DESIGN AND IMPLEMENTATION OF MULTI-STANDARD AUDIO DECODER

Kong Ji, Liu Peilin
Department of Electronic Engineering, Shanghai Jiaotong University, Dongchuan Road 800#, Shanghai, China

Deng Ning, Fu Xuan, Zhang Guocheng, He Bin, Liu Qianru
Fujitsu R&D Center Co.,LTD, Shanghai, China
Department of Electronic Engineering, Shanghai Jiaotong University, Shanghai, China

Keywords: Multi-Standard, Algorithm Analysis, IMDCT, Software/Hardware Co-Design, FILTERBANK.

Abstract: In this paper, a design and implementation for Multi-Standard Audio Decoder is presented. The architecture of the decoder is designed to support MPEG-2/MPEG-4 AAC LC Profile (ISO/IEC 13818-7 2006) (ISO/IEC 14496-3 2006), Dolby AC-3 (ATSC 1995), Ogg Vorbis (Xiph.org Foundation 2004), Windows Media Audio (WMA) (Microsoft 2006) and MPEG-1 Layer 3 (MP3) (ISO-IEC/JTC1 SC29 1991). Based on the analysis of algorithms of these multi-standards, software/hardware co-design method is used to implement the audio decoder in which a module called FILTERBANK is designed as a hardware engine. The FILTERBANK which can support IMDCT (Inverse Modified Discrete Cosine Transform) process of different standards is configured by CPU according to the decoded information. Compared with the solutions of DSP/RISC or ASIC multi-standard decoders, our Multi-Standard decoder has achieved a balance between software's flexibility and hardware's high efficiency. Also it meets the requirement of low cost, low power and high audio quality. The implementation results on FPGA are given and the performance of the decoder is evaluated.

1 INTRODUCTION

Nowadays, there are various audio compression standards such as AAC, MP3, AC-3, Vorbis and WMA etc. Besides, new audio standards will be issued in the future. The question is, are today's audio decoders capable of processing so many existing standards and ready for handling the new comers?

Audio decoders are usually implemented on DSP, microprocessor or ASIC. While a solution based on DSP or microprocessor possesses certain flexibility, it entails somewhat higher frequency and higher power, which can hardly meet the increasing demand of low power applications.

On the other hand, an ASIC solution is often characterized by lower frequency, lower power, but with less flexibility. Once new standards are issued or upgrading needed, the whole system will be re-designed.

In order to solve these problems, software hardware co-design method is used in our Multi-Standard decoder, which reduces the consumption of power, area, cost, and meanwhile retaining re-configurability and excellent performance.

Main features of the proposed Multi-Standard Audio Decoder are as follows:

- Low frequency for real-time decoding (required clock frequency < 20MHz for AAC decoding)
- Low power consumption
- Compatibility for AAC, MP3, AC-3, Vorbis and WMA
- Expansibility for new standards or program updating
- Pipeline architecture between software/hardware
- Multi-Standard FILTERBANK (straightforward interface for SOC applications as an IP macro)
- An architecture considering the shared arithmetic modules and memories
- Broad applications: personal audio player, set-top box, audio system on motorcar, PDA/portable terminals and Internet multimedia.

Ji K., Peilin L., Ning D., Xuan F., Guocheng Z., Bin H. and Qianru L. (2007).
DESIGN AND IMPLEMENTATION OF MULTI-STANDARD AUDIO DECODER.
In *Proceedings of the Second International Conference on Signal Processing and Multimedia Applications*, pages 301-304
DOI: 10.5220/0002137803010304

This paper presents the analysis of algorithms of various audio standards, and then introduces the software/hardware co-design method which is used to implement multiple audio standards into a universal architecture. After that the architecture of the Multi-Standard FILTERBANK and its operation flow are presented. At the end, the implementation results on FPGA with the performance evaluation and the conclusion are given.

2 ALGORITHM ANALYSIS & MULTI-STANDARD IMPLEMENTATION

2.1 Algorithm Analysis of Multiple Audio Standards

Most of the existing audio standards are based on Psycho-acoustic Model to achieve high compression ratio and ensure high audio quality. The algorithms of main stream decoders are mainly composed of Lossless Decoding, IQ (inverse quantization) and IT (inverse transform).

Lossless Decoding, usually realized by VLD (variable length decode), is used to decompress the information which is originally compressed without any distortions, and the most popular algorithm of VLD is Huffman Decoding.

The IQ algorithms in various audio standards are mostly nonlinear. For example, in AAC standard, the IQ algorithm is realized by a nonlinear function of 4/3 power.

The IT algorithm is used to transform spectral data to the time domain, and in most audio standards IMDCT is used as the realization method.

2.2 Complexity Analysis

For complexity analysis on various audio standards (include AAC, MP3, AC-3, Vorbis and WMA), some typical standards such as AAC, MP3 and Vorbis are profiled to find out the key operation modules.

The AAC LC decoder for the complexity analysis is mainly composed of IMDCT, Huffman Decoding, Inverse Quantization and some other modules. The profile is based on MetaWare IDE for ARC Version7.4.3, which simulates the ARC600 core. The result of AAC profile is given in Figure 1, which shows that the IMDCT module consumes 47.98% of decoding time.

Besides, Vorbis Standard is profiled on Sun Sparc RISC CPU with a typical Vorbis Decoder referenced by Vorbis Library. The largest time consumption part of Vorbis decoder is IMDCT transform with 34% of total time.

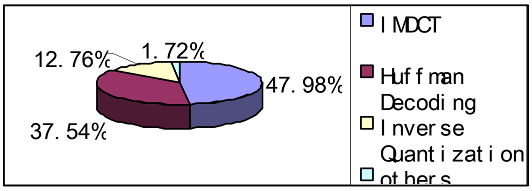

Figure 1: Profile of AAC Decoder.

The conclusions from the algorithm and complexity analysis of multiple audio standards are:

a) The filter-bank module in each standard consumes the largest portion of decoding time.
b) Most of the filter-banks are realized by Radix-2 points IMDCT. (36-point or 12-point IMDCT for MP3).
c) Other arithmetic tools are usually nonlinear or composed of conditional branches, and have significant distinctions among different audio standards.

2.3 Software/Hardware Co-Design Method for the Implementation of Multi-Standard Decoder

The filter-bank modules of various standards are realized by IMDCT, while other modules, which may have significant distinctions among different standards, usually introduce conditional branches or de-multiplex processes. Therefore, these modules are much more suitable to be implemented by CPU. On the other hand, filter-bank modules of different standards are quite similar except for the number of points of IMDCT or the window shape. And filter-bank module is also the key part in each standard. So a common hardware engine implementation will be much more efficient and convenient. As shown in Figure 2, the decoding flows of AAC & MP3 standards are divided into software and hardware. In hardware part, the Multi-Standard FILTERBANK is configured to support different standards (AAC or MP3) by parameter settings. More details about the structure of Multi-Standard FILTERBANK will be given in section 3.

Besides, the software on CPU and the hardware engine are executing in pipeline. Figure 3 shows the decoding process of MP3. After the software on CPU decodes the first granule of a frame (including left and right channel), CPU starts FILTERBANK and continue to decode the second granule of this frame. After software finished the second granule of this frame, it will wait for a completion signal from FILTERBANK to start a new frame.

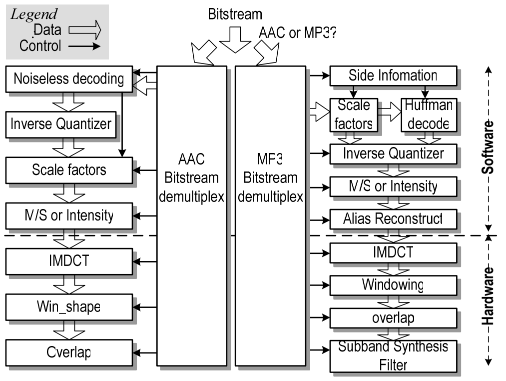

Figure 2: Hardware/Software Co-Design for AAC & MP3.

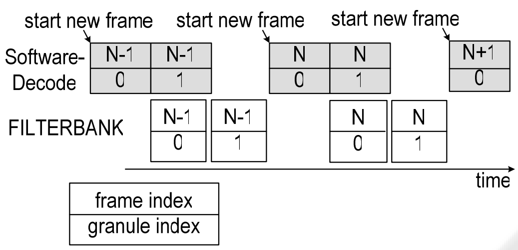

Figure 3: Pipeline of Software/Hardware for MP3.

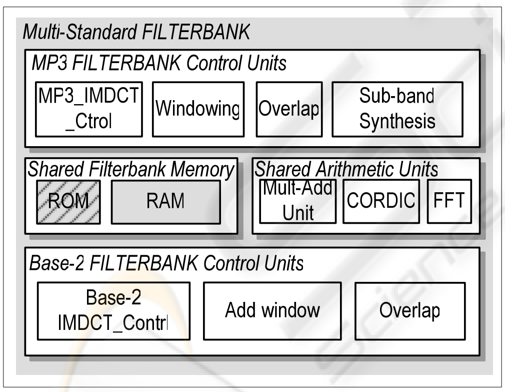

Figure 4: Multi-Standard FILTERBANK Structure.

3 MULTI-STANDARD FILTERBANK

3.1 The Architecture of Multi-Standard FILTERBANK

The architecture of Multi-Standard FILTERBANK is shown in Figure 4. In this architecture, two different strategies, coarse granule sharing and fine granule sharing, are used to share the resources. For Radix-2 points IMDCT, FFT (R. Gluth 1991) module as a coarse granule is used for IMDCT implementation for different standards. On the other hand, fine granule, which is to realize IMDCT with CORDIC (Coordinate Rotation Digital Computer) (Despain 1974), is used for 36-point or 12-point IMDCT in MP3 standard.

The CORDIC module can calculate the FFT coefficients and Sine Window coefficients online without any ROM to store cosine function tables and is free of multipliers, which is a major consumer of resources.

3.2 Operation Flow of Multi-Standard FILTERBANK

The Multi-Standard FILTERBANK works under different modes to decode bit streams for different standards. The operation modes are related to the decoded information such as block length, which can be different among various standards. Table 1 gives the mode to be configured by CPU. Once started by the CPU, the Multi-Standard FILTERBANK receives decoding information and is re-configured to the corresponding mode.

In AAC bit stream, the block length in each frame is constant (1024 or 128), so the FILTERBANK is configured only at the beginning of decoding each frame (including both left and right channel) and sends the 'IMDCT_Complete' signal to CPU after the IMDCT process of the whole frame is done.

In AC-3 mode, FILTERBANK works similarly to AAC mode. But in WMA or Vorbis mode, the current block length may be different from that of pre-block, therefore the FILTERBANK must be configured at the beginning of each block.

Table 1: Operation Modes and Decoded Information of Multi-Standard FILTERBNAK.

Standard	Block Length	Configure FILTERBANK	Window Shape
AAC	1024/128	once/frame	Sine/KBD
MP3	36/12	once/frame	Sine
AC-3	256/128	once/frame	KBD
Vorbis	64-8192	once/block	Sine/KBD
WMA	64-2048	once/block	Sine

4 IMPLEMENTATION RESULTS

The Multi-Standard Audio Decoder is implemented on Altera Stratix FPGA to evaluate the performance. The size of each module is given in Table 2 on the basis of unit of LE (logic element). These data are the synthesis result of Quartus II Version 5.1, and the whole audio decoder uses about 7645 LEs.

Table 2: Size of Each Block.

Module name	Size/LE
CPU & BUS	2156
Bass-2 FILTERBANK Control Units	1669
MP3 FILTERBANK Control Units	1878
Shared ALU	144
Shared CORDIC	1362
Shared FFT	436

To evaluate the performance of our Multi-Standard Audio Decoder, it is set to the operation mode of AAC and MP3 separately, and the minimum real-time decoding frequency is obtained for several sequences. Table 3 shows the average decoding cycles for each AAC frame and the frequency requirement for several sequences, and Table 4 for each MP3 frame. The frequency needed for real-time decoding is calculated by equation (1).

$$\text{frequency\_needed} = \frac{1}{\text{N}}(\text{cycles\_per\_frame} \times \text{sample\_frequency})$$

$$\text{N} = \begin{cases} 1024 & \text{for AAC} \\ 1152 & \text{for MP3} \end{cases} \qquad (1)$$

5 CONCLUSIONS

A design and implementation for Multi-Standard Audio Decoder is presented in this paper. Based on the analysis of algorithms of these audio standards, software/hardware co-design method is used to implement the audio decoder, in which a module called FILTERBANK is designed as a hardware engine. The FILTERBANK, which supports IMDCT process of different standards, is configured by CPU according to the decoded information. Compared with the solutions of pure DSP/RISC or ASIC multi-standard decoders, our multi-standard decoder has achieved a balance between software's flexibility and hardware's high efficiency. Also it meets the requirement of low cost, low power and high audio quality. The presented decoder is capable of real-time decoding for MP3/AAC/AC-3/Vorbis/WMA standards, and is easy to be extended to new standards or updated to new program versions.

Table 3: AAC frequency test results.

Test AAC sequences	Bit rate /kbps	Sample rate /kHz	Cycle per frame	Frequency Needed /MHz
Hot	96	48	169513	8.0
Diamond	288	48	211769	9.3
Love	128	44.1	246046	10.6
Once you	224	44.1	329723	14.2
Mayday	160	44.1	290996	12.6
Liang Zhu	192	44.1	324050	14.0
Believe	320	44.1	372695	16.9

Table 4: MP3 frequency test results.

Test MP3 sequences	Bit rate /kbps	Sample rate /kHz	Cycle per frame	Frequency Needed /MHz
Test case	128	44.1	296205	11.3
Air	192	44.1	310800	11.9
David	192	44.1	335832	12.8
Happy	250	44.1	379191	14.5
Drum	320	44.1	421035	16.1

ACKNOWLEDGEMENTS

This work is supported by Fujitsu RD Center Co.,LTD. and Fujitsu Laboratories LTD.

REFERENCES

ISO/IEC 13818-7, 2006, Generic Coding of Moving Picture and Associated Audio Information C Part7: Advanced Audio Coding

ISO/IEC 14496-3, 2006, Information technology - Coding of audio-visual objects - Subpart 4: General Audio Coding (GA) C AAC, TwinVQ, BSAC

ATSC, 1995, Digital Audio Compression (AC-3) Standard

Xiph.org Foundation, 2004, Vorbis I specification.

Microsoft, 2006, http://www.microsoft.com/windows.

ISO-IEC/JTC1 SC29, 1991, CD 11172-3 Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 MBIT/s Part 3 Audio Contents

R. Gluth, 1991, Regular FFT-related transform kernels for DCT/DST-based polyphase filter banks. In ICASSP

Despain, 1974, A.M., Fourier Transform Computations Using CORDIC Iterations, IEEE Transactions on Computers, Vol.23,1974,pp.993-1001