SGS-THOMSON DSP D950-CORE AUDIO COMPRESSION

A.M. Alvarez, S. George, Yang H. and K. Das
DSP R&D Centre
SGS-Thomson Microelectronics
28 Ang Mo Kio Industrial Park 2
Singapore 569508
Tel: +65 4805726; Fax: +65 4820240; E-mail: antonio.alvarez@st.com

ABSTRACT

Dolby Laboratories provides a detailed flow of operation of the Composite multichannel audio compression algorithm, from receiving a frame in the input data buffer all the way to outputting PCM audio signals, in the form of a C-like floating-point pseudocode. This pseudocode helps the licensees in understanding the details in implementing a real-time AC-3 decoder. A set of bitstreams or test vectors are designed to exercise different aspects of the AC-3 algorithm and are used for design verification. This paper describes how these tools were used in order to implement the AC-3 decoder algorithm on a DSP-core.

1 INTRODUCTION

The C-like pseudocode specifies the decoding process only and suitable for use in digital audio transmission and storage applications. The coded representation may convey from 1 to 5 full bandwidth channels, along with a low frequency enhancement channel. A wide range of encoded nominal bit-rates is supported and varies from 32 to 640 kbps. Two distinctive sections are considered for the implementation: a) bit decoding (stripping) and TDAC (time domain aliasing cancellation) processing. The way how these two sections are related is function of the information available in a synchronisation frame. Each frame contains 6 coded audio blocks, each of which represents 256 new audio samples. The synchronisation information header keeps synchronisation, the bit stream information informs about the transmission mode and options, the auxiliary data field may include dummy data or user data. The field is adjusted by the encoder such that the crc2 element falls on the last 16-bit word of the frame, the cyclic redundancy check (crc1) is checked after 5/8 of the frame has been received and (crc2) is checked after the complete frame has been received.

An audio block contains information of the block splitting into 2 sub-blocks during the transformation from the time domain to the frequency domain. A long transform length is more suitable for input signals whose spectrum remains stationary or quasi-stationary. This provides a greater frequency resolution, improved coding performance and a reduction of computing power required. Two short transform length, possessing greater time resolution, is more desirable for rapidly changing signals with time. The computer power for two short transforms should be higher than for only one transformation.

This approach is very similar to behaviour known to occur in human hearing. The manner this function is performed is not described in this paper but can be obtained from [1].

Dither, dynamic range, coupling function, channel exponents, bit allocation function, gain, channel mantissas and other parameters are also contained in each block. However, they are represented in compressed format, therefore, unpacking, setting-up tables, decoding, expansion, calculations and computations have to be performed before the 256 PCM audio samples can be recognised.

2 DECODER IMPLEMENTATION

The input bit stream will typically come from a transmission (HDTV, CTV) or a storage system (CD, DAT, DVD). This data can be transmitted in a continuous way or in a burst fashion. Two distinctive sections are considered for the implementation: bit decoding (stripping) and TDAC (time domain aliasing cancellation) processing.

The bit decoding section synchronises the frame, stores up to 2/3 of data before start processing, unpacks the synchronisation (SI) and bit stream information (BSI) only once per frame. The audio blocks (AB0, .. AB5) are unpacked one by one; at this moment the length of each block containing 256 new audio samples may not have the same

number of bits. Once the audio blocks are decoded, each audio block will have the same length. The audio block AB0 contains not only the new 256 PCM audio samples but also extra information which concerns the complete frame. The rest of the audio blocks may contain a smaller number of bits.

The bit decoding section performs basically an unpacking and decoding function which final product will be the frequency transform coefficients of each channel involved, in a floating-point format (exponents and mantissas).

The D950-core DSP has been used to strip the bit stream and produce the coefficients block-by-block. The unpacking and bit stripping consumes 26.32 MIPS (million of instruction per second) when 6 input channels are transmitted and compressed and the required output corresponds to 6 channels. There is no different computing power requirement when only two channels are required as output. However, when only two channels are transmitted the computing power required reduces to 11.32 MIPS. The figure 2.1 shows the detailed requirements for the bit decoding section.

Bit Decoding (Stripping)

Performance Measurements

Unpacking & Bit Stripping

- **26.32 MIPS** for 6 channel input/6 channel output
- **26.32 MIPS** for 6 channel input/2 channel output
- **11.32 MIPS** for 2 channel input/2 channel output

Routine	No. of cycles (Manual)	No. of cycles (Music Test Vector)	Program Memory (Words)	Repetition in a Frame
Unpack SI & BSI	506	369	300	1
Set-up Down-Mixing Table	359	281	551	1
Unpack Audio Blocks (AB0-5)	3525	1913	1172	1
Decode Exponents	5540	5165	101	5.1
Expand Delta Allocation	1100	72	62	5.1
Compute Dynamic Range Factors	63	55	247	5
Calculate Gain	98	98	91	5
Compute PSDs & Excitation Function	4907	4536	479	5.1
Compute Masking Curve	700	594	59	5.1
Compute Bit Allocation Values	2202	2109	37	5.1
Decode & Dequantize Mantissas	6706	6625	296	5.1
Denormalize Mantissas	1249	1203	99	5.1
Find Minimum Exponents	271	262	16	5.1
Compute Transform Coefficients	1784	1742	31	5.1
Re-Matrix	4916	0	240	1
Main	5564	5510	976	1
Total	140415.9 (26.32 MIPS)	122608.8 (22.99 MIPS)	4757	

Figure 2.1 Performance Results for the Bit Decoding (Stripping) Section.

TDAC Processing

Performance Measurements

Complex IFFT, Downmixing & Windowing based on double precission arithmetic (32/16)

- **35.24 MIPS** for 6 channel input/6 channel output
- **31.07 MIPS** for 6 channel input/2 channel output
- **11.27 MIPS** for 2 channel input/2 channel output

Routine	No. of Cycles Block Switched	No. of Cycles Non-Block Switched	Program Memory (Words)	Data Memory (Words)
Data Splitting & Bit Reversing	2937	2908	748	1088
1st Complex Multiply	3642	3624		
Complex IFFT	13111	15371		
2nd Complex Multiply	2607	2589		
Downmixing (6 to 2)	7632	7632	138	4608
Windowing Per Channel	5565	5565	179	3072
Total	174804 (32.77 MIPS)	187974 (35.24 MIPS)	1065	8768

Figure 2.2Performance Results for the TDAC Signal Processing Section.

The TDAC processing section first receives the transform coefficients one block at a time. In normal operation, when signals are stationary in nature and have been frequency domain transformed, only 256 coefficients are transmitted. However, the block switch flag is disabled and its value is also transmitted, and the TDAC uses a 512 point inverse fast frequency transform (IFFT) to obtain 512 time domain samples. When fast changing signals are considered, the block switch flag is enabled and frequency domain transformed differently, though 256 coefficients are also transmitted

3 SPECIFIC APPLICATION BOARD DEVELOPMENT

For the AC-3 decoder, a one-chip two-core solution was adopted. In order to evaluate the chip design, a specific application board was developed. Two single D950-core DSPs with 25 nsec instruction cycle time were used.

and the corresponding memory. Three sampling rates are furnished: 32, 44.1 and 48 kb/s. In this board only two output channels are considered for DVD application.

4 DOLBY LABS TEST RESULTS (C-MODEL)

The testing of IC implementations are done at two stages of development. Initially, it is recommended that the chip or code design be modelled by modifying the reference decoder "C" language code, and that the suite of test vectors be used to validate the new reference model. Any differences in output between the Dolby reference decoder and the new implementation should be noted. Minor differences in the PCM level may not cause a problem that is likely to become noticeable in the SNR test otherwise. After verification this "C" model becomes the new implementation reference model and the IC implementation should comply this model bit-by-bit. A test fixture with a sample chip must be provided to Dolby Laboratories Licensing Corporation for evaluation.

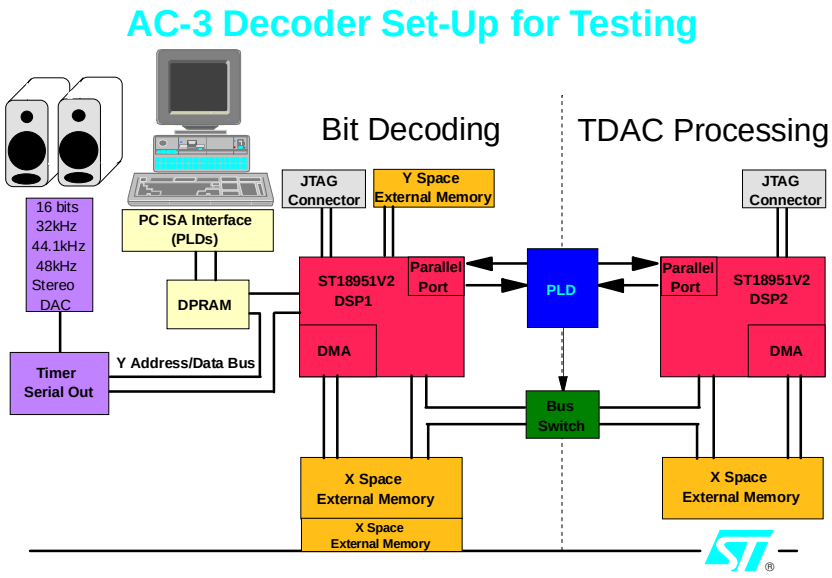

Figure 3.1 System Configuration for AC-3 Decoder (6 input channels - 2 output channels)

In principle, while the board and eventually the chip may acquire AC-3 data in any of several forms, means to allow the device under test (D.U.T.) to receive IEC-958 format AC-3 data must be provided. In addition, while the board/chip may be designed to drive DACs directly, means must be provided for outputting the PCM decoded data in a pair of channels (L and R) in IEC-958 format.

Figure 3.1 shows a block diagram of the AC-3 decoder sent to Dolby Labs for evaluation. It consists of two D950-core DSPs, some glue logic, dual port RAM for data transfer to/from PC, common memory for data transfer between DSPs

The results of the "C" model implementation are presented in figure 4.1 (a) and (b). (a) represents a 20 Hz output for the left channel using Dolby reference double precision decoder; the SNR corresponds to 88.2746 dB. (b) represents SGS-Thomson double precision decoder solution; the SNR is 87.1691 dB, only 1.1055 dB difference. Although no psychoacoustics phenomenon is verified in this test, its performance could be concluded to be very similar to the Dolby solution. Before sending the chip implementation, an application specific board for the AC-3 decoder has been prepared to verify the performance of the

implementation. And finally, with figure 3.1, the chip evaluation will be provided [3].

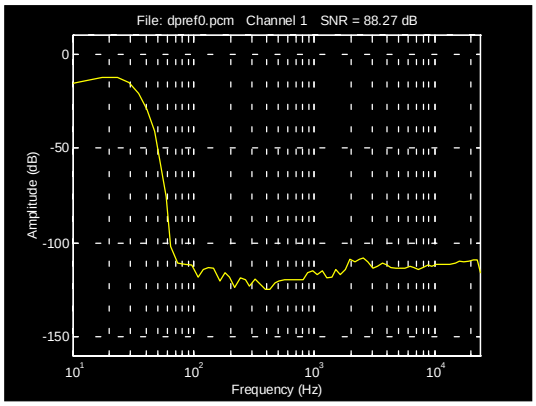

Figure 4.1 (a) Dolby Labs Reference Left Channel.

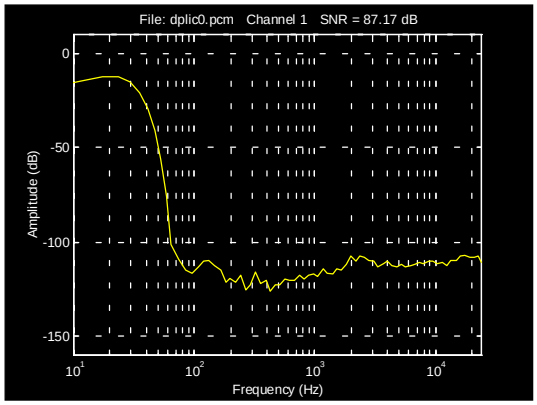

Figure 4.1 (b) SGS-Thomson Left Channel.

6 CONCLUSIONS

The Dolby AC-3 decoder has been developed in C and assembly code and results have been presented. The C-model was evaluated by Mr Uttam Suri from Dolby Laboratories Licensing Corporation at San Francisco, USA and a preliminary classification was provided giving SGS-Thomson Microelectronics enough confidence to continue with the final implementation.

The authors would like to express their gratitude to SGS-Thomson Microelectronics for the permission to present and publish the results of this exciting project.

7 REFERENCES

[1] J.P. Princen and A.B. Bradley, "Analysis/Synthesis Filter Bank Design Based on Time Domain Alaising Cancellation", IEEE Trans Acoustics, Speech and Signal Processing, vol ASSP-34, no. 5, pp 1153-1161, Oct 1986.

[2] M. Bosi, "Independent Channel and Composite Audio Transform Coding: an Overview of AC-2 and AC-3', Procc 6th Tirrenia International Workshop on Digital Communications, Italy, Sept, 1993.

[3] A. Alvarez et al, "DSP D950-Core One Chip Solution for the Dolby AC-3 Decoder Algorithm", DSP^{x} 96 Conference, USA, Mar 1996.