者	修改说明
2018.09	V1.1	Cherry.Chen	初始版本

目录

1	概述.....	1
2	寄存器配置方法	1
2.1	查看寄存器	1
2.2	寄存器配置	1
3	常用寄存器说明	1
4	调试参数说明	2
4.1	sound_thd.....	2
4.2	gain	3
4.3	vad_con_thd	4

插图目录

图 1 RK3308 VAD 模块对应寄存器..... 1

图 2 RK3308 VAD 寄存器配置示例..... 1

图 3 VAD 输入信号放大前后对比 3

图 4 原始信号噪音幅度和最小语音幅度对比 3

图 5 放大 15.898 信号噪音幅度和最小语音幅度对比 4

图 6 语音信号 vad_cnt 示例 4

表格目录

表 1 RK3308 VAD 寄存器 0x005c 说明 1

表 2 RK3308 VAD 寄存器 0x0060 说明..... 2

1 概述

本文档主要针对 RK3308 语音检测模块（VAD, Voice Activity Detection）常用寄存器对应参数说明和寄存器配置方法。

2 寄存器配置方法

2.1 查看寄存器

查看 VAD 相关寄存器可以使用命令 `cat /sys/kernel/debug/vad/reg` 查看当前 VAD 寄存器值，如下图：

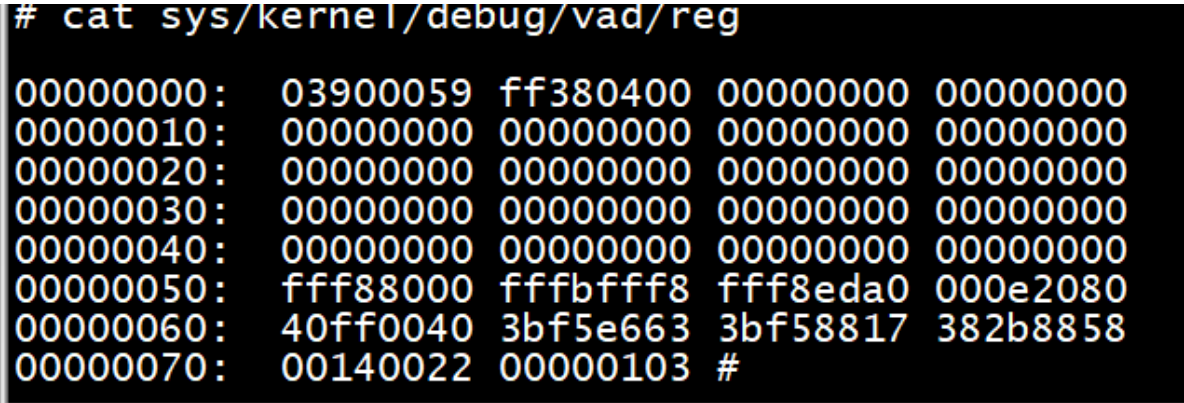


图 1 RK3308 VAD 模块对应寄存器

2.2 寄存器配置

寄存器配置示例如下图：

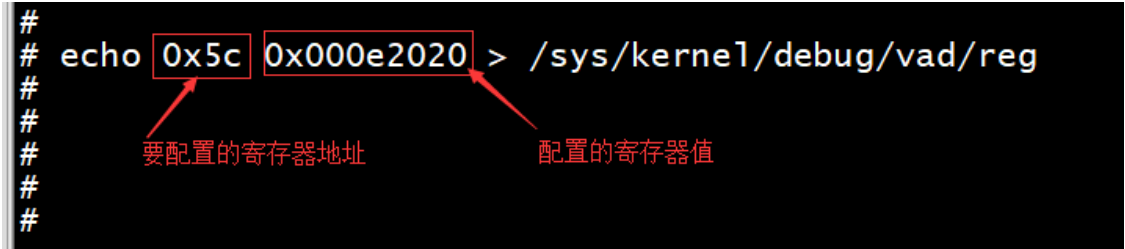


图 2 RK3308 VAD 寄存器配置示例

3 常用寄存器说明

表 1 RK3308 VAD 寄存器 0x005c 说明

Address: Operational Base + offset (0x005c)

Bit	Attr	Reset Value	Description
31:24	RO	0x0	reserved
23:16	RW	0x02	vad_con_thd When continuous sample number(>=vad_con_thd) exceed threshold, then assert the vad_det interrupt, the value N means N+1. When this value is higher, the voice detect condition is more strict
15	RO	0x0	reserved

14:12	RW	0x4	noise_level Noise level, valid value is 0x1~0x6 when this value is higher, the voice detect condition is more strict
11:10	RO	0x0	reserved
9:0	RW	0x020	gain The gain control of voice data, because of truncation 5 bits of result when the value is 32, the gain is 1 when the value is 64, the gain is 2 when is value is 16 the gain is 1/2 ...

表 2 RK3308 VAD 寄存器 0x0060 说明

Address: Operational Base + offset (0x0060)

Bit	Attr	Reset Value	Description
31	RO	0x0	reserved
30	RW	0x1	min_noise_find_mode Minimal noise value find mode 1'b0: Always find the value at the range of noise_frm_num 1'b1: When receive N frame: if N is less than noise_frm_num, find the value at the range of N; if N is more than noise_frm_num, find the value at the range of noise_frm_num
29	RW	0x0	clean_noise_at_begin 1'b0: The noise will be clean only at the begin of the first time vad is enable after reset 1'b1: The noise will be clean every time at the begin of vad is enable
28	RW	0x0	force_noise_clk_en Force noise calculate clk enable 1'b0: The clock will be auto gating for low power 1'b1: The clock will be always enable
27:26	RO	0x0	reserved
25:16	RW	0x0ff	noise_sample_num The number of sample in one frame to calculate the noise, the value N means N+1 sample. When this value is higher, the voice detect condition is more strict
15:0	RW	0x0064	sound_thd Initial sound threshold when this value is higher, the voice detect condition is more strict

4 调试参数说明

表 1 和表 2 分别是寄存器 0x5c 和 0x60 说明，在调试 vad 参数的时候，根据应用场景，选取最优的参数值。其中，几个比较重要的调试参数说明如下：

4.1 sound_thd

sound_thd 对应寄存器地址为 0x60 的 0~15bit，表示在应用场景下，纯净语音（理论上这个值是无

噪声环境下的语音幅度值，但是无噪声环境实现难度较大，因此一般选取安静环境下的语音（的最小幅度值的绝对值。`sound_thd` 的取值范围为：0~2¹⁵。

例如，应用场景为 5 米外说“灵犀灵犀”唤醒，那么 `sound_thd` 的调参方法为：在安静环境下 5m 距离，抓取“灵犀灵犀”的 RAW PCM 数据，“灵犀灵犀”声音起点的最低幅度值，由于抓取的语音数据不可能是纯净语音，因此，`sound_thd` 的选取可适当的减去当前环境噪声值。

4.2 gain

`gain` 对应寄存器地址为 0x5c 的 0~9bit，表示对 vad 输入信号的线性放大。放大倍数为 `g` 计算公式：

$$g = \frac{gain}{32}, \quad -512 \leq gain \leq 511 \quad (1)$$

放大倍数 `g` 的取值范围为，-16~15.9689。

当 vad 的输入信号幅度较小时，在同等信噪比语音下，噪声与声音信号的落差较小，因此需要对信号进行线性放大，增大噪声与声音信号的落差值。

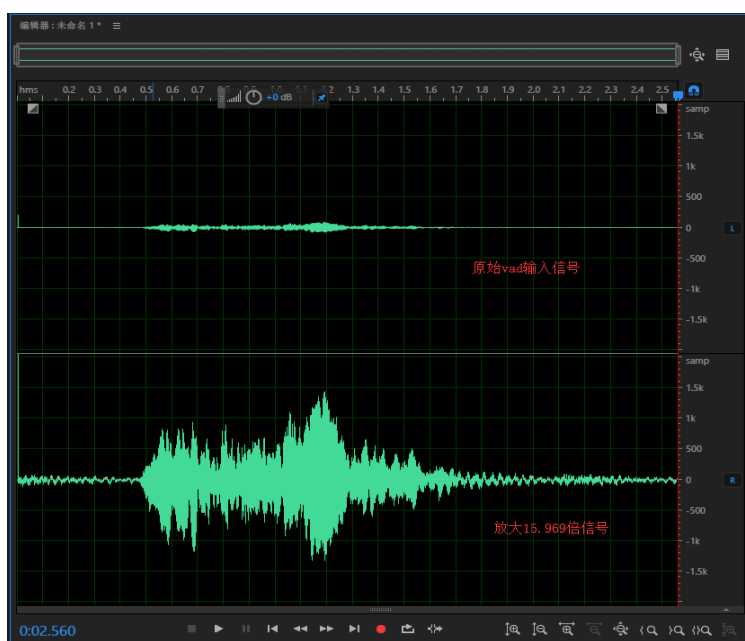


图 3 VAD 输入信号放大前后对比

如图 3 所示，上面为输入信号的原始波形，下面的波形为放大 15.969 倍以后的波形，图 4 为原始信号噪音幅度和最小语音幅度的对比，从图中可知，原始信号幅度和最小语音信号幅度的落差值小于 10，导致 `sound_thd` 的可调范围很小，增大 vad 判断难度。图 5 为放大 15.969 倍之后的噪声信号幅度和最小语音信号幅度对比，从图中可知，噪声信号和语音信号的幅度落差比放大 15.898 倍，增大了 `sound_thd` 的可调整范围。因此，在 vad 输入信号幅度太小的情况下，可以对信号进行适当的幅度放大，增加 vad 的灵活性。

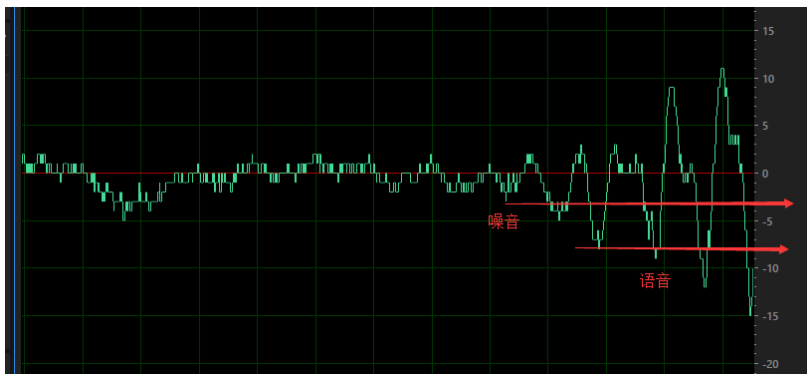


图 4 原始信号噪音幅度和最小语音幅度对比

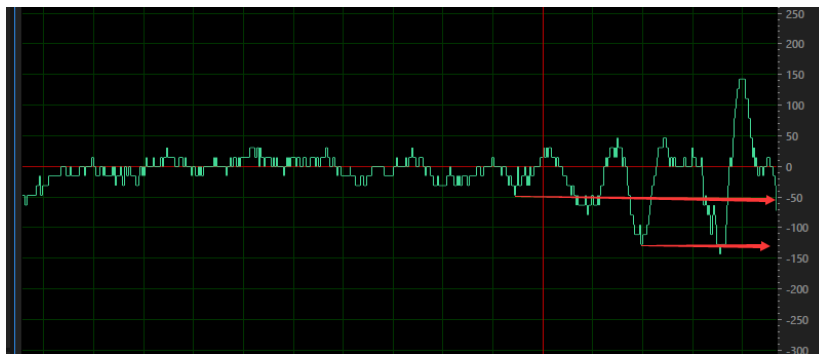


图 5 放大 15.898 信号噪音幅度和最小语音幅度对比

4.3 vad_con_thd

vad_con_thd 对应于寄存器 0x5c 的 16~23bit, 取值范围为: $0 \sim 2^7$, 表示连续超过 vad 门限值的样本个数 vad_cnt 大于一定的门限值 vad_con_thd 时, vad 中断触发, 即:

```
if(vad_cnt > vad_con_thd)
    vad_flag=1;//有语音
```

增大 vad_con_thd 值能够滤除一些毛刺影响, 但是过度增大可能导致部分不能检测到频率较高的声音。如图 6 所示, 图中蓝色为声音信号, 红色的线为阈值, 其中第一个波峰连续超过门限的样本数 vad_cnt=6, 第二个 vad_cnt=15, 若设置 vad_con_thd=8, 则, 在第一个波峰的毛刺声音能被滤除, 而第二个声音播放能触发 vad。

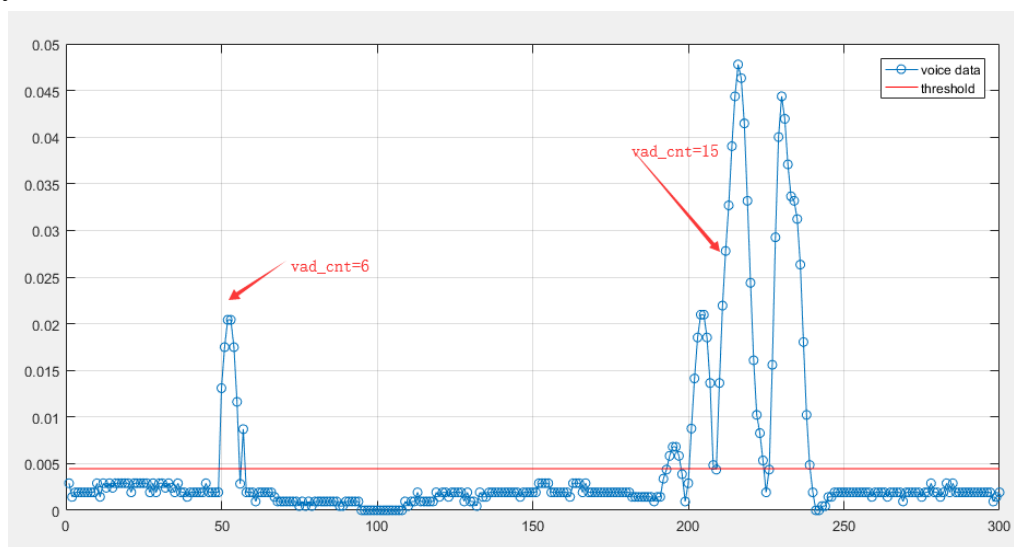


图 6 语音信号 vad_cnt 示例