

MPEG 2 / MPEG 4 の基礎とHD規格

YITコンサルティング

石川 安則

MPEGフォーマットは、デジタルHDコーデックなどの圧縮フォーマットとして利用されるほか、DVD、デジタル放送、ネットワーク配信など、コンシューマユースレベルでも広範に普及している。MPEGは国際標準圧縮方式であり、こうしたさまざまな利用（アプリケーション）レベル間でのデータ交換に際して高い互換性をもつ一方、幅広い応用への適応のため技術仕様が重くなりがちでコスト面での問題がある。このため、全体規格のサブセットを規定して特定の用途向けに最適化する手法が用いられている。また、MPEG技術の世代交代の影響もあって、利用場面では複数のMPEG圧縮フォーマットが並立することもある。

ここでは、まずMPEG技術の変遷と種類について概説したあと、現在主流となっているMPEG 2とMPEG 4についてビデオ制作における適用を説明し、最後に最新の圧縮技術であるH.264について簡単に触れることにする。

MPEG技術の変遷

よく知られているように、MPEGとは Moving Picture Expert Groupという動画圧縮（符号化）の国際標準を

策定する機関の略称であり、これがそのまま圧縮方式の名称として用いられている。動画圧縮技術は、テレビジョン放送の実用化初期段階から連綿と研究開発が続けられてきたが、デジタルビデオ技術の実用化が視野に入ってきた1980年代後半から本格的な技術標準の策定作業が始まった。図1は、それ以降の技術変遷についての流れを示したものである。

図1には、静止画符号化の標準系列も含まれるが、Motion-JPEG、Motion-JPEG2000などに見られるように動画と関係は無視できない。ITU-Tは、国際電気通信連盟（ITU）の下部組織で主に有線通信に関わる技術標準が所掌範囲であるのに対し、ISO（国際標準化機構）は、より汎用的な技術標準の作成が使命であり、通信・放送・蓄積メディアなど広い応用範囲をカバーしている。

実際には、2つの機関を兼務する検討委員も多く、また検討会合自体が合同で行われることも多い。静止画のJPEG（Joint Photographic Expert Group）は、両者の合同作業グループのことであり、またH.264の検討作業は、JVT（Joint Video Team）という名称の合同組織で行われている。

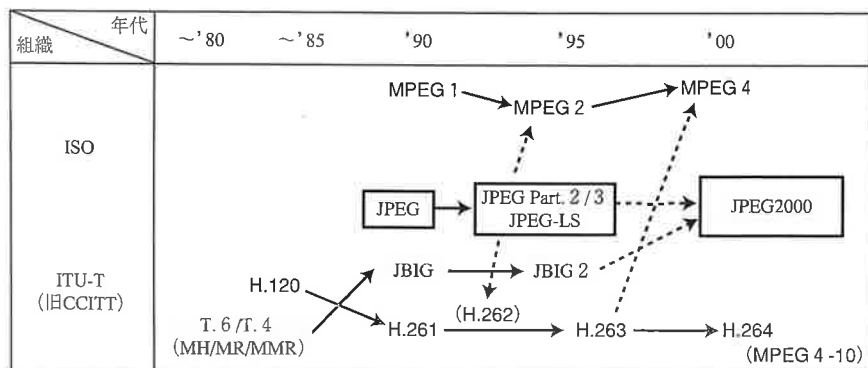


図1 画像符号化国際標準方式の変遷

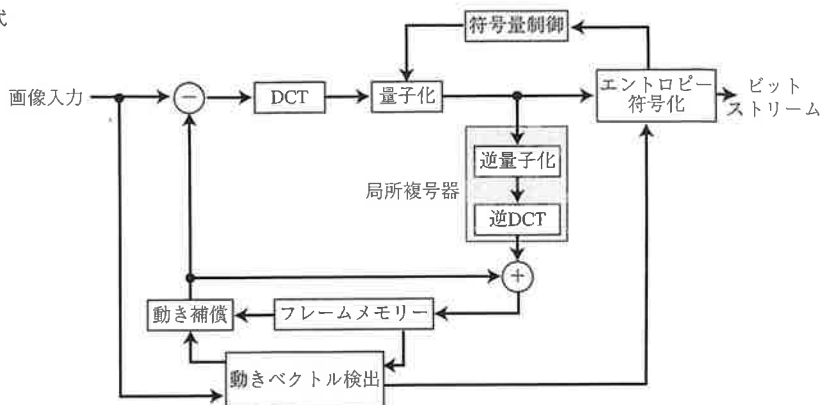
□：静止画（2値） そのほか：動画

■：静止画（多値）

ISO：International Standardization Organization

ITU-T：International Telecommunication Union
(Telecommunication Standardization Sector)

図2 動画像符号化の基本方式



■MPEGの基本技術

上述のように、動画画像圧縮技術の系列は、MPEGとH.26Xの2つの流れがあるが、両者は密接に関連している。共通の基本技術となる符号化方式の枠組みは、最初の動画画像符号化標準方式であるH.261以来大きな変化はない。図2は、基本方式におけるエンコーダーのブロック図である。符号化方式の枠組みとは、以下の基本技術から成り立っている。

①動き補償フレーム間予測

時間的に変化する動画像の各フレーム間で各画素値の差分値のみを符号化する。動きのある部分は、さらに動きベクトルを求めて動き量を補償する。予測モードは、過去のフレームのみを参照するPピクチャー、過去および未来のフレームを参照するBピクチャーと、フレーム間予測を行わないIピクチャー(すなわち静止画の符号化と同じ)がある。

②離散コサイン変換 (DCT=Discrete Cosine Transform)

フレーム内の画素値は、相関をもっておりこの相関(冗長度)を除去することにより圧縮が可能となる。また、人の視覚特性に基づき歪みを知覚しにくい信号成分(主に高周波数成分)の除去が可能である。このため、DCTにより周波数成分(DC係数、AC係数)への変換が行われ、それぞれ量子化をすることによってこれらの特性を得ることができる。

③エントロピー符号化

量子化された各周波数成分値は、値の出現確率にもとづく可変長符号が割り当てられる。すなわち、出現確率の高い値には短い符号を、低い値には長い符号を割り当てることにより、全体として効率の良い符号割当が可能となる。通常は、処理計算量が少ないハフマン符号が用いられる。

■MPEG技術のバリエーション

H.261の主たる応用分野は、テレビ電話・テレビ会

議システムであり、音声符号化方式(G.711、G.726)やデータ多重化方式(H.221)を含めて、最終構成としてH.320というシステム標準がITU-Tで制定されている。これに対して、最初のMPEG規格であるMPEG 1は、VHS品質のビデオをCD-ROMなどの蓄積媒体に記録・配布するモデルに基づき策定された。

二者の大きな違いは、H.261がリアルタイムのエンコード・デコードが要求されるのに対して、MPEG 1は、エンコードは非リアルタイム、デコードはリアルタイムという非対称性が前提となっていることである。

第2世代の動画像符号化国際標準であるMPEG 2は、DVDやデジタル放送の圧縮フォーマットとして確立された基盤技術となっている。MPEG 1と異なり、応用分野として通信・放送・蓄積系のすべてを網羅する標準となっている。このためMPEG 1に比べ多くの圧縮技術要素を含んでおり、これらの選択的な組み合わせにより、種々の応用分野に対応できる規格としている。プロファイル(圧縮技術要素の選択)とレベル(画像サイズ、フレームレートの選択)という概念を導入して、応用分野ごとに適用するプロファイル/レベルを規定する方式をとる。

第3世代の規格は、まずH.263がH.261の発展形として標準化された(実際にはH.263→H.263+→H.263++というように順次改良が加えられた)。MPEG 4は、MPEG 2に続いて、当初は移動体端末を主ターゲットとした超低ビットレート動画像符号化方式としての企画立案から始まったが、審議過程で多くの有用な提案がなされるに従い、非常に多様な圧縮技術要素(圧縮ツールと呼ばれるようになった)が採用されるに至った。こうしてMPEG 4標準は、超低ビットレート符号化からHDTV用の高画質符号化までカバーする非常に汎用的な動画像圧縮標準の形態をとることになった。現在のMPEG 4規格は、MPEG 1、MPEG 2の適用範囲

表1 MPEG技術の比較

技術	方式	MPEG 1	MPEG 2	MPEG 4 (自然画像)
入力画像 フレーム構造		プログレッシブ	インターレース プログレッシブ	インターレース プログレッシブ
サブサンプル		4:2:0	4:2:0 4:2:2	4:2:0 4:2:2 4:4:4
ビット長/画素		8ビット	8ビット	8ビット (4~12ビット)
予測モード		フレーム	フィールド フレーム適応	フィールド フレーム適応
ピクチャー構造		I、P、B	I、P、B	I、P、B
動きベクトル精度		半画素	半画素	半画素 1/4画素
DCTブロック構成		フレーム (8×8画素)	フィールド フレーム適応 (8×8画素)	フィールド フレーム適応 (8×8画素)
イントラ係数予測		DC係数のみ	DC係数のみ	適応DC/AC係数予測
DCT係数スキャン		ジグザグ	ジグザグ オルタネート※1	ジグザグ オルタネート※2
係数符号化方式		2次元VLC※3	2次元VLC	3次元VLC
ビットレート		~1.856Mbps	~300Mbps	~1.8Gbps

※1 垂直方向のみ

※2 水平、垂直方向

※3 VLC: Variable Length Code=可変長符号

を完全に包含する規格となっている。表1に各MPEG規格における技術要素の比較を示す。

なお、MPEG規格の構成は表2に示すPart.1~Part.5までの基本部分と、MPEG2/4ではさらにPart.6以降に独自の規格が規定されている。MPEG4では、現在Part.19までのプロジェクトがあり、精力的な審議活動が続けられている。

MPEG3規格は、当初MPEG2に続くHDTVレベルの画像サイズおよびビットレートに特化した規格として標準化制定される予定であったが、プロファイル/レベル概念の導入によりMPEG2がHDTVレベルまで包含する規格として制定されることになったため、ISO内でプロジェクトが取り下げられ、規格として消滅してしまった。

表2 MPEG規格の基本構成

Part. 1	システム
Part. 2	ビデオ (ビジュアル)
Part. 3	オーディオ
Part. 4	互換性テスト
Part. 5	参照ソフトウェア

HD圧縮フォーマットにおけるMPEG技術

このように種々のフォーマットが存在するMPEGであるが、つぎにHD圧縮フォーマットに関係の深いMPEG技術に絞って解説する。

■MPEG2 システム (TS/PS)

MPEGのシステム規格とは、ビデオとオーディオデータおよびその他のデータ (文字データなど) の同期を確保し多重化するための規格である。MPEG2に準拠した圧縮フォーマットは、すべてMPEG2システム規格のビットストリームで構成される。

MPEG2のシステム規格には、トランスポートストリーム (TS) と呼ばれる放送・通信系の規格と、プログラムストリーム (PS) と呼ばれる蓄積系 (DVDなど) の規格がある。

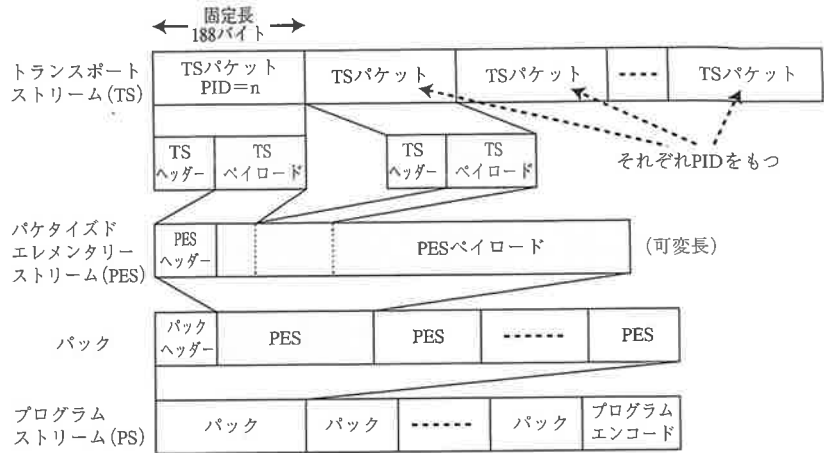
PSはMPEG1システム規格の拡張であるのに対して、TSはデジタル放送を主なターゲットとして規定され、現在世界各国のデジタルテレビ放送は、MPEG2システム規格を採用している。表3に2つのMPEG2システム規格の比較を示す。TSおよびPS規格のビットストリームは、図3に示すような構造となっている。

ビデオおよびオーディオのそれぞれのデータは、エレメンタリーストリーム (ES) と呼ばれる。各ESは、可変長のPESパケットに組み込まれ、これがTSおよびPSの基本単位となる。PSでは、複数のPESからパックと呼ばれる可変長のパケットが構成され、パックのストリームがPSとなる。TSは、逆にPESを分割してTSペイロードに組み込む。TSパケットは、188バイトの固定

	適用対象	特長
トランスポートストリーム (TS)	ビット誤りの発生する環境における伝送・蓄積 適用例 デジタル放送 (衛星、地上波)	・固定長パケット (188バイト) ・PESパケットを分割してTSパケットに乗せる ・セクションデータの送信 ・複数プログラムを含む
プログラムストリーム (PS)	ビット誤りの発生しない環境における伝送・蓄積 適用例 DVD	・複数のPESパケットからなるパックで構成される ・単一プログラムのみ

表3 MPEG2 TSとPSの比較

図3 TSとPSの構成



長なので分割したPESストリームが短い場合はアダプテーションフィールドという挿入ビットで埋める。なお、ペイロードとは一般に伝送・蓄積されるべき本来のデータのことをいい、ヘッダーは制御情報の意味である。またTSでは、PESのほかにセクションというデータも搭載でき、これは放送番組情報などを含むPSI (Program Specific Information) などに利用される。

MPEG 2 システムでは、このほかにビデオ・オーディオなどのデータ同期を確保するための情報(タイムクロック、タイムスタンプなど)について規定されている。

■MPEG 2 ビデオ(4 : 2 : 2 プロファイル)

MPEG 2 からプロファイル/レベルの概念が導入されたことは述べた。表4にMPEG 2 ビデオにおけるプロファイル/レベルの適用表の抜粋を示す。現在デジタル

放送の配信では、SD規格ではMP@MLが、HD規格ではMP@HLが主流となっている。CATVなどでは、SP@MLも利用されている。

MPEG 2 のプロファイルは当初はHPまでであったが、ビデオ制作現場における高画質要求から、4 : 2 : 2 プロファイルが追加された。MPの色差フォーマットが4 : 2 : 0に限定されるのに対して、HPではITU-R BT.601規格に準拠する4 : 2 : 2が利用できる。しかし、HPはスケーラビリティ機能(表4)が必須であり、装置のコストアップとなる。このため、HPからスケーラビリティ機能を除き、かつビットレートの最大値を300Mbpsまで上げた規格として、4 : 2 : 2 プロファイル(422P)が制定された。ビットレートの最大値を上げたことで、P、Bピクチャーを用いないIピクチャーのみのモードも利用されるようになった(これ

表4 MPEG 2 のプロファイル/レベル(抜粋)

プロファイル		Simple (SP)	Main (MP)	High (HP)	4 : 2 : 2 (422P) ^{※4}
プロファイルの機能	色差フォーマット	4 : 2 : 0	4 : 2 : 0	4 : 2 : 0 4 : 2 : 2	4 : 2 : 0 4 : 2 : 2
	ピクチャー	I, P	I, P, B	I, P, B	I, P, B
	スケーラビリティ ^{※1}	なし	なし	SNR、空間	なし
レベル ^{※2}	High [1920×1080] (HL) [60fps]		MP@HL [80Mbps]	HP@HL [100Mbps]	422P@HL ^{※4} [300Mbps]
	High-1440 [1440×1080] (HL1440) [60fps]		MP@HL1440 [60Mbps]	HP@HL1440 [80Mbps]	
	Main [720×576] (ML) [30fps]	SP@ML [15Mbps]	MP@ML [15Mbps]	HP@ML [20Mbps]	422P@ML ^{※4} [50Mbps]
	Low [382×288] (LL) [30fps]		MP@LL [4 Mbps]		

※1 複数の解像度(空間)や複数の画像品質(SNR)の画像ストリームを、1つのビットストリームで階層的に送信する機能

※2 各レベルの欄右側の数値は、最大解像度[単位なし]、最大フレームレート[単位fps]を示す

※3 各Profile@Level の下の数値は、最大ビットレートを示す

※4 MPEG 2 規格では規定されない

表5 MPEG 4 スタジオプロファイルの比較

プロファイル名	Simple Studio	Core Studio
最大画像サイズ	2048×2048	2048×2048
色差フォーマット	4 : 2 : 2 4 : 4 : 4	4 : 2 : 2 4 : 4 : 4
ビット長/画素	最大12ビット	最大10ビット
ピクチャー	I	I, P
スプライト符号化*	なし	あり
最大ビットレート	1.8Gbps	900Mbps

※スプライト(シーンのすべての背景をカバーする大画面)を一度に伝送しておき、各フレームの背景はスプライトから切り出した画面を幾何変換して用いる高効率な符号化手法

は、4 : 2 : 2 プロファイルの標準規定内である。たとえば、MPEG IMX方式などがある。

■MPEG 4 ビジュアル(スタジオプロファイル)

MPEG 4 の特徴の 1 つにオブジェクトベース符号化がある。オブジェクトベースとは、各種マルチメディアデータの 1 つ 1 つをオブジェクトとして、これらを自由に組みあわせて符号化ストリームを形成する考え方である。ビデオオブジェクトとしては、自然画に加えCG、アニメーションなどの合成画像からなる動画・静止画、テキストなどさまざまな画像データが対象となる。また、従来の矩形形状のほか任意形状のオブジェクトが扱え、さらにアルファマッピング機能によりオブジェクトの重ね合わせに透過率の指定が可能である。これらの機能により、画像の個々の部品を組みあわせて 1 つの画面を形成する操作が可能となり、この結果ビデオ制作における生産性向上が図れる。

MPEG 4 ビジュアル規格(ビデオでなくビジュアルと称する)は、最初に 9 つのプロファイルをもつ Version. 1 規格が制定され、次いで 6 つのプロファイルを追加した Version. 2 規格が続き、さらに Version. 3 および 4 規格が制定されつつある。この中で、表 5 に示すスタジオプロファイル規格は、すでに規格化が完了している。これは、4 : 2 : 2 に加え 4 : 4 : 4 フォーマットの規定が追加されプライマリーカラー(原色)の RGB データも扱え、また任意形状符号化とアルファマップ機能がある。このうちシンプルスタジオプロファイルは、I ピクチャーのみの規定で最大 1.8Gbps のビットレートまで可能な高画質化指向のプロファイルである。コアスタジオプロファイルは、P ピクチャー機能を追加して高圧縮率指向に対応している。なお、MPEG 2 / 4 : 2 : 2 プロファイルとの互換性を考慮し、符号化ビットストリームから直接変換可能な符号化形式となっている(符号化テーブルの交換のみで変換可能)。

次世代圧縮フォーマット H.264/MPEG 4 Part.10 の概要

MPEG 技術は進化を続けており、これまでに MPEG 4 ビジュアルの審議において主にコスト面で採用が見送られてきた技術を集大成し、現時点での究極の圧縮効率を追求する規格として、H.264 (AVC=Advanced Video Coding) が制定された。もともと ITU-T における H.263 の後継規格として、H.26L というコード名で検討されてきたが、前述のように ISO との合同検討グループ JVT へと発展し、H.264/MPEG 4 Part.10 として規格化された。

前述した MPEG の基本フレームワークの上に、以下のような技術改良が行われている。

①動き補償フレーム間予測

従来の 16×16 画素または 8×8 画素ブロック単位の予測に対し、16×8、8×4、4×4 など 4 画素単位で縦横比可変ブロックの予測が追加された。また、ブロックごとに参照するフレームを切り替えることができる(従来は、同一のフレームだけしか参照できなかった)。

②フレーム内画素予測

従来は、DCT 変換後の DC および AC 係数の予測は行っていたが、H.264 では画素値そのものをフレーム内の符号化済み画素から予測するモードが追加された。

③整数値 DCT

DCT は、本来実数値変換であるが、これを近似的に整数値の変換に置き換える。ブロックサイズも、4×4 が基本となった(従来は 8×8 画素 DCT)。処理の高速化に加え、ブロック歪み減少の効果があるとされる。

④エントロピー符号化

CAVLC (Context Adaptive Variable Length Coding) という方式が採用された。DCT 係数の値に対し、符号割当の効率を高くできる。また JPEG2000 で採用された算術符号を用いる CABAC (Context Adaptive Binary Arithmetic Coding) も新たに追加された。

以上のような技術改良により、MPEG 4 ビジュアル規格に対して 1.5～2 倍の圧縮効率向上が見込まれている。ただし、圧縮ツールの処理が大幅に複雑化しており、デコーダーで 1.5 倍程度、エンコーダーは 15 倍程度処理量が増加すると考えられる。

※参考文献

デジタル放送教科書 (IDG ジャパン)

MPEG 4 スタジオプロファイル(映像メディア誌/Vol.55 No.6)
H.264/AVC 教科書(インプレス)

■注目の圧縮技術「ウェーブレット技術」とは

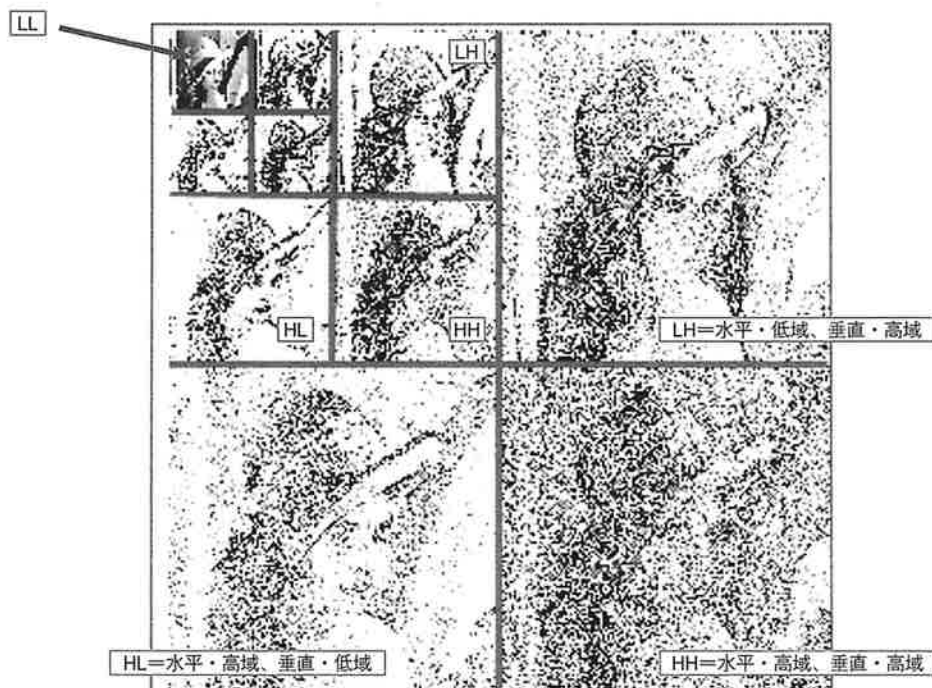
画像圧縮の分野ではウェーブレット技術が注目されるキーワードの1つとなってきた。ウェーブレット (Wavelet) とは、「小さな波」あるいは「さざ波」を意味する言葉であるが、ウェーブレット波形を用いた時間周波数解析技術として、1980年台初頭にフランスの石油探索技師J. Morletが、鉦脈探索に利用する人工地震波解析のために考案した技術がもとになっている。

周波数解析技術であるため、フーリエ変換やコサイン変換などの技術範疇に含まれるが、フーリエ変換などがすべての周波数帯域について同じ解像度となるのに対し、ウェーブレット変換は高い周波数では高い解像度で、低い周波数では低い解像度で解析ができるという特徴がある。

画像圧縮 (画像符号化) への応用でも高い空間周波数の領域 (絵柄の細かい部分) は解像度を細かく、低い周波数領域 (絵柄が平坦な部分) は解像度を粗くして圧縮することができ、人の視覚特性によく合う変換方式であるといわれている。このほかFBIで用いられている指紋検索システムや最近では経済時系列解析による株式投資への応用など、幅広い分野への適用が考えられている。

静止画圧縮符号化の最新標準方式であるJPEG2000では、この特徴を利用するため離散ウェーブレット変換 (DWT) を利用している (従来のJPEGは、離散コサイン変換=DCTを利用)。この結果、特に圧縮率が高いときにJPEGより優れた画像品質が得られるとされ (ブロック歪みが発生しない)、JPEG2000の最大の利点の1つとなっている。JPEG2000に使われているDWTは、オクターブ分割フィルターという方式であり、図Aに示すよう水平・垂直方向とも低域フィルターをかけた結果 (LL) は、1/2の解像度にダウンサンプルされ、これを繰り返すことで低周波領域ほど解像度が小さくなる。

なお、MPEG 4にも静止画オブジェクトの規定があり (Still Picture Object)、このモードではDWTが使われている。MPEG 4規格検討の初期段階では動画へのDWT適用が検討されたが、結論としてMPEG4ビジュアル規格の動画 (自然画) としてはDCTが採用され、DWTは静止画モードへ適用されるという棲み分けが行われる形となった。しかし、ノンリニアメーカー独自規格のHDコーデック圧縮方式としてDWTの採用がいくつかあり、今後のビデオ圧縮技術として期待がもたれる。



図A DWT (オクターブ分割)