

■2群（画像・音・言語）-5編（画像符号化）

11章 国際標準

■2群 - 5編 - 11章

11-1 MPEG/JPEG/H.26X

(執筆者：鈴木輝彦) [2010年2月 受領]

画像符号化の国際標準では、MPEG, JPEG, H.26X (H.261, H.262, H.263, H.264) が有名で広く使用されている。これらの規格は、国際標準化専門委員会である ISO/IEC JTC1 SC29 及び ITU-T SG16 で標準化された。

11-1-1 MPEG

MPEG (Motion Picture Experts Group) は ISO/IEC が設立した標準化組織 ISO/IEC JTC1 SC29/WG11 の略称である。もともとは標準化団体の名称であったが、最近では規格そのものの名称としても使われている。JTC1 (Joint Technical Committee for information technology) は、国際標準化団体である ISO (International Organization for Standardization) と IEC (International Electrotechnical Commission) が情報処理分野の標準化を担当する技術委員会である。SC29 (Sub-Committee) はその中で、音声、画像、マルチメディア、ハイパーメディア、情報符号化に関する標準化を担当する副委員会である。WG11 (Working Group) は SC29 傘下で、動画像符号化の標準化を担当する専門家グループである。

MPEG は動画像符号化の標準化で有名だが、MPEG-7 ではメタデータの標準化、MPEG-21 ではコンテンツ配信の枠組みの標準化などを行っている。また、MPEG-A と呼ばれるアプリケーションフォーマットの標準化や、次世代システム技術の標準化を推進している。

MPEG で標準化された動画像符号化の標準には表 11・1 のものがある。

表 11・1 MPEG で標準化された画像符号化方式

MPEG 標準	特 徴
MPEG-1 (ISO/IEC 11172-2)	Video CD を想定した動画像符号化方式。Video CD の読み出しビットレートである 1.5 Mbps に最適化されている。動き補償と離散コサイン変換 (DCT) を組み合わせたハイブリッド符号化方式が採用されている。
MPEG-2 (ISO/IEC 13818-2)	DVD、放送などエンターテインメント分野で汎用に使われることを目的に規格化された動画像符号化方式。MPEG-1 をベースにテレビで使用されているインタレース方式に対応した。ITU-T SG16 と共同で標準化され、ITU-T では勧告 H.262 として出版されている。
MPEG-4 (ISO/IEC 14496-2)	携帯端末用のビデオコーデックとして当初は標準化された。その後、ネットワーク配信や放送プロダクション用の規格まで含むように拡張された。ITU-T で携帯用の動画像符号化方式として標準化された。勧告 H.263 をベースに機能拡張や符号化効率の改善が図られた。
AVC (ISO/IEC 14496-10)	AVC は Advanced Video Coding の略。MPEG-4 規格のパート 10 として標準化された。MPEG-2 のほぼ倍の符号化効率を達成している。Blu-Ray Disc など HDTV 製品で使用されている。AVC は ITU-T SG16 と共同で標準化された。MPEG と ITU-T で JVT (Joint Video Team) を設立し、JVT で標準化作業が進められた。

11-1-2 JPEG

JPEG (Joint Photographic Experts Group) は ISO/IEC が設立した標準化組織 ISO/IEC JTC1 SC29/WG1 で作成された規格の略称で、もともとは標準化団体名であった。WG1 は SC29 傘下で、静止画像符号化の標準化を担当する専門家グループである。静止画の符号化方式だけでなく、Motion JPEG など動画像符号化方式も標準化している。SC29/WG1 で標準化された静止画像符号化の標準には表 11・2 のものがある。

表 11・2 SC29/WG1 で標準化された画像符号化方式

JPEG 標準	特 徴
JBIG-1 (ISO/IEC 11544)	2 値画像の符号化方式。ITU-T では勧告 T.82 として出版されている。
JPEG (ISO/IEC 10918)	静止画の符号化方式で、デジタルカメラや、インターネットなどで広く使用されている。離散コサイン変換 (DCT) を採用した符号化方式である。可逆符号化に対応した JPEG-LS も標準化されている。ITU-T SG16 と共同で標準化されており、ITU-T では勧告 T.81 として出版されている。
JBIG-2 (ISO/IEC 14492)	JBIG の符号化効率を改善した、2 値画像符号化方式。ITU-T では勧告 T.88 として出版されている。
JPEG-LS (ISO/IEC 14495)	可逆符号化に対応した静止画符号化方式。ITU-T では勧告 T.870 として出版されている。
JPEG 2000 (ISO/IEC 15444)	JPEG の符号化効率を改善した静止画符号化方式。離散ウェーブレット変換 (DWT) を採用している。ITU-T では勧告 T.800 として出版されている。

11-1-3 H.26X

H.26X は ITU-T SG16 で規格化された動画像符号化方式の一群である。H.261, H.262, H.263, H.264 などがある。ITU (International Telecommunication Union) は無線、電気通信の標準化のために設立された国際連合の専門機関であり、ITU-T は電気通信に関する標準化を担当する部門である。SG16 (Study Group) はマルチメディア符号化、システム、アプリケーションの標準化を担当する専門家委員会である。ITU-T SG16 と ISO/IEC JTC1 SC29 では、多くの国際標準を共同で策定している。ITU-T で標準化された動画像符号化の標準には以下のものがある。

表 11・3 ITU-T SG16 で標準化された画像符号化方式

ITU-T 勧告	特 徴
H.261	TV 会議を想定して標準化された動画像符号化方式。動き補償と離散コサイン変換を用いたハイブリッド符号化方式を採用した最初の方式。
H.262	MPEG-2 と同一規格
H.263	携帯端末を想定した動画像符号化方式。
H.264	AVC と同一規格

■参考文献

- 1) ISO/IEC 11172-2, “Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s: Part 2 Video”.
- 2) ISO/IEC 13818-2 | ITU-T Recommendation H.262, “Generic coding of moving pictures and associated audio

- information: Part 2 Video”.
- 3) ISO/IEC 14496-2, “Coding of audio-visual objects: Part 2 Visual”.
 - 4) ISO/IEC 14496-10 | ITU-T Recommendation H.264, “Coding of audio-visual objects: Part 10 Advanced Video Coding”.
 - 5) ISO/IEC 11544 | ITU-T Recommendation T.82, “Coded representation of picture and audio information -- Progressive bi-level image compression”.
 - 6) ISO/IEC 10918 | ITU-T Recommendation T.81, “Digital compression and coding of continuous-tone still images”.
 - 7) ISO/IEC 14492 | ITU-T Recommendation T.88, “Lossy/lossless coding of bi-level images”.
 - 8) ISO/IEC 14495 | ITU-T Recommendation T.870, “Lossless and near-lossless compression of continuous-tone still images”.
 - 9) ISO/IEC 15444 | ITU-T Recommendation T.800, “JPEG 2000 image coding system”.
 - 10) ITU-T Recommendation H.261, “Video codec for audiovisual services at p x 64 kbit/s”.
 - 11) ITU-T Recommendation H.263, “Video coding for low bit rate communication”.

■2群 - 5編 - 11章

11-2 2値画像符号化

(執筆者：小野文孝) [2009年12月受領]

11-2-1 2値画像とその符号化

2値画像は大きく、文字やラインアートなどの文書画像と写真などの多値画像を2値で表現したハーフトーンに分かれる。2値画像の1次元符号化方式としては連続する同一色の画素数を符号化するランレングス符号化が古くから知られ、その後2次元相関を利用する方式の研究に進んだ。その中から、文書画像の符号化に適した変化点着目方式と、対象画像を選ばないマルコフモデル方式とが生まれ、国際標準としては1次元符号化方式のMH方式、2次元変化点着目方式のMR/MMR方式がITU-T標準となり、マルコフモデル方式のJBIG/JBIG2がJTC1ならびにITU-Tの共通標準となっている。

11-2-2 MH符号化

MH (Modified Huffman) 符号化はG3ファクシミリ必須方式であり、ITU-TのT.4勧告に規定される1次元ランレングス符号化である。ランの長さをメークアップ符号 $64 \times M$ ($0 \leq M \leq 40$) とターミネーティング符号 N ($0 \leq N \leq 63$) との和で表現するという工夫により符号の種類を減らしたのが特徴であり、(2560 + 63)以上のランに対してはメークアップ符号の連続で対処している。

11-2-3 MR/MMR符号化

MR (Modified READ : Modified Relative Element Address Designate) 符号化はITU-TのT.4勧告に規定されるG3ファクシミリの2次元符号化オプションである。直上ラインでの参照変化点に対応する変化点が±3画素以内のいずれで存在する場合は垂直モード符号(−3から+3までの7通り)を送り、対応変化点が存在せず参照変化点を進める場合はパスモード符号を送付し、それ以外は水平モード符号に続き2色のランの長さのペアをMH符号で符号化する。2次元符号では伝送路エラーがあると回復の保証がないため、Kラインに1ラインは1次元のMH符号で符号化し、エラー伝播を阻止する。一方、MMR (Modified MR) 符号はMR符号化でKを無限大としEOL (ライン同期符号) を省略したもので、ITU-TのT.6勧告に規定され、G4ファクシミリやG3ファクシミリ誤り訂正モードなど、誤りが存在しない環境で用いられる。典型的な画像ではMHによる圧縮性能が1/7、MRが1/10、MMRが1/15、後述のJBIGが1/18程度とされている。

11-2-4 JBIG符号化

JBIG標準とはJTC1とITU-Tの協同チームであるJBIG (Joint Bi-level Image Experts Group) により作成された最初の標準であるITU-T T.82 | ISO/IEC 11544 (Progressive bi-level image compression) をさす¹⁾。また、後述のJBIG-2の登場後はJBIG-1と呼ばれることがある。JBIG標準の効能としては、これまでの逐次的 (シーケンシャル) 符号化方式に加え、ソフトコピー通信に適した階層的 (プログレッシブ) 符号化方式への適用が可能なことや、従来のファクシミリ用符号化標準 (MH, MR, MMR) が不得意とするハーフトーン画像の符号化

性能向上などがあげられる。

JBIG のプログレッシブ符号化では原画に対し解像度を水平・垂直共に 1/2 にする低解像度化 (縮小) 処理を繰り返し行い, 作成された最低解像度画像をまず送信する。次に水平・垂直共にその 2 倍の解像度の画像を差分レイヤ情報 (既送信画素を参照したマルコフモデル符号化情報) として送り, 受信側では順に高い解像度の画像を得ていく。このほかに, 目的とする解像度の画像を直ちに送るシーケンシャルモードもある。JBIG ベースシステムを図 11・1 に掲げる。この中で標準の縮小処理 (PRES) は, 縮小対象領域の 4 画素, その周辺 5 画素, 縮小画像での周辺 3 画素の合計 12 画素の情報に基づきテーブル参照形式で縮小結果を得る。縮小関連の予測として, 縮小方式に依存する決定的予測 (DP : Deterministic Prediction) と, 依存しない典型的予測 (TP : Typical Prediction) とがあり, 効率向上と符号化時間短縮効果を有している。モデルテンプレートで認識するコンテキスト数は最低解像度レイヤでは 2^{10} , 差分レイヤでは 2^{12} であり, QM 符号という JPEG と共通の算術符号でエントロピー符号化する。適応テンプレートは周期的ハーフトーン画像に効果を発揮し, 算術符号はマルチコンテキスト情報源を高効率で符号化する効能を有する。なお, JBIG は G3 ファクシミリオプション符号 (ITU-T T.85) としても採用されており, JBIG 搭載機は一般にスーパーG3 と呼ばれている。

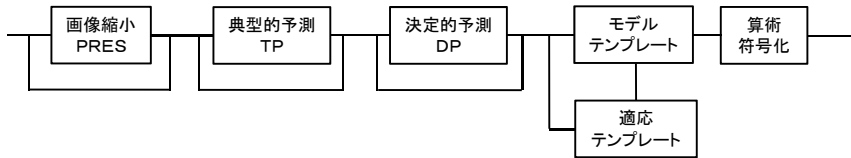


図 11・1 JBIG ベースシステムブロック図

11-2-5 JBIG-2 符号化

JBIG-2 標準とは, JBIG コミティが JBIG 標準に続き標準化した規格で ITU-T T.88 | ISO/IEC 14492 (Lossy/lossless coding of bi-level images) をいう。JBIG でのシーケンシャルモード符号化より多数のコンテキストを認知するほか, パターンマッチングを利用した文字のロッシー符号化機能, ソフトパターンマッチングを利用した同ロスレス符号化機能を有し, 一定領域内の黒画素数を送信するハーフトーン対応ロッシー符号化も可能である。なお, 算術符号として QM 符号とは若干異なる MQ 符号を使う。

■参考文献

- 1) 小野文孝, 渡辺 裕, “国際標準画像符号化の基礎技術,” コロナ社, 1998.

■2群 - 5編 - 11章

11-3 静止画符号化 JPEG

(執筆著：小野文孝) [2009年12月 受領]

11-3-1 JPEG 標準の位置づけ

JPEG (Joint Photographic Experts Group) は、静止画像符号化の国際標準化を行う JTC 1 と ITU-T の協同チームの名称であり、同時にその作成規格群の愛称としても用いられている。正式な名称は ITU-T Rec. T.81 | ISO/IEC 10918-1, Digital compression and coding of continuous-tone still images : Requirements and guidelines に始まるシリーズである。JPEG は静止画像符号化の最初の国際標準で、デジタルカメラ、Web ブラウザなどのアプリケーションで現在最もよく使用されている¹⁾。JPEG は、DCT (Discrete Cosine Transform) を利用した非可逆符号化が基本であり、最小限の機能を規定する基本システムと、算術符号 (QM 符号) や階層的伝送が使える拡張システムに分かれるが、更に DCT を使用せず可逆符号化が可能な独立機能も定義されている。以下では、基本システムについて解説する。

11-3-2 JPEG 基本システム

基本システムのブロック図を図 11・2 に示す。まず符号器では入力画像を 8×8 画素のブロックに分割し、2次元 DCT を行う。その結果、やはり 8×8 の DCT 係数が得られ、係数位置ごとに規定されたステップサイズをもつ量子化テーブルで線形量子化する。量子化テーブルの値は、自由に設定できるが標準書に掲載の例に比例係数を乗じて使われることが多い。ここで 8×8 の左上角の 1 係数がブロック内の平均レベルを表す DC (直流) 係数であり、残りの 63 個が AC (交流) 係数である。

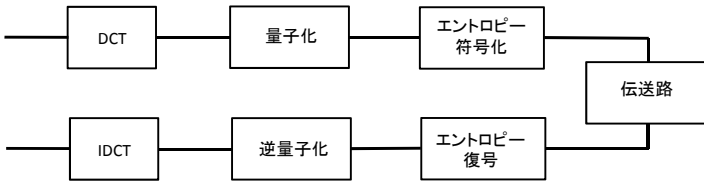


図 11・2 JPEG 基本システムブロック図

基本システムでは量子化された DCT 係数をハフマン符号で符号化する。DC 係数は同一色成分の DC 係数の直前値との差分データをグループ化し、そのグループ番号に対するハフマン符号にグループ内での個別値を特定する固定長の付加ビットを送付する。63 個の AC 係数は、低周波から高周波にジグザグ状に取り出し、連続する 0 (無効係数) のランレングスと、これを終端する非 0 係数値の属するグループ番号とを組み合わせた要素に対し構成されたハフマン符号を送り、続いて非 0 係数グループ内での個別値を特定する固定長の付加ビットを送付する。なお、ブロック内に非 0 係数がなくなった時点で EOB (End of Block) を送り、該当ブロックの符号化を終了する。また、デフォルトのハフマン符号は規定されておらず、使用ハフマン符号を受信側で生成するための情報として何ビット長の符号が何個あるかというテーブルと発生頻度順に並べた符号化要素テーブルとが送られる。受信側ではハフマン復

号のあと、逆量子化、IDCT（逆 DCT 変換）が行われる。

■参考文献

- 1) W. B. Pennebaker and J. L. Mitchell, "JPEG Still Image Data Compression Standard," New York: Van Nostrand Reinhold, 1993.

■2群 - 5編 - 11章

11-4 静止画符号化 JPEG 2000

(執筆著：小野文孝) [2009年12月 受領]

11-4-1 JPEG 2000 の概要

JPEG コミティでは JPEG の標準化の後、ロスレス及びニアロスレス符号化を規定する JPEG-LS (ITU-T Rec. T.87|ISO/IEC 14495-1 他) を標準化し、続いて JPEG の後継標準として、より高い圧縮性能、非可逆/可逆統一アルゴリズム、解像度/画質などに応じた段階的伝送、などの実現を目的に JPEG 2000 (ITU-T Rec. T.800|ISO/IEC 15444-1: JPEG 2000 image coding system: Core coding system ほか) を標準化した¹⁾。その基本機能はパート 1 に規定され、他に、拡張システム、動画像対応フォーマット、適合性試験、参照ソフトウェアなど各種の支援・応用標準パートが引き続き策定されている。JPEG 2000 は JPEG の基本システムより一般に 30% 以上少ない符号量で同等の画質 (SNR) が実現でき、デジタルシネマのほか各種のデータベース・アーカイブへの採用が進んでいる。

11-4-2 JPEG 2000 の変換方式

JPEG 2000 システムのブロック図を図 11・3 に示す。JPEG 2000 では、低ビットレートでの画質性能の高さから、画像変換方式としてウェーブレットを採用した。具体的には 1 次元ウェーブレット変換を画像の縦横各方向に独立に施して画像を四つのサブバンドに周波数分割し、最低周波数成分を担うサブバンドを更に再帰的に 4 バンドに分割する Mallat 分割を繰り返す。ウェーブレット係数の量子化で画質が制御され、基本機能でのウェーブレット変換には、最終的には可逆符号化が可能な (5, 3) フィルタと、非可逆で終わるがより高い効率の (9, 7) フィルタとがある。ROI (Region of Interest) とは関心領域の優先伝送機能である。また JPEG 2000 には各種エラー耐性処理が用意されている。

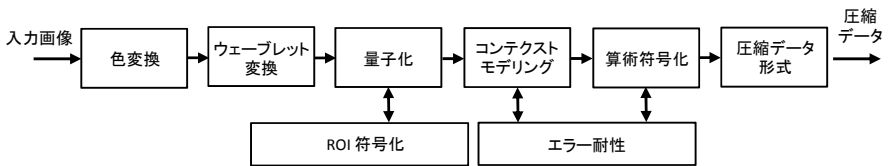


図 11・3 JPEG 2000 システムブロック図

11-4-3 JPEG 2000 のコンテキストモデリング

JPEG 2000 では、ウェーブレット変換係数を各サブバンドで固定サイズの符号化ブロック (縦横のサイズは 4 以上 1024 以下、面積は 4096 以下) に分割し、係数の絶対値を自然 2 進ビットプレーン表現し、JBIG-2 と共通のテーブル参照型 2 値算術符号である MQ 符号で符号化する。係数のコンテキストモデリングは EBCOT (Embedded Block Coding with Optimized Truncation) 方式²⁾ と呼ばれる。EBCOT では、符号化ブロック内の全ビットが 0 である上位のプレーン数がヘッダ情報として送信されたあと、各ビットプレーンが、垂直幅 4 画素のストライプに分割され、各ストライプ内を左上から垂直下方に進み、次に一つ右の列の最も上

から垂直下方に進む走査を繰り返す。各ビットプレーンデータは、**significance propagation** パス（有意な係数を周囲にもつまだ非有意な係数）、**refinement** パス（既に有意な係数）、**cleanup** パス（上記以外の係数）の3つのパスに分かれ、複数のコンテキストで符号化される。なお、**significance propagation** パスでは非0が符号化されると続いてサインビットが符号化される。また、**cleanup** パスには、条件により連続4変換係数がすべてゼロであるか否かを符号化するランモードが導入されている。

符号化ブロックごとに生成された符号データは再生画像の **SNR** への寄与度などの指標に応じて更に複数のレイヤに分類され、4要素（レイヤ、解像度、ブロック位置、色成分）の組合せに従う各種プログレッシブ伝送が実現される。

■参考文献

- 1) 小野, 野水, 原, “JPEG 2000 のすべて,” 電波新聞社, 2006.
- 2) D. Taubman, “High performance scalable image compression with EBCOT,” IEEE Trans. IP, vol.9, no.7, pp.1158-1170, Jul. 2000.

■2群 - 5編 - 11章

11-5 動画符号化 H.261

(執筆者：山田悦久) [2009年11月 受領]

11-5-1 H.261 符号化方式の概要

H.261¹⁾は1990年にITU-T(当時はCCITT)で勧告化された(最新版は1993年に出版)。1980年代にサービスが開始されたデジタルネットワーク「ISDN(Integrated Services Digital Network)」を利用した双方向映像サービスを目指したものである。ISDNは64kbit/sのデータチャンネルを基本構成としていたため、勧告のタイトルにはその整数倍を想定して「 $p \times 64$ 」という表記がとられており、 p の値は1~30とされている。

11-5-2 符号化方式の特長

テレビ電話・テレビ会議といった双方向のリアルタイム映像通信を実現するために開発された方式であり、1990年当時のH/W(LSI)において実時間処理が可能でかつ大きな処理遅延をとまわらない符号化技術の組合せによって構成されている。動き補償予測(MC)と離散コサイン変換(DCT)を組み合わせたハイブリッド符号化方式の原型である(図11・4)。

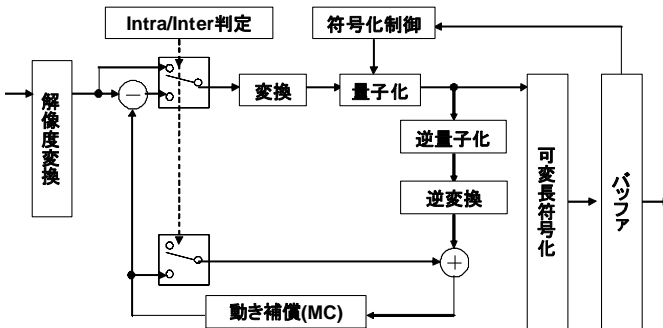


図 11・4 H.261 符号化器のブロック図 (MC+DCT ハイブリッド符号化)

NTSC/PAL/SECAM といったテレビ方式の異なる国同士であっても映像通信を実現するために、CIF (Common Intermediate Format) / QCIF (Quarter CIF) の映像フォーマットを策定し (QCIF は必須, CIF はオプション), 符号化されている映像フォーマットを世界共通にした (表 11・4)。これにより, 映像フォーマット変換・逆変換処理はエンコーダ・デコーダの外側に置かれ, コーデック機能そのものは世界統一化された。CIF/QCIF の解像度の違いによるコーデック内部の構成の差異を小さくするために, 33 個の MB (Macro Block) から構成される GOB (Group of Block) の組合せによりピクチャを構成する階層構造をとっている (図 11・5)。また, MB 単位で量子化パラメータを設定できるため, MB 単位での符号量・画質の制御が可能である。符号化雑音を軽減してフレーム間予測符号化の効率を高めるためにループ内フィルタが設けられている。

表 11・4 QCIF と CIF の映像フォーマットパラメータ

名称	水平画素数	ライン数	フレーム周波数	色空間
QCIF	176	144	29.97Hz	4:2:0
CIF	352	288		

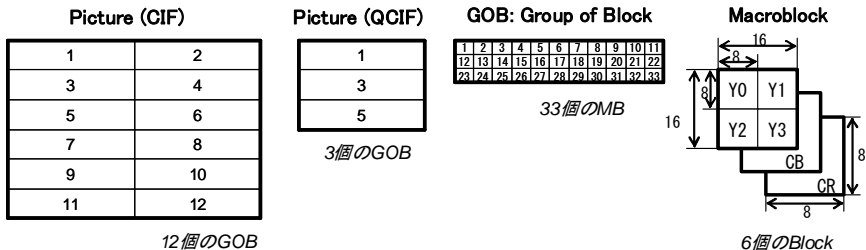


図 11・5 H.261 のピクチャ階層構造 (Picture, GOB, Macroblock)

H.261 では、圧縮された符号化データ (ビットストリーム) のシンタックス (パラメータのビット長とその並び順)・セマンティクス (パラメータの値の意味) と復号処理の手順のみを決め、符号化処理の多くの部分については標準を決めていない。そのため、符号化処理のかなめとなる符号化制御は各メーカーの独自色を出すことが可能となり、フレームレートや画質などの符号化性能をはじめ、処理遅延・装置規模などに対して様々な工夫が行われ、標準成立後も符号化技術が研究・開発されることにつながった。

H.261 の標準化作業の際に行われた、参照モデル (RM : Reference Model) を作成し、参照モデルに対して新しい符号化技術を提案して性能向上を図ったり、参照モデルの不備を発見して規格をより成熟させていくという手法は、その後の ITU-T・MPEG の標準化作業にも引き継がれている。H.261 の標準化作業の際には、ソフトウェアだけではなくハードウェアも試作し、国際間接続による標準方式の確認作業を行った。

映像のシーンによって発生符号量が大きく変動する映像データを、遅延やゆらぎが発生するネットワークを用いながら安定して双方向で通信するために、仮想バッファの仕組みが取り入れられている。送信側では、受信側のバッファの状態を想定してバッファがアンダーフロー・オーバーフローしないように発生符号量を制御するよう、規格で定められている。

■参考文献

- 1) ITU-T Recommendation H.261 (1993), "Video codec for audiovisual services at p x 64 kbit/s," 1993.

■2群 - 5編 - 11章

11-6 動画符号化 MPEG-2/H.262

(執筆著：山田悦久) [2009年11月受領]

11-6-1 MPEG-2 符号化方式の概要

MPEG-2 video/H.262¹⁾はMPEG-1²⁾を拡張する形で規格化が進められたもので、ISO/IECとITU-Tの共同作業により、1994年にISO/IECより標準化、ITU-Tより勧告化された(最新版は2000年に出版。その後も機能拡張は継続されている)。1990年代前半に検討されていたテレビジョン放送のデジタル化や高品質映像パッケージメディア、高速ネットワークを用いた映像通信といった多様なアプリケーションへの利用を想定し、タイトルにもあるように汎用的な符号化方式を目指して策定された。MPEGでは、SDTV解像度のをMPEG-2、HDTV解像度のをMPEG-3として標準化する計画が立てられていたが、MPEG-2の方式の大枠が固まった段階でHDTVでも十分な性能が得られることが確認されたため、MPEG-3を吸収する形でMPEG-2がHDTV解像度まで拡張された。DVDの映像圧縮方式や日米欧でのデジタル放送方式として、国際的に広く普及している。

11-6-2 MPEG-1 符号化方式の概要と特長

MPEG-1は1993年に成立したもので(標準化作業の主たる時期は1989~1991年頃)、タイトルから類推できるように、コンパクトディスクにAVデータを記録する符号化方式を目的としたものであり、再生専用装置を想定して標準化された。そのため復号処理のリアルタイム性や、早送り・逆方向再生といったような再生装置として必要な機能が取り入れられ、H.261では採用されなかった両方向予測・半画素動き補償予測・GOP(Group of Picture)構造(周期Iピクチャ)などが採用されている(図11・6)。

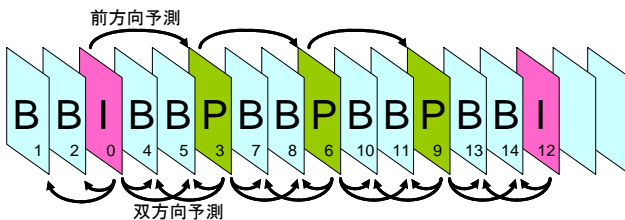


図 11・6 GOP 構造 (I/P/B ピクチャ。図中の番号は符号化順を表す)

11-6-3 MPEG-2 符号化方式の特長

MPEG-2がMPEG-1から拡張されている最大のポイントは、インタレース映像信号に対する符号化機能、ツールの導入である。1991年秋の横須賀市久里浜で開催された会合において、各機関から提案された符号化手法による圧縮映像に対して主観評価実験が行われた。性能が良いと評価された符号化ツールを組み合わせた参照モデル(TM: Test Model)を作成し、参照モデルのバージョンアップにより手法の改善が進められた。参照モデルはMPEG-2 part 5³⁾として標準化されている。

インタレース信号に対する符号化ツールとして、ピクチャ構造(フィールドピクチャ、フ

レームピクチャ), 動き補償予測(フィールド予測, フレーム予測, デュアルプライム予測), DCT(フィールドブロック, フレームブロック), 量子化係数のスキャン(zigzag, alternate)などがあげられる。映画のようなフィルム素材のものや動きのないシーンではフレームピクチャを, 動きのあるシーンではフィールドピクチャを, といったようにシーン単位で選択することがエンコーダでは可能である。シーン内に動きのある領域と動きのない領域が混在するような場合には, マクロブロック単位で動き補償予測やDCTに対してフレーム, フィールドを選択することが可能である。図11・7に8×8のDCTを行う際の, フレームブロック, フィールドブロックの構成を示す。

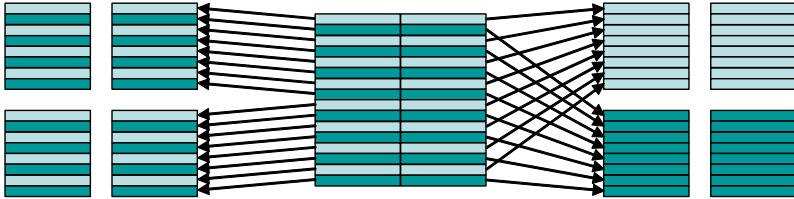


図11・7 DCTブロック構成(中:マクロブロック, 左:フレームブロック, 右:フィールドブロック)

MPEG-1で採用された半画素動き補償予測は, 画素と画素の間の半画素位置に対して補間処理により求められた仮想画素を予測に使用するもので, 精細な画像に対する予測効率の向上や, 補間処理で使われるフィルタ機能により符号化雑音が低減され符号化効率が向上するなどの効果が得られる。

MPEG-2では汎用方式を目指したものであるため, 様々な符号化ツールが採用されているが, アプリケーションによっては不要な符号化ツールもある。例えば, リアルタイム双方向通信を行う場合には, 遅延が発生してしまう両方向予測は使用したくない, というような場合があげられる。そこでMPEG-2では「プロファイル」が導入され, それぞれの用途に合わせて最適な符号化ツールを備えたH/W構成が取れるようにされている。また, 扱う映像の解像度などに応じて最適な処理速度やメモリ量を備えたH/W構成が取れるように「レベル」が導入されている(表11・5)。

表11・5 MPEG-2 videoのプロファイルとレベルの概略

プロファイル	ピクチャタイプ	色空間	レベル	最大解像度(例)
Simple	I,P	4:2:0	High	HDTV相当(1920×1088×60i)
Main	I,P,B	4:2:0	Main	SDTV相当(720×480×60i)
4:2:2	I,P,B	4:2:0 or 4:2:2	Low	SIF相当(352×240×30P)

■参考文献

- 1) ISO/IEC 13818-2:2000 | ITU-T Recommendation H.262 (2000), "Information technology - Generic coding of moving pictures and associated audio information: Video," 2000.
- 2) ISO/IEC 11172-2:1993, "Information technology - Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s - Part 2: Video," 1993.
- 3) ISO/IEC TR 13818-5:2005, "Information technology - Generic coding of moving pictures and associated audio information - Part 5: Software simulation," 2005.

■2群 - 5編 - 11章

11-7 動画符号化 H.263

(執筆者：山田悦久) [2009年11月受領]

H.263¹⁾は1996年にITU-Tで勧告化された(その後拡張され、最新版は2005年に出版されている)。勧告書のタイトルが示すように、低レートでの映像通信を想定して策定された。H.263は第三代携帯電話(3GPP)の必須符号化方式として、オプション符号化方式のMPEG-4 visualとともに広く普及している。

映像フォーマットとしては、H.261で定義されたCIF/QCIFのほかに、Sub-QCIF/4CIF/16CIFの合計5種類が定義され、QCIFとSub-QCIFの復号処理を必須とし、あとはオプション扱いとなっている。

低レートを主な対象としたことから、量子化係数の可変長符号化処理に対してゼロランとレベルに加えてブロック内の最後の係数であるか否か(ラスト)を加えた3次元ハフマン符号化を取り入れ、少ないデータをより効率良く符号化できるように工夫されている。

基本機能のほかに、符号化性能の向上やコーデックの機能を高めるために18のオプション機能が設けられている。これらのオプション機能は、映像通信を開始する前に送信機・受信機間でお互いに処理が可能な機能を情報交換することによって利用が可能となる。例えば、動き補償予測の機能としては、画面外参照、マクロブロック4分割、PBピクチャなどが定義されている。

■参考文献

- 1) ITU-T Recommendation H.263 (2005), "Video coding for low bit rate communication," 2005.

■2群 - 5編 - 11章

11-8 動画像符号化 MPEG-4

(執筆著：中條 健) [2009年12月受領]

MPEG-4は、誤り耐性、オブジェクト符号化、自然画像と人工画像（CG）との融合符号化など、単なる動画像符号化のみならず、これまでになかった多くの機能を実現している。基本的な符号化技術としては、先行して標準化が行われたITU-T H.263の影響を強く受けている。H.263が特に数+kbps程度の低いレートでの符号化において、ヘッダ情報やモード情報などのオーバーヘッドのビット量が無視できなくなることを考慮して注意深く設計されていることを踏襲して、MPEG-4では、H.263の基本部分に加えていくつかの符号化ツールを改良するような構成になっている。

MPEG-4には多くのプロファイルが存在するが、基本的な符号化技術に誤り耐性技術が加わったシンプルプロファイル（Simple Profile）が最も実用化が進んでいる。現状では、それについて、後から規格化されたアドバンスドシンプルプロファイル（ASP：Advanced Simple Profile）が普及していると言われている。アドバンスドシンプルは、シンプルプロファイルに加えて、インタレース信号対応や、双方向画像予測構造などのいくつかの符号化効率改善ツールが追加されている。

MPEG-4の特徴的な技術としては、ビデオオブジェクトと呼ばれる概念を導入して、任意形状の映像の符号化が実現できる点があげられる。ビデオオブジェクトとは、映像中の特定の意味の単位であり、例えば、映像中の人物の部分だけを一つのビデオオブジェクトすることができる。任意形状符号化では、人物のみの映像を符号化して、他の背景映像と合成して表示するようなことが可能となる。そのため、図11・8のように通常の画素値の信号に加えて、形状情報を別途加えて符号化する枠組みが用意されている。このオブジェクトという考え方は、オーディオやシステムを含めたMPEG-4標準全体に適用されている。

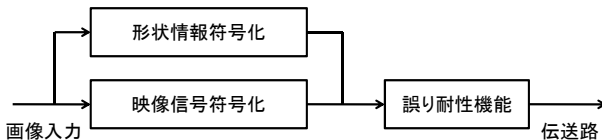


図 11・8 MPEG-4の符号化の流れ

また、第三世代の携帯電話サービスやインターネットをも含めたグローバルなネットワークでの使用を考慮した場合の通信路は、誤り率が高いだけでなく、様々な状況が想定される。そこで、システム側で訂正しきれなかった各種の誤りパターンを想定し、このような環境の下でも映像品質を確保できるような、動画像符号化としての誤り耐性技術が採用されている。

MPEG-4で採用されている誤り耐性技術を、すべてを使って符号化データの構造として図示すると図11・9のように表すことができる。具体的には、画面の単位より細かい単位でビデオパケット（Video Packet）と呼ばれる構造を作り同期を確保し、その後の情報を正しく復号することが可能となる再同期マーカ（Resynchronization Marker）と双方向に復号可能なリバーシブル可変長符号（Reversible Variable Length Code）が採用されている。この符号は、誤

りを検出した場合にも、次の再同期マーカから逆方向に復号することにより、復号できずに廃棄されてしまうビット数を減らすことができる。なお、ここで使われている符号は、自然数に対する符号化を応用したもので符号化効率の低下は、悪くとも数%以内であることが知られている。また、重要情報の二重化は、ヘッダ拡張符号 (Header Extension Code) というフラグを立てて、この後にヘッダに含まれる情報を二重化情報として付け加えることができる構造である。データパーティショニング (Data Partitioning) は、ビデオパケット内のマクロブロックの符号化データをまとめて、階層化することで、情報の優先度をつけ、低優先度のデータに誤りが入って復号できない場合に、エラーコンシールメントが実現できる仕組みである。上記二つのツールは、誤りの修復がしやすくなるような技術と言える。

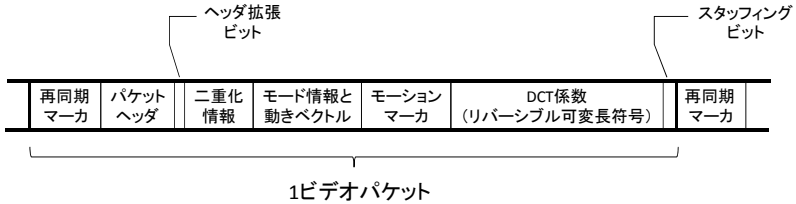


図 11・9 誤り耐性技術の符号化データ構造

ビデオオブジェクトの構想は必ずしも実用化の面で成功しているとは言えないが、MPEG-4 は、PC やインターネットでの動画像のアプリケーション、それから、モバイル、特に第三代携帯電話という新たな市場を作り出した功績は大きい。

■2群 - 5編 - 11章

11-9 動画像符号化 AVC/H.264

(執筆著: 中條 健) [2009年12月受領]

AVC/H.264は、ITU-TのVCEGのH.26Lプロジェクトがその始まりである。2001年末にMPEGと共同でJVT (Joint Video Team)を組織して規格化された。したがって、双方の組織の規格名を合わせてもっている。この規格では、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。

MPEG-4では、多くの機能を実現することに努力を払ってきたのに対して、この規格の最大の狙いは、符号化効率の向上である。概ねMPEG-2の2倍の符号化効率があると言われている。ちなみに、2倍の符号化効率とは、同じ画質で半分のデータ量に圧縮できるという意味を示している。AVC/H.264では、ハイブリッド符号化の考え方をより高度化し、多くのツール群の集合体となっている。図11・10は、AVC/H.264の骨格となる五つのプロファイルのツール群を示す。これだけ多くの符号化ツールを取り込むことができたのは、標準化の参照ソフトウェアにレート-ひずみ最適化の枠組みが実装され、それに基づいてツールの性能評価をしてきたという背景がある。一方、この最適化技術は、符号化効率を稼ぐことができる反面、仮の符号化を繰り返してその結果を評価するため、演算量が多いという問題点がある。そのため、特にエンコーダの実装上においては、この規格の潜在能力をいかに維持しつつ、演算量を削減することができるか大いに工夫の余地がある。

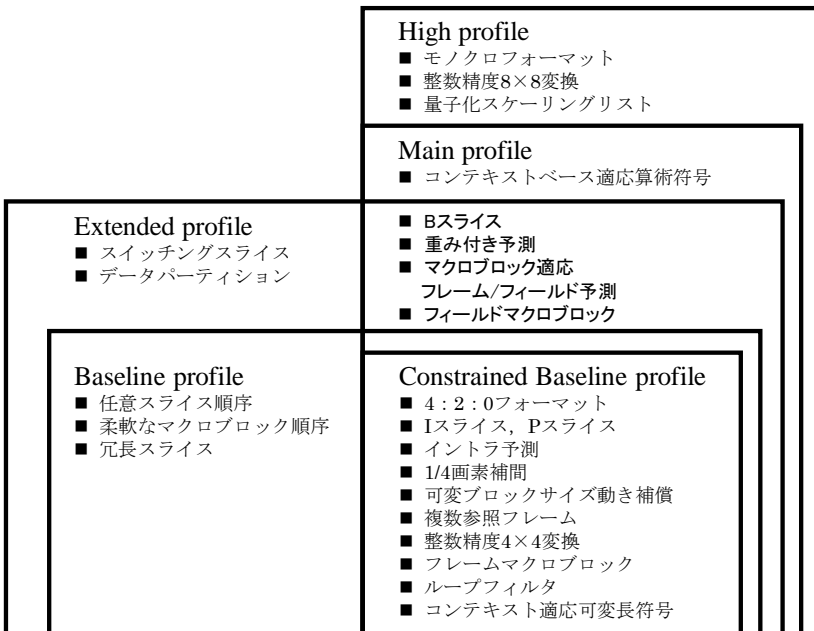


図11・10 AVC/H.264の骨格となるプロファイルのツール群

当初策定された三つプロファイルのうち、ベースラインプロファイル (Baseline Profile) は携帯電話やテレビ会議などの通信系でデコーダの演算量が少なくすむようなアプリケーションを想定している。一方、メインプロファイル (Main Profile) は放送蓄積メディアを想定していた。メインでは、ベースライン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)、インタレース関係のツール、コンテキストベース適応算術符号化 (CABAC : Context-based Adaptive Binary Arithmetic Coding) が追加される。なお、その後、実際の AVC/H.264 の使用状況に合わせる形で、ベースラインプロファイルとメインプロファイルとサブセットの部分に制約付きベースライン (Constrained Baseline Profile) として使えるようにしている。

ハイプロファイル (High Profile) は、FRExt (Fidelity Range Extensions) として追加されたプロファイルの一つで、メインプロファイルに 8×8 変換や量子化スケーリングリストが追加されている。この二つのツールは、高精細テレビ (HDTV) の符号化効率の向上と映画素材におけるフィルムグレインノイズの再現性に寄与している。FRExt に関しては、プロフェッショナル用途向けに更に拡張し、プロフェッショナル用途のフォーマットへの対応やイントラのみプロファイルなどが定められている。

その後、標準化が行われたスケーラブル符号化 (SVC : Scalable Video Coding) は、映像の時間、空間、画質のスケーラビリティを実現するために、映像を階層的に符号化する技術である。階層的に符号化することで、例えば、空間スケーラビリティならば、符号化データの一部を取り出すことで、携帯電話で再生するような小さな画面サイズの映像から、HDTV のような大きな画面サイズの映像までスケーラブルに再生することができる。これまでのスケーラブル符号化はその機能を実現するために、符号化効率が大幅に低下するという欠点を抱えていたが、H.264/AVC の SVC では、この点に関して改善が図られ、時間スケーラビリティならば、ほぼオーバーヘッドなしに、空間スケーラビリティと SNR スケーラビリティならば二階層のとき、約 10% のオーバーヘッドで、実現できると言われている。

マルチビュー符号化 (MVC : Multi View Coding) は、複数のカメラを配置し、同時に多視点で撮影された映像を符号化する技術である。AVC/H.264 の MVC では、符号化の基本的な部分に変更せずに、符号化構造の上位レベルの変更でその機能を実現している。

AVC/H.264 は、符号化効率の面で飛躍的な性能向上を果たした結果、HDTV 解像度の映像のビデオ映像の記録・再生をする光ディスクやメモ리카ードに記録するビデオカメラなどの新しいアプリケーションを生み出してきている。また、携帯電話でデジタル TV 放送が楽しむことができる地上波デジタル 1 セグメント放送も実現した。今後、実用面での更なる発展が期待されている。

■2群 - 5編 - 11章

11-10 動画像符号化標準方式の比較

(執筆著者：中條 健) [2009年12月 受領]

動画像符号化の国際標準の歴史を作ってきたのは、ISO (国際標準化機構) と ITU-T (国際電気通信連合-電気通信標準化部門) の二つの組織である。より具体的には ISO の下部組織 (JTC1/SC29/WG11) の MPEG (Moving Picture Experts Group) と呼ばれている組織と ITU-T の下部組織 (SG16/Q6) の VCEG (Video Coding Experts Group) と呼ばれている組織である。MPEG は、通称 MPEG-1, 2, 4 と呼ばれる規格を作成してきており、VCEG は、H.261, 2, 3, 4 と呼ばれる規格を作ってきている。時には MPEG-2/H.262 や AVC/H.264 のようにお互い協力して同一文書の規格を作成してきた。

表 11・6 動画像符号化標準方式の比較

符号化ツール	H.261	MPEG-2 (MP@ML)	MPEG-4 (Simple Profile)	H.264/ AVC (Baseline Profile)	H.264/ AVC (High Profile)
変換	8×8 DCT	8×8 DCT	8×8 DCT	整数精度 4×4 DCT	整数精度 4×4 DCT, 整数精度 8×8 DCT
量子化	ミッドトレッド型一様量子化器	デッドゾーン付ミッドトレッド型一様量子化器, 量子化マトリクス	ミッドトレッド型一様量子化器	ミッドトレッド型一様量子化器	ミッドトレッド型一様量子化器, 量子化スケールングリスト
イントラ予測	なし	DC 係数予測	適応 DC/AC 予測	4×4 単位で 9 通り, 16×16 単位で 4 通りの予測	4×4 単位で 9 通り, 8×8 単位で 9 通り, 16×16 単位で 4 通りの予測
予測構造	I, P	I, P, B	I, P	I, P 複数参照フレーム	I, P, B 複数参照フレーム 重み付予測
動き補償ブロックサイズ	16×16 画素単位	16×16 画素単位	16×16, 8×8 画素単位	16×16, 16×8, 8×16, 8×8, 8×4, 4×8, 4×4 画素単位	16×16, 16×8, 8×16, 8×8, 8×4, 4×8, 4×4 画素単位
動きベクトル精度	1 画素精度	1/2 画素精度	1/2 画素精度	1/4 画素精度	1/4 画素精度
画面外の動き補償対応	なし	なし	あり	あり	あり
ループフィルタ	ローパスフィルタ	なし	なし	適応デブロッキングフィルタ	適応デブロッキングフィルタ
エントロピー符号化	2次元 VLC	2次元 VLC	3次元 VLC	CAVLC	CAVLC, CABAC
インタレース対応	なし	フレーム/フィールド適応動き補償/DCT, Dual Prime 予測	なし	なし	マクロブロック適応フレーム/フィールド予測

方式的には、ITU-T H.261 によって確立された動き補償 (MC : Motion Compensation), 離散コサイン変換 (DCT : Discrete Cosine Transform), ミッドトレッド型一様量子化, エントロピー符号化を組み合わせたいわゆるハイブリッド符号化のアーキテクチャを踏襲している。

H.261 は, 技術的にその後の規格のベースとなったという点で重要である。表 11・6 が示すように基本的には新しい規格ほどより高度な符号化ツールへと進化していつている。

MPEG-2 は, 最も成功した動画像符号化の国際標準規格であると言える。標準化の過程で多くのアイデアが出されたが, 現在最も使われているメインプロファイル/メインレベル (MP@ML) 及びメインプロファイル/ハイレベル (MP@HL) では TV で用いられている飛び越し走査方式のインタレース信号への対応が重要なポイントとなっていた。また, MPEG-1 で採用された双方向画像予測構造 (B Picture) や 1/2 画素精度まで動きベクトル精度をもたせた。

MPEG-4 の基本部分は, 技術的には先行して標準化が行われた ITU-T H.263 の影響を強く受けており, 8×8 の動き補償ブロックサイズや, 画面外の動き補償対応は, H.263 を引き継いでいるが, イントラ予測として適応 DC/AC 予測を採用し, また DCT 係数のエントロピー符号化に改良が加えられた。

AVC/H.264 では, 更に多くの改善が図られている。ベースラインプロファイルにおいて特徴的なのは, 整数精度 4×4 DCT, 複数参照フレーム, 最小 4×4 画素単位の動き補償ブロックサイズなどである。そのため, これまでの標準方式とは圧縮時の視覚的な絵柄がやや異なる傾向があった。ハイプロファイルにおいて, 整数精度 8×8 DCT, 量子化スケールリングリストなど従来の規格においても用いられてきたツールが追加されたことにより, 同様な傾向の絵作りが可能になった。

■2群 - 5編 - 11章

11-11 SMPTE VC-1, VC-2

(執筆著：中條 健) [2009年12月受領]

VC-1は、Windows上の動画像ツールのWindows Media Video Version 9をもとにSMPTE(米国映画テレビ技術者協会)でSMPTE 421Mとして標準化したものである。表11-7のようにシンプル(Simple)、メイン(Main)、アドバンスド(Advanced)の三つのプロファイルが存在し、機能的にはオニオンリング状に拡張されるような構成になっている。シンプルは、モバイル用、メインはInternet用、アドバンスドがパッケージメディア用という位置づけになっている。

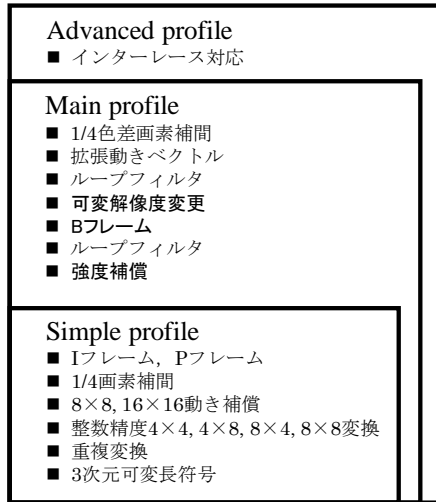


図11-11 VC-1のプロファイルとツール群

Windows Media Videoは、もともとMPEG-4のシンプルプロファイルからスタートした経緯があり、方式的には、MPEG-4を拡張し、かつAVC/H.264と同様な方向で進化した方式になっており、特に少ない計算量で実装できることを意識している。例えば、動きベクトル精度は、1/4画素精度までであるが、補間フィルタはAVC/H.264の6タップに対して4タップである。

VC-2は、BBC(英国放送協会)がオープンソースで開発していたDiracという名前のWaveletベースの動画像符号化方式のうち、Dirac Proと呼んでいるIntra部分に関して、プロフェッショナル用途にSMPTE 2042として標準化したものである。10ビット画素ビット長と4:2:2フォーマットと、ロスレスとビジュアルロスレスの符号化をサポートする。