

一种高效 MPEG-2/4 AAC 音频解码硬件实现

王 婷,田 泽,范飞虎

(中国航空计算技术研究所,陕西 西安 710119)

摘 要:目前 MPEG-2/4 AAC(Advanced Audio Coding,先进音频编码)音频处理广泛采用 DSP 核与特定音频格式编解码的软件实现,使得 MPEG-2/4 AAC 编解码技术在嵌入式系统中面临集成难度大、功耗高等问题。为了完成智能化、通用化、小型化的实时解码,文中提出了一种高效 MPEG-2/4 AAC 音频解码的硬件实现,介绍了体系结构设计、控制器、存储设计和软硬件协同工作等关键技术。通过仿真、FPGA 验证和流片测试证明该系统使用硬件实现音频解码,软件辅助流程控制,完成了音频实时解码,可移植性强,易集成,取得了较好的效果。

关键词:AAC;状态机;存储设计;软硬件协同工作

中图分类号:TP39

文献标识码:A

文章编号:1673-629X(2015)05-0010-03

doi:10.3969/j.issn.1673-629X.2015.05.003

Hardware Realization of an Efficient MPEG-2/4 AAC Audio Decoder

WANG Ting, TIAN Ze, FAN Fei-hu

(Aeronautics Computing Technique Research Institute, Xi'an 710119, China)

Abstract: It is normally adopted DSP and special software to realize MPEG-2/4 AAC audio decoder, so MPEG-2/4 AAC encoder and decoder has been faced with the problem of hardly integrated and high power in embedded system. In order to complete the real-time decoder with intelligence, generalization and miniaturization, propose the hardware realization of an efficient MPEG-2/4 AAC audio decoder. Introduce the key techniques like architecture design, memory, storage design, and the hardware/software co-design. Through the simulation, FPGA verification and test, the results show that the system utilizes the hardware to implement the audio decoding, software to aid the process control, finishing the real-time audio decoding, with strong portability and easy integration, which has good effects.

Key words: AAC; state machine; memory design; hardware/software co-design

0 引 言

MPEG-2/4 AAC 数字音频压缩标准是一种感知音频编解码算法,是当今最流行和广泛使用的编解码技术,具有音质好、压缩比高等特点,在数字声音的存储、网络音视频多媒体的传输和数字音视频产品领域得到了广泛的应用^[1-2]。通常采用优化了音频处理能力的 DSP 核与特定音频格式编解码的软件实现 MPEG-2/4 AAC 编解码,不利于小型化设计,集成难度大,功耗高^[3]。

文中基于某高清音视频解码芯片的项目实践,结合 MPEG-2/4 AAC 音频解码技术,提出了一种高效的 MPEG-2/4 AAC 音频解码硬件实现方案,给出了体系结构设计,详细介绍了控制器、存储设计和软硬件协同工作等关键技术。采用硬件实现音频解码,软件

辅助流程控制,可移植性强,为芯片实现及项目实践提供了有力的保证。

1 MPEG-2/4 AAC 音频编解码技术

MPEG-2/4 AAC 音频编码通过去除音频信号的客观和主观冗余达到压缩音频信号的目的,将输入的 PCM 音频源压缩编码成 ADTS 码流,每 1 024(单声道)或 2 048(双声道)个 PCM 压缩成一个 ADTS 帧,压缩比约为 11:1^[4-5]。反之,将 ADTS 码流解码输出 PCM 二进制序列就是解码过程。PCM 是模拟音频信号经过模数转换之后的二进制序列。ADTS 码流即音频传输流,是 AAC 的封装形式之一,其结构如图 1 所示。音频经压缩后以原始帧形式存在,原始帧的长度是可变的,对原始帧采用 ADTS 封装就形成了 ADTS 帧。一

收稿日期:2014-06-16

修回日期:2014-09-19

网络出版时间:2015-04-22

基金项目:“十二五”微电子预研(51308010601,51308010711);总装预研基金(9140A08010712HK6101)

作者简介:王 婷(1985-),女,硕士,中级工程师,主要从事数字集成电路设计和验证等方面研究;田 泽,博士,研究员,中国航空工业集团首席技术专家,研究方向为 SoC 设计、嵌入式系统设计、VLSI 设计等。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20150422.1113.036.html>

个 ADTS 帧由帧头和帧净荷组成;帧头定义了音频采样率、声道数、帧长度等关键信息,用于帧净荷的解析和解码;帧净荷由 1 至 4 个原始帧组成^[6-7]。

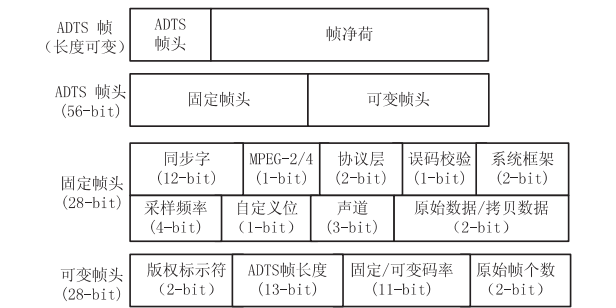


图1 ADTS 帧结构

2 MPEG-2/4 AAC 音频解码电路体系结构设计

MPEG-2/4 AAC 音频解码电路采用 Verilog 实现了 ISO 13818-7 和 ISO 14496-3 协议,能够完成将 MPEG-2/4 AAC 格式的压缩码流解码成 16-bit PCM 音频数据并通过 I²S 接口输出,体系结构如图 2 所示。

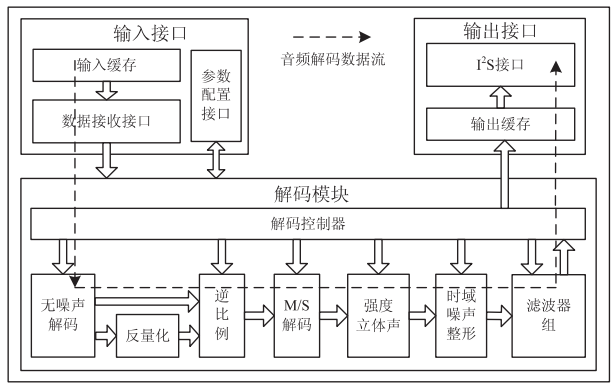


图2 音频解码电路体系结构

音频解码电路按照 ADTS 帧逐帧进行解码,图中虚线部分表明了解码的数据流方向。输出接口中 I²S 接口有三条信号线:位时钟 DABCK、左右声道选择时钟 DAWS 和数据线 DADATA^[8-9]。

3 关键技术介绍

3.1 解码控制器

解码控制器通过两个状态机完成输入输出接口以及解码各子模块之间数据流的控制。状态机 1 完成解码模块对外的状态控制,包括音频初始化和音频源输入;状态机 2 完成对 ADTS 帧解码过程的控制。

3.1.1 状态机 1

状态机 1 设计见图 3。音频解码开始时处于 IDLE 状态,之后根据不同的触发条件分别进入写寄存器、读寄存器或者音频源输入状态。其中,写寄存器优先级最高,读寄存器次之,最后是音频源输入。每一个状态

结束,状态机 1 都会回到 IDLE 状态,等待下一次触发条件。音频源输入条件满足后,解码控制器从输入接口读取一组固定大小的 ADTS 数据写入解码模块待解码,写完一组数据后状态机重新回到 IDLE 状态。

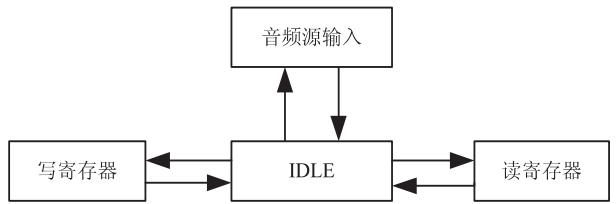


图3 状态机 1

3.1.2 状态机 2

状态机 2 如图 4 所示。音频解码模块按照 ADTS 帧逐帧进行解码,状态机 2 表示了一帧 ADTS 的解码过程,当一帧解码完成后状态机 2 重新回到 IDLE 状态。

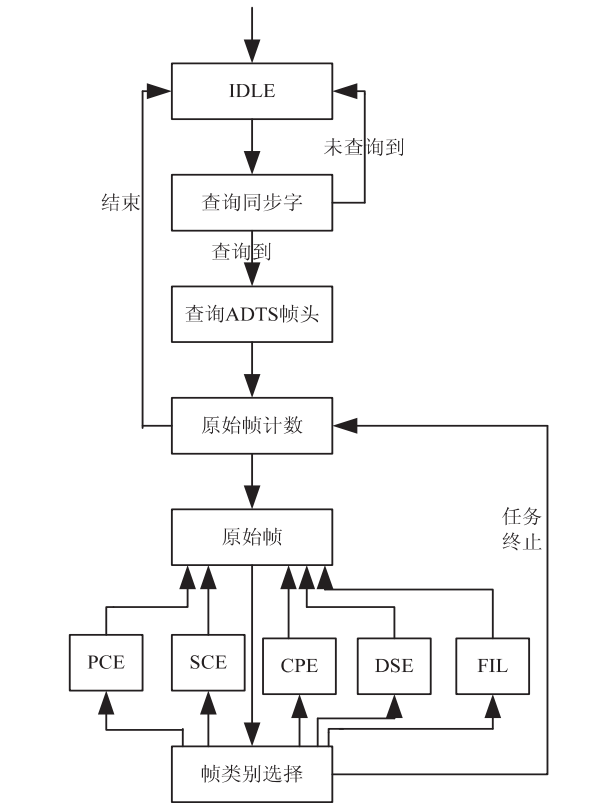


图4 状态机 2

状态机 2 通过对 MPEG-2/4 AAC 标准及其语法元素的分析,获取解码参数及控制信息,并通过发出启动信号和接收结束信号来监控其他解码功能模块(包括无噪声解码,反量化,逆比例,M/S 解码和强度立体声等功能),完成音频数据流在各个解码功能模块之间的逐级解码。

状态机 2 首先找到 12-bit 的同步字;接着查询解析 ADTS 帧头,获得解码参数并把这些参数写入参数寄存器组;然后对帧净荷中原始帧(通常是 1 到 4 个原始帧)进行解析,文中涉及的音频解码电路支持五种

常见原始帧类型的解码,包括,PCE(program_config_element),SCE(single_channel_element),CPE(channel_pair_element),DSE(data_stream_element)和 FIL(fill_element);对于原始帧的解码是通过各原始帧子状态机控制。当检测到原始帧不属于上述五类原始帧类型时,表示该解码电路不支持此类帧类型的解析,跳过解码电路不支持的原始帧,原始帧计数加 1,继续下一个原始帧解码。所有原始帧解析完成后,重新回到 IDLE 状态,进行下一帧解析。

3.2 存储设计

为了满足音频实时解码,输入端需要有足够的 ADTS 码流供解码电路解码,并同时完成解码电路和外部电路的时钟域转换;输出端存储 PCM 数据,需要保证输出的连续性。

3.2.1 输入缓存

对音频解码而言,因为传输的 ADTS 码流速度不确定,需要设定一个缓冲区来作为码流的暂时存放点,否则码流传输太快容易引起信息丢失,太慢则引起系统等待码流输入而浪费时间达不到实时解码要求,很难提高解码质量和声音的连续性。缓冲区大小由公式(1)计算得出^[10]:

$$\text{缓冲区大小} = (\text{每帧采样数} * \text{量化长度} * \text{声道数}) / \text{压缩比} \quad (1)$$

该系统测试码流为单声道和双声道,每帧采样数为 1 024,压缩比约为 11,按双声道计算缓冲区大小约为 372-bit。为了计算方便和留出一定的余量,取 $128 * 32\text{-bit}$ 作为缓冲区的大小,确保音质和最小计算复杂度。

此外,通常情况下音频源输入时钟频率较快,和外部输入电路有关,与音频解码系统时钟存在差异,所以输入缓存还需要完成音频源不同时钟域的时序转换。

综合上述两点因素,该设计使用一个位宽为 32-bit,深度为 128 双口 DPRAM 作为输入缓存。

DPRAM 的 A 口为音频源输入时钟,只使用写功能,当 DPRAM 为空时,写入 ADTS 码流,以 32-bit 为单位写入,写入 128 个 32-bit 后,数据接收接口将写数据完成信号发送给解码控制器,表示当前 DPRAM 中数据准备好。

DPRAM 的 B 口为音频解码系统时钟,只使用读功能,通过 B 口读取 DPRAM 中的 ADTS 码流;当输入缓存的 DPRAM 数据准备好且解码模块需要数据输入时,解码控制器将 DPRAM 中的 128 个 32-bit 数据分八组读入解码模块,每组 16 个 32-bit 数据。

3.2.2 输出缓存

输出缓存用于存储解码后的 PCM 数据,需要保证 PCM 数据的连续输出;每一个 ADTS 帧解码输出 PCM

的个数是固定的,且解码精度是 16-bit,即双声道时一个 ADTS 帧解码输出 2 048 个 16-bit PCM 数据,单声道时一个 ADTS 帧解码输出 1 024 个 16-bit PCM 数据;此外,通过考虑 I²S 接口数据传输速度和音频解码速度以及实时解码的要求,文中的输出缓存采用乒乓存储,输出缓存由四个深度为 1 024,宽度为 16-bit 的 DPRAM 组成,每两个 DPRAM 为一组,乒乓操作缓存 PCM 数据,可用于存储两个 ADTS 帧解码后的 PCM 数据。

当 PCM 为双声道时,每组中的 DPRAM 分别存储左声道和右声道,当选中的一组 DPRAM 中两个 DPRAM 都存满时,表示一帧解码完成,通知 I²S 接口从选中组的两个 DPRAM 中轮换读取左右声道数据;当 PCM 为单声道时,每组中只使用一个 DPRAM 存储数据,当选中组的一个 DPRAM 存满时,表示一帧解码完成,通知 I²S 接口读取数据,单声道时可认为是左右声道的数据相同,所以读取一次 DPRAM,同时在 I²S 接口的左右声道输出。

3.3 软硬件协同工作

在音频解码过程中,基本的数据传输处理都是由硬件完成,而传输状态不能由硬件显性地表现出来,因此可以考虑使用软件来实现传输过程中的控制以及状态监测^[11-12]。软件主要完成的工作有:初始化操作、中断查询及处理、解码状态查询。

初始化操作主要完成音频解码电路的初始配置。中断的设置用于硬件逻辑向处理器报异常事件,当解码异常时,硬件逻辑向处理器报中断,处理器接到中断信号后,复位硬件逻辑并进行重新配置,配置完成后,处理器查找到下一个有效的 ADTS 帧帧头,写入音频解码电路,重新开始解码。在解码过程中,处理器可以通过查询解码状态,确保音频解码电路正常工作。

4 结束语

通常采样率为 48 kHz 时,基于 ARM7 的解码方案使用未优化的定点 C 代码,在系统时钟 60 MHz 下解码,每帧需要时间 232 ms,不能满足实时解码需求;基于 DSP 的解码方案系统时钟为 40 MHz,解码每帧需要时间 17.2 ms,满足实时解码要求 24 ms^[13-14]。文中设计通过对音频编解码技术的分析,采用自顶向下的设计思想,提出了一种高效 MPEG-2/4 AAC 音频解码硬件实现方案,并对设计中的控制器、存储设计和软硬件协同工作等关键技术进行详细的介绍,解码电路采用满足 I²S 总线传输要求的时钟。通过仿真、FPGA 验证和流片测试,重建音频质量好,若采样率为 48 kHz,音频解码系统时钟仅为 12.288 MHz,解码一帧需要约

4 结束语

文中主要利用极限学习机在学习过程中无需迭代调整隐层参数,具有学习速度快、泛化能力强等优点,提出一种基于极限学习机的图像压缩算法。与基于 BP 神经网络的图像压缩方法相比,该方法结构简单且易于实现。

实验结果表明,将极限学习机应用于图像压缩可以取得更好的图像压缩重建效果,并且具有较快的学习速度,比基于 BP 神经网络的图像压缩方法快 70 多倍。未来的研究方向在于如何在高压压缩比率下,进一步提高基于极限学习机的图像压缩方法的重建效果。

参考文献:

- [1] 吴乐南. 数据压缩原理与应用[M]. 北京:电子工业出版社,2003.
- [2] Gonzalez R C, Woods R E. 数字图像处理[M]. 北京:电子工业出版社,2005.
- [3] 许 锋,卢建刚,孙优贤. 神经网络在图像处理中的应用[J]. 信息与控制,2003,32(4):344-351.
- [4] Sicuranza G L, Ramponi G, Marsi S. Artificial neural network for image compression[J]. Electronics Letters, 1990, 26(7):477-479.
- [5] Mougeot M, Azencott R, Angeniol B. Image compression with back propagation:improvement of the visual restoration using different cost functions[J]. Neural Networks, 1991, 4(4):467-476.
- [6] Reddy K S N, Vikram B R, Rao L K, et al. Image compression and reconstruction using a new approach by artificial neural network[J]. International Journal of Image Process-

ing, 2012, 6(2):68-85.

- [7] Gaidhane V H, Singh V, Hote Y V, et al. New approaches for image compression using neural network[J]. Journal of Intelligent Learning Systems and Applications, 2011, 3(4):220-229.
- [8] Seiffert U. ANNIE-artificial neural network-based image encoder[J]. Neurocomputing, 2014, 125:229-235.
- [9] 张建宏. 基于混沌神经网络的图像压缩算法[J]. 煤炭技术, 2010, 29(5):167-168.
- [10] 钱海军. 基于 BP 神经网络的图像压缩的 Matlab 实现[J]. 电脑开发与应用, 2011, 24(12):77-79.
- [11] 张 尧,胡 骏,宋美荣,等. 神经网络在彩色图像压缩中的应用[J]. 计算机技术与发展, 2012, 22(11):9-12.
- [12] 张福威,高振亮,李 军. 一种结合改进遗传算法和 BP 神经网络的图像压缩算法[J]. 长春理工大学学报:自然科学版, 2013, 36(6):136-139.
- [13] Huang Guangbin, Zhu Qinyu, Siew C K. Extreme learning machine:theory and applications[J]. Neurocomputing, 2006, 70(1):489-501.
- [14] Huang Guangbin, Chen Lei, Siew C K. Universal approximation using incremental constructive feedforward networks with random hidden nodes[J]. IEEE Trans on Neural Networks, 2006, 17(4):879-892.
- [15] Zong W W, Huang G B. Face recognition based on extreme learning machine[J]. Neurocomputing, 2011, 74(16):2541-2551.
- [16] Yuan Q, Zhou W D, Li S F, et al. Epileptic EEG classification based on extreme learning machine and nonlinear features[J]. Epilepsy Research, 2011, 96(1):29-38.

(上接第 12 页)

20 ms, 完成了智能化、通用化、小型化的实时解码,可移植性强,易集成,取得了较好的效果。

参考文献:

- [1] 张树华, 窦维蓓, 杨华中. MPEG-2/4 AAC 音频编码器的低复杂度优化[J]. 电声技术, 2010, 34(4):71-74.
- [2] Yen C H, Lin Y S, Wu B F. A low-complexity MP3 algorithm that uses a new rate control and a fast dequantization[J]. IEEE Trans on Consumer Electronics, 2005, 51(2):571-579.
- [3] 夏宇闻. Verilog 数字系统设计教程[M]. 北京:北京航空航天大学出版社, 2003:304-305.
- [4] 陈树娟. MPEG-4 AAC 实时编解码系统研究[J]. 煤炭技术, 2011, 30(12):212-213.
- [5] 王 婷, 田 泽, 许宏杰, 等. MPEG-2/4 AAC 音频编码模块的验证[J]. 计算机技术与发展, 2012, 22(7):57-59.
- [6] ISO/IEC 13818-7: MPEG-2 AAC, second edition[S]. [s.

l.]:ISO/IEC, 2003.

- [7] ISO/IEC 14496-3: part 3: audio[S]. [s. l.]:ISO/IEC, 2004.
- [8] Philips Semiconductor. I2S bus specification[S/OL]. 1996. <http://www.Semiconductor.Philips.com>.
- [9] 邵富杰, 张国利, 周 勇. 基于 I2S 总线实现嵌入式语音采集与回放[J]. 微计算机信息, 2011, 27(6):72-74.
- [10] 范飞虎. 基于 FPGA 的 MPEG-2/4 AAC 音频解码器的设计与验证[D]. 西安:西安电子科技大学, 2013.
- [11] 田 靖, 田 泽. AFDX-ES SoC 虚拟仿真平台的构建与应用[J]. 计算机技术与发展, 2010, 20(8):192-194.
- [12] 刘 浩, 田 泽. FC-AV 协议及实现方法研究[J]. 计算机技术与发展, 2012, 22(7):1-4.
- [13] ISO/IEC 11172-4 information technology-coding of moving pictures and associated audio for digital storage media at up to 1.5 Mbits/s[S]. [s. l.]:ISO/IEC, 1993.
- [14] 张轩磊, 窦维蓓, 董 明. 面向 SoC 的数字音频解码系统设计方案[J]. 电声技术, 2008, 32(8):72-76.

一种高效MPEG-2/4 AAC音频解码硬件实现

作者：[王婷](#)，[田泽](#)，[范飞虎](#)，[WANG Ting](#)，[TIAN Ze](#)，[FAN Fei-hu](#)

作者单位：[中国航空计算技术研究所, 陕西 西安, 710119](#)

刊名：[计算机技术与发展](#)

英文刊名：[Computer Technology and Development](#)

年，卷(期)：2015(5)

引用本文格式：[王婷](#).[田泽](#).[范飞虎](#).[WANG Ting](#).[TIAN Ze](#).[FAN Fei-hu](#) 一种高效MPEG-2/4 AAC音频解码硬件实现[期刊论文]-[计算机技术与发展](#) 2015(5)