

知っておきたいキーワード

H.264って何ですか

正会員 八島由幸†

† 日本電信電話株式会社 NTTサイバースペース研究所

"What's H.264?" by Yoshiyuki Yashima (NTT Cyber Space Laboratories, Yokosuka)

キーワード：H.264, MPEG-4 part10 AVC, 画像符号化

H.264の位置づけ

H.264は動画像圧縮符号化国際標準方式の一つです。ここで、「H.264」という呼び方はITU-T標準として付与された名称であって、同じ内容のものがISOにおいて「MPEG-4 part10 AVC」として標準化されています。図1はH.264標準化の経緯を示しています。

動画像符号化のルーツは、テレビ電話やテレビ会議向けに標準化されたH.261です。MC（動き補償）とDCT（離散コサイン変換）を組合せて圧縮する枠組みがH.261で確立されました。そしてこのMC+DCT方式は、以降の

MPEG-2, MPEG-4, H.263にも継承されました。MPEG-2では、放送、通信、蓄積に汎用的に利用できるような拡張がなされ、またMPEG-4ではオブジェクト単位で符号化できる新しい枠組みが取り入れられています。そしてH.264は、MPEG-2や従来のMPEG-4（MPEG-4 Part-2, Simple ProfileやAdvanced Simple Profileが相当する）に比べて2倍以上の効率を狙った超高压縮をターゲットとして標準化が進められました。標準化はITU-T（SG16/VCEG）とISO（SC29/WG11（MPEG））の共同検討チームJVT（Joint Video Team）で進められ2003

年に勧告されています。さらにその拡張方式として、4:2:2や10ビット映像といったプロフェッショナル向けを指向する、より高い機能を備えたH.264 FExt（Fidelity Range Extension）が標準化されました。H.264でもこれまでと同様に基本的にはMC+DCTの枠組みが使われています。図1にはこれまでの各標準方式に取り入れられた技術の変遷を示していますが、H.264では基本的枠組みは変えずに、種々の最適化を図れるしくみを導入しているところが特徴です。

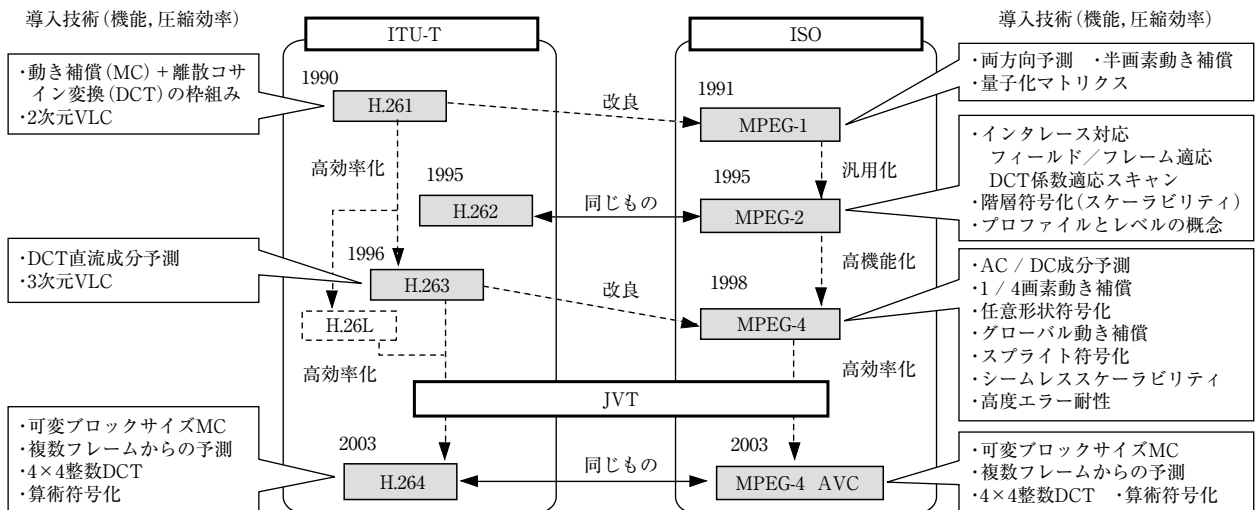


図1 H.264の位置づけ

H.264への新規導入技術

H.264に新しく導入されている技術は、要素技術という観点から見ると必ずしも目新しいものではありません。映像符号化標準化の過程では、ある要素技術を導入するかどうかを判断するとき、符号化効率とともにその要素技術を実行するための計算の複雑さも重要なポイントとなります。従来は、計算機のCPU速度やLSI/コーデック装置実現性などを考慮して、圧縮効率は向上してもあまり複雑な方式は採用を見送ってきたという経緯がありました。一方H.264では、計算をできるだけ複雑にしないという条件をある程度

緩やかにして、符号化効率の向上する技術はどんどん取り入れられました。「デコードできる道具はたくさん用意したからエンコーダは上手に利用してね」という設計思想です。符号化ソフトウェアやハードウェア装置設計者（特にエンコーダ設計者）から見れば、道具（符号化ツール）選択の余地が多く、どのように最適化すればよいかといううれしい難題をふっかけられたこととなります。H.264では非常に多くの拡張がなされています。表1にH.261、MPEG-2との主な技術の比較を示します。詳細は文献1)などに譲るとして、ここではいくつかのキーポイントだけを説明します。

- (1) フレーム内予測：MPEG-2など従来は画素値そのものをDCTしていたものを、フレーム内予測してからDCTすることにしました(図2)。予測モードは13種類、従来符号量が大きかった1フレームの符号量が劇的に減ります。
- (2) 可変ブロックサイズ動き補償：従来は基本的に16x16のブロックサイズごとに動きを検出したものを、表1に示す7種類のサイズを適応的に用いることができるようにしました(図3)。
- (3) 予測参照フレーム：過去未来の任意の2枚からの予測が可能です(図4)。
- (4) 整数DCT：演算が簡単で、かつミスマッチ対策がいらぬなどの利点があります。
- (5) 算術符号：あらかじめ決められた可変長符号ではなく、信号の発生確率に随時適応するため、映像の種類や符号化パラメータが変化しても、それに追従した最適に近い符号割当てが可能となります。

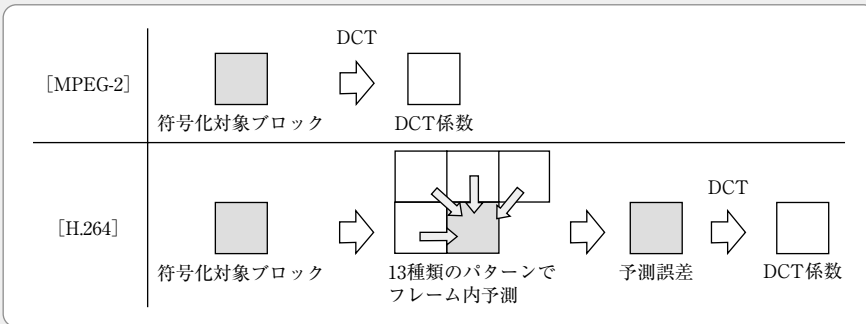


図2 フレーム内予測

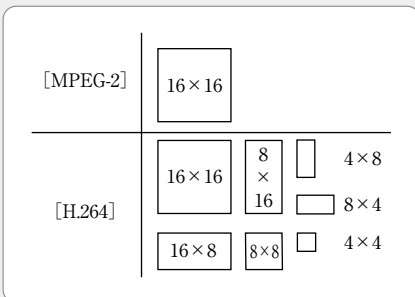


図3 可変ブロックサイズ動き補償

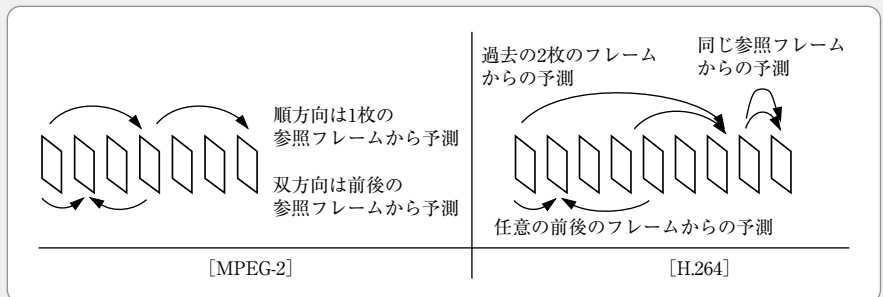


図4 参照フレーム選択の拡張

表1 標準化に採用されている符号化ツールの比較

比較項目	H.261	MPEG-2 Main Profile	H.264/AVC Main Profile	H.264/AVC FRExt
入力信号	プログレッシブ	インタレース	インタレース	インタレース
色信号	4:2:0	4:2:0	4:2:0	4:2:0, 4:2:2, 4:4:4
フレーム内符号化	DCT	DCT	フレーム内予測 + 整数DCT	フレーム内予測 + 整数DCT
フレーム間符号化	フレーム間予測 + DCT	フレーム間予測 + DCT	フレーム間予測 + 整数DCT	フレーム間予測 + 整数DCT
フレーム間予測	順方向	双方向	双方向	双方向
予測参照フレーム	過去1枚	過去未来各1枚	任意の2枚	任意の2枚
重みつき予測	なし	なし	あり	あり
動き補償ブロックサイズ	16x16	16x16	16x16, 8x16, 16x8, 8x8, 4x8, 8x4, 4x4	16x16, 8x16, 16x8, 8x8, 4x8, 8x4, 4x4
動き補償精度	整数画素	1/2画素	1/4画素	1/4画素
小数画素作成フィルタ	なし	2タップ	6タップ, 2タップ	6タップ, 2タップ
動きベクトル符号化	左との差分	左との差分	近傍三つの中央値差分	近傍三つの中央値差分
DCTブロックサイズ	8x8	8x8	4x4	4x4, 8x8
DCTミスマッチ対策	代表値奇数化	最高次係数符号反転	必要なし	必要なし
符号割当て	Fixed VLC	Fixed VLC	CAVLC, CABAC	CAVLC, CABAC
ループ内フィルタ	あり (ON/OFF制御)	なし	あり (フィルタ強度適応)	あり (フィルタ強度適応)

H.264の符号化性能

H.264はどのくらいの符号化効率を達成できるのでしょうか。MPEG-2とH.264を実際の動画像を使って試してみました。MPEG-2 TM5およびH.264 JM10.1といういずれも標準化テストモデルソフトウェアを利用しました。JMでは高速動き検出モードを使用しています。図5に、横軸をビットレート、縦軸を画質としてその結果を示します。画質はSNR比で比較しました。図5から、同じ画質を達成するためのビットレートが、H.264はMPEG-2に比べると低レートでは7割～5割程度で済んでいることがわかります。ただし、TMやJMというのはあくまでもモデルであって、工夫の余地がたくさん残されていることには注意しなければなりません。

H.264の最大性能を引き出すときには、RDO (Rate-Distortion

Optimization) という概念が重要になります。符号化の過程では、16×16単位のマクロブロックで、イントラ/インタ判定、フレーム内予測パターン、動きベクトル、ブロック分割情報などをどのように決めるかという「符号化モード」が決められます。ある符号化モードで符号化したときの歪み量をD、必要となる情報量をRとします。RDOとは、いろいろな符号化モードで符号化してみて評価値 $J=D+\lambda R$ を最小にするような符号化モードを選択することを意味します。ここで、通常Dは原信号と復号信号との二乗誤差で定義されます。また、RはDCT係数の符号化に必要となる情報量のほか、動きベクトルや符号化モードの符号化に必要となる情報量もすべて含みます。 λ は未定乗数であり、例えばJMでは、量子化の粗さの関数として定められています。RDOの概念はH.264に限ったものではありませんが、H.264では先

に述べたように、選択できる符号化モードがMPEG-2に比べて格段に多くなっているため、RDOにより最適値が得やすくなるという特徴があります。なお、すべての符号化モードに対してRDOを行うことが理想ですが、演算量が膨大になってしまいます。このため、限られたモードのみで最適化する方法や、符号化をいちいちやるのではなく最適モードをうまく予測する方法の考案が重要となります。歪みDを二乗誤差だけでなく、人間の視覚特性パラメータを入れて定義することも重要です。

アプリケーションによって最適化の方法や作りが違ってきますので一概には言えませんが、高品質映像サービスに適用する場合、標準テレビで2～3Mbps、HDTVで6～10Mbpsというところがターゲットレートになってくるでしょう。

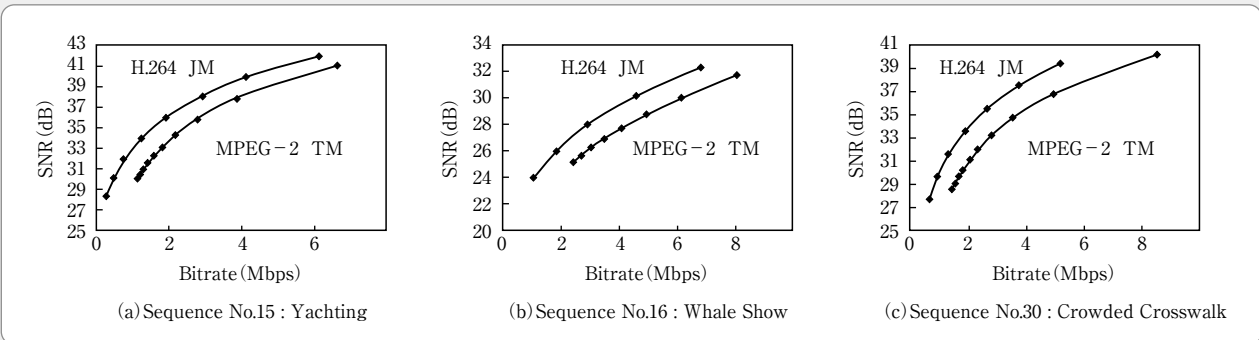


図5 H.264の符号化性能 (映像情報メディア学会標準テスト画像利用)

H.264の使われ方と今後

符号化効率が非常に良いことから産業界でもH.264の利用が広がってきました。2006年4月からはじまった、携帯・移動体機器向けデジタル放送(ワンセグ放送)の映像符号化方式としてH.264が採用されています。また、次世代DVDでも、HD-DVDおよびブルーレイいずれもH.264が搭載必須コ

ーデックの一つとなっています。また、IPTV、すなわちIPネットワークを利用した放送・映像配信のシステム構築においても、ネットワークリソースの効率的利用の観点からH.264の利用が検討されています。

さらに、今後の映像符号化標準化にも影響を与えています。MPEGで検討されている、時空間およびSNRスケラビリティを同時に実現する新しい

スケーラブル符号化SVCでは、そのベースレイヤ符号化方式としてH.264が使われます。また、マルチカメラからの多視線映像を効率的に符号化するMultiview Video Codingでも基準画像はH.264とする方向で標準化が議論されています。今後の映像サービスを支える映像符号化キーテクノロジーとして、H.264は確たるポジショニングを築きつつあります。

参考文献

- 1) 大久保榮監修：H.264/AVC教科書，インプレス



やしま よしゆき
八島 由幸 1983年、名古屋大大学院工学研究科電子工学専攻修士課程修了。同年、日本電信電話公社(現NTT)入社。以来、画像信号の高効率符号化、MPEG関連システムの研究開発に従事。現在、NTTサイバースペース研究所画像メディア通信プロジェクト映像符号化技術グループリーダー、主幹研究員、工学博士、正会員。