

HD DVD に用いる動画像符号化技術

Video Coding Technologies for HD DVD-Video

山影 朋夫

■ YAMAKAGE Tomoo

中條 健

■ CHUJOH Takeshi

古藤 晋一郎

■ KOTO Shinichiro

HD DVD ビデオの規格では、MPEG-2 (Moving Picture Experts Group-phase 2) に加え、MPEG-4 AVC (Advanced Video Coding) /H.264 及び VC-1 という新しい動画像符号化技術 (Advanced CODEC (COder-DECoder)) が採用された。これらの新しい符号化技術と HD DVD ディスクを組み合わせることで、デジタル放送の画質をしのぐ映像コンテンツを提供することが可能になる。一方、これらの新しい符号化技術は符号化時の処理量が増加するため、実用化するうえで処理量を削減することが必要である。

東芝は、MPEG-4 AVC/H.264 の処理量削減技術として、高速モード判定技術及び高速符号量推定技術を開発した。

Advanced video coding technologies such as MPEG-4 AVC/H. 264 and VC-1 have been adopted as HD DVD-Video specifications in addition to MPEG-2. The combination of advanced video coding technologies and HD DVD discs makes it possible to provide superior picture quality compared to existing digital broadcasting. However, because of the increase in computational complexity, especially for encoding, it is essential to reduce such computational complexity in implementing the encoder.

Toshiba has developed a fast coding mode decision method and a fast bit-rate estimation method, which achieve complexity reduction for MPEG-4 AVC/H. 264 encoding.

1 まえがき

DVD フォーラムでは2004年6月に、HD DVD ビデオ規格に用いる必須の動画像符号化技術として、現行DVDで採用しているMPEG-2 (Moving Picture Experts Group-phase 2) に加え、新しい符号化技術 (Advanced CODEC (COder-DECoder)) であるMPEG-4 AVC (Advanced Video Coding) /H.264 及び VC-1 を正式に承認した。VC-1 は SMPTE (Society of Motion Picture and Television Engineers)⁽¹⁾ で規格化が行われている。HD DVD ディスクの制作者にとっては、前述の3種類から自由に方式を選んで使うことが可能であり、再生装置にとっては3種類すべての方式の再生をサポートすることが義務づけられる。

ここでは、Advanced CODEC を採用することのメリットを述べるとともに、その一つである MPEG-4 AVC/H.264 の規格の概要及び研究開発動向について述べる。

2 Advanced CODEC のメリット

HD DVD ビデオ規格は、DVD フォーラムで承認した物理規格にのったディスクに映像データを記録するための論理 (アプリケーション) 規格である。赤色レーザーを用いた物理規格として4.7 Gバイト (片面1層) と8.5 Gバイト (片面2層) が、また、青色レーザーを用いた物理規格として15 Gバイト (片

表1. DVD/HD DVD ディスク1枚への映像の記録時間

Recording time of audiovisual data on DVD/HD DVD

| 平均ビット転送レート (Mbps) | | 記録時間 (h) | | | |
|-------------------|---------|----------|----------|---------|---------|
| | | DVD | | HD DVD | |
| ビデオ | オーディオなど | 片面1層 | 片面2層 | 片面1層 | 片面2層 |
| 1 | 0.5 | 4.7 Gバイト | 8.5 Gバイト | 15 Gバイト | 30 Gバイト |
| 8 | 0.5 | 7 | 13 | 24 | 47 |
| 12 | 1 | 1.3 | 2.4 | 4.2 | 8.4 |
| | | 0.9 | 1.6 | 2.8 | 5.5 |

面1層) と30 Gバイト (片面2層) が存在し、これらのディスクに、高精細 (HD) 映像だけでなく、標準解像度 (SD) の映像も記録することを想定している。

符号化映像の平均ビット転送レート (以下、レートと略記) とディスクへの記録時間の関係を表1に示す。一般に、ハリウッドで制作される映画の90%以上が132 min以下の長さで、半分が90 min程度と言われている。DVDの特長として、ボーナスコンテンツを提供できることが挙げられるが、本編とボーナスコンテンツを合わせて1枚のディスクで提供できることが重要である。

Advanced CODEC の符号化効率は、MPEG-2 の2~3倍である (この特集のp.2~8参照)。したがって、Advanced CODEC の12 Mbps は、15 Gバイト又は30 Gバイトのディスク1枚で、BS デジタル放送 (HD 映像) の画質をしのぐ映像を

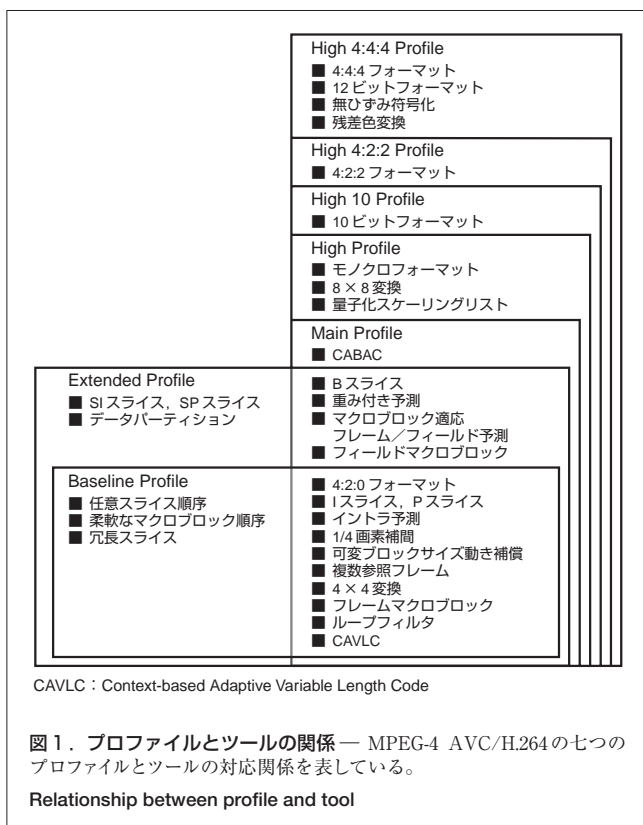
提供できることになる。一方、Advanced CODECの8Mbpsは、地上デジタル放送(HD映像)と同等以上の画質となり、短編映画であれば8.5Gバイトのディスク1枚で提供できることになる。更に、SD映像では1Mbps(MPEG-2の2~3Mbps相当)で、30Gバイトのディスク1枚に47hもの長時間の映像を記録することができるようになり、連続ドラマ及びそのメイキングビデオを1本のディスクに収録することが可能になる。

3 MPEG-4 AVC/H.264

MPEG-4 AVC/H.264は、当初、ITU-T(国際電気通信連合-電気通信標準化部門)のH.26Lプロジェクトがスタートであり、途中からMPEGと共同でJVT(Joint Video Team)を作り、規格化された。したがって、双方の組織の規格名を合わせて持っている。この規格では、図1のように、多くの符号化ツールを目的に応じて組み合わせ、プロファイルを定義している。

当初策定されたプロファイルは、Baseline, Main, Extendedの三つである。それぞれ、Baseline Profileはテレビ電話や携帯電話を、Main Profileは放送蓄積メディアを、また、Extended ProfileはIP(Internet Protocol)ストリーミングを想定している。

Main Profileでは、Baseline Profileの誤り耐性関係の



ツール以外と、Bスライス、重み付き予測(Weighted Prediction)、インターレース関係のツール、コンテキストベースの適応算術符号化であるCABAC(Context-based Adaptive Binary Arithmetic Coding)が追加される。

この中で、重み付き予測は、Pスライス、Bスライスの動き補償予測のときに用いられる手法で、通常は、過去に符号化した画像を予測に用いるのに対して、この手法は、画素信号値に重みを付ける手法で、映画などのコンテンツで多用されるフェードやディゾルブといった、映像表現で有効な手法である。重み付き予測には、重み係数を明示的にデコーダ側に送る明示的(Explicit)モードと、参照画面の時間的な位置関係から重み係数を計算する暗黙的(Implicit)モードが存在する。東芝は、このツールに関するいくつかの提案を行い、規格策定に貢献している⁽²⁾。

2003年3月に最終規格草案(FDIS)を発行した後は、当初、スタジオ機器などのプロフェッショナルな用途向けの色フォーマットへの規格拡張作業を行っていた。しかし、途中から目的を修正して、Main Profileを包含して新たな符号化ツールを加えるという形で、Fidelity Range Extensionsとして、2004年7月に最終補正草案(FDAM)を発行した。ここでは、High, High 10, High 4:2:2, High 4:4:4という新しいプロファイルが四つ定義された。

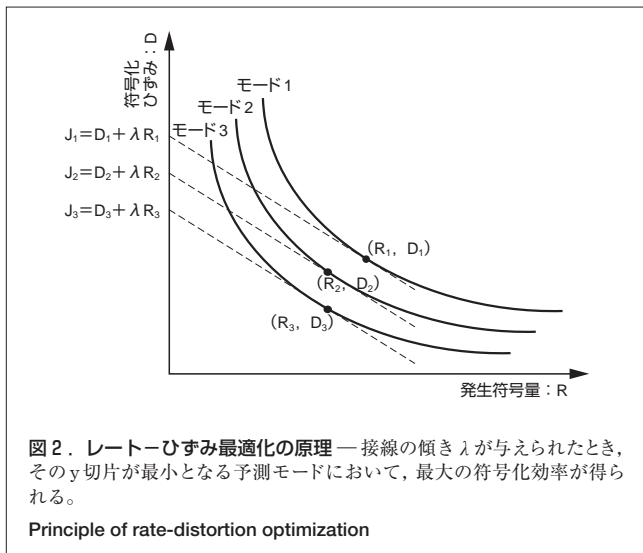
この四つのプロファイルは、図1に示すように互いに包含関係になっている。新たに追加された技術として特に重要なものは、High Profileに入っている8×8変換と量子化スケールリングリストである。8×8変換は、これまでの4×4変換に加えて適応的に用いることで、符号化効率の改善に寄与している。量子化スケールリングリストとは、JPEG(Joint Photographic Experts Group)やMPEGでこれまで用いられてきた量子化マトリクスのことである。特にこの二つのツールは、高精細テレビ(HDTV)の映画素材における符号化効率とグレイノイズの再現性の向上に寄与している。

4 MPEG-4 AVC/H.264 高画質符号化技術

4.1 レート-ひずみ最適化

MPEG-4 AVC/H.264では、画像ブロック単位に多数の予測モードの中から最適なモードを選択して、予測符号化される。最適な予測モードは、画像ブロックの性質によってそれぞれ異なり、適切なモードを選択することで、高い符号化効率を得ることができる。

複数の予測モードから最適なモードを選択する手法として、ラグランジュの未定乗数法に基づくレート-ひずみ最適化手法⁽³⁾が知られている。レート-ひずみ最適化手法では、予測モード*i*で符号化した際の符号化ひずみ D_i (原画像と再生画像との2乗誤差)と、発生符号量 R_i から計算される



符号化コスト J_i が最小となるように、予測モードの選択をする。符号化コストは(1)式で計算される。 λ は、量子化パラメータに基づいて決定される定数である。

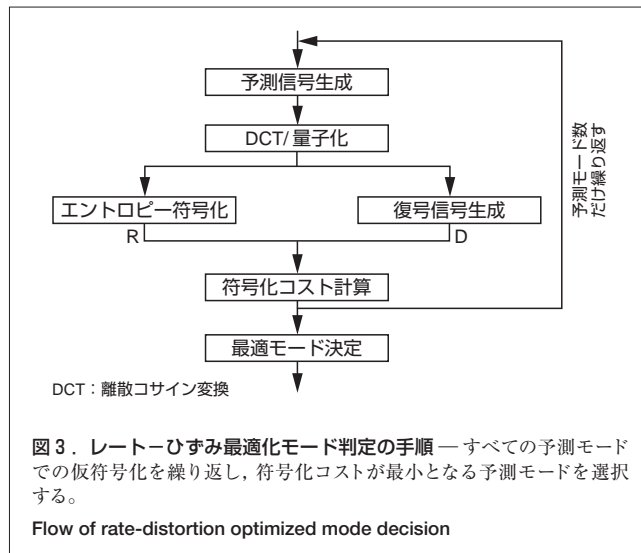
$$J_i = D_i + \lambda \cdot R_i \quad (1)$$

レート-ひずみ最適化の原理を模式的に図2に示す。符号化ビットレートを調整するための量子化パラメータを変化させると、それに応じて画質が変動する。発生符号量と符号化ひずみとの関係(以下、R-Dカーブと略記)は、下に凸の単調減少関数となる。

R-Dカーブは、符号化対象の画像の性質や予測モードによって、それぞれ異なる特性を示す。ある画像ブロックに対して、与えられた量子化パラメータを用いて、予測モード1から3で符号化した場合の発生符号量と符号化ひずみを、それぞれ (R_i, D_i) , $1 \leq i \leq 3$ として計測し、(1)式から符号化コスト J_i , $1 \leq i \leq 3$ を得る。 λ は、R-Dカーブの接線の傾きの大きさに相当し、符号化コストはR-Dカーブの接線のy切片に相当する。

画像符号化での符号化効率とは、いかに少ない符号量で、いかに画質劣化の少ない符号化ができるかということの意味する。つまり、符号化コスト J_i が小さいほど符号化効率が高くなる。したがって、与えられた量子化パラメータに対して J_i が最小となる予測モード(図2の (R_3, D_3))を画像ブロックごとに選択することで、符号化効率を最大化することができる。

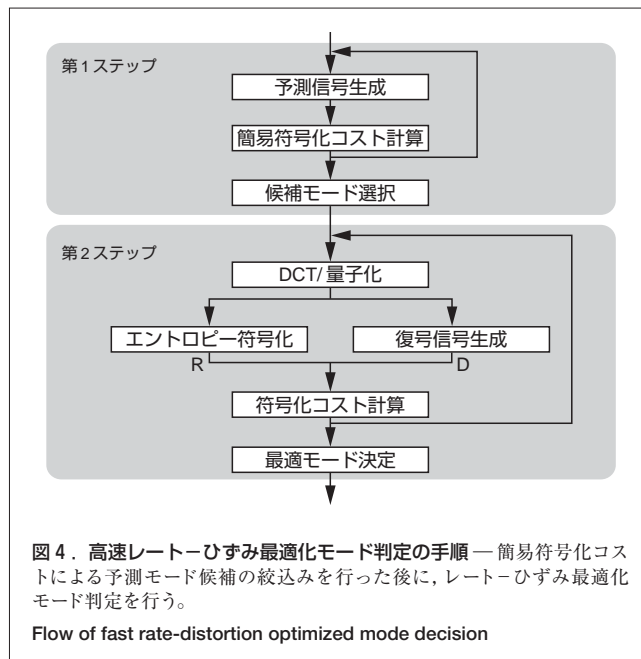
レート-ひずみ最適化手法では、図3に示すように、画像ブロックごとに、取り得るすべての予測モードでの仮符号化及び復号化を行って、発生符号量 R 及び符号化ひずみ D を計測する必要がある。したがって、MPEG-4 AVC/H.264のように予測モード数が多数あると、符号化時に膨大な演算量が必要となる。



4.2 高速モード判定技術

レート-ひずみ最適化手法による符号化効率の高さを維持しつつ、符号化時の演算量を大幅に削減することが、MPEG-4 AVC/H.264の高画質リアルタイム符号化を実現するための重要な課題となる。当社は、適応的かつ階層的にモード判定手法を切り替えることで、符号化効率を落とさずに高速に符号化することができる、階層モード判定手法⁽⁴⁾を提案している。また、更なるモード判定の高速化のために、高精度な発生符号量推定手法⁽⁵⁾を提案している。

4.2.1 階層モード判定による高速化 図4に示すように、まず画像ブロックごとに取り得る予測モードすべての予測信号を生成し、予測誤差信号の大きさなどの評価尺度に基づく簡易符号化コストを計算し、候補となる予測モード



を選択する。候補数は、画像ブロックごとに可変とし、簡易符号化コスト値のばらつきや量子化パラメータなどの指標に基づいて、候補中に最適な予測モードが含まれる確率が十分高くなるように適応的に決定する。次に、選択された候補の中から、レート-ひずみ最適化手法により最適な予測モードを決定する。

4.2.2 発生符号量推定による高速化 MPEG-4 AVC/H.264のCABAC符号化器では図5に示すように、まず、符号化すべき多値のシンタックスデータエレメントを、可変長の2値化シンボル系列に変換する。次に、符号化すべきデータの周囲の情報や既に符号化された2値化シンボル値に基づいて、2値化シンボル1ビットごとにコンテキスト計算を行う。そして、計算されたコンテキスト情報に応じて、算術符号化の確率テーブルを選択し算術符号化処理を行うとともに、確率テーブルの更新を行う。

このように、CABAC符号化では、2値化シンボル1ビットごとに複雑な処理が必要となり、また、現ビットの符号化が終わらないと次のビットの符号化に進めないため、並列化も困難である。

レート-ひずみ最適化モード判定では、仮符号化でのCABACの発生符号量計測の処理量が大半を占めてしまう。そこで、図6に示すように、仮符号化時は実際の算術符号化は行わずに、2値化シンボルのシンボル長の総和から算術符号化の発生符号量を推定することで、仮符号化での演算量が大幅に削減される。

予測モード*i*での発生符号量推定値 \hat{R}_i は、2値化シンボル長 N_i と、直前の符号化済み画像の2値化シンボル長と発生符号量との比率 a を用いて、(2)式に示す1次近似式を用いて推定する。

$$R_i = a \cdot N_i \quad (2)$$

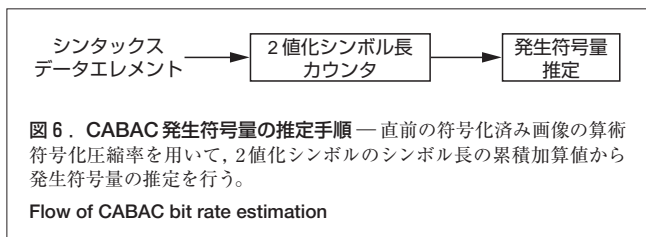
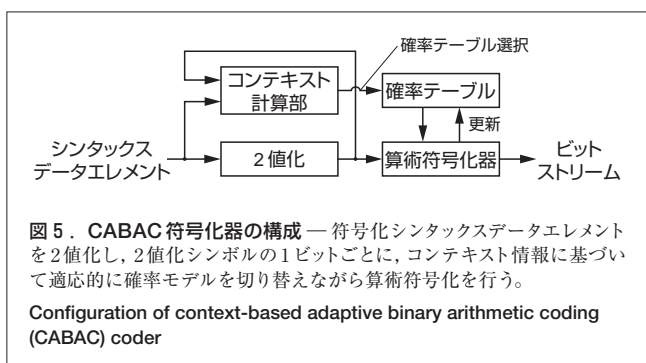
以上の高速化手法の組合せにより、レート-ひずみ最適化モード判定と同等の性能を維持しつつ、モード判定の処理時間を従来の1/10~1/4程度まで短縮した。

5 あとがき

HD DVDを支える物理規格と論理規格の両方が、適切なタイミングで実用化されることは、コストパフォーマンスの高い製品を開発するうえで不可欠である。MPEG-4 AVC/H.264に対する処理量削減による符号化処理の高速化への取組みは、HD DVDの成功に貢献できるものと考えている。

文献

- (1) SMPTE org. <http://www.smpte.org/>, (accessed 2004-12-02).
- (2) Koto, S., et al. "Adaptive bi-predictive of video coding using temporal extrapolation". Proceedings IEEE International Conference on Image Processing 2003 (ICIP2003). Spain, 2003-09, IEEE Signal Processing Society. 2003, p.829 - 832.
- (3) Wiegand, T. Rate-Constrained Coder Control and Comparison of Video Coding Standards. IEEE Trans. Circuits Syst. Video Technol. **13**, 2003, p.688 - 703.
- (4) Tanizawa, A., et al. "A Study on Fast Rate-Distortion Optimized Coding Mode decision for H.264". Proceedings of IEEE International Conference on Image Processing 2004 (ICIP2004). Singapore, 2004-10, IEEE Signal Processing Society. 2004, p.793 - 796.
- (5) 浅野 渉, ほか. "CABAC圧縮率推定を用いたH.264におけるレート-歪み最適化モード判定の高速化". 第3回情報科学技術フォーラム(FIT2004)講演論文集. 京都, 2004-09, (社)電子情報通信学会. 東京, 2004, p.233 - 234.



山影 朋夫 YAMAKAGE Tomoo

研究開発センター マルチメディアラボラトリー主任研究員。動画画像符号化の研究・開発に従事。電子情報通信学会会員。Multimedia Lab.



中條 健 CHUJOH Takeshi, D.Eng.

研究開発センター マルチメディアラボラトリー研究主務、工博。動画画像符号化に関する研究・開発に従事。電子情報通信学会、映像情報メディア学会、IEEE会員。Multimedia Lab.



古藤 晋一郎 KOTO Shinichiro

研究開発センター マルチメディアラボラトリー研究主務。動画画像符号化の研究・開発に従事。映像情報メディア学会会員。Multimedia Lab.