

F-29

# 16bit DSP 用低演算量 MPEG-2/4 AAC ステレオエンコーダの開発

## MPEG-2/4 AAC Stereo Encoder on a 16-bit DSP

高見沢 雄一郎† 奥村 剛† 野村 俊之† 池川 将夫†  
Yuichiro Takamizawa Tsuyoshi Okumura Toshiyuki Nomura Masao Ikekawa

### 1. はじめに

近年、PDA や携帯電話に代表される携帯デジタル機器市場が急激に拡大しつつあり、市場競争が激化しつつある。そのような市場で製品シェアを拡大していくためには他製品との差別化機能の提供が重要であり、オーディオ再生機能もその一つとして注目されている。

オーディオ再生機能は、消費者のみならず製造者にとっても魅力的な機能である。なぜなら、従来から携帯デジタル機器が内蔵している DSP やメモ리카ードスロットなどを利用して低コストに機能追加を実現できるためだ。さらにはオーディオ録音(エンコード)機能も実現できれば理想的であるが、携帯デジタル機器に搭載されている DSP のエンコード処理実現は演算量やメモリ量などの面で困難を伴う。本論文では低コストにエンコード機能の追加を可能とするために開発したメモリ・演算量削減手法と、開発したエンコーダの性能評価結果について述べる。

### 2. AAC エンコード方式

今回開発したオーディオエンコーダは MPEG-2/4 AAC LC Profile[1] (図 1)を圧縮アルゴリズムとして用いた。エンコーダでは入力オーディオ信号にまず MDCT を施して MDCT 係数に変換する。この MDCT 係数は、TNS とステレオ符号化処理によって冗長度が取り除かれた後に量子化される。心理聴覚分析はこの量子化処理を量子化雑音が聴感的に最小となるように制御する。量子化された MDCT 係数はハフマン符号化されてビットストリームに多重化される。

AAC エンコーダを小型 DSP に実装する場合、三つの問題に直面する。一つ目は、その大きな演算量である。エンコーダはデコーダの数倍程度の演算量を必要とする。携帯デジタル機器でエンコード機能をデコード機能並みに普及させるためには、その演算量をデコーダと同程度にまで削減する必要がある。

二つ目は、小型 DSP の内部メモリ容量が小さいことである。外付けメモリの付加はコスト的にも消費電力的にも好ましくなく、内部メモリのみで動作するようにエンコーダの必要メモリ量を削減する必要がある。

三つ目の問題は、オーディオ信号処理で一般的に用いられる 24bit 演算あるいは浮動小数点演算用命令を小型 DSP が持たないことに起因する。広いダイナミックレンジを持つオーディオ信号を小型 16bit DSP で扱えるよう、エンコーダの各処理を注意深く最小演算誤差で 16bit 演算に変換する必要がある。

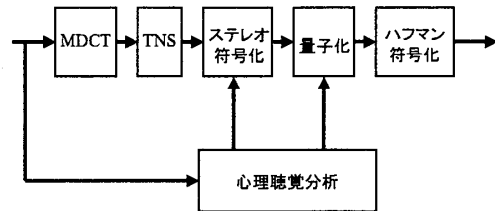


図 1: AAC エンコーダブロック図

次章ではこれらの問題点を解決するために用いた手法を、AAC エンコーダの各ブロックについて述べる。

### 3. 実装手法

#### 3.1 MDCT

AAC では下式に示す 2048 点の MDCT を用いている。

$$mdct[k] = 2 \sum_{n=0}^{2047} pcm[n] win[n] \cos\left(\frac{2\pi}{2048}(n+512.5)(k+0.5)\right)$$

ここで  $mdct[k]$  は MDCT 係数、 $pcm[n]$  は入力 PCM 信号、 $win[n]$  は窓係数である。この計算を小型 DSP に実装するには処理の固定小数点化が必要となる。ここで、全ての値を 32bit 整数として扱えば演算精度は向上するが、32bit 数同士の乗算を多く必要としてしまう。これら 32bit 数同士の乗算は、16bit 数同士の乗算命令しか持たない小型 DSP にとって必要演算量を大きく増加させる要因となる。しかしながら 16bit 整数ではダイナミックレンジが広い MDCT 係数の表現には足りず、音質劣化を招く。開発したエンコーダではこの演算量と音質のトレードオフを考慮し、 $\sin$  や  $\cos$  関数値が主に格納されているテーブル値を 16bit 整数として、その他の値を 32bit 整数として扱うこととした。これにより全ての 32bit 数同士の乗算を、32bit 整数と 16bit 整数の乗算に変換することができた。演算誤差については、簡単な主観音質評価試験によって、本変換による音質劣化が無いことを確認できた。

MDCT では入力オーディオ信号の特性に応じて、より変換長が短いショートブロックを選択することができる。これは、時間的に特性変化の激しい打楽器のようなオーディオ信号を符号化する際に用いられ、2048 点 MDCT の代わりに 256 点 MDCT が用いられる。しかしながら、今回開発したエンコーダでは以下の 2 つの理由からショートブロックを用いず、常に 2048 点 MDCT を用いることとした。

†NEC マルチメディア研究所

†NEC マイクロシステム

一つ目の理由は、データ用メモリ量の増加を防ぐことである。ショートブロックの使用は事前に知る必要があり、そのために常に1フレーム先の信号特性を調べる必要性が生じる。これにより入力オーディオ信号用バッファとして4kバイトが新たに必要となってしまう。

二つ目の理由は、命令用メモリ量の増加を防ぐことである。MDCT、心理聴覚分析、量子化、ステレオ符号化、TNSなどはショートブロック用に別のプログラムコードが必要となり、その結果、必要となる命令用メモリ量が2倍となってしまう。

### 3.2 心理聴覚分析

従来の実装[1],[2]では、入力信号にFFTを施し、FFT係数上でトーン性計算やマスキング計算を行っていた。しかしながら、AACで用いられている2048点MDCTは心理聴覚分析に十分な周波数解像度を持ち、そのために従来のFFTをMDCTで置き換え可能であると考えた[3]。これによりFFTを不要とし、音質を劣化させることなく演算量を削減することができた。

### 3.3 量子化

量子化処理では下式により量子化値(Q)を得る。

$$Q = \text{int}(M \times 2^{(0.25\text{scf})} \times 0.75)$$

すなわち、MDCT係数(M)をスケールファクタ(scf)で増幅し、0.75乗演算により非線形量子化する。DSPは一般にこのようなべき乗計算を簡単に出来ないため、以下の2ステップから計算を行った。ステップ1では

$$M \times 2^{(0.25\text{scf})}$$

について、指数部(0.25scf)をその整数部であるexponentと、小数部であるmantissaに分割し、

$$M \times 2^{\text{mantissa}} \times 2^{\text{exponent}}$$

と表す。

ステップ2では上式の0.75乗である以下の計算をする。

$$(M \times 2^{\text{mantissa}})^{0.75} \times 2^{0.75\text{exponent}}$$

式右半分( $2^{0.75\text{exponent}}$ )は再度その指数部(0.75exponent)を整数部(exponent')と小数部(mantissa')に分割し、

$$2^{\text{exponent}'} \times 2^{\text{mantissa}'}$$

と表す。式左半分は近似計算によって算出され、この近似計算結果に( $2^{\text{mantissa}'}$ )を乗じ、exponent'ビットシフトすることで量子化値を得ることができる。

### 3.4 ハフマン符号化

ハフマン符号テーブルが必要とするメモリ量を削減するために、ハフマン符号長とハフマン符号を一つの16bit値にまとめてメモリに格納している。ここで、ハフマン符号長は5bit、ハフマン符号は11bitとした。ハフマン符号が11bitを越える場合はハフマン符号の先頭に連続する'1'を取り除きメモリに格納している。

## 4. 性能評価

これまで述べてきた様々な手法を用いてAACエンコーダを2種類の低消費電力16bit DSP ([4], 表1)に実装した。開発したエンコーダが必要とするメモリ量(ワード数)は命令RAMが4,712、データRAMが13,011、データROMが7,970であり、外部メモリを必要とせずに動作する。

表2にCDクォリティのリアルタイムエンコードに必要な演算量を示す。必要演算量はサンプリングレートとビットレートに比例して増加するが、CDのエンコードで一般的に用いられるであろう44.1kHz、128kbps/stereoという条件下では37.3MIPS、消費電力にして20mW( $\mu$ PD77213)でCDクォリティのリアルタイムエンコードが可能である。

## 5. おわりに

小型低消費電力16bit DSPに適したAACエンコーダを開発した。開発したエンコーダは37.3MIPS、20mWでCDクォリティのリアルタイムエンコードを外部メモリ無しに実現できる。これにより携帯デジタル機器において、低コスト、低消費電力なAACエンコーダの実現が可能となる。

表1: DSP仕様

	$\mu$ PD77115	$\mu$ PD77213
Max freq. [Mhz]	75	120
命令 RAM	11.5k x 32b	15.5k x 32b
命令 ROM	-	64k x 32b
データ RAM	32k x 16b	36k x 16b
データ ROM	-	64k x 16b
コア電圧 (V)	2.0~2.7	1.5
パッケージ	80 pin FBGA (9mm <sup>2</sup> ) 80 pin QFP (12mm <sup>2</sup> )	161 pin FBGA(10mm <sup>2</sup> ) 144 pin LQFP(20mm <sup>2</sup> )
消費電力	1.25 [mW/MIPS]	0.5 [mW/MIPS]
プロセス	0.25 $\mu$ m	0.13 $\mu$ m
周辺回路	Audio Serial IF SD Memory Card IF	

表2: リアルタイムエンコードに必要な演算量

サンプリング 周波数 [kHz]	ビットレート [kbps/stereo]	必要 MIPS 数 [MIPS]
32	64	23.9
	96	27.7
	128	28.3
44.1	64	29.6
	96	33.1
	128	37.3

## 参考文献

- [1] ISO/IEC 14496-3, "Information technology - Coding of audio-visual objects - Part 3: Audio," Dec. 1999.
- [2] J. D. Johnston, "Transform Coding of Audio Signals Using Perceptual Noise Criteria," IEEE Journal on Selected Area in Communications, Vol. 6, No. 2, Feb. 1988.
- [3] T. Nomura, and Y. Takamizawa, "Processor-Efficient Implementation of a High Quality MPEG-2 AAC Encoder," 110th AES Convention, May 2001.
- [4] <http://www.ic.nec.co.jp/dsp/english/index.html>