

圧縮入門

目次

- 1 圧縮についての基礎知識
- 3 圧縮の基礎を理解する
- 3 圧縮の種類
- 4 フレームの種類と圧縮
- 4 カラーモード
- 5 色深度
- 5 カラーサンプリング
- 6 オーディオ圧縮
- 7 圧縮効率
- 7 後工程での圧縮に備えたビデオ制作
- 10 ビデオのキャプチャ：形式および接続方法の選択
- 13 適切なキャプチャコーデックの選択
- 14 ビデオの書き出し
- 14 前処理
- 16 クロップ
- 16 スケール
- 17 ノイズリダクション
- 17 イメージ処理
- 19 オーディオ処理
- 19 エンコーディング
- 21 フレームレートの調整
- 23 ビデオの配信方法
- 24 圧縮形式の選択
- 25 Windows Media 形式への書き出し
- 26 Windows Media Player
- 27 Windows Media ビデオのコーデック
- 28 Windows Media オーディオ (WMA) コーデック
- 29 RealMedia 形式への書き出し
- 29 RealMedia ビデオコーデック
- 29 RealMedia オーディオコーデック
- 30 QuickTime 形式への書き出し
- 31 QuickTime ビデオコーデック
- 35 QuickTime オーディオコーデック
- 37 MPEG-1 形式への書き出し
- 38 MPEG-1 ビデオ / オーディオコーデック
- 38 DVD/MPEG-2 形式への書き出し
- 39 MPEG-4 オーディオコーデック
- 40 MPEG-4 形式への書き出し
- 41 MPEG-4 オーディオコーデック
- 42 AVI 形式への書き出し
- 42 AVI ビデオコーデック
- 43 AVI オーディオコーデック

圧縮についての基礎知識

この入門では、高品質な圧縮されたビデオを準備する際に参考となる内容を全編にわたって掲載しています。最良の結果を得るにはまず、圧縮を行う理由とその方法を理解することが重要です。圧縮の数学的基礎は非常に複雑ですが、基本原理は簡単に理解できます。LAN やインターネット経由でデータを効率的に転送するには、非圧縮ビデオファイルはサイズが大きすぎ、データレートも高すぎるため、圧縮する必要があります。

ビットストリームの種類	データレート (キロビット / 秒)	1 枚の 700MB CD-ROM での時間
非圧縮 HD (1920 × 1080 29.97fps)	745750	7.5 秒
非圧縮 SD (720 × 486 29.97fps)	167794	33 秒
DV25 (miniDV、DVCAM、DVCPRO)	25000	3 分 44 秒
標準的 DVD	5000	18 分 40 秒
VideoCD	1167	80 分
ブロードバンド Web ビデオ	100-2000	3 時間 8 分 (500Kbps)
モデム Web ビデオ	18-48	48 時間 37 分 (32Kbps)

DV メディアの種類ごとのデータレートと必要となる記憶領域

ビデオの制作に使用されるデータレートと圧縮ビデオの配信に使用されるデータレートとの間には、大きな開きがある場合があります。例えば、デジタルビデオ (DV) から DVD に変換する場合、通常、5:1 の圧縮が必要となり、非圧縮ソースをモデム互換ビデオに変換する場合は、約 9000:1 の圧縮が必要となります。つまり、元のデータの約 0.1% しか残りません。Web 上のビデオの画質や音質の高さに驚くかもしれません。

また、圧縮ビデオの種類によってビットレートは異なります。例えば Web ビデオの場合、ローエンドのダイヤルアップモデムレートとハイエンドのブロードバンドレートでは、約 100 倍の差があります。

ビデオの種類 (形式) と圧縮の種類は別物です。1 つのビデオ形式で複数の圧縮の種類と圧縮率がサポートされる場合があり、1 つの圧縮の種類が複数の形式で使用される場合があります。ただし、多くの標準的なビデオ形式の要件には、圧縮の種類とデータレートが含まれています。

DVD 形式

米国で最も多く鑑賞されているビデオのメディアは DVD の圧縮ビデオです (その次に多く鑑賞されているのは、CATV や衛星で配信されるデジタル放送です)。

最新の DVD-R レコーダでは、1 枚のディスクに 4.7GB のデータを書き込めます (DVD-5 と呼ばれます)。プレス生産された DVD は、2 層形式を使用して 8.54GB のデータを記録できます (DVD-9 と呼ばれます)。DVD には数多くの種類があり、より大容量の形式が次々と公開されています。DVD-5 ディスクには 5GB 近いデータを記録できます。これは高品質の 2 時間のムービーには十分な容量です。現在では、民生用の DVD レコーダでより大容量のディスクへの書き込みも行えるようになっています。

DVD プレーヤーで再生可能な DVD の種類は、厳密には DVD-Video と呼ばれるものです。また、DVD-ROM もあります。これは、再生にコンピュータが必要となるコンピュータ用にフォーマットされたディスクで、実際には大容量の CD-ROM と言えます。ハイブリッドタイプの DVD-Video/DVD-ROM ディスクもあります。これは、通常の DVD プレーヤーで再生できるだけでなく、コンピュータで再生したときのみ利用可能な情報も含まれています。

DVD-Video ディスクは、NTSC 方式 (National Television System Committee 方式の略で米国および日本で使用されている方式) または、PAL 方式 (Phase Alternating Line 方式の略でヨーロッパおよび多くのアジア諸国で使用されている方式) でマスタリングされます。一方の方式を再生できるように設計された汎用プレーヤー (家庭用 DVD プレーヤーなど) は、通常もう一方の方式のディスクを再生できませんが、コンピュータでは通常どちらの種類も再生できます。DVD にはリージョンコードを設定することもできます。この場合、特定の地域向けの DVD プレーヤーかソフトウェアでしか再生できません。リージョンコードの設定はオプション機能です。

VideoCD 形式

VideoCD は古いディスクベースの形式です。VideoCD の解像度およびデータレートは DVD に比べてはるかに低いため、提供される品質は低くなります。ただし、標準の CD メディアを使用するため、大部分のパーソナルコンピュータなど様々な機器で再生可能です。プレス生産された VideoCD は DVD プレーヤーでも再生できます。

DVD と同様に、VideoCD ディスクは、NTSC または PAL になります。PAL を採用している国、特にアジア諸国では、VideoCD は非常に人気があります。DVD と同様にソフトウェアプレーヤーは両方の方式をサポートしていますが、セットトッププレーヤーの多くはいずれか片方しかサポートしていません。

CD-ROM 形式

1990 年代前半から、CD-ROM ビデオが人気のメディアになりました。インタラクティブコンテンツ用のメディアとしては Web がほぼ CD-ROM に取って代わりましたが、コンピュータゲームやキオスクなど多くのアプリケーションでは現在も CD-ROM ビデオが使用されています。Web ビデオを比較すると、CD-ROM はブロードバンド接続よりはるかに高いバンド幅を提供するため、高速のコンピュータでは DVD の品質を実現できます。

インターネット形式

Web は、圧縮技術の開発が最も盛んなメディアです。ディスクに比べて Web で使用可能なバンド幅ははるかに小さいため、すべてのビットを有効に使用することが重要です。また、DVD などの民生用電子機器とは異なり、Web を使用すれば開発者はプレーヤーソフトウェアをいつでもバージョンアップできます。そのため、Web は短期間で最良の技術を激しく競い合う分野になりました。

Web ビデオは、リアルタイムストリーミングモードまたはプログレッシブダウンロードモードのいずれかを使用できます。リアルタイムストリーミングファイルは専用のストリーミングサーバーソフトウェアが必要になり、サーバーとエンドユーザー間で使用できるバンド幅によって品質が制限されます。プログレッシブダウンロードファイルは、標準的な Web サーバーから配信されるため、ダウンロードに時間がかかる場合がありますが、より高品質で信頼性の高いビデオおよびオーディオ品質を提供できる場合があります。これらのモードの違いとそれぞれの長所および短所を理解することがこの文書の主要なテーマの 1 つです。

圧縮とモバイル機器

携帯情報端末 (PDA) や最新の携帯電話などのモバイル機器は、圧縮ビデオでは比較的新しいメディアです。ラップトップコンピュータと比べてこれらの機器の画面は小さくプロセッサも低速ですが、非常に軽量の機器でビデオを再生できることは大きな価値を持つ可能性があります。

Adobe® Flash® Video、MPEG-4、Microsoft® Windows Media、RealMedia など、圧縮ビデオの大部分の形式はいくつかのモバイル機器で再生できます。Kinoma などのモバイル専用形式も使用できます。

多くの点で、モバイル機器、特に携帯電話のビデオは技術的にまだまだ発展段階であり、デスクトップコンピュータのビデオの黎明期を思い起こさせます。現在、モバイル機器には汎用の形式や設定はなく、少数のデバイスクラスや、場合によっては特定の製品モデルに対して、コンテンツを最適化する必要があります。ただし、最適化がうまくいけば、すばらしい結果を得られる可能性があります。

圧縮の基礎を理解する

圧縮ツールおよび圧縮形式は圧縮に関する複雑な数学的処理によって実現されていますが、これはユーザから見えないところで行われています。ただし、最良の結果を得るには、圧縮ツールおよび圧縮形式の動作原理の基礎を理解しておく必要があります。

視覚と聴覚の仕組み

人間の視覚と聴覚は非常に優秀ですが、それには限界があり、それを利用して圧縮が行われます。人間の目と脳は、激しい動きと細部を同時に感知したり、激しい動きと色を同時に感知する能力に長けています。同様に、聴覚も非常に優秀ですが、また同様に限界があります。静かな環境では簡単に音を聞くことができますが、騒々しい環境では音は聞き取りにくくなります。人間の耳は人間の声の周波数に合わせて最適に調整されており、それより高い音や低い音の認識は苦手です。圧縮技術では、人間の感覚のこうした長所と短所を利用して、人間が敏感に感知できないような部分のデータを削減しています。

圧縮の仕組み

基本的に、圧縮は、連続したイメージおよび音声ができる限り効率的に要約することで機能します。圧縮済みのビデオのユーザに対する表示は、圧縮前のファイルとは同一ではありません（同一にするにはデータレートが非常に高くなってしまいます）。そうではなく、エンコードされたファイルは、コンテンツの重要な細部の表現ができる限り簡略化すると同時に、人間の知覚システムが元のコンテンツをできる限り正確に体験できるような再生を行います。

圧縮は、コンプレッサとデコンプレッサの一致するペアによって実行されます（コーデックという名前がよく知られています）。コンプレッサは圧縮処理の一部を担うもので、ビデオの保存に必要なデータ量を削減します。デコンプレッサはデータの受信側で動作し、圧縮データをデコードしてビューアで表示可能にします。デコーダ（デコンプレッサ）で、エンコーダ（コンプレッサ）のデータの要約方法が認識される必要があるため、このペアが一致していることが重要です。

圧縮の種類

ビデオコーデックは、空間圧縮、時間圧縮またはその両方の組み合わせを使用できます。

空間圧縮

空間圧縮（イントラフレーム圧縮とも呼ばれます）は、一度に1つのフレームにしか影響を与えません。多くのコンプレッサは空間圧縮しか使用しないため、フレーム間は相互に依存しません。この種類の圧縮では、ビデオ内の任意の時間位置にランダムアクセスが可能で、編集も容易です。

空間圧縮は、フレーム内の冗長性（単色または近い色が集まった領域か、フレームのほかの領域と似ている領域）を検出し、これらの類似領域を（各ピクセルすべての色属性を定数として記述するのではなく）参照によってエンコードすることによって領域を節約します。

多くのコーデックは、離散コサイン変換（DCT）という技術に基づいています。こうしたコーデックは、イメージ内の滑らかなグラデーションを処理するのは非常に得意ですが、ランダムな細部やシャープなエッジを低ビットレートでエンコードする処理は苦手な場合があります。低すぎるデータレートでイメージをエンコードしようとすると、イメージの表示品質が落ちる場合があります。この品質低下により、斑点と呼ばれる表示上の問題が発生します。斑点の最も一般的なものには2種類あります。共鳴（シャープなエッジにハローが出たりにじんだりする場合があります）とブロックノイズ（滑らかなエッジがぎざぎざに表示される場合があります）です。問題がよく発生する領域の1つは、シャープなエッジを持つ小さいテキストです。ウェーブレットのような一部のコーデックは、まったく異なる技術を基盤としているため、シャープなエッジをより適切に処理できます。

時間圧縮

時間圧縮（インターフレーム圧縮とも呼ばれます）は、現在のフレームで他のフレームを参照する機能を追加するものです。ビデオでは、各フレームは直前のフレームと類似している傾向があります。時間圧縮は、フレーム間で変化した部分だけをエンコードします（以下で説明するキーフレームは例外です。キーフレームはフレーム内のデータの完全な表現が含まれており、直前のフレームを参照していません）。

時間圧縮は、研究上の注目を集めており、今後さらに向上の余地があります。

フレームの種類と圧縮

ファイル内の各フレームの種類を見れば、圧縮に関する多くの基本概念が理解できます。コーデックおよび使用されているオプションに応じて、1～3種類のフレームがファイル内に存在する可能性があります。

キーフレーム (I フレームとも呼ばれます) は、イントラフレーム圧縮を使用した自己完結型のイメージです。I フレームは、ファイル内のほかのどのフレームにも基づいていません。ムービーの最初のフレームは必ず I フレームになります。

デルタフレームは、インターフレーム圧縮を使用して圧縮され、前後のフレームとは異なる部分だけが含まれます。2つのフレームが数学的にまったく同一で、2番目のフレームが P フレームの場合、エンコードするビットの数はほんの少しで済みます。逆に、2つのフレームが半分以上異なるものである場合、多くのコーデックは自動的に I フレームを新規に生成します。通常のコンテンツの場合、I フレームと同じ品質の同じデルタフレームのサイズは、I フレームの 10～20%にしかなりません。

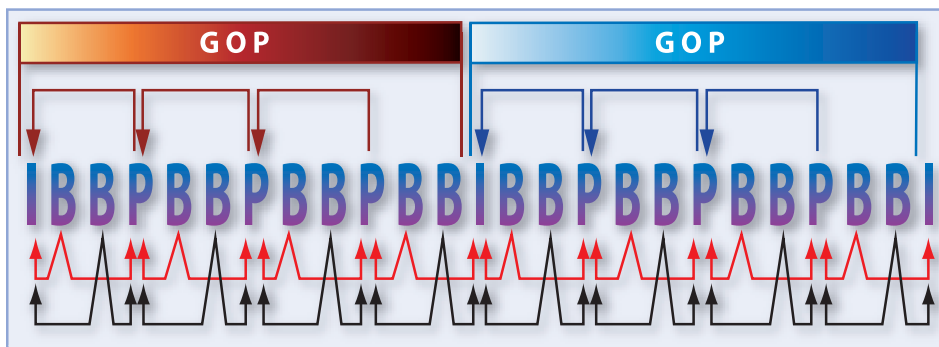
デルタフレームには主要な 2つの種類があります。それは P フレームと B フレームです。P フレーム (Predictive frame の略で予測フレームという意味) の内容は、前のフレームに基づいており、B フレーム (Bi-directional frame の略で双方向フレームという意味) の内容は、前と後のフレームに基づいています。

I フレームは、単独でデコードできる唯一のフレームであるため、ビデオのランダムアクセスを行う点として重要です。デルタフレームにアクセスするには、そのフレームだけでなく、デルタフレームが参照している I フレームもデコードする必要があります。I フレームの出現頻度が高くなると、ランダムアクセス操作の際にデコードする必要のあるフレームの数は減少します。

B フレームは主に MPEG 形式で使用されます。B フレームでは、その前後のフレームのイメージ要素を使用できるため、圧縮効率が向上します。ただし、B フレームでは他の B フレームのイメージ要素を使用できません。この制限により、フレームが、最初のフレームを基準とするフレームを基準とすること (不可能な再帰ループ) を回避できます。

モーション JPEG や DV など、多くの制作コーデックは I フレームしか使用しません。他のフレームを基準とする必要なしに、フレームを再構築できるため、I フレームを使用した編集は容易です。

コーデックによってフレーム間の変化をエンコードする方法が異なります。例えば、フェードアップとフェードダウンを非常に苦手とするコーデックもあれば、新しいコーデックの中にはフェードを処理するためのコードが組み込まれているものもあります。



NTSC DVD で使用される標準の 15 フレームの Open GOP。Open GOP では、末尾の B フレームは次の GOP の最初の I フレームを参照できます。

カラーモード

難解に思えるかもしれませんが、実際、コーデックが色を表現する様々な方法は大きな問題となります。こうしたことがビデオの世界の用語を分かりにくくしているかもしれません。

現在のコーデックに 3種類の基本的な色があります。

- **モノクロ** モノクロモードでは、ルミナンス (輝度) だけが問題となります。モノクロは、白黒ムービーに最適です。多くのビデオコーデックには明示的なグレースケールモードはありません。注目すべき例外は、Cinepak と JPEG です (ただし、すべてのツールで色空間制御が公開されているわけではありません)。
- **RGB (Red, Green, Blue)** コンピュータの画面およびテレビは RGB がネイティブのモードです。つまり、CRT の場合は赤、緑、青の蛍光体が塗布され、LCD の場合は Planes が使用されています。RGB はビデオコンテンツにはあまり適していないため、RGB コーデックはブロードキャストビデオ配信には使用されていません。しかし、RGB は画面の記録やモーショングラフィックに広く使用されています。

- **Y'CbCr** (YUVとも呼ばれますが、YUVは厳密にはNTSCビデオでの色の使用方法だけを指す用語です) これらのコーデックはルミナンス(輝度)と2つのカラー差チャンネルをエンコードします(Adobe Photoshop® LABモードで使用されている技術と同じです)。色は3つあるため、Y'CbCrをRGBに変換して表示できます。クロマチャンネルとルミナンスチャンネルを分離することで、重要なルミナンス情報を最大解像度で維持し、クロマ情報により低い解像度を使用できます。この処理により、データレートと処理要件を大幅に削減できます。RGBとY'CbCr間である程度レベルの変換を行うことはできますが、一方のモードで使用できてももう一方のモードで使用できない色が常にあるため、変換の一部として常に丸め誤差が生じる可能性があります。Adobe After Effects®などの1チャンネル当たり16ビットのRGBをサポートするツールで作業すれば、この変換における品質が大幅に向上します。

色深度

色深度は色の精度を定義し、輝度(明度)の階調数を決定します。色深度は、チャンネルあたりのビット数(bpc)で表され、次のカテゴリに分類されます。

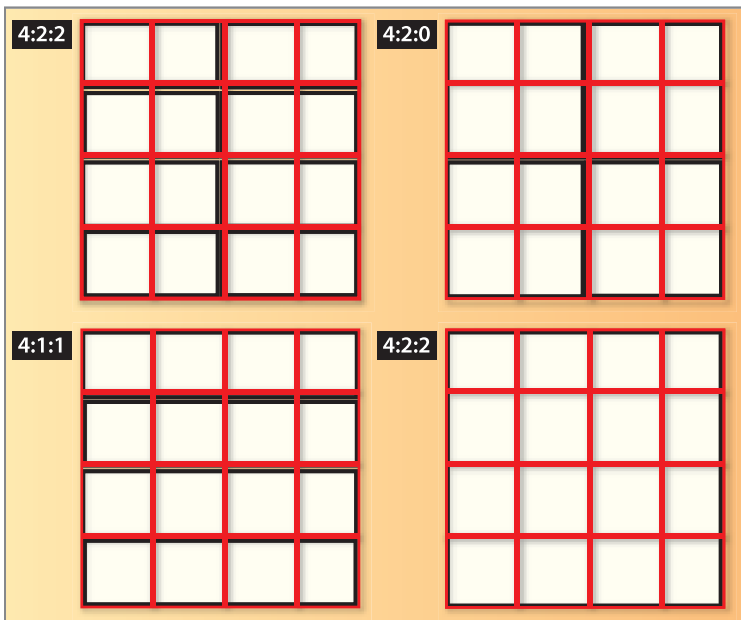
- **8ビット(インデックス)カラー**(通常8ビットまたは256色と呼ばれます)初期のカラーコンピュータではインデックスカラーが使用されていました。各ピクセルが、赤、緑、青の測定値を持つのではなく、各ピクセルにコードが割り当てられ、ルックアップテーブルでそのコードを実際のRGB値に変換してコンピュータで表示していました。4色または16色など、さらに低いモードをサポートして、イメージの保存に必要なビット数を削減することで、非常に単純な圧縮技術を提供するコーデックもありました。ビデオでは最も単純な場面でも256色を超える数の色を使用するため、8ビットの色深度はビデオには適していません。主要な形式で、現在でも8ビットインデックスカラーを必要とするのはGIFだけです。
- **5bpc**(数千色またはHigh Colorとよく呼ばれます)8ビット時代の後、最も一般的な民生用のビット数は5ビットになりました。Macintosh®コンピュータでは、赤、緑、青のそれぞれに5ビットずつが割り当てられ、残りの1ビットはアルファチャンネル用に予約されていました。Microsoft Windows®では、残りの1ビットは緑に割り当てたため、ルミナンス(輝度情報)の質は向上しましたが、アルファチャンネルのための余地はなくなりました。5ビットカラーモードは一部の非常に初期のビデオ圧縮コーデックに使用されていました。現在、5bpcが主に使用されているのは、8bpcよりも優れた圧縮およびパフォーマンスを、極端な品質低下なしに実現できる可能性のあるスクリーン記録アプリケーションです。
- **8bpc**(数百万色またはTrue Colorとも呼ばれます)24ビットモードまたは32ビットモードとも呼ばれる場合もあります。32ビットモードは8bpcですが、アルファチャンネルが追加されています。
- **10bpc**このモードは、ハイエンドY'CbCr制作システム(デジタルベータカムや、D1など)および一部のアナログキャプチャカードなど使用されています。このモードは256階調ではなく1024階調を提供するため、特にエフェクトを繰り返し生成した場合のバンディング(グラデーションが滑らかでない状態)を削減できます。このモードは配信形式では使用されていませんが、一部の制作コーデックではこのモードがサポートされています。After Effectsでの10ビットY'CbCrモードと16bpc RGBとの間の変換は、8ビットカラーの4倍の精度があります。
- **16bpcおよび32bpc**一部のハイエンドアプリケーションは、16bpcや32bpcを使用してより多くの色の詳細を保存します。これらのモードは、複数ステップの制作プロセスがある場合、ビジュアルエフェクトをイメージに適用する必要がある場合、Y'CbCrとの間で変換を行う場合に特に便利です。After Effects 6およびそれ以降のバージョンでは、16bpcカラーの処理が行えます。After Effects 7.0では、32bpcカラー機能が導入されました。16bpcまたは32bpcで保存されたビデオは、現状ではエンドユーザーには配信されていません。

カラーサンプリング

Y'CbCr形式では、カラーサンプリングがサポートされています。カラーサンプリングが記録するクロマ(カラー)情報のピクセル数はルミナンス(輝度)情報よりも少なくなります。カラーサンプリングの用語ではX:Y:Z形式が使用されるため、わかりにくくなっています。最初の値は、表現されるルミナンス値サンプルの数で、たいいていの場合この値は4です。2番目の値は、1ライン目、3ライン目、5ライン目および後続の奇数ラインのルミナンス値サンプルごとのクロマ値サンプルの数です。したがって、この値が2の場合、4つのルミナンス値サンプルにつきクロマ値サンプルが2つずつ水平方向に存在しています。最後の値は、2ライン目、4ライン目、6ライン目および後続の偶数ラインのルミナンス値サンプルごとのクロマ値サンプルの数です。この数が0の場合、最初の行のクロマ値サンプルが使用されます。

- **4:4:4** サブサンプリングを行わずに、各ピクセルでクロマ値がサンプリングされます。この形式は多数のビットを使用しますが、品質が顕著には向上しないため、Y'CbCr配信コーデックではこの形式を使用しません。制作アプリケーションによっては、内部的にこの形式を使用するものもあります。RGBは常に4:4:4です。

- **4:2:2** この形式は、DV50、デジタルベータカム、DI、モーション JPEG など、ハイエンドビデオ制作に使用される形式です。少数ですが、H.263 などの配信コーデックでもこの形式が使用されています。
- **4:2:0** 大部分の配信コーデックはこの形式が使用されます。2 × 2 ピクセルのブロックにつき 1 つのクロマ値サンプルが存在しています。このクロマ値のサンプルの表示品質は 4:4:4 と同等のレベルですが、フレームごとのビット数は半分です。この形式は PAL DV でも使用されます (PAL DVCPRO では使用されません)。
- **4:1:1** NTSC DV および PAL DVCPRO では、この形式が使用されます。この形式にはフレームごとに 4:2:0 と同数のビットが含まれますが、4 × 1 ピクセルのブロックにつき 1 つのサンプルが存在します。この形式は、フィールドごとに独自のサンプルが存在しているため、インターレースコンテンツに適していますが、飽和度の高い領域ではクロマ斑点が発生する可能性があります。そのため、ブルースクリーンやモーショングラフィックスソースで問題が発生する可能性があります。
- **YUV-9** この形式では、4 × 4 ピクセルのブロックにつき 1 つのクロマ値サンプルが存在しています。この形式では著しく品質が低下するため、最新のコーデックでは使用されていません。この形式を使用する古いコーデックには、Indeo ファミリーや Sorenson Video のオリジナル版などがあります (Sorenson Video 3 では、4:2:0 が使用されています)。



最も一般的な種類のカラーサンプリング

オーディオ圧縮

オーディオは圧縮ファイルのビットをほんのわずかしか使用しないものの、ユーザに提供する内容の半分を担うものです。

サンプリングレート

オーディオの最も基本的な要因はサンプリングレートであり、単位はヘルツ (Hz) またはサイクル / 秒です。オーディオ CD は 44100Hz (44.1KHz) を使用し、優れた再生品質を提供しています。

サンプリングレート	用途	十分な音質を提供できる対象
8KHz	電話	音声
11KHz	古いマルチメディア標準	より品質の高い音声
22.050KHz	同上 (古いマルチメディア標準)	聞き取れるという程度の音楽
32KHz	一般的なブロードバンドビデオ	最低品質の音楽
44.1KHz	オーディオ CD	音楽
48KHz	DVD、DAT	音楽
96KHz	ハイエンドオーディオレコーディング	音楽のマスタリング

一般的に使用される標準サンプリングレート

ビット数

ビット数は、各サンプルでの各チャンネルのビット数です。ビット数の値が高いほど、オーディオの品質が高くなります。

- **8ビットサンプリング** もともと、マルチメディアオーディオは8ビットサンプリングを使用していました。つまり、各サンプルごとにボリュームを0～255の値で測定します。このビット数（初期のCD-ROMタイトルのざらついた音質の原因はこのビット数にあります）では、高品質のオーディオは望めません。8ビットオーディオを配信する理由はもはやありません。最新の16ビットコーデックは、8ビットコーデックと比べて、より小さいファイルサイズでより高い品質を実現できます。
- **16ビットサンプリング** このビット数は、オーディオ配信の現在の標準となっています。大部分の最新のコーデックは16ビットサンプリングを使用しています。また、オーディオCDも16ビットサンプリングです。
- **20ビットサンプリング** このビット数はプロフェッショナルオーディオレコーディングで使用され、16ビットよりも優れたダイナミックレンジを提供します。24ビットオーディオ（後述）をサポートしているコーデックは20ビットソースの利点を活用できます。
- **24ビットサンプリング** このビット数は、16ビットに比べてより多くの音の詳細を提供できるため（エラーを発生させずにより多くのエフェクトやレイヤーを使用できるため）、急速にプロフェッショナルオーディオレコーディングの標準になりつつあります。現在、24ビットをサポートする配信コーデックもわずかながら存在し、最も有名なものでは、Microsoft Windows Media Audio (WMA) があります。16ビットサンプリングと24ビットサンプリングの差は、一般的な家庭用コンピュータのスピーカシステムでは簡単には聞き分けられませんが、パソコンをホームシアターに統合すると、問題が出てくる可能性があります。

チャンネル

コンピュータのオーディオのほとんどは、単一チャンネルのモノラルか、2チャンネルのステレオです。低ビットレートでは、一般的に必要なビット数の少ないモノラルが使用されます。業界では徐々にマルチチャンネル（5.1や7.1）オーディオサポートも使用されるようになってきました。マルチチャンネルサポートは、現在Microsoft WMA9 Pro および RealNetworks RealAudio 10 に含まれています。もちろん、DVD は以前から Dolby Digital を使用したマルチチャンネルがサポートされています。

知覚的オーディオ圧縮

サンプルレートとチャンネルの数を削減することでデータレートを削減できますが、高い圧縮率を得るにはそれ以上のことが必要です。大部分の最新のコーデックは知覚的エンコーディング技術を使用しています。この技術では、人間の聴覚システムの内部モデルによってオーディオソースの重要部分を決定し、その部分に対して使用可能なビットを割り当てます。低データレートでは、オーディオの音がソースとまったく異なるように聞こえる場合があります。「十分に良い」品質に必要なビットの数はここ数年で急激に減少しています。

圧縮効率

それでは、コーデックを比較するにはどうすればよいでしょうか。これは、基本的には圧縮効率を測定することで比較します。圧縮効率は、特定の品質レベルに達するために必要なビットの数によって決まります。圧縮効率を高めるには、品質を変えずにデータレートを下げるか、データレートを変えずに品質を上げるか、またはその中間のポイントを選択するかのいずれかになります。

単一の数値だけで圧縮効率を測定することはできません。相対的パフォーマンスはソースとデータレートによって大きく変化します。通常、データレートが高いほど品質の差は小さくなります。極端なデータレートと複雑なコンテンツを使用すると、圧縮効率に差が出ます。

後工程での圧縮に備えたビデオ制作

高品質な圧縮を行うための鍵の1つは、適切に圧縮できるコンテンツを作成することです。こうした種類のコンテンツを作成できない場合もあります（古いコンテンツはそのまま使用するほかありません）。しかし、圧縮に適したコンテンツを作成する場合は、様々な技法が存在し、これらを適用することでより良い結果を得ることができます。これらの技法は、プリプロダクション、プロダクションまたはポストプロダクション中に適用できます。

プリプロダクション

プリプロダクションでは、プロジェクトの様々な要素について計画を立てることになります。プロジェクトの初期段階で圧縮のアイデアを取り入れて、チームのメンバー全員がそれを考慮できるようにすることをお勧めします。

プロダクション

プロダクションで、ビデオが効率よく圧縮されるようにする技法があります。

- **インターレーススキャン対プログレッシブスキャン** ビデオカメラで撮影するコンテンツの場合、インターレーススキャンビデオとプログレッシブスキャンビデオのどちらを使用するかは、基本的な決定事項の1つです。従来、ビデオは常にインターレース方式でした。インターレースフレームでは、奇数ラインと偶数ライン（フィールドと呼ばれる）に、フレームの半分ずつのイメージを時間をずらして取り込みます。各ビデオフレームは2つのフィールドで構成されます。

ビデオにはなぜフィールドがあるのでしょうか。最初のテレビは、1つの電子銃を使用して、ブラウン管の内面に安定した電子ビームを照射していました。この電子が、ブラウン管の内側にコーティングされている蛍光体を励起させ、イメージを表示させていました。電子銃は、画面の端から端に移動して、走査線と呼ばれる1本の線を発光させます。次に1ライン下がって、再度端から端に移動します。電子銃は、イメージを構成する486ラインに対してこの動作を行うこととなります。ここで問題となるのは、電子銃が画面の下半分を走査し始めているときには、上の方のラインは既に暗くなり始めているということでした。

この問題を解決するために、486ラインをodd（奇数）とeven（偶数）と呼ばれる（つまり奇数ラインと偶数ラインで構成される）2つのフィールドに分割しました。これにより、電子銃は1ラインおきに走査することで、1つのフィールドなら上から下まで完全に表示できるようになりました。次に、空いているラインを上から埋めていきます。その結果、一度に消えるイメージは半分だけであるため、人間の目には完全なイメージに見えます。この奇数ラインと偶数ラインの「インターレース」は、インターレーススキャンまたはフィールドスキャンモードと呼ばれるようになりました。

NTSC ビデオのフレームレートは 29.97 フレーム / 秒 (fps)（または約 30fps と表現される場合もあります）で、2 フィールド / フレーム (59.94 フィールド / 秒) です。PAL ビデオでは 25fps (50 フィールド / 秒) です。その結果、2 番目のフィールドには、NTSC では約 1/60 秒後、PAL では約 1/50 秒後のイメージが表示されます。ビデオで動きが発生した場合、一方のフィールドのイメージはもう一方のイメージと少しずれることとなります。これは、VCR (ビデオデッキ) で再生を一時停止すれば確認できます。多くの VCR は、フレーム上で一時停止し、そのフレームを構成する 2 つのフィールドを交互に繰り返します。このことにより、フィールドジッターと呼ばれる現象が発生します。

フィールド間にこうした差があることで、圧縮プロセスはより複雑になります。優れたビデオ圧縮ツールは、フィールドを読み込んで理解し、独立して圧縮し、圧縮形式でフィールドとして再生できるようにする必要があります。一般に、Web 用の圧縮では、コンピュータモニタがフィールドベースではないため、フィールドを表示しません。高度なコンピュータ技術を使用すると 1 回の走査で画面全体を表示できます（通常、ブラウン管ディスプレイでは複数の電子銃を使用します）。液晶モニタでは電子銃は使用されず、瞬時に画面全体を表示します。この処理はプログレッシブスキャンと呼ばれます。

インターレースビデオは、プログレッシブスキャンモニタではきれいに表示されません。通常、Web 用の圧縮技術ではインターレーススキャンをプログレッシブスキャンにビデオを変更しています。この変更を行う最も基本的な方法は、一方のフィールドを破棄し、もう一方のフィールドを複製して完全なフレームを作成する方法です。この処理では元のデータが 50%削減されるため、最終品質に悪影響を与える可能性があります。より複雑な方法では、エンコーダが両方のフィールドを読み込み、フィールド間の時間的な差を滑らかにするようにピクセルを変更して 1 つのフレームに補完（ブレンド）します。この方法を使用すると、ビデオからの元のデータをすべて使用しつつ、より高品質のプログレッシブスキャンモニタ用の出力を作成できます。

- **カメラの動き** カメラの動きの量および種類はコンテンツの圧縮性に大きな影響を与えます。カメラが動くとき全フレームの全ピクセルが変更されます。もちろん、カメラの動きがないのが最も単純ですが、最新のコーデックはドリリーを使用して作成したようなカメラの動きを非常に適切にエンコードできます。チルトやパンはさらに難しくなります。カメラを回転させること（ダッチ）にほとんど影響を受けないコーデックもあれば、圧縮効率が極端に低下するコーデックもあります。コーデックにとって最も難しい動きは不規則に揺れる手持ちカメラの動きです。

圧縮処理では（特に低いデータレートでは）、カメラの動きの処理より、ビューの変更の処理の方がはるかに容易です。ポストプロダクションの際に、After Effects のモーションスタビライザツールを使用して圧縮処理前に不要なカメラの動きを取り除き、出力ビデオファイルのサイズを小さくすることができます。

- **シャッタースピード** これはカメラの絞りが開放される時間、つまり、光がフィルムまたは CCD に当たる時間を制御するものです。シャッタースピードを遅くすると、より多くの光が通過し、粒状性を抑えることができるため、圧縮効率が高くなります。また、シャッタースピードを遅くすると、イメージの移動中の部分がぶれるため、圧縮がより容易になります。

大部分のカメラのデフォルトのシャッタースピードはフレームレートの半分、つまり、NTSC では 1/60 秒、PAL では 1/50 秒、フィルムでは 1/48 秒です。大部分の圧縮用ビデオについては、デフォルトで良い結果が得られます。こうしたデフォルトの例外は、アクションおよびスポーツのビデオ撮影であり、ビデオ撮影者ははっきりとしたスローモーションイメージを得るためにモーションブラーの発生を避けます。圧縮テストでは、これらの例外に注意してください。

- **背景** コーデックにとって難しい動きに、風に揺れる木の葉があります。木の葉は緑色であり、ルミナンスの大きい部分を占めるため、圧縮で多数のビットが割り当てられます。木の葉には、効率よくエンコードするにはビットを多く必要とするシャープで不規則な形のエッジや細部があります。また、風に揺れる木の葉の動きは概してランダムで、木の葉は絶えず重なり合ったり、現れたりするため、すべての動きを動き検出で処理するのが非常に難しくなります。
- **被写界深度** イメージの複雑さを制御するための有効で巧みな手法として、被写界深度があります。被写界深度を制御することで、重要でないオブジェクトを焦点外の背景や前景に追い出しながら、実際の被写体をはっきりと表示することができます。被写界深度は、映画撮影や写真撮影で非常によく使用されるため、見た目は自然に映ります。

オブジェクト間の空間的距離が十分でない場合、カメラを後ろに下げて望遠レンズを使用することで、合焦範囲を狭めることができます。

ポストプロダクション

ポストプロダクションの技法は、プロダクションの技法と同じくらの影響を圧縮性に与えることがあります。ポストプロダクションの技法はプロセスの後の方で使用します。そのため、多くの場合、テスト圧縮の結果に基づいてその技法を調整することが可能であることから微調整はずっと容易です。

- **静止画対動画** プロダクションの場合と同様に、極端に動画が多いと、高い率の圧縮を計画しているコンテンツには適しません。このようなことは DVD ではそれほど重要ではありませんが、Web や CD-ROM では重要です。
- **モーションブラー** 現実世界のシャッタースピードが、圧縮に役立つモーションブラーを発生させるように、レンダリングアプリケーションで仮想シャッターを使用できます。モーションブラーをサポートしているエフェクトの場合、ソースのシャッタースピードと一致するシャッタースピードでそのエフェクトをレンダリングします。このスピードはデューレーションとして表現されたり、After Effects でのように仮想シャッターの角度で表現される場合があります。デフォルトは 180° で、これはフレームレートの半分のシャッタースピードに相当します。
- **アンチエイリアシング** このレンダリング技法は、レンダリングされた要素の領域のエッジにあるピクセルを中間値にして、そのオブジェクトのエッジをソフトにします。イメージの部分間の境界を滑らかにすることで、圧縮性と外観が向上します。

アンチエイリアシングは、単独で制御できる場合と、レンダリング設定グループ全体の一部となっている場合があります。例えば、After Effects では、Best モードのエンコード時にアンチエイリアシングが自動的にオンになります。



これらのサンプルでは、Adobe Illustrator® ファイル (左上) がアンチエイリアシング前の Photoshop ファイル (PSD) に書き出され、次にアンチエイリアシングされて (左下)、Adobe ImageReady® で JPEG に変換されています (右下)。これらのファイルは同じサイズに変換されていますが、アンチエイリアシングされたイメージは斑点が減少しているのがはっきりと分かります。

- **プログレッシブレンダリング** インターレースのビデオソースを扱う場合、圧縮マスター作成時には、モーショングラフィックスをプログレッシブスキャンとしてレンダリングするのが適している場合があります。プログレッシブスキャンモードでプロジェクトをレンダリングすることで、インターレースソースをインターレース解除できますが、グラフィカル要素は完全なプログレッシブのままになります。

ビデオのキャプチャ：形式および接続方法の選択

ビデオのキャプチャとは、VTR やカムコーダーなどの外部ソースからコンピュータにコンテンツを取り込む方法のことです。コンテンツのキャプチャ用に選択した、ビデオの種類 (形式) および接続方式 (影響の程度は小さいですが) がコンピュータ上のビデオの品質に劇的に影響を与える可能性があります。常に、可能な限り最高品質のビデオおよびオーディオ信号を使用することから始めることをお勧めします。ほとんどすべての場合、圧縮によってビデオの見た目が元のソースよりも良くなることはありません。使用するソースの品質が高いほど、最終的な圧縮結果の品質が高くなります。この逆もまた真実です。したがって、使用可能なビデオやオーディオ形式とそれらの品質の違いについて基本を理解しておくことをお勧めします。

Adobe Premiere® Pro などの現在の編集ソフトウェアおよび多数の圧縮ソフトウェアツールは、様々な方法でコンテンツをキャプチャできます。アナログソースまたはデジタルソースからビデオをキャプチャできます。どちらも様々なビデオおよびオーディオ形式を使用でき、様々な物理接続方式を使用できます。ほとんどの場合、キャプチャを行うには、コンピュータに適切な物理コネクタの付いたビデオキャプチャカードを組み込む必要があります。

ビデオ形式を理解するには、まず、ビデオ信号の構成要素を理解する必要があります。ビデオの最初の説明では、輝度 (ルミナンス) 信号についてのみ説明しました。最低輝度から最高輝度までの輝度差を利用して、白黒テレビが生まれました。信号に対して 2 つのカラーチャンネル、つまりクリミナンス (クロマ) 値を生成することで、色が追加されました。まだ使用されていた何百万もの白黒テレビに対応するため、白黒テレビが色を無視してルミナンス (白黒) の画像を表示できるような色の追加方法が発明されました。テレビ放送はアンテナ経由で送信されるため、ルミナンスと 2 つのカラーチャンネルが 1 つの信号として一緒に送信されます。現在使用されている様々なビデオ形式は、この技術を使用して作成されました。

一般的なビデオ信号形式と、それらを送信するための一般的なケーブルおよびコネクタの種類を以下に示します。

- **アナログ形式** アナログビデオは、変動する電圧および周波数信号を使用してイメージデータを送信します。アナログビデオの品質には非常に大きな範囲がありますが、本質的に電力線など電子ノイズからの視覚的干渉を受けやすく、ケーブル配線が長い（通常 300 メートル以上）と減衰（信号の強度と品質の低下）が生じます。アナログ信号をサポートしている機器は一般に安価であり、現在でも有効な古い機器が長期にわたって使用されてきたため、アナログビデオ形式が今でも最も一般的に使用される形式となっています。

アナログから DV へのコンバータ

デバイスによっては、アナログビデオ形式を DV25 にキャプチャ用にその場で変更できるものもあります。通常そのような機器は、コンポジット、S-Video、およびアンバランスアナログオーディオを DV 形式に変換してコンピュータに保存します。DV の 25Mbps のデータレートは、カムコーダーで制作したコンテンツには十分ですが、下位機種のコピーではアナログノイズが目立つソースの変換に問題が発生する可能性があります。

コンピュータではデジタルデータしか使用しないため、アナログソースをキャプチャするには、デジタルコピーに変換する必要があります。これをビデオテープなどのアナログメディアに送り返すと、アナログ信号に変換しなおされます。この変換はエンコーディングとデコーディングと呼ばれますが、アナログ信号とデジタル信号間の変換は完全ではないため、小さいエラー（量子化誤差）が発生する可能性があります。適切な形式、優れた機器、ケーブル、接続方法を使用することで、このような量子化誤差を最小限に抑えることができます。

- **コンポジット** コンポジットビデオ信号は、ルミナンスチャンネルとクロミナンスチャンネルを 1 つの信号で伝送します。コンポジットビデオの最大の問題点は、ルミナンス情報がクロマにリークする可能性とその逆の可能性があることです。このリークが原因で、クロマクロールと呼ばれるノイズが発生し、このノイズがビデオキャプチャプロセスによって増幅され、品質全体を低下させる可能性があります。一般に、コンポジットビデオは使用可能な最低品質のビデオ信号であると理解されており、コンテンツ制作（ビデオ編集、DVD 制作、Web 用のエンコーディング）のソースとしてよりも、ビデオ配信（放送）に適しています。コンポジットビデオの一般的なソースは、VHS (1/2 インチテープ) VCR およびカムコーダー、テレビ放送信号、古い U-Matic 3/4 インチプロフェッショナル VTR です。

コンポジットビデオに共通のコネクタは、民生用機器では RCA または Cinch、業務用機器では BNC (Bayonet Neill-Concelman の略で、業務用放送ビデオ機器に共通のツイストロックコネクタ) です。使用する RCA や BNC ケーブルの品質を上げると、ノイズや減衰を低減できます。

- **S-Video** S-Video 形式の導入によって、民生用および業務用のビデオエディタの品質が大きく向上しました。S-Video では、ルミナンスチャンネルとクロマチャンネルが 2 つの別個のワイヤーペアに分割されており、これによりコンポジットのノイズの斑点が約半分が排除されます。2 つのクロマチャンネルは結合したままであるため、お互いに干渉しあい、信号に多少ノイズが残ります。Hi8 カムコーダー、S-VHS VCR など、多くの業務用および民生用グレードのビデオ機器の製造元は、S-Video コネクタを提供しています。もはや真の業務用グレードとみなされることはありませんが、S-Video はコンポジットからの大きな進歩であり、編集および圧縮用の優れたビデオキャプチャの最低限の標準とみなすことができます。

S-Video (つまり Y/C。Y はルミナンス、C はクロマ) ビデオ信号は主に S-Video ケーブルで特有の 4 ピンコネクタを使用します。これらのケーブルは短距離用（通常 30 メートル以内）です。業務用および放送用機器によっては、2 本の BNC ケーブルを使用して（1 本をルミナンス、もう 1 本をクロマに使用）、S-Video 信号を送信するものもあります。この設定を使用すると、接続をより遠くまで配線し、ケーブルが外れて信号が失われないように接続を固定できます。

- **コンポーネントアナログ** コンポーネントアナログは、YUV (Y はルミナンス、U は一方のクロマチャンネル、V はもう一方のクロマチャンネル) と呼ばれていますが、長年にわたって業務用および放送用の標準となっており、現在も広く使用されています。コンポーネントアナログを使用すると、ルミナンス信号と 2 つのカラー信号がすべて専用のケーブルで送信されます。3 つの構成要素（ビデオ信号の 3 つのチャンネル）すべてが独立して送信されるため、信号の品質が非常に良くなります。ノイズは非常に少なく、ビデオの色はより豊かかつ正確になります。放送および業務用環境では、コンポーネントアナログの多くはデジタルビデオ形式に取って代わられてきています。とはいえ、S-Video およびコンポジット形式と比較して、信号のノイズが少なくなるため、家庭用 DVD プレーヤーを新しいテレビに接続するための推奨形式としてコンポーネントアナログは消費者市場では高い人気になりつつあります。

業務用環境では、接続に長い配線を引き回したりコネクタを固定したりする必要があるため、コンポーネントアナログ信号は 3 本の BNC ケーブルで伝送されます。新しい家庭用 DVD プレーヤーでは、DVD プレーヤーを正しく接続しやすいように、使用する RCA または Cinch ケーブルが赤、緑、青に色分けされているのが一般的です。

- **アンバランスオーディオ** このタイプのオーディオ接続は、信号を伝送する単線とその周囲のアースされたシールドによって構成されます。これは、接続や回路が複雑にならないことから、消費者市場向けのオーディオ製品に一般的に使用されています。アンバランスオーディオの欠点は干渉を受けやすいことで、業務用アプリケーションではあまり使用されません。

アンバランスオーディオコネクタには3つの基本的な種類があります。消費者にとっても最もなじみのあるものは、RCAまたはCinchジャック（通常、赤が右チャンネル、白または黒が左チャンネル）を使用する、おなじみのアンバランスオーディオです。ステレオ信号をRCA接続するには、各チャンネルにRCAケーブルを1本使用しても、2本の単線ケーブルが1つになっているモールドステレオケーブルを使用することもできます。

このほかにも単線のアンバランスコネクタがあります。1/4インチコネクタは、ハイエンドのヘッドホンでは一般的で、ほとんどのホームステレオ機器での標準となっています。これとよく似ていますが、より小さいものに1/8インチコネクタがあります。これはポータブルオーディオ機器やコンピュータサウンドカードの標準となっています。

- **バランスオーディオ** この種類の接続は、2本の線（平衡信号線）とアースで構成されます。バランス接続は、干渉をはるかに受けにくいので、ケーブル配線が長い場合でも高い品質を維持できます。通常、バランスオーディオはXLRコネクタ（3ピンのロックコネクタ）を使用します。これは、アンバランスオーディオに使用されるコネクタと比べて格段に強固なものです。またこれも、長いケーブル配線とロックコネクタによって業務用接続の基本的な標準を満たしています。

デジタル形式

デジタルビデオは、コンピュータが使用するのと同様のデジタルデータを使用して、元のビデオ信号をテープに記録します。すべての輝度および色のデータがこの非常に正確なコンピュータスタイルの方式で保存されるため、ビデオ信号のチャンネル間で干渉やノイズの問題が発生しません。そのため、デジタル信号はアナログソースよりもクリアでシャープになります。デジタル形式を使用すると、アナログビデオをキャプチャするときに必要となる変換プロセスを経ずに、デジタルビデオをテープからコンピュータのハードディスクにキャプチャできます。

ビデオ業界には、複数のデジタル形式が存在しています。こうした形式の多くはデジタルカメラによる入力や撮影ではなく圧縮に関係しています。この節では、人気の高いデジタルビデオの入力形式についてのみ説明します。

- **DV25 (DV)** この形式は、通常使用する物理接続規格（IEEE-1394またはFireWire）で知られていますが、この形式はIEEE-1394接続規格で使用可能なデータレートおよび速度のごく一部しか使用していません。この形式は、ビデオ編集業界に革命を起しました。DVカムコーダーはサイズが小さく比較的コストなため、DV25は最近のビデオエディタの中では最も急速に成長している形式であり、ニュース用のデファクトスタンダードとなりました。信号のデータレートを25Mbpsに下げたため、カムコーダーへの入力時または記録時に5:1の割合で圧縮され、品質はS-Videoアナログ形式より少し高くなるだけです。ハイエンド寄りのDVカムコーダーでは、より高度な記録技術を使用しているため、同じDVデータレートでコンポーネントアナログビデオとほぼ同じ品質のDV信号を生成できます。

DV形式はコンピュータスタイルの1と0の羅列でアナログではありません。これによりビデオデータをDVテープからFireWire接続を装備しているコンピュータに直接コピーできます。コンピュータ上のデータはテープ上のデータと完全に同一になるため、その結果完全なコピーが作成されます。Adobe Premiere Proなどの多くのノンリニアビデオエディタは、この形式をネイティブで読み込み、編集できます。完成したビデオは元のソース映像と事実上同一です。処理エラーはほとんど発生しないか、まったく発生しません。

オーディオとビデオの信号は、単一のFireWire（IEEE 1394、ソニーのものはi-Linkと呼ばれます）ケーブルで伝送されます。このケーブルは、コンピュータでDVカムコーダーを制御するためのデバイスコントロール信号も伝送します。

DV25に加えて、DV50やDV100など（それぞれデータレートが50Mbpsと100Mbps）他にもよく使用されるDV形式は多くあります。DV50はパナソニックのDVCPProおよびソニーのDVCPAMで、DV100はDVCPProHDなどのHD（高精細）デジタルビデオ形式で使用されています。

Adobe Premiere ProではDV25およびDV50を標準のFireWireポート経由でサポートしています。これらのポートはほとんどの新しいコンピュータには通常装備されています。また、安価な FireWire 拡張カードがほとんどのコンピュータに装着可能です。4ピンと6ピンのDVコネクタが使用できますが、これらには互換性があります。通常、DVカムコーダーでは、より小さいコネクタである4ピンコネクタが搭載されています。また、データ転送には4ピンしか必要ありません。6ピンコネクタの残りの2つのピンは電源供給用で、一般的には FireWire ハードディスクなどで使用されています。

- **SDI (Serial Digital Interface)** SDIはハイエンドの業務用デジタル信号形式です。SDI形式は専門用語でCCIR-601または601と呼ばれています。この形式はコンポーネントアナログビデオに比べて品質が高く、業務用および放送用のテレビ制作環境の標準となっています。SDI形式はデジタルで非圧縮のため、何度も編集とコピーを行っても、品質劣化は検出できないレベルです。もともとSDIはSD用でしたが、HDが一般的になるにつれて、そのHD版であるHD-SDIの人気の高まっています(Adobe Webサイトの入門用ドキュメント『高精細ビデオの技術と利用』を参照してください)。通常、SDIはハイエンドの機器でしかサポートされていないため、(特にSDIビデオモニタの場合は)非常に高価になる場合があります。

ビデオ(場合によってはオーディオも)を1本のBNCケーブルで伝送できます。

- **デジタルオーディオ** 多様なデジタルオーディオ形式が存在します。民生用グレードのデジタルオーディオは、電子接続ケーブル(S/PDIFなど)または光接続ケーブル(TOSlinkや丸型ミニプラグなど)で伝送されます。デジタルビデオは、デジタルオーディオ信号をデジタルビデオ信号に埋め込んで一緒に伝送します。最も一般的な業務用グレードのデジタルオーディオ形式は、AES/EBU(Audio Engineering Society/European Broadcasting Union)です。AES/EBUのステレオチャンネルは、1本のBNCケーブルまたは1本のXLR接続ケーブルで送信できます。SDIビデオ機器で埋め込まれたAES/EBUデジタルオーディオを伝送できる場合もありますが、別個のAES/EBUまたはバランスオーディオステレオペアをSDIビデオ信号と一緒に使用する方が一般的です。

適切なキャプチャコーデックの選択

コンテンツのキャプチャをする際には、適切なコーデックを選択する必要があります。この選択は、使用するキャプチャハードウェア、キャプチャ元の形式および必要となるストレージによって決まります。3つの基本的な種類のコーデック、つまり、DV、モーションJPEG、および非圧縮から選択します。

DVコーデック

DV25デッキからキャプチャする場合は、DV25形式が最適です。FireWireコネクタ経由で、DV25を使用してネイティブなビットをテープからハードディスクにキャプチャして、完全なコピーを作成できます。より高いデータレートで異なる形式を使用すると、結果のビデオは若干劣化します。

DVブリッジを使用して、アナログソースからDV25に変換できますが、品質は低下します。DV25のデータレートは、アナログノイズがあるソースから最適なエンコーディングを行う場合には不十分です。また、4:1:1の色空間も品質劣化を伴います。

モーションJPEGコーデック

モーションJPEGは、ここしばらくの間、ハイエンドビデオ編集システムの有力な形式となっています。TargaやPinnacleコーデックなど、キャプチャカードの多くのベンダがサポートする独自のコーデックは、モーションJPEGのバリエーションであり同じ特性を持っています。

モーションJPEGは4:2:2であり、その実装の多くでデータレートや品質レベルを指定できます。実装によっては、数学的に可逆のモーションJPEGを作成できるものもあります。通常、50Mbps以上のデータレートがあれば、放送品質とみなされます。

非圧縮コーデック

4:2:2の非圧縮ビデオをサポートするカードが増えています。非圧縮コーデックのデータレートは大きく、品質は可逆圧縮と変わりません。いくつかの非圧縮コーデックでの重要な利点の1つに、10ビット/チャンネルのキャプチャをサポートしていることがあります。これは、アナログソース、特にSDIソース用に適しています。After Effects Professionalなどのハイエンドビデオ編集アプリケーションでは、8ビット/チャンネル超でビデオを処理できるため、品質が向上します。ハードウェアベンダによっては、Adobe Premiere Proの10ビットレンダリングをサポートしている場合もあります。非圧縮コーデックはデータレートが高いため、RAID(Redundant Array of Independent Disks)システムなどの高速ドライブを使用する必要があります。

注意: 非圧縮4:2:2と非圧縮RGBコーデック(Noneとよく呼ばれます)を混同しないでください。

ビデオの書き出し

テープからのキャプチャは、多くのプロジェクトで必要になりますが、プロジェクトによっては、すべてのコンテンツを電子的に作成したり、多くの制作過程をコンピュータでの編集および合成で行ったりする場合があります。このような場合、テープに戻っていると作業に集中できません。After Effects や Adobe Premiere Pro などの多くのツールは、タイムラインから圧縮ファイルに書き出すことができ、ファイルが1つだけの場合に有効です。しかし、複数の出力形式へのボリュームエンコードなど、他のエンコーディングツールの機能が必要になる場合は、タイムラインから中間ファイルに書き出し、このファイルをその後の処理に使用すると便利です。

ファイルへの書き出し

後で処理するために書き出されたファイルのことを中間ファイルと呼びます。中間ファイルを作成する目的は、圧縮ノイズを発生させることなしに、最も使いやすい形式で高い品質を実現することです。

前処理を行うか、または、レンダリングを回避するかの選択

最初に判断すべき事項は、前処理(次の節で説明)をいつ適用するか(レンダリング段階で適用するか、またはそれ以降に適用するか)ということです。一般に、レンダリング時に前処理を行うとファイルは小さくなりますが、レンダリング時間は長くなります。また、逆テレシネなど、特殊な操作にエンコーディングツールを使用できなくなります。

基本的に、レンダリングを回避するには、ソースビデオと同じコーデックとファイル形式を使用する必要があります。前処理を行わない場合は、変更されていないフレームを復元して再圧縮する必要はなく、フレームをソースから出力にコピーします。ツールによって、この処理をサポートする方法は異なります。Adobe Premiere Pro での処理をサポートするには、ムービー書き出し設定ダイアログボックスの「ビデオ」の項目で「再圧縮」オプションをオフにします。この処理が機能するためには、ビデオのフレームレート、フレームサイズ、ピクセルアスペクト比などが全く変更されておらず同一である必要があります。ビデオに変更が加えられていないことを確認する最も簡単な方法は、読み込みに使用したのと同じプロジェクトプリセットを書き出しに使用することです。これは、Adobe Premiere Pro でのデフォルト動作です。

ソースのコーデックを使用して書き出す、ということには例外が1つあります。ソースが DV25 で、多数のモーショングラフィックスや合成が含まれる場合です。こうした合成要素は正確な色を持っているため、NTSC DV25 の 4:1:1 ではなく、モーション JPEG などの 4:2:2 コーデックに書き出しをすることで、色の詳細情報を2倍にできます。

前処理を行う場合、最高のフレームレートと最大のフレームサイズを選択します。Web コーデックを対象としている場合は、正方形ピクセル(ピクセルアスペクト比が1)を使用します。

レンダリング用のファイル形式の選択

レンダリング用には、さまざまなファイル形式を選択することができます。最も一般的なファイル形式は、QuickTime と AVI で、どちらも高品質の中間コーデックを提供します。高いデータレートでは、MPEG-2 も使用できます。After Effects や Adobe Premiere Pro などの最新ツールでは、Windows Media 9 形式に書き出せます。重要なことは、エンコーディングツールがそのファイル形式をサポートしていることを確認することです。

書き出し用のコーデックの選択

中間ファイルを書き出す場合、圧縮ノイズを発生させないコーデックおよびデータレートを選択します。前処理を行わない場合は、通常ソースと同じコーデックまたはモーション JPEG を使用します。モーション JPEG を使用する場合、プログレッシブスキャンとして書き出すなら単一フィールドで使用し、インターレーススキャンとして書き出すならインターレースで使用します。

Windows の AVI ファイルの場合、Huffyuv コーデックを使用するのが一般的です。

前処理

前処理は、ビデオ圧縮の中間ステップであり、ソースビデオのデコード済みフレームを、コーデックに渡す最適化フレームに変換する処理が含まれます。前処理の目的は、ビデオを最終的な形式に変換することです。

前処理が重要な理由

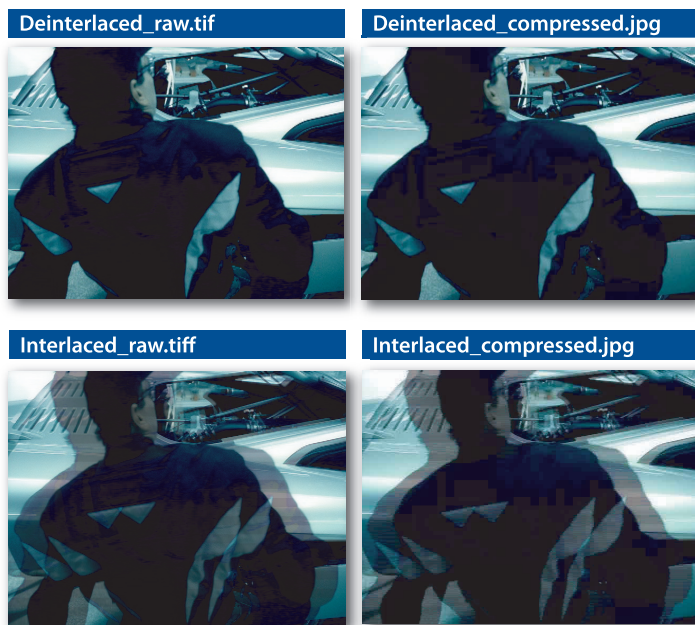
前処理は加減が難しい圧縮手順の1つで、多くの場合、最も時間を要する難しい手順となります。また、前処理は品質上、非常に重要であり、前処理を適切に行うことで最終的なビデオの品質を劇的に向上させることができます。適切な前処理が行われたビデオは、不適切な前処理を行った同一のソースに比べて、見栄えはより良くなりますが、エンコードのデータレートが2倍になる場合もあります。

前処理は Web ビデオでは非常に重要ですが、一般に DVD や SVCD 用の MPEG-2 などのビデオ形式に変換する場合、一般的には必要ありません。これらの形式は、コンピュータの画面で再生されるのではなく、ビデオとして再生されるため、現状のままとしておくことができます。

また、前処理を行わず、コンピュータでの再生用にレンダリング済みのグラフィックをネイティブに作成することもできます。

ビデオのインターレース解除

最も基本的で重要な前処理は、インターレース解除です。大部分のビデオはインターレースビデオとして再生および編集されます（とはいえ、プログレッシブスキャンも増加してきています）。しかし、ほとんどすべての Web ビデオはプログレッシブスキャンとしてエンコードされます。ビデオのソースフレームにインターレースラインをそのまま残すと、その出力にはスタックラインが表示されます。このラインは見栄えが悪だけでなく、エッジがシャープなため、エンコードが難しくなります。そのため、同じビットレートでは、インターレース解除されたビデオの方がインターレースビデオに比べて表示されるノイズは少なくなります。



圧縮前(左)と圧縮後(右)の同じビデオフレームを、プログレッシブ方式(上)とインターレース方式(下)で比較したものです。ファイルサイズが同じでも、インターレースフレームのほうが圧縮ノイズが多く発生しています。

注意: DVD や SVCD 用にエンコードされた MPEG-2 ファイルは、インターレース解除する必要はありませんし、解除しないようにすべきです。MPEG-2 は特有のインターレースモードでフレームをエンコードできます。

プログレッシブ出力用に制作する場合、プログレッシブモードで作業すると、After Effects や Adobe Premiere Pro など、大量のグラフィックやその他の合成エフェクトをビデオの前面に追加できるツールから書き出す際に良い結果が得られます。こうしたツールは、インターレースソースを自動的にインターレース解除すると同時に、プログレッシブモードですべての追加処理を行い、書き出す際にすべてをインターレース解除するよりも高い品質を実現できます。

テレシネと逆テレシネ

プログレッシブスキャンの特殊なケースとして挙げられるのは、フィルムソースとフィルムライクな 24P 制作です。従来の映画フィルムは、正確に 24 フレーム / 秒 (fps) のプログレッシブスキャンで再生されますが、これは NTSC の 59.94fps や PAL の 50fps には変換されません。

フィルムから PAL への変換は簡単です。ビデオの速度を 4% 上げて 25fps にし、プログレッシブ PAL に変換します。

NTSC の場合は少し複雑になります。まず、フィルムの速度を 23.976fps まで下げ、次に 3:2 ブルダウンを使用して 59.94fps に変換します。フィルムの最初のフレームは 3 フィールドのビデオになり、その次のフレームは 2 フィールドのビデオになり、その次のフレームは 3 フィールド、というように続きます。この処理の結果、5 フレームのうち 2 フレームには、異なるソースフレームのフィールドが含まれることになります。

幸いなことに 24P ビデオもその制作の標準となっており、24P で制作し、必要に応じてその汎用マスターから PAL、NTSC、フィルムまたは Web 形式に変換できる編集ツールが増えてきています。可能であれば、読み込み時に 24P に変換し、24P モードでプロジェクトのすべての作業を行うことをお勧めします。

クロップ

クロップとは、ソースフレームの領域を削除して、出力フレームに含まれないようにする処理のことです。通常、ビデオ用に作成されたコンテンツでは、重要な要素をいわゆるセーフエリアに限定します。セーフエリアは画面の端の部分以外の部分で、一部のテレビ（特に旧式のもの）ではセーフエリア外は表示されなかったり、歪んで表示される場合があります。

デジタルビデオ編集アプリケーションによっては、フィルタを使用してイメージをクロップできるものもあります。After Effects などのほかのツールでイメージをクロップする最も簡単な方法は、不要な端の部分をフレームから除外するようにビデオを拡大する方法です。非対称のサイズ変更を行う場合は、ビデオのアンカーポイントを必要に応じて移動します。

720 ピクセル幅のソースを、VideoCD 用に 352 ピクセル幅の出力に変換する場合を例にとってみます。指定したサイズ変更方法では、ソースの左右から 8 ピクセルをクロップして、704 ピクセル幅のフレームを残し、これをちょうど 50% 縮小して 352 ピクセル幅にします。

別の特殊なサイズ変更例としては、NTSC での 486 ラインから 480 ラインのソース間での変更というものがあります。モーション JPEG では一般に 720×486 が使用され、MPEG-2 では一般に 720×480 が使用されます。ライン数が異なる原因はサイズ変更の違いによるものではなく、486 ラインモードでキャプチャされるビデオ信号のライン数が 6 本多いためです。486 ラインから 380 ラインに変換する場合（DVD 制作などの場合）、6 ライン分をクロップする必要があります。一般に、上から 4 ライン、下から 2 ライン分をクロップすることをお勧めします。逆に、480 ラインから 486 ラインに変換する場合は、6 ライン分追加する必要があります（これも同様に上に 4 ライン、下に 2 ライン追加することをお勧めします）。奇数を使用すると、フィールドの順序が切り替わるため、大きな問題が発生します。



内側の境界線がタイトルセーフエリアで、外側の境界線がアクションセーフエリアです。

スケール

スケールとは、クロップされたソースの矩形を適切な出力フレームサイズになるようにサイズ変更することです。例えば、 720×480 ソースフレームを Web ビデオ用に 320×240 出力フレームに変換する、といったものです。

どちらか一方の軸に基づいて拡大すると、ソースにないデータがビデオ編集アプリケーションによって作成されることになります。そのため、可能であれば常に両方の軸に基づいて縮小をすることをお勧めします。

アスペクト比補正

スケールの重要な要素の1つに、アスペクト比補正があります。DV など大部分の制作形式では非正方形ピクセルが使用されています。例えば、NTSC の DV は 720×480 であり、ピクセルを単独で見ると、アスペクト比は 3:2 になります。しかし、DV は 3:2 ではなく、その制作方法によって 4:3 または 16:9 になります。4:3 モードでは、ピクセルアスペクト比は正方形よりも幅が狭くなり、16:9 モードでは正方形よりも幅が広がります。非正方形ピクセルは、DVD 用の MPEG-2 や SVCD 用の MPEG-1、(形は異なりますが) SVCD 用の MPEG-2 で使用されています。

逆に大部分の Web 形式では正方形ピクセルが使用されるため、計算は簡単です。例えば、 320×240 は正方形ピクセルで 4:3 です。

非正方形ピクセルのソースから正方形ピクセルのソースに変換する場合に、アスペクト比補正を正しく行うためのキーは、出力のフレームサイズをソースのアスペクト比と一致させることです。つまり、4:3 のソースを使用する場合は、ソースのフレームサイズが 720×480 、 720×486 、 640×480 、 352×480 などであっても、 160×120 、 320×240 、 400×300 、 512×384 、 640×480 などの 4:3 のフレームサイズが適したサイズになります。



この図は、4:3 (上) と 16:9 (下) での 720×480 のイメージを示しています。それぞれ 4:3 と 16:9 出力に変換されています。このように、出力のアスペクト比の補正は非常に重要です。

ノイズリダクション

ノイズリダクションには、イメージからノイズを除去して(特に粒状性を抑えて)圧縮を容易にする多様な技法が含まれます。非常に単純なノイズリダクションアルゴリズムはブラーで、粒状性を抑えることができますが、同時にイメージが非常にソフトになる場合があります。一般に使用されているミディアンフィルタは、ノイズリダクションに適したブラーです。イメージ内の粒状性を含む部分だけを、フレーム間の違いを利用してぼかすという高度なアルゴリズムもあります。

After Effects Professional などのツールでは、自動または手動によるスクラッチ除去など処理により、劣化したソースの改善や修復を実行できるものもあります。通常このようなツールはスタンドアロンのアプリケーションまたはビデオツールのプラグインで、ワークフローの初期段階で使用する必要があります。

ノイズリダクションの適用は必須というわけではありません。非常にノイズの少ないソースの場合は、ノイズリダクションによって品質が改善することではなく、ノイズリダクションによってソースオーディオが品質が損なわれてしまう可能性もあります。通常、ソースの品質の低さを改善するために最低限必要なノイズリダクションだけを使用することをお勧めします。

イメージ処理

イメージ処理とは、イメージ編集アプリケーションに見られるようなコントロールを使用してイメージ全体を変更する処理のことです。コーデックによって使用される色空間は異なるため、イメージ処理は多くのプロジェクトで必要とされていました。現在では、ほとんどの部分で、最新の形式およびツールによって自動的にイメージ処理が行われます。そのため、イメージ処理は主として圧縮に対するコンテンツの適性を高めるために行われるようになり、多くのプロジェクトでは必要がなくなってきています。

ルミナンスレンジの拡大

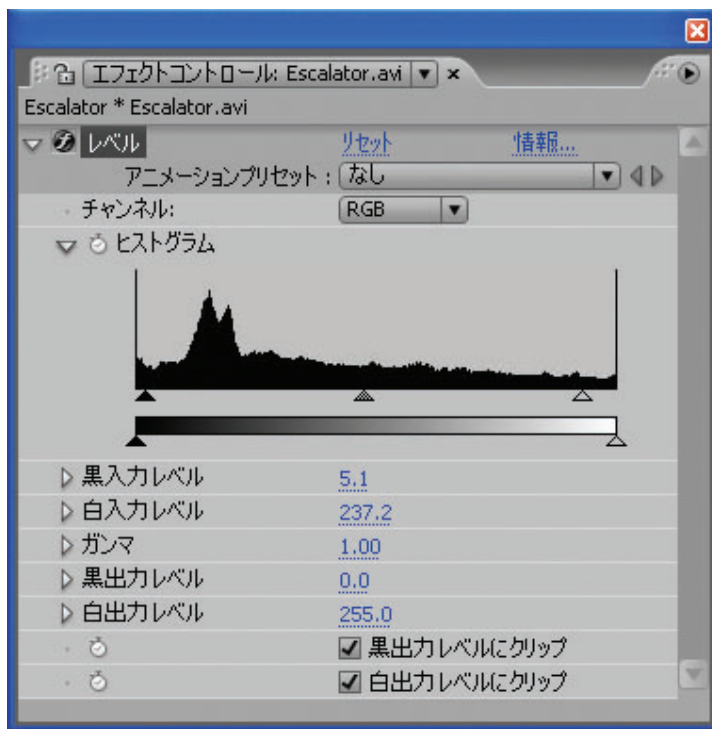
以前は、ビデオ圧縮における主要な問題は、確実に黒を真っ黒、白を真っ白にすることでした。ツール、形式およびコーデックによっては、黒を 0、白を 255 と見なすものもあれば、ITU 規格を使用して黒を 16、白を 235 とするものもありました。一般に、最新のツールではこれらが自動的に統一されます。ただし、黒潰れや白飛びをしている場合、制作のプロセスでルミナンスレンジの不一致が発生した可能性があります。通常、すべての RGB コーデックでは 0 ~ 255 の範囲が使用され、Y'CbCr コーデックでは 16 ~ 235 の範囲が使用されます。ビデオツールでは、自動的に Y'CbCr または RGB から他へ自動的に変換されます。16 ~ 235 の範囲も使用するスタジオ RGB を使用するコーデックには注意してください。配信コーデックに変換する際にコントラストを追加する必要がある場合があります。

黒レベル

ルミナンスを制御するもう 1 つの目的は、ソースの黒と白を出力でも数字的に黒と白にすることです。デジタルレンダリングされたタイトルなどの場合は自動的にそうなるため必要ありませんが、アナログテープからキャプチャされたビデオには必ずノイズが入っており、これによりピクセルのルミナンス値がランダム化されるため、一部の黒が純粋な黒より高い値になります。

輝度を下げるか、コントラストを上げることによって、値を黒に落とすことはできますが、一般的に一方の技法だけを使用するよりも、両方の技法を使用するほうがより微妙な表現が可能です。ルミナンスレンジを調整する前にノイズリダクションを適用すると、より明るい黒ピクセルの一部を黒に近づけることができるため、それほど極端なルミナンス拡大を使用せずに済みます。

ヒストグラムを使用する After Effects のレベルフィルタは、こうした処理に非常に有効です。黒であるべき領域を黒として処理できるようにすることが重要です。輝度とコントラストが十分でない場合、黒または白の復元フィルタを使用することもできます。これらのフィルタはしきい値を設定するフィルタであるため、しきい値未満のピクセルは黒に、しきい値以上のピクセルは白になります。



After Effects のレベルエフェクトの標準的な設定。ルミナンスの黒と白のレベルをノイズの下限と上限のレベルを超えるように設定しています。

ガンマ補正

ガンマとは、0 ~ 255 のルミナンスの範囲が、表示時にどのような明るさに変化するかを測定するものです。ガンマ値が異なる場合、0 と 255 の表示は変わりませんが、その中間の値の表示は異なります。プラットフォームが異なると (Mac OS、SGI、Windows など)、デフォルトのガンマも異なります。その結果、ファイルによっては、Windows を実行しているコンピュータ上での表示よりも、Macintosh コンピュータでの表示の方が、中間階調が明るく表示される場合があります。ただし、大部分の最新のビデオコーデックでは、再生時に適切なガンマに自動変換できます。

残念なことに、ガンマ補正が実行されない古いスタイルのコーデックが、現状でも多く使用されています。制作に使用するプラットフォーム以外のプラットフォームでコンテンツを正しく再生する必要がある場合、そのプラットフォームの組み合わせを使用してサンプルをスポットチェックし、正しく表示されることを確認することが重要です。

オーディオ処理

オーディオはビデオを構成するビットをほんのわずかしが使用しないものの、ユーザに提供する内容の半分を担うものです。一般にオーディオ処理はビデオ処理よりも簡単ですが、オーディオとビデオの両方を正しく処理することが重要です。

オーディオの再サンブル

低データレート（特にモデムを対象としたデータレートの場合）では、一般にオーディオを 22.050KHz 以下の低いサンブルレートにする必要があります。最新のビデオ編集アプリケーションでのレートの再サンブルでは、数年前に比べてはるかに高い品質を提供するアルゴリズムが使用されます。

ボリュームの設定

瞬間最大ボリュームは、最大許容ボリュームより少し下に設定する必要があります。非常に低いデータレートの場合、コンプレッサやリミッタを使用してダイナミックレンジを下げること、オーディオ品質を向上できることもあります。一般的に、テレビ放送用オーディオはエンコードが容易です。オーディオがコンプレッサやリミッタで処理済だからです。

チャンネルとビットレートの調整

低いデータレートでは、モノラルのコンテンツの方がステレオコンテンツよりもエンコード品質が高くなります。そのため、低ビットレート（通常 32Kbps 未満）を対象とする場合は、オーディオをステレオからモノラルに変換することをお勧めします。

低サンブルレートでオーディオをステレオに設定するか、高サンブルレートでオーディオをモノラルに設定するか、という選択を行う必要がある場合、コンテンツの音質を考慮するなら、ステレオにするよりもサンブルレートを高くする方が重要であるため、モノラルを選択することをお勧めします。

5.1 や 7.1 のディスクリートサラウンドサウンドオーディオをサポートしているシステムも多くあります。一般に、このようなサラウンドサウンドでは高いデータレート（オーディオだけで 128Kbps 以上）が必要なため、リアルタイムストリーミングでの使用は困難ですが、ディスクベースでの再生では非常に有効である可能性があります。

Dolby Pro Logic でエンコードされたサラウンドサウンド（サラウンドサウンドをステレオペア内に収めることができます）は、多くのオーディオコーデックのエンコードプロセスをそのままの形では通過することはできません。プロジェクトで Pro Logic エンコーディングが必要となる場合、慎重にテストして Pro Logic のデータがエンコード処理を通過できるかどうかを確認します。コーデックによっては、データレートを上げるか、ジョイントステレオなどのオプションをオフにすることが有効である場合もあります。コーデックによっては、Pro Logic をエンコードできないものもあります。

ノイズの低減

オーディオ用のノイズリダクションフィルタもありますが、通常このようなフィルタは圧縮ツールの一部ではなく、Adobe Audition™ などの業務用オーディオツールの一部となっています。オーディオソースをクリーンで歪みがないものにし、ヒスノイズやポップノイズをオーディオソースから除去することで圧縮効率を向上させます。

エンコーディング

前処理が圧縮の芸術であるなら、エンコーディングは圧縮の科学といえます。エンコーディングとは、データレート、フレームサイズ、フレームレートなどのパラメータを正しく定義して、プロジェクトの最適化を図ることです。多くのトレードオフが存在するため、これらのパラメータのバランスを取ることは難しい場合があります。ビデオを圧縮することの最終目標を理解していることが、常に重要です。この節では、はじめにすべてのコーデックに適用されるパラメータについて説明し、続いてビデオとオーディオに固有のパラメータについて説明します。

データレートモードの選択

コーデックや形式によって使用できるデータレートモードは異なります。コーデックによっては 1 つのデータレートモードしかサポートしないものもあれば、多数のオプションが使用できるものもあります。多くの場合、特定の用途に対する最適なモードは 1 つです。

データレートモードには、以下のものがあります。

- **平均制限**は最も一般的なデータレート制限です。このモードでは、平均データレートを指定すると、コーデックは、そのファイルが要求された平均値となるように処理を行います。コーデックは、映像内の複雑な変化に対応できるようにレートを増減します。

CD-ROM および QuickTime の大部分のコーデックはこの方法で動作します。平均制限エンコーディングは、ファイルサイズとデータレートが制限要因となる多くのプログレッシブダウンロードファイルで使用されます。

- **ピーク制限**エンコーディングは、一般的にデコーダ（CPU、メモリまたはディスク）の速度が制限要因となるハードウェアデバイスや高データレートコンテンツに適しています。特定のピーク制限データレートを指定しても、実際にファイルのピークが指定したビットレートになるわけではありません。ピーク制限データレートは、サポートされる最大レートですが、コンテンツに適している場合、エンコーダではこれより低いレートを自由に使用できます。

ピーク制限は、高解像度で高データレートの Web ファイルや CD-ROM ファイルをローエンドコンピュータで再生する必要がある場合にも有効です。コーデックによっては、平均制限とピーク制限の両方を使用するエンコーディングをサポートしているものもあります。

- **バッファ制限**エンコーディングはリアルタイムストリーミングに適した方法です。バッファ制限エンコーディングでは、平均データレートがファイル全体に適用されるのではなく、ビデオの任意の部分に適用されます。そのため、3秒のバッファを使用してビデオをエンコードすると、ファイルの任意の3秒分のデータレートは目標のデータレート以下になります。多くのストリーミングエンコードファイルは、平均制限とバッファ制限の両方を使用します。バッファ制限エンコーディングは、固定ビットレートまたは CBR とよく呼ばれます。
- **品質制限**エンコーディングはデータレートを指定しないため、他の方法とは異なっています。その代わりに、品質制限では目標品質を指定し、各フレームはこの目標品質を達成するために必要な数だけのビットを使用します。このモードはリアルタイムストリーミングとの互換性はありませんが、コンテンツをアーカイブする場合や、ミッションクリティカルなビデオをダウンロードする場合など、ファイルサイズやダウンロード時間よりも品質の保証が重要になる場合に非常に有効です。例えば、ビデオ販売会社は、コンテンツの画質ができるだけ良くなるように、また、不必要に高いデータレートでビットを浪費しないようにしてコンテンツのエンコードの容易性を高めるように、短い Web サンプル用に品質制限エンコードファイルを使用できます。

多くの品質制限コーデックは 0 ~ 100 の範囲で品質を指定しますが、同じ値を指定しても、コーデックによって生成される実際の表示品質は大きく異なります。コーデックによっては、平均制限と品質制限を同時にサポートするものもあります。この場合、品質制限はフレームごとの最低品質を設定します。使用可能なビットレートに基づいてフレームが品質制限以下になると、品質を満たすようにそのフレームのビットレートが上げられます。通常、低品質フレームが発生すると、その後の 1 つ以上のフレームをコマ落ちさせて平均ビットレートを維持します。このモードは、動きのスムーズさよりもイメージの品質を優先します。

エンコーディング速度か品質かの選択

多くのコーデックでは、エンコーディング時間と品質間のトレードオフを指定するコントロールが使用できます。通常、高速モードでは、高負荷の動き検出は行われません。一般にエンコーディング速度を大きく変更しても品質はそれほど改善しません。8 倍時間のかかるエンコード処理を指定しても、圧縮効率が 20% しか向上しない場合もあります。速度と品質間の正しいトレードオフは、プロジェクトによって異なります。DVD のプロジェクトの場合は、データレートが非常に高いため、高品質を得るために低速のエンコーディングを選択しても、1 時間未満のタイトルではあまりメリットはありません。Web 配信コンテンツの場合は、多くの場合にバンド幅が Web ビデオの制限要素となるため、低速のエンコーディングにすることのメリットがあります。

フレームサイズの設定

フレームサイズ（解像度とも呼ばれることが多い）は、最終的なビデオのフレームを構成する縦横のピクセル数のことです。データレートはフレームの面積（縦×横）に比例します。そのため、320×240 から 640×480 に変更すると、データレートは 4 倍必要になりますが、192×144 に変更するとデータレートは元の 1/3 になります。この結果はこの数式から考えられるほどの強い影響を持ちません。ピクセル数が少なくなれば、ピクセルの品質はより重要になるからです。

大部分のコーデックおよび形式では、ビデオの縦と横のピクセル数が 2 で割り切れる必要があり、縦と横のピクセルが 16 で割り切れる場合に圧縮効率は最高になります。

アスペクト比の指定

アスペクト比は、フレームサイズで重要となる観点です。大部分の Web および CD-ROM 形式はデフォルトで正方形ピクセルであり、多くの形式は正方形ピクセルしかサポートしていません。逆に、ディスクベースの VCD、SVCD、および DVD 形式では、非正方形ピクセルしか使用されません。まず、対象のアスペクト比を把握し、次に、コンテンツを配信する画像の形状で、選択したフレームサイズが正しいアスペクト比を実現できるようにすることが重要です。

4:3	16:9
160 × 112	208 × 120
192 × 144	256 × 144
256 × 192	336 × 192
320 × 240	432 × 240
384 × 288	512 × 288
512 × 384	688 × 384
640 × 480	848 × 480

正方形ピクセルで作業する場合の 4:3 および 16:9 アスペクト比の推奨フレームサイズ

フレームレートの調整

フレームレートとは、ビデオを再生するときの 1 秒間のフレーム数のことです。フレームレートが高いと表示はより滑らかになり、フレームレートが低いと表示はぎくしゃくした感じになります。20fps 未満になると滑らかに表示されず、10fps 未満の場合動画というよりもばらばら漫画のように表示されます。

フレームレートを増減させても、フレームサイズの場合と比べてデータレートに与える影響はずっと小さいものです。これは、フレームレートを下げると、フレームごとに 2 倍の動きをエンコーディングする必要があり、これにより、フレームごとに必要なビット数が増えるためです。また、各フレームの表示時間が 2 倍になるため、それぞれのノイズも長く表示され、気づきやすくなります。そのため、同じ知覚的品質を実現するには、フレームごとの画質を上げる必要があります。

したがって、可能であればソースのフレームレートを使用し、ダイヤルアップのモデムユーザに対するリアルタイムストリーミングを対象としない限り、ソースの半分以上のフレームレートを使用することをお勧めします。一般に、ユーザの鑑賞品質を向上させるには、フレームサイズを上げるよりも、フレームレートを上げた方が効果的です。

NTSC のフレームレートは整数よりも 0.1% 低くなります。そのため、「30fps」は実際には 30 / 1.001 であり、通常 29.97 に四捨五入されます。NTSC ソースを処理する場合は、29.97fps の出力の方が 30fps よりも正確です。また、インターレースビデオにはフレームごとに 2 つのフィールドがあります。

フレームレートの変換で非常に重要なルールがもう 1 つあります。それは、秒あたりの出力フレーム数は秒あたりのソースフレーム数の整数除算である必要があるということです。そのため、出力は 5 分の 3 などではなく、2 分の 1 や 3 分の 1 になります。各種ソースに適したフレームレートを次の表に示します。

フィルム	PAL	NTSC
59.94	23.976/24	50
29.97	12	25
15	8	12.5
10	6	8.67
7.5	N/A	6.25
6	N/A	5
5	N/A	N/A

ソースフレームレートを割り切れるフレームレートを使用することが重要です。

フレームドロップ (コマ落ち) のネゴシエート

コーデックによっては、難しいフレームの画質を維持するために、必要に応じて (または自動的に) 目標平均データレートを維持するようにフレームレートを下げるものもあります。この処理は一般に、フレームごとに最低画質を設定するコントロールの形を取ります。コーデックによって通常指定したフレームに割り当てられるビットで目標品質を達成できない場合、そのフレームのデータレートを品質確保に最低限必要なデータレートにまで上げられます。次に、目標の平均データレートを維持するために後続のフレームを 1 つ以上コマ落ちさせます。

この処理が適切かどうかはコンテンツで決まります。映画や特に音楽ビデオなどのエンターテインメントコンテンツの場合、ビデオとオーディオをスムーズに同期させることが重要になるため、フレームドロップ処理は避けることをお勧めします。逆に、プレゼンテーションビデオ、特に背景にホワイトボードがあるものは、画面のテキストを判読可能な状態にするために、多くのフレームドロップを行った方が良い場合もあります。

キーフレームレートの働き

キーフレームレートは、ビデオに挿入されるキーフレーム (I フレーム) の数を制御するものです。ツールや形式に応じて、キーフレームレートは、1 キーフレームあたりのフレーム数、または、1 キーフレームあたりの秒数で表されます。MPEG-1 および MPEG-2 では、多数のキーフレーム (通常 1 秒間に 2 キーフレーム) が必要となります。Web 形式では、10 秒間にキーフレームが 1 つしかない場合もあります。

大部分の形式では、シーンのカットなどビデオが劇的に変化する場合や、爆発シーンなどで極端に早い動きがある場合に、ナチュラルキーフレームと呼ばれる追加キーフレームが挿入されます。ナチュラルキーフレームによって「キーフレーム単位の」で指定されている値がリセットされるため、ビデオに速いカットが多数あり、自動キーフレームが 100 フレームごとや、1000 フレームごとに設定されていても関係ありません。この場合、キーフレームがすべてナチュラルキーフレームになります。ナチュラルキーフレームも挿入するコーデックの場合、「キーフレーム単位の」値は実際には、キーフレーム間の最大間隔を設定しているに過ぎません。

キーフレームでの 1 パスおよび 2 パスアルゴリズムの使用

従来、大部分のコーデックは 1 パスで機能していました。つまり、フレームごとにソースファイルを読み込み、圧縮データを書き出しました。この処理は機能していましたが、データレートメカニズムがその後のコンテンツの構成を推測する必要がありました。この処理のために多くの優れた発見的 (ヒューリスティック) アルゴリズムが開発されましたが、どのアルゴリズムも場合によっては予測を誤る場合があります。最悪の状況としては、1 パスアルゴリズムで、ビデオがさらに複雑になると解釈され、品質を一定に保つためにデータレートが上げられる場合や、現在のフレームの数秒後により複雑な別のシーケンスが続く場合があります。ビデオが複雑になった瞬間に、以前に上げたレートを埋め合わせるため、コーデックはデータレートを下げなければなりません。

2 パスアルゴリズムでは、コーデックによってまずビデオが分析され、各フレームにおける相対的なエンコードの複雑さを示すテーブルが作成されます。次に、可能な限り最高の平均品質を実現するために、必要な実際のフレームあたりのビット数を計算し、2 番目のパスでこのビットを適用します。最終的には、正しい品質値によって目標データレートに達している品質制限エンコードの場合と同じになります。

2 パスアルゴリズムファイルの利点の 1 つは、品質が非常に一貫している (簡単な部分と複雑な部分の品質差が少ない) ということです。この一貫性によって、品質がシーンによって大きく異なるビデオに比べて、鑑賞品質はより向上します。

形式およびコーデックでサポートされている場合は、ビットレート制限、ピーク制限、およびバッファ制限ファイルで 2 パスエンコーディングを使用できます。バッファ制限ファイルを使用すると、バッファを大きく取ることで、アルゴリズムはより柔軟にビットを分配できるようになります。

コーデックの中には、少数ながら 3 パスまたは N パスシステムもあります。これらのシステムでは、コーデックによって 2 番目のパスの結果が分析されて新しいデータレート割り当てテーブルが作成され、ファイルのエンコードが行われます。これによって通常、より良い結果が得られます。N パスシステムでは、コーデックはこの処理を複数回繰り返す場合がありますが、その後のパスでは結果の品質向上度合いは低下します (一般的に、適切な 2 パスを実装すれば最適な結果に非常に近くなります)。

Windows Media や RealMedia などの最新の形式では、2 パスがユーザインタフェースに統合されています。2 パスをサポートしている AVI および QuickTime コーデックでは、2 パスエンコーディングを動作させるのはより難しくなります。AVI の場合、通常、コーデックは 2 つの異なるエンコードファイル (ダミーファイルと完全なファイル) を 1 つずつ作成する必要があります。

サンプルレートがオーディオの品質に与える影響

サンプルレートとは、1 秒間に実行されるオーディオ測定回数の中で、キロヘルツで表されます。CD の場合は 44.1KHz、DV なら 48KHz、アナログ電話なら 8KHz です。

32kHz 未満では、音質はソースと比べて顕著に劣化します。大部分のブロードバンドや CD-ROM では、最低 32KHz を使用する必要があります。

対象となるコンピュータの仕様が入手できる場合を除けば、コンピュータでの再生用途では 44.1KHz を使用します。この方が 48KHz を選択するより安全です。多くのサウンドカード (特に古くて安価なモデル) では、内部的に 44.1KHz しか再生できません。そのため、48KHz を使用すると、ソフトウェアでのサンプルレート変換が必要となります。この変換は、圧縮時に適用できる高品質フィルタほど良くないため、最終的な出力品質はさらに悪くなります。オーディオソースが 48KHz 以上で、出力コンピュータまたは出力機器が 48KHz をネイティブでサポートしている場合は、48KHz をお勧めします。コーデックによっては、少数ですが 96KHz などのより高いサンプリングレートでエンコード可能なものもあります (その代表的なものとしては WMA9 Professional があります)。

チャンネル

大部分のコーデックでは、チャンネルはモノラルかステレオのいずれかになります。モノラルのみのコーデック、ステレオのみのコーデック、5.1 や 7.1 などのマルチチャンネルモードが利用できるコーデックなどがあります。

IMA などの古いコーデックでは、モノラルからステレオに変換するとデータレートが 2 倍になるものもあります。最新のコーデックでは、チャンネル間の冗長性を低減できるものもあります。これにより、ステレオでの必要ビットレートは、同一の知覚的品質のモノラルのビットレートと比較して 20 ~ 50% 程度（コーデックやチャンネル間の差がどの程度あるかによって異なります）しか増加しません。

サンプルレートや知覚可能なノイズと比べて、チャンネル数は知覚的品質にとっては重要ではありません。32KHz のモノラルの方が、22.050KHz のステレオより、知覚的品質は高くなります。一般的に、最低でも 32KHz、理想的には 44.1KHz 以上を使用できる場合にのみ、ステレオの使用をお勧めします。

マルチチャンネル (5.1 および 7.1) オーディオコーデックも存在します。これらのコーデックで最も一般的なものは、DVD 用の Dolby Digital (AC-3) と WMA9 Professional です。RealAudio 10 ではマルチチャンネルがサポートされています。

ビデオの配信方法

コンテンツを圧縮する方法は、そのコンテンツの配信方法によって大きく異なります。ビデオが使用される様々な環境と、各環境で必要となるトレードオフを理解しておくことは有益です。

ディスクベースの配信

従来、CD-ROM はデジタルビデオ配信の主要な形式でした。しかしその後、Web ビデオは極めて小さなスライドショーからフルスクリーン、フルモーションのビデオへと進化しました。今日では、携帯電話などのモバイル機器上でビデオを視聴できるようになっています。業界の関心はネットワーク配信に移行していますが、その一方で、ディスクによる再生がその質を向上し続けていることは好材料です。

ディスクベースの配信は有効に機能します。また、使用可能なデータレートは Web ビデオよりも劇的に高くなります。8 倍速の IDE CD-ROM でさえ、8000Kbps のデータレートを持続的に提供できるのです。CD-ROM ディスクの最大サイズは、今日では 800MB に達します。DVD ではさらに膨大な容量を使用できます。DVD-5 ディスクで 4.7GB、DVD-9 ディスクでは最大 8.4GB のデータを記録できます。

最新型の高速プロセッサを搭載したパソコンを対象にすると、ディスクベースの再生では素晴らしい結果を得ることができます。2002 年に購入されたほとんどのコンピュータでは、1280 × 720 のフレームサイズで圧縮ビデオを再生できます。2003 年のハイエンドコンピュータであれば、1920 × 1080 のビデオを再生できます。数年前のコンピュータでも、特に、古い、パフォーマンスチューニングされたコーデックを使用すれば（それに対応して圧縮効率は低下します）640 × 480 のビデオを簡単に再生できます。

ダウンロード型ファイル

Web 上の最も基本的な種類のビデオは、単純なダウンロード型ファイルです。このファイルは、他のファイルと同様、ユーザのハードディスクにダウンロードされます。

ダウンロード型ファイルの利点は、ダウンロード中にユーザがリアルタイムの再生を期待していないことです。ファイルのサイズが（多くの場合はダウンロード時間も一緒に）、ダウンロード前にブラウザからユーザに通知されます。また、ファイルがユーザのコンピュータにコピーされるため、各ユーザ自身が互いにファイルを受け渡すことができ、その分だけ視聴者あたりのバンド幅コストは小さくなります。欠点は、ダウンロードが完了するまでファイルを一切表示できないことです。このため、最初の 10 秒で不要とわかるファイルを長時間かけてダウンロードするということが起こる可能性があります。

ダウンロード型ファイルは、マーケティングや非営利の用途でよく使用されます。ダウンロード型ファイルは、ビデオゲームのサンプルで非常に一般的です。この用途では、十分な品質でゲームプレイをエンコードするために、非常に高いデータレートが必要になります。

プログレッシブダウンロードファイル

プログレッシブダウンロード中、ファイルは FTP (File Transfer Protocol) や HTTP (HyperText Transfer Protocol) などの標準的な Web プロトコルを使用して送信されますが、ユーザは既に送信された部分を、残りがまだダウンロード中でも表示できます。

大多数の形式では、ダウンロードの残り時間がクリップのデュレーションより短くなると、自動的にビデオの再生が始まります。この形式を使用すると、クリップを再生しながらダウンロードを完了できます。ただし、場合によっては、ビデオの再生がダウンロードされたデータに追いつき、再生が一時的に停止することもあります。

プログレッシブダウンロードでは、HTTP、FTPなどのプロトコルを使用して全データパケットの到達を保証するため、リアルタイムストリーミングに関連するビデオやオーディオの欠損の影響を受けることはありません。しかし、プログレッシブダウンロードでは、信頼性のあるリアルタイムパフォーマンスを提供できません。200Kbpsの接続上で400Kbpsのファイルを再生するには、1分のビデオに対して1分近くのバッファ時間が余分に必要となります。当然、200Kbpsの接続上で同じ400Kbpsのファイルをストリーミングすることはできません。

プログレッシブダウンロードは、制作および展開が最も簡単で、最も安く、そして最も気軽に利用できる種類のWebビデオです。Webビデオ初心者のユーザが始めるのに最適な方法です。

リアルタイムストリーミング

次のレベルのWebビデオはリアルタイムストリーミングです。この種類の送信では、ほんの数秒のバッファの後にビデオが見られるという従来どおりの体験を提供しますが、画像の欠損、低品質や、データレートが表示の際にクライアントで利用可能なバンド幅に制限されるというリスクがあります。

では、なぜこうした制限をあえて受け入れるのでしょうか。それは、リアルタイムストリーミングによって、リアルタイムパフォーマンスと、ファイルの様々な部分へのランダムアクセスが可能になるからです。プログレッシブダウンロードは映画の予告編には有効ですが、映画以外のコンテンツ、特にユーザが最後の20分までスキップしたい場合には向いていません。リアルタイムストリーミングでは、ファイルの長さに関係なく1分足らずで再生が始まり、ユーザは目的の部分にスキップできます。また、大半のリアルタイムストリーミング形式でスケーラビリティが提供されており、ストリームを動的にユーザの接続速度に適合させることができます。

わずか数年前と比べても、リアルタイムストリーミングの鑑賞品質は大幅に向上しました。コーデック、サーバとも大きく改善されています。ユーザがサーバを意識することはめったにありませんが、サーバとネットワーク技術の向上が、ユーザに大きな利益をもたらしてきました。最新のサーバではバッファを積極的に管理できるようになっており、フレームが必要になる数分前に送信します。これにより、ドロップパケットを再送信し、変化するネットワークの状態に基づいてバンド幅を調整する時間を十分に提供できます。

共通の質問は、いつリアルタイムストリーミングを使用し、いつプログレッシブダウンロードを使用するのかということです。選択が明確な場合もあれば、それほど明確でない場合もあります。検討の対象となるルールはデュレーションです。3分未満のクリップにはプログレッシブが向いており、3分を超えるクリップには、多くの場合ストリーミングが適しています。

最新のリアルタイムストリーミングの重要な機能の1つがスケーラビリティです。インターネット上のデータレートには信頼性がありません。ケーブルモデム経由の接続性は、1日を通じて大きく変化する可能性があります。数多くのユーザが職場へのT1接続を共有しているため、ユーザあたりの使用可能なバンド幅が急激に変化します。

ストリームの開始時にデータレートを計測するだけでは不十分です。そうではなく、プレーヤとサーバがリアルタイムで通信して、ある時点において受信できる以上のデータを送信することなく、使用可能なバンド幅をすべて利用できるようにする必要があります。

現在のスケーラブルな形式はすべて、ビデオおよびオーディオのデータの複数バージョンを同じファイル内に別々のストリームとして格納することで機能しています。こうしたスケーラブルなソリューションは、一般にマルチビットレートまたはMBRと呼ばれます。その詳細は様々ですが、優れたストリーミング体験を提供する上でコーデック自体と同じくらいの重要性を持つ場合があります。

圧縮形式の選択

プロジェクトによって、適切な形式の選択は、簡単な場合もあれば判断に悩む場合もあります。顧客から形式を指定されることもありますが、圧縮を実行する担当者が形式を推奨することが必要になる場合もあります。特定の形式でしか使用できない機能をプロジェクトが必要とする場合もあります。要件が決まっていない場合もあるでしょう。形式を選択したら、その形式をサポートするプレーヤのバージョンを決定する必要があります。この節では、形式を選択する際のガイドとなるいくつかの疑問について説明します。

- **機能** 多くの場合、プラットフォームを決定する最大の要因は、ユーザが必要とする機能です。ある形式がプロジェクトの要件を満たすことができない場合、当然その形式は選択肢から外れます。すべてのプロジェクトで、機能によって重要度が異なります。圧縮効率は、ほとんどすべての場合において重要です。リアルタイムストリーミングのスケーラビリティはもう1つの主要機能です。

- **ユーザベース** プロジェクトとオーディエンスによっては、ユーザベースが非常に重要になる場合があります。ほとんどの広告コンテンツの場合、マーケティングクリップを見るためだけに、ユーザが新しいプレーヤやコーデックをダウンロードすることはまず考えられません。しかし、本当に魅力的なコンテンツであれば、ユーザは手間を惜しまない可能性があります。例えば、Apple 社では、非常に人気のある映画の予告編をリリースした後 72 時間で、数百万件の QuickTime のダウンロードを記録しました。
- **プラットフォーム** 形式によっては、1つのプラットフォームでしか使用できないものあれば、オペレーティングシステムの種類やバージョンに応じて機能が異なるものもあります。最新のメディアプレーヤの多くは、少なくとも Mac OS® と Windows で使用できます。Linux® やそのほかのオペレーティングシステムで利用できるものもあります。幅広い互換性が必要になる場合は、MPEG-1 などの古い標準に基づく形式が適していることもあります。

多くの場合、特にインターネット用のコンテンツを作成する場合は、ユーザが選択できるように複数の形式でコンテンツを配置できれば最適です。

それでも、デジタルメディアのほとんどの消費者は、主要な形式のプレーヤを既にインストールしています。4つのデータレートで5つの形式のファイルを提供する誘惑にかられるかもしれませんが、一般的には、できるだけ少ない形式に集中し、低品質のファイルを20本用意するよりも最適化したファイルを3本作成した方が有効です。

Windows Media 形式への書き出し

Windows Media は、Microsoft 社のデジタルメディア技術です。当初は NetShow として導入されたこの技術は大幅に改良され、2003 年にリリースされた Windows Media 9 プラットフォームは、ポピュラーな独自仕様のデジタルメディアプラットフォームとなっています。

Windows Media を使用したストリーミング配信

Windows Media はもともとリアルタイムストリーミングに焦点を合わせていました。時間とともに、この技術はより多様なものとなりましたが、リアルタイムストリーミングが中核であることには変わりありません。Windows Media 9 は、MBR 技術、サーバーおよびコーデックが卓越した、優れたストリーミングアーキテクチャを備えています。

Windows Media 形式の MBR ソリューションは、インテリジェントストリーミングと呼ばれています。Windows Media 9 より前のインテリジェントストリーミングには制限が多く、ファイルごとに最大5つのビデオデータレート帯域を処理できましたが、これらすべてのデータレート帯域で同じ解像度とオーディオトラックを共有する必要がありました。

インテリジェントストリーミングは Windows Media 9 で大きく改善されました。現在では、ビデオとオーディオの両方について10個以上のデータレート帯域を使用できます。ビデオ帯域では解像度を変更することもでき、低速モデム用のオーディオのみから、ローカルエリアネットワーク (LAN) 用の高精細再生まで対応できる単一のリンクを提供することが可能です。インテリジェントストリーミングに対する唯一の大きな制約は、すべての帯域で同じコーデックとエンコーディングモード (例えば、Windows Media 9 と 2 パス CBR) を使用する必要があることです。

インテリジェントストリーミングのオプションを使用するには、Microsoft Windows Server™ 2003 上で実行される Windows Media 9 サーバーでコンテンツをホストする必要があります。ビデオファイルを Windows Media Server 4.1 for Microsoft Windows NT® 上でホストする必要がある場合は、すべてのビデオ帯域に同じ解像度を使用し、オーディオ帯域を1つだけ使用します。(Windows Media 9 より前の) 古いプレーヤでは、利用可能な機能のサブセットを使用することも必要になります。

Windows Media を使用したプログレッシブダウンロード

Windows Media 7 とそれ以前のバージョンはプログレッシブダウンロードに適していませんでしたが、バージョン 7.1 でこの状況は変わりました。現在、プレーヤは部分的にダウンロードされた素材を正確に再生でき、エンコーダは、2 パスビットレート制限 VBR など、ダウンロードに最適化されたモードをサポートしています。

Windows Media を使用したディスクベース配信

Windows Media Player または別の Microsoft DirectShow ベースのプレーヤが使用されている限り、Windows Media は CD-ROM や DVD-ROM にも利用できます。ただし、ディスクではデータレートの制限がはるかに小さいため、コンピュータのプロセッサがデコードできるピークのデータレートによって再生のパフォーマンスが制限される可能性があります。ピーク制限 VBR モードは、こうした場合に最低限の再生プラットフォーム上でのパフォーマンスを保証するために使用できます。

Macromedia Director の CD-ROM オーサリング環境では、ビデオ再生に直接 DirectShow を使用できません。Director で Windows Media ファイルを使用するには、サードパーティのプラグインが必要です。

DVD-ROM ディスクで高精細の Windows Media 9 コンテンツを使用したマルチメディアタイトルの数が増えています。ディスクの再生には高速なコンピュータが必要になりますが、Windows Media 9 が実用的な最初の高精細ディスク形式であることが証明されつつあります。

拡張子

Windows Media の形式には、ファイルの種類と拡張子が多数あります。最も一般的な拡張子とその意味を次に示します。

.asf

Windows Media ファイルで使用される形式は、Advanced Streaming Format と呼ばれ、元の拡張子は .asf でした。しかし、この拡張子はオーディオのみのファイルとビデオファイルとを区別しなかったため、Windows Media 7 で拡張子 .wmv と .wma が導入されました。現在、.asf ファイルは Microsoft MPEG-4v3 ビデオコーデックを使用する古いコンテンツである場合が一般的です。ただし、ファイル形式自体は変更されていません。

.wmv

Windows Media Video ファイル (.wmv) は、ビデオを収めた .asf ファイルです。

.wma

Windows Media Audio ファイル (.wma) は、.asf ファイルと同じものですが、オーディオしか含まれません。

.asx

.asx ファイルは、Windows Media Streaming のメタファイルです。.asx ファイルは、Web サーバーに配置され、ストリーミングサーバー上のメディアを参照します。.asx ファイル形式自体は変更されていませんが、WM7 の出現により、.wvx と .wax の拡張子がより一般に使用されるようになりました。

.wvx

.wvx ファイルは、ビデオと、オプションでオーディオが含まれるメタファイルです。

.wax

.wax ファイルは、オーディオだけが含まれるメタファイルです。

Windows Media Player

いろいろなプラットフォームで異なる様々な Windows Media Player が使用できるため、互換性の確保が困難になる場合があります。一般的なオペレーティングシステムの中にも、最新版のプレーヤが常に入手できるとは限らないものがあり、複数バージョンのコンテンツを提供するか、または一部のプラットフォームを除外せざるをえないという状況があります。以下の一覧では、最も一般的に使用されているプレーヤについて説明します。

Windows Media Player 6.4 (Windows および Windows NT)

Windows Media Player 6.4 は Windows 98 および Windows 2000 に付属しており、Windows 95 と Windows NT で利用できる最新のバージョンです。Windows 95 には、デジタルビデオ消費者に対する有意なマーケットシェアはありませんが、Windows NT は企業のデスクトップでは現在でも一般に使用されており、配信用のプラットフォームとして重要です。一般に Windows Media Player 6.4 では、パフォーマンスと機能の両面において大きな制約があります。同じ接続上で同じコンテンツに接続した場合、Windows Media Player 6.4 は、新しいバージョンと比べてバッファ処理により長い時間がかかり、再生の信頼性も低くなります。

Windows Media Player 6.4 には、最先端のコーデックとして Microsoft MPEG-4 バージョン 3 が含まれていますが、Windows Media 9 およびそれ以降では、プレーヤは自動的にコーデックを更新できます。この更新には、コンピュータに対する管理者権限と、ダウンロードを許可するファイアウォールが必要です。また、更新されたコーデックをネットワーク上のシステムイメージで提供するために使用できる Enterprise Deployment Pack も存在します。通常、Windows NT システムは古いハードウェアで実行されているため、こうしたシステムで最新のコーデックを再生するとパフォーマンスの問題が発生することがあります。

Windows Media Player 6.4 は、Windows Media Player 9 のような非正方形ピクセルやインターレースビデオに対応していませんが、これらの機能には通常、多くの Windows NT コンピュータが動作している環境よりも高速の CPU が必要になります。

Windows Media Player 9 以降 (Windows)

Windows Media Player 9 は、Windows 98 Second Edition (SE) と同時期の古い Windows オペレーティングシステムで動作します。Windows Media Player 10 は、Windows XP でのみ動作します。これらのプレーヤは、Windows Media Video 9 (WMV9) の完全なスイートをサポートしており、マルチチャンネルオーディオ、インターレースビデオおよび非正方形ピクセルなどの新機能をサポートする唯一のバージョンです。

Windows Media Player 9 for Mac OS X (10.1 以降)

Mac OS 用のプレーヤは、WMV9 コーデックと Windows Media Audio 9 (WMA9) コーデックをサポートしています。このサポートがなければ、この更新は、非正方形ピクセルとインターレースビデオのサポートがない、古い Mac OS プレーヤと変わりません。低解像度コンテンツのパフォーマンスは非常に優れていますが、最速の G5 コンピュータでも高精細の Windows Media Player 9 コンテンツを再生できません。古いとはいえまだ一般的な ACELP.net スピーチオーディオコーデックをサポートする Mac OS X プレーヤは存在しません。

Microsoft 社では、以後このプレーヤを更新する計画はありません。ただし、Flip4Mac™ を使用した Windows Media Components for QuickTime をインストールして、QuickTime Player またはブラウザで直接 Windows Media コンテンツを再生できます。

Windows Media Player Mobile

モバイルプレーヤは、Pocket PC、Smartphone、パームサイズ PC などのモバイル機器用に数多くのバージョンがあります。現在のバージョン 10 のモバイルプレーヤは、バージョン 9 を通じて Windows Media Audio および Video のコーデックをサポートします。マルチチャンネルオーディオなど、機能によっては、完全にはサポートされないものもありますが、サラウンドサウンド対応の Smartphone を所有しているユーザはわずかです。

ほとんどのモバイル機器のプロセッサやビデオパフォーマンスは低速であるため、広範な機器を対象にする場合、解像度、フレームレートおよびデータレートは、コンピュータで再生する場合よりもずっと低くする必要があります。理想的には、ユーザが理想的な再生の設定を調整できるように、特定のモデルを対象にすることをお勧めします。新しい Smartphone では、かなり低いデータレートによる、携帯電話接続を経由したビデオのストリーミングに対応しています。ダウンロード可能な音楽ファイルのデータレートは、もっと高くなる可能性があります。ただし、モバイルユーザは接続時間に基づいて料金を支払っており、数分を超えるダウンロードは敬遠される可能性があることに注意してください。

ハイエンドの Pocket PC の場合、WMV9 の妥当な設定は、ネイティブフレームレートおよび 300Kbps のデータレートで 320 × 240 です。48KHz のステレオまでであれば、WMA9 オーディオが良好に再生されます。

通常、Pocket PC の再生はローカルで行われますが、Pocket PC のプレーヤでは、ワイヤレスのインターネット接続を使用したリアルタイムストリーミングも可能です。

Windows Media Player for Pocket PC 2002

Windows Media Player for Pocket PC 2002 は、WMV8 でしかサポートされていません。利用できる環境が低パフォーマンスのハードウェアだったため、実現するパフォーマンスも低くなります。

WMP for Pocket PC 2000

Windows Media Player for Pocket PC 2000 は、2002 版のプレーヤと同様のものですが、サポートされるのは WMV7 だけで、WMV8 はサポートされません。また、Pocket PC 2000 用の機器は 2002 用の機器よりも低速となる傾向があるため、さらに慎重な圧縮設定が必要になります。

Windows Media ビデオのコーデック

Windows Media ではコーデックの広範なセットが利用できますが、最新のコーデックが最適であることは言うまでもありません。

Windows Media Video 9 (WMV9)

WMV9 は、モデムから高精細までの様々な用途で優れた結果が得られる主要なビデオコーデックです。WMV9 は、エンコードプロセスとデコードプロセスの両方で高い圧縮効率と優れた処理速度を実現します。

現時点では、WMV9 の最大の欠点は、Windows および Mac OS の最新バージョンの Windows Media Player にしかプレインストールされていないということです。

WMV9 は、インターレースビデオと非正方形ピクセルビデオをサポートする、Windows Media プラットフォームで最初のビデオコーデックです。Web アプリケーションでの使用には適していませんが、このコーデックによって、家電製品やビデオオーサリングワークフローに対する Windows Media の適合性が大幅に高まります。

Windows Media Video 8 (WMV8)

WMV8 では、WMV9 の圧縮効率は達成されず（約 20% 低くなります）、非正方形ピクセルモードやインターレースモードもありません。しかし、WMV8 には、WMV9 の様々な 2 パスモードと VBR モードがあります。プレーヤのバージョンが事前にわからない場合、ほとんどの一般的な Windows Media プロジェクトに対して WMV8 は適切な選択となります。

WMV8 の利点は、ほぼすべての主要なプラットフォームでサポートされていることです。こうしたプラットフォームには、Windows 95 および Windows NT、Mac OS 9 および Mac OS X、Pocket PC 2002 などがあります。このコーデックは、Mac OS プレーヤ、Pocket PC 2002 以降および Windows Media Player for Windows 7.1 以降にプレインストールされています。ただし、これらの古いプラットフォームをサポートする必要がある場合は、WMV9 コーデックの高品質を利用することができます。

Windows Media Video 7 (WMV7)

WMV7 をサポートする大半のプレーヤは WMV8 もサポートしています。唯一の例外は Pocket PC 2000 で、これは WMV8 にアップグレードできません。

Microsoft MPEG-4 バージョン 3

古い Microsoft MPEG-4v3 コーデックは、圧縮効率の点で WMV8 および WMV9 よりも遅れていますが、このコーデックは、標準的な Windows Media Player 6.4 でデフォルトで使用できる最新のコーデックです。通常、このコーデックは最新のエンコーダで使用できる最も古いコーデックです。

ISO MPEG-4 バージョン 1

International Organization for Standardization (ISO) MPEG-4 コーデックは、MPEG-4 の標準的なシンプルビジュアルプロファイルを実装したものです。しかし、Windows Media 形式は MPEG-4 のファイル形式や、ストリーミングまたはオーディオのコーデックをサポートしていません。WMV7 には、MPEG-4 へのアクセスを可能にする MPEG-4 コーデックが含まれています。

Windows Media Video 9 Screen

WMV9 Screen コーデックは、ソフトウェアのチュートリアルで使用されるコンテンツのような画面の記録に最適化されています。他のベンダーによるスクリーンコーデックとは異なり、このコーデックでは明示的にデータレートを制御できるため、ストリーミングに適しています。

スクリーンコーデックは、スクリーンアニメーションのようなソース、言い換えると、各フレームの大部分が前のフレームとまったく同じであるような RGB ソースに最適です。また、一部の種類のフラットカラーモーショングラフィックスにも使用できます。

Windows Media Screen バージョン 7

Windows Media Screen バージョン 7 は、WMV9 Screen の古いバージョンです。圧縮効率は多少低くなりますが、そのほかの点では WMV9 Screen と同様に機能します。

Windows Media オーディオ (WMA) コーデック

Windows Media には、各種の用途向けに最適化された様々なオーディオコーデックも用意されています。

Windows Media Audio 9

Windows Media Audio 9 コーデックは、WMA エンコーダの最新バージョンであり、その出力は以前の WMA2 エンコーダと下位互換性があります。現在使用されている Windows Media Player は、すべて WMA9 との互換性があります。名前の先頭に WMA9 が付いたコーデックが他にも存在するため、元の WMA は現在 WMA Standard と呼ばれています。ただし、エンコーディングツールに表示されることはありません。この資料では、この標準コーデックを WMA9 で表します。

現在、WMA9 は WMA9 ビットレートモードをすべてサポートします。ただし、Windows Media Player 6.4 との互換性が必要な場合は、CBR モードを使用する必要があります。ビットレート VBR モードでは、データレートを 48 ~ 192Kbps に制限することができます。CBR モードでは、5 ~ 320Kbps の範囲で選択できます。品質制限 VBR モードでは、完全な 0 ~ 100 の範囲は提供されませんが、ドロップダウンメニューで選択できる数多くのオプションが用意されています。

WMA9 CBR 用の 0Kbps オプションもあります。このオプションは、オーディオなしのストリーミングファイルに、Windows Server 2003 より前に使用されていた古い Windows Media Services 4.1 との互換性を持たせるためのものです。このサーバーのインテリジェントストリーミングでは、各ファイルに常に 1 本のオーディオトラックが必要になるため、この 0Kbps オプションによって、オーディオなしのファイルを使用する手段を提供しています。

ACELP.net

ACELP.net は、低ビットレートのスピーチコーデックです。このコーデックの音質は、24Kbps 未満のスピーチコンテンツの場合、WMA9 よりも良好とされています。ただし、プラットフォームによっては、デコーダが利用できない場合があります。特に、Mac OS X と Pocket PC では使用できません。

Windows Media Audio 9 Voice

WMA9 Voice は、ACELP.net に代わる、低ビットレートのスピーチコーデックです。

WMA9 Voice は、音楽モードまたは混合モードで再生される場合に、非常に低いビットレートの音楽コンテンツで良好に機能します。WMA9 Voice は WMA 9 Narrowband と呼ばれています。一般に、20Kbps 以下の音楽の場合は、WMA9 よりも WMA9 Voice を選択する方が適切です。

WMA9 Voice はモノラルで、最大サンプルレートが 22KHz、最大データレートが 20Kbps の CBR しかありません。このコーデックが適しているのは、低ビットレートのアプリケーションだけです。

Windows Media Audio 9 Professional

WMA9 Professional は、マルチチャンネルおよび高分解能オーディオ用の新しいコーデックです。最低 128Kbps のデータレートが必要になるため、ほとんどの Web コンテンツには適しません。このコーデックは、3 つ以上のオーディオチャンネル、48KHz よりも高いサンプルレート、16 ビットより大きいビット数をサポートする主要アーキテクチャの 1 つにおける最初のコーデックでした。WMA9 Professional は、高データレート要件として上記いずれかの機能を必要とするソースでのみ使用してください。高ビットレートの WMA9 は、48KHz 以下のステレオ 16 ビットソースに有効です。

マルチチャンネルの WMA9 Professional は、ステレオスピーカしかないシステムでも、チャンネルを自動的に結合して再生できます。他にも、7.1 から 5.1 に、あるいは、左、右、中央のゲーム用システムに切り替えるなど、より少ない再生チャンネルに変換することができます。それでも、数多くのリスナーがマルチチャンネルオーディオ再生を行うことが予想される場合はマルチチャンネルが最適です。

WMA9 Professional は、2 パス CBR および 3 種類の VBR すべてを含む、WMA9 のすべてのデータレートモードをサポートしています。他の例と同様、ストリーミングには 2 パス CBR、ファイルベースの再生には VBR を使用する必要があります。

Windows Media Audio 9 Lossless

WMA9 Lossless コーデックは、オーディオのアーカイブに使用する、数学的に可逆なコーデックです。可逆圧縮では特定のデータレートを保証できないため、1 パス VBR だけになります。WMA9 Lossless が機能するプレーヤは、WMA9 Professional の場合と同じです。Lossless は、配信には過剰なデータレートであり、使用するのはアーカイブやマスタリングだけにしてください。また、このコーデックは、オーディオ CD コンテンツをデジタルアーカイブする目的でも広く使用されています。

RealMedia 形式への書き出し

RealMedia は、最も古いリアルタイムストリーミング Web 形式です。先駆的な RealAudio は 1995 に出荷されました。RealNetworks 社は機能強化を続け、2004 年 1 月には更新された Real 10 プラットフォームが発表されました。

Windows Media とは異なり、大多数の RealMedia ツールは合計データレート（これからオーディオデータレートを差し引いたものがビデオデータレートになります）を指定します。ただし、ツールによっては、この方法を使用しないものもあります。特定のツールが使用するモードを知っておくことが重要です。

RealMedia を使用したストリーミング配信

RealMedia は最初のストリーミングプラットフォームでした。RealMedia は、ストリーミング環境で実力を発揮します。RealMedia の MBR ソリューションは SureStream と呼ばれています。Windows Media とは異なり、SureStream では異なる複数のストリームでコーデックを混合および適合させることができます。すべてのストリームが同じフレームサイズを共有する必要があるため、スケラビリティは制限されることがあります。

RealMedia を使用したプログレッシブダウンロード

RealMedia は主としてストリーミング形式と考えられていますが、ここ数年の間に、優れた .rmvb 2 パス VBR プログレッシブモードをサポートするようになりました。これらのファイルは実際には標準の .rm ファイルとは異なる拡張子を使用しており、このファイルをストリーミングサーバーに配置しても正常に機能しないことを示しています。

RealMedia を使用したディスクベース配信

.rmvb ファイルは、CD-ROM や DVD-ROM から使用することもできます。Microsoft PowerPoint など、ほとんどのソフトウェアツールでは、RealMedia ファイルを直接埋め込むことはできませんが、このファイルは RealOne プレーヤーでディスクから問題なく再生できます。また、RealMedia は優れた高精細再生ソリューションでもあります。RealMedia のコーデックのデコード要件は WMV9 よりも軽量なため、ローエンドのコンピュータで高解像度の再生が可能です。また、新しい Real 10 にはマルチチャンネルオーディオサポートも含まれています。

RealMedia ビデオコーデック

RealMedia では、他の形式と比べ、サポートされるビデオコーデックの数が少なくなっています。常に最新のコーデックを選択することをお勧めします。

RealVideo 10

新しい RealVideo 10 は、RealVideo 9 に対する下位互換性がある拡張です。古いコーデックと比較すると、圧縮効率と、圧縮品質とエンコード時間を調整するための制御が大幅に向上しています。

RealVideo 9

RealVideo 9 コーデックは、優れた汎用コーデックです。このコーデックには、ローエンドではモデムのデータレートおよびモバイル機器による再生から、ハイエンドでは高精細の再生にまで対応できる、優れた拡張性があります。

RealVideo 9 は、RealOne Player (Windows および Mac OS X) に標準で付属しており、RealPlayer 8 (Mac OS 9 での最新バージョン) 用の自動コーデックアップグレード版です。

RealVideo 8

RealVideo 8 は、以前のバージョンの RealVideo です。RealVideo 9 は、RealVideo 8 と同様に広く使用されており、より高い圧縮効率を実現できるため、通常は RealVideo 9 を代わりに使用する方が適切です。

RealVideo G2

RealVideo G2 は、現在の技術に基づく最初の RealVideo コーデックでした。このコーデックを必要とするのは、古い RealPlayer のインストール環境だけです。

RealMedia オーディオコーデック

他の形式とは異なり、RealAudio コーデックは、コーデック、データレート、サンプルレートおよびチャンネルの組み合わせによって定義されます。ユーザはコーデックと他のパラメータを選択するのではなく、包括的な組み合わせを選択します。RealMedia オーディオサポートのもう1つの特徴は、多くのツールで、ビデオと音楽のコンテンツに対して、つまりビデオとオーディオのファイルとオーディオだけのファイルの両方に対して、オーディオコーデックを指定する必要があることです。ユーザは、ファイルのエンコードに使用されたモードに基づいてオーディオコーデックを選択します。実際に使用されているモードに対する正しい設定を選択することが重要です。

RealAudio 10

RealAudio 10 は、128Kbps 未満のビットレートには古い RealAudio 8 Stereo Music コーデックを使用し、これより高いビットレートには AAC-LC を使用します（以前のバージョンでは高いビットレートに Sony の ATRAC3 が使用されていました）。このコーデックは、低ビットレートでは RealPlayer 8 と下位互換性がありますが、高ビットレートではプレーヤーで更新が必要になります。AAC-LC の詳細については、MPEG-4 オーディオコーデックを参照してください。

RealAudio 10 Lossless

このコーデックでは RealPlayer 10 が必要になります。44.1KHz ステレオの場合、データレートは 700Kbps と推定されるため、このコーデックは Web やストリーミングのソリューションにはなりません。

RealAudio 10 Multichannel

このコーデックは、RealPlayer 10 を必要とする 5.0 または 5.1 チャンネルコーデックです。高いデータレートが必要になるため、ストリーミングソリューションにはなりません。

RealAudio 8 ステレオミュージック

RealAudio 8 ステレオミュージックは、RealMedia 形式の主要な汎用オーディオコーデックです。このコーデックの圧縮効率はモノラルの RealAudio ミュージックコーデックよりも高いため、このコーデックは、モノラルオーディオでも 12Kbps の低データレートで問題なく機能します。

データレートによっては、このコーデックの「高周波」版を使用できるものもあります。この高周波版は、より広い周波数範囲を提供します。ただし、複雑なオーディオの場合、圧縮ノイズが増える可能性もあります。実験を行って、様々な種類のコンテンツで適切に機能する設定を見つける必要があります。

RealAudio 8 Surround

RealAudio 8 Surround コーデックは、RealAudio 8 ステレオミュージックコーデックのバリエーションです。このコーデックは、実際には 5.1 または 7.1 のシステムではありません。代わりに、このコーデックはステレオペアを格納しつつも、Dolby Pro Logic サラウンドサウンドエンコーディングを維持できます。これは、ベースとなる RealAudio 8 ステレオミュージックコーデックではできないことです。このコーデックは完全なマルチチャンネル体験を提供するものではありませんが、広範な機器との互換性があります。

このコーデックは、ソースが Dolby Pro Logic サラウンドサウンドミックスの場合に適しています。通常のステレオミュージックの場合は、RealAudio 8 ステレオミュージックコーデックの方が、より高い品質と柔軟なデータレートオプションを提供します。

RealAudio Speech

RealAudio スピーチコーデックは、比較的標準的なスピーチコーデックです。最低のデータレートでもスピーチが明瞭で、高データレートでの音質も優れています。大部分のスピーチコーデックとは異なり、このコーデックは、オーディオブックなどのコンテンツに対して高い忠実度を持つ 32Kbps と 64Kbps のオプションを提供します。

RealAudio ステレオミュージック

RealAudio Stereo Music コーデックは、G2 リリースの最初の RealAudio ステレオコーデックです。このコーデックは、RealAudio 8 ステレオミュージックコーデックよりも圧縮効率が低くなります。このコーデックは幅広い下位互換性を提供します。

RealAudio ミュージック

RealAudio ミュージックコーデックは、G2 リリースの最初の RealAudio モノラルコーデックです。モノラルコンテンツの場合でも、RealAudio 8 ステレオミュージックコーデックの方が高い品質を実現します。しかし、RealAudio ミュージックコーデックは、RealAudio 8 ステレオミュージックコーデックよりも低いデータレートで使用できるため、場合によっては、低ビットレートでも有効です。

QuickTime 形式への書き出し

QuickTime は主要なデジタルメディア形式の中で最も古いものです。QuickTime が発表されたのは 1991 年です。インターネットが現在のように広く普及する前、CD-ROM ドライブがまだ標準的なコンピュータ機器ではなかった時代からあった技術ですが、成熟して幅広いアプリケーションをサポートするようになりました（もともと、リアルタイムストリーミングについては、RealMedia や Windows Media の先進性には及びません）。QuickTime の最大の強みは、リッチインタラクティブティとリッチメディアをサポートしていることです。

デフォルトでは、QuickTime のデータレートは毎秒の合計キロバイト数として設定され、RealMedia の毎秒のキロビット数や Windows Media のストリームあたりの毎秒のキロビット数とは異なります。しかし、複数の形式をサポートするツールの多くは、Kbit 単位でデータレートを提供します。一貫性を保つため、以下ではデータレートを Kbit 単位で示します。

この節では、MPEG-4 や AVI など、他の多くの形式を制作および再生できる QuickTime アーキテクチャではなく、QuickTime 形式 (.mov) について説明していることに注意してください。

代替ムービー

QuickTime のスケーラビリティに対するアプローチは、Windows Media や RealMedia の形式とは大きく異なります。QuickTime では、データレートが異なる複数のバージョンのデータが含まれる単一ファイルを作成する代わりに、データレートが異なる複数のファイルを使用します。これらのファイルは、ムービーの再生が始まる時に切り替わります。

このアプローチには、いくつかの欠点があります。第 1 に、切り替えはファイル再生の開始時にしか行われなため、再生中の切り替えはできません。第 2 に、制作作業がより複雑になります。Adobe Media Exporter などの数少ない一部の高度な圧縮用製品しか、エンコード時の代替ムービーの制作をサポートしていません。第 3 に、データレートを決定する QuickTime の環境設定を正しく設定する役割を単純にユーザに依存しています。したがって、ユーザが職場のブロードバンド接続を使用してラップトップ上の QuickTime を設定し、その後、ホテルの部屋からダイヤルアップモデムを使用して接続すると、QuickTime はブロードバンドバージョンのファイルを送信しようとするため、リアルタイムストリーミングであれば全く使用に耐えないか、またはプログレッシブファイルの場合はダウンロード時間が非常に長くなります。

これらの理由から、一般的な代替ムービーのリンクを提供している場合でも、各メディアファイルへの直接リンクを提供することをお勧めします。

代替ムービーには、ユニークな機能がいくつかあります。第 1 に、代替ムービーは、ストリーミングファイルだけでなく、プログレッシブファイルにも対応しています。もう 1 つの機能としては、データレート以外のパラメータによるフィルタリングができます。Mac OS および Windows のユーザ用に、様々な言語を使用しているコンピュータ用に、また、様々な CPU 速度のコンピュータ用に、異なるバージョンを提供できます。

QuickTime を使用したリアルタイムストリーミング

QuickTime には、SureStream やインテリジェントストリーミングのようなリアルタイムスケラビリティソリューションはありません。QuickTime は、データレートが多様なため、インターネット上のリアルタイムストリーミングにはあまり適していない可能性があります。ただし、QuickTime はイントラネット環境には適しています。1つの利点は、専用のストリーミングサーバーが、無償で利用できるオープンソースの Darwin Streaming サーバであることです。この製品は非常に広範なプラットフォームで利用できます。

QuickTime にストリーミングサポートが追加されたのは、開発の比較的後の段階でした。ファイルが常にストリーミング可能な他の形式とは異なり、QuickTime では、ストリーミングするファイルにヒントトラックを追加する必要があります。この追加データが、データのストリーミング方法に関する指示をサーバに与えます。なお、このデータ自体は送信されないため、ヒントを付加したムービーファイルのサイズは、クライアントに実際に送信されるデータの合計よりも大きくなります。

QuickTime を使用したプログレッシブダウンロード

QuickTime はプログレッシブダウンロードの先駆者であり、この形式は現在でも QuickTime で Web ビデオを配信するための最もポピュラーな手段です。

プログレッシブ QuickTime ムービーの複雑さの1つは、即再生可能なムービーである必要があることです。すべての QuickTime ムービーにムービーヘッダーがあります。このヘッダーにはファイルの構造に関する情報が含まれており、再生の際に必要となります。ファイルを変更するたびに、QuickTime によってムービーヘッダーがファイルの末尾に移動されます。つまり、ムービーはすぐには開始されず、ムービー全体がダウンロードされるまでは、QuickTime のロゴだけが再生されます。ムービータイトルなどのメタデータを編集した場合でも、この変更が強制的に行われる可能性があります。ヘッダーをファイルの先頭に移動させる最も単純な方法は、QuickTime Player Pro の中で「別名で保存」を実行することです。

プログレッシブファイルと一緒にヒントトラックを使用しないでください。ヒントトラックは使用されておらず、また、ファイルのサイズを増加させるためです。代わりに、プログレッシブファイルでは圧縮ムービーヘッダーモードを使用します。このモードを使用するとヘッダーが小さくなります。

QuickTime を使用したディスクベース配信

QuickTime は CD-ROM やその他のディスクベースのアーキテクチャに最適のソリューションです。QuickTime のユーザインタフェイスを使用すると、非常に簡単に QuickTime 再生をアプリケーションに追加できます。また、QuickTime は、Director 内で最もサポートされている形式です。ただし、Windows の PowerPoint は、ネイティブでは QuickTime をサポートしていません。

QuickTime ビデオコーデック

柔軟性およびサードパーティのコーデックベンダーに対するサポートという QuickTime の長年にわたる遺産を受け継いでいるため、QuickTime でサポートされるコーデックは、他の形式よりも広範です。ただし、多くの場合、最適なコーデックは無償では提供されていません。

Sorenson Video 3

QuickTime で使用可能な最高品質のコーデックは、Sorenson Media 社の Sorenson Video 3 です。このコーデックは非常に柔軟なコーデックであり、プログレッシブダウンロード環境とストリーミング環境の両方で機能します。基本バージョンには、圧縮効率とストリーミングパフォーマンスを向上させる様々な拡張機能は含まれていません。

QuickTime には、Sorenson Video 3 の無償の基本バージョンが付属します。ただし、プロフェッショナル用途には、Sorenson Video 3.1 Professional アップグレード版が利用できます。Sorenson Video 3.3 Pro では、圧縮速度と品質が向上します。前者はデュアルプロセッサコンピュータで、後者はリアルタイムストリーミングプロトコル (RTSP) ストリーミングで特に有効です。商用の QuickTime エンコーディングには、常に Pro バージョンが必要になります。Pro エンコーダでは、ファイルと付属デコーダとの完全な互換性が確保されます。ファイルを表示するためにユーザアップグレードは必要ありません。

Sorenson Video 3 Pro のオプション

Sorenson Video 3 は、2001年の夏にリリースされたバージョン 5.0.2 以降の QuickTime に付属しています。

- **Quick Compress** このオプションは、品質をわずかに低下させることで、エンコーダを若干高速化します。どちらの効果も大きくはないため、ほとんどの場合、このオプションは選択解除の状態にしておくことができます。

- **Bidirectional Prediction** Sorenson Video 3 で B フレームエンコーディングを使用し、圧縮効率とスケーラビリティを付加するには、このオプションを選択します。スケーラビリティについては、プロセッサの速度がすべてのフレームを再生するには十分でない場合や、データレートが RTSP を使用してすべてのフレームをストリーミングするのに十分でない場合、B フレームがドロップされます。このオプションでは、フレームレートを半分に落とすと代わりに、フレームを自動的にドロップし、CPU 要件を最低でも 50%、データレートを約 25%それぞれ削減します。B フレームがなければ、バンド幅に負荷がかかるとキーフレーム以外のすべてのフレームがサーバによってドロップされるため、このオプションはストリーミングで特に有効です。

したがって、すべての場合において双方方向予測を使用することが理想的です。ただし、B フレームには、オーディオに対してビデオが 2 フレーム遅れるという問題が 1 つあります。29.97fps では、この遅延は少し気になる程度ですが、これより低いフレームレートでは、ビデオと音声明らかにずれる可能性があります。また、QuickTime 5 より前のバージョンでは、B フレームが原因で動作が不安定になる場合もあります（この問題は、QuickTime 6 で修正されたようです）。

既にエンコードされたファイルの同期を修正する Mac 用の AppleScript があります。

- **Rate Control Mode** このオプションのラベルなしのレートコントロールボックスには、3 つの追加オプションがあります。最初のオプションは、標準的な CBR エンコードを実行します。

「Sorenson One-Pass VBR」オプションはデータレートを緩和します。このモードは、RTSP のエンコーディングと、ピークデータレートが問題にならない大部分の非 RTSP エンコードファイルに最適です。「Streaming」オプションを選択した場合、VBR ウィンドウが、ストリーミングに適した値に制限されます。

「Minimum Quality」オプションは、フレームごとに最低品質しきい値を 0 ~ 100 の範囲で設定します。「Drop Frames To Maintain Data Rate」オプションを選択した場合、品質しきい値を超過するとフレームレートが下がり、データレートが維持されます。このオプションを選択しない場合、最低品質が満たされなくなるとデータレートが上昇します。つまり、最終的にデータレートが高くなりすぎることを意味します。

- **Image Smoothing** このオプションはデブロッキングフィルターを有効にします。これにより画像のブロックノイズを低減します（ブロックノイズが発生している画像には、陰影のない色の長方形またはブロックが含まれます）。このオプションは、高データレートではそれほど効果がないため、コンピュータの速度がこの操作を実行するのに十分でない場合、再生時に自動的に無効になります。通常、このオプションは有効にしてください。
- **Media Key** 「Media Key」オプションはビデオにパスワードを割り当てます。ビデオを再生するには、プレーヤアプリケーションからこのパスワードを入力する必要があります。
- **Streaming** RTSP ストリーミングされるすべてのファイルについて、このオプションを選択する必要があります。ほとんどのプロジェクトでデフォルト設定が適切です。

「Slice Picture Into Packets Of Bytes」オプションは変更しないでください。このオプションは、ヒンテイング設定と同期して変更する必要があり、一部のプライベートネットワークを除いて、他の値は有効ではありません。

「Force Block Refresh Every」オプションは、指定された間隔でビデオのすべての 16 × 16 ブロックがリフレッシュされるようにします。この更新は、エラーリカバリーに役立ちます（パケットの欠落による欠損が少なくともその頻度で消去されます）。このオプションを使用することで、ファイル内のキーフレームの数を減らすことができます。このオプションにより、静止画での圧縮効率は若干下がりますが（これはそれほど重要ではありません）、動きによってブロックが自然にリフレッシュされる）速い動きにはそれほど影響を与えません。

- **Masking** ビデオにリアルタイムアルファチャンネルを含めるには、このオプションを使用します。コーデックは、提供されているアルファチャンネルを使用するか、内部クロマキーヤーに基づいてアルファチャンネルを生成することができます。内部クロマキーヤーの機能は限定的なため、After Effects などのプロフェッショナルツールを使用してアルファチャンネルを生成することで、はるかに良い結果が得られる可能性があります。ソースファイルでアルファチャンネルを使用するには、色を数百万色 + に設定する必要があります。
- **Watermarking SV3.1 Pro** ビデオにリアルタイム合成ウォーターマークを追加するには、このオプションを使用します。このウォーターマークは、ビデオソースにウォーターマークを追加する場合に比べてはるかに高い品質を実現できます。これは、多くの場合、ソースウォーターマークの細部が圧縮で失われてしまうためです。

ウォーターマークを不透明にしたり、色を付けたりすることができます。なお、ウォーターマークのファイル形式には注意が必要です。コーデックのマニュアルで、ウォーターマークの作成方法を説明している節をよく読んでください。Photoshop を使用してアルファチャンネルを Sorenson ウォーターマークに変換する場合、出発点として既存のアルファチャンネルを 50% のグレー背景に適用し、続いてイメージをフラットにするとよいでしょう。

Sorenson Video

QuickTime は古い Sorenson Video コーデック (番号は記載されていませんが、バージョンは 2 です) も搭載しています。このコーデックは圧縮効率と品質が低いため、現在の Sorenson Video 3 とはまったく異なるコーデックです。バージョン 5 より前の QuickTime のバージョンとの互換性が必要な場合に、このコーデックを使用できます。

Sorenson Video 2.2 Developer Edition と呼ばれる最初の Sorenson Video コーデックのアップグレード版が存在しましたが、現在は市販されていません。

MPEG-4

QuickTime は、MPEG-4 コーデックを搭載しています。このコーデックは、Mac OS のエンコーダでデフォルトで使用できます。ただし、Windows では、開発元がライセンス料を支払ってこのコーデックを有効にした少数の Windows エンコーダでしか使用できません。

QuickTime は再生に関して Simple Visual 規格をサポートしており、サードパーティの MPEG-4 エンコーダを使用して、QuickTime を再生するファイルを作成できます。通常、サードパーティのエンコーダは、2 パス VBR エンコーディングなどの技術を使用して、より高い圧縮効率を実現します。

Sorenson Video 3.1 Pro は、MPEG-4 よりも高い圧縮効率とストリーミングパフォーマンスを実現します。QuickTime における MPEG-4 の最大の利点はデコーダ速度であり、Sorenson Video 3 のデコーダ速度よりも高速です。

MPEG-4 は、2002 年の夏に出荷された QuickTime 6 以降の QuickTime に搭載されています。

H.263

H.263 は、QuickTime 4 以降の QuickTime に搭載されている標準的なテレビ会議用コーデックです。このコーデックは良好なストリーミングを行い、ストリーミングパフォーマンスも優れています。このコーデックでは、ストリーミングに関して Sorenson Video 3 の無料バージョンよりも (まれに) 良好な結果が得られる可能性があります、Sorenson Video 3.1 Pro ほどは優れていません。

H.263 のユニークな機能の 1 つは、リストされている 3 つの解像度しか内部的に使用しないことです。別の解像度を選択された場合、このコーデックは、その次に低い解像度を内部的に使用し、リアルタイムでビデオを正しい解像度にスケールアップします。必要に応じて特定の解像度を指定することもできますが、ほとんどの場合、この指定は不要です。

Cycle Intra Macroblocks 機能は、Sorenson Video 3 の Block Refresh と同様の機能を実行するもので、圧縮に伴うデータ損失が問題となるネットワークで特にストリーミングパフォーマンスを向上します。この機能は、RTSP でのみ有効にしてください。

Cinepak

Cinepak は、1990 年代の多くの期間で主要な CD-ROM コーデックでした。このコーデックは、10 年以上前のバージョン 1.5 以来、QuickTime に搭載されてきました。Cinepak の特徴は非常に高速なデコーダであることです。320 × 240 のビデオが 60 MHz のコンピュータで良好に再生されました。ただし、エンコーダが単純であるために、圧縮効率は現在のコーデックの圧縮効率を大きく下回ります。Cinepak は、白黒のソースでより高速なデコードパフォーマンスを実現する、ネイティブグレースケールモードをサポートした数少ない配信コーデックの 1 つです。

最近では、Cinepak は、K-12 教育期間でよく見られるような、古くて低速なコンピュータで高解像度の再生を行う目的で主に使用されています。データレートが最低 5000Kbps であっても、640 × 480、29.97fps のビデオを 5 年前のコンピュータで問題なく再生できます。

VP3

On2/The Duck Corporation 社は、10 年間にわたって商用コーデックを開発してきました。会社による最新の VP5 および VP6 のコーデックは、商用製品としては入手できませんが、VP3 コーデックはオープンソースであり、QuickTime では無償で入手できます。VP3 コーデックは標準インストールに含まれていませんが、コンポーネントダウンロードを通じて入手できます。つまり、このコーデックを使用するコンテンツが検出されると、コーデックをダウンロードするかどうかの確認が QuickTime によって自動的に行われます。インストールを強制するには、QuickTime Updater ユーティリティを実行します。VP3 コーデックは、サードパーティオプションの 1 つです。

VP3 コーデックには、ネイティブパケタイザがないため、このコーデックは RTSP ストリーミングに適したツールではありません。また、VP3 コーデックは数年前のものであり、圧縮効率が Sorenson Video 3.3 Pro よりも低く、データレートも予測不可能であるため、Web 配信には適しません。ただし、デコーダは Sorenson Video 3 よりも高速であるため、VP3 コーデックは高解像度の CD-ROM およびキオスク用途で実行可能なオプションになっています。

VP3 コーデックはオープンソースであるため、実装によって違いがあります。例えば、Windows で AVI ファイルを作成するための、このコーデックの DirectShow 版では、Mac OS エンコーダでは現在使用できないオプションがいくつかあります。ただし、Mac OS プレーヤはこれらの機能を使用してエンコードされたファイルを再生できます。

VP3 コーデックの「Allow Dropped Frames」オプションは、データレートを維持するためにフレームをドロップするようにエンコーダに指示します。VP3 コーデックの最低品質はかなり高いため、このオプションを選択しない場合、目標データレートを大幅に超過する可能性があります。しかし、フレームのドロップは気になるものであるため、VP3 コーデックは高い圧縮効率で使用するための最適なコーデックではありません。

Quick Compress モードはエンコードを大幅に高速化しますが、品質の改善は最小限にとどまります。他のコーデックとは異なり、デフォルトでは Quick Compress モードで VP3 コーデックを実行するのが安全です。さらに時間がある場合は、このモードを無効にしてエンコードを行うことにより、できる限り高品質で高いデータレートが保証されます。

VP3 には、大部分のコーデックよりも複雑なキーフレーム挿入モードがあります。しきい値を設定できるだけでなく、キーフレームの最小数と最大数を指定できます。Threshold および Minimum のデフォルト設定は、ほとんどのプロジェクトに適しています。ビデオに非常に高速な編集が含まれる場合は、より低い最小値を設定できます。また、非常に高いフレームレートで作業する場合は、高い最大値を設定できます。このダイアログボックス以外のエンコーディングアプリケーションで設定されたキーフレームレートが、フレームの最大数よりも大きいことを確認してください。

Animation

Apple 社のアニメーションコーデックは、インターフレーム Run-Length Encoded (RLE) RGB コーデックです。したがって、高い圧縮効率を実現するため、このコーデックはフレームごとに全く同じ色の長い水平方向の線に依存します。Animation コーデックにはデータレートコントロールはなく、品質スライダーしかありません。したがって、このコーデックは、隣接する 2 つのピクセルが全く同じになることがめったにない自然画像から制作したビデオのエンコーディングには適していません。このコーデックは、記録済みのスクリーンキャプチャを適切に処理できますが、そのためにはワークフローの高度な技巧がいくつか必要になります。これについては、この文書では割愛します。

Animation コーデックは、品質スライダーが 100% に設定されている場合は可逆です。これより低いレベルに設定すると、ほとんど同一の水平方向の線が単純に作成されます。このプロセスによって、多くのソースで品質が低下し、目に見えるバンドリングが発生する可能性があります。名前にもかかわらず、Animation コーデックは、照明効果によって等価のピクセルで構成される領域が発生しない 3 次元最適化のエンコーディングにはあまり適していません。

Animation コーデックを使用して中間ファイルを保存することもできます。Animation コーデックは、After Effects などの RGB ネイティブツールで RGB ワークフローを使用する場合、特に合成画像の場合に、中間ファイルを適切に保存します。しかし、Animation コーデックによるデータレートは、自然画像の場合、非圧縮データレートと比較してそれほど向上しません。また、圧縮ツールによっては、インターフレームエンコードファイルで正常に機能しないものもあるため、このような場合にはすべてのフレームでキーフレームを使用する必要があります。Portable Network Graphics (PNG) コーデックと Microcosm コーデックは、このような場合に同じ品質でより高い圧縮効率を実現します。

Graphics

Graphics コーデックは、Animation コーデックと同様に可逆 RGB コーデックです。違いは、Graphics コーデックの方がより高い圧縮効率を実現する点ですが、8 ビットカラーでしか動作しません。Graphics コーデックは優れたスクリーンレコーディングを実現しますが、これは最初にファイルを 8 ビットに変換した場合だけであり、これは手間のかかるプロセスです。

PNG

PNG は、もともと GIF のパテントフリーの代替物として開発された RGB 専用コーデックです。しかし、このコーデックは進化して、より高い圧縮効率を実現し、8 ビット以上のインデックスカラーをサポートするようになりました。PNG コーデックは、スクリーンショットなど、フラットカラーの領域がある人工コンテンツのエンコーディングで効率が非常に高くなります。PNG コーデックは、インターフレーム専用であり、データレートコントロールモードがないため、動画ビデオの配信コーデックには適していません。PNG は、数百万色 + モードをサポートしているため、アルファチャンネルを伝送できます。

PNG のインターレースモードはインターレースフィールドには適していません。その代わりに、このモードは画像のプログレッシブデコーディングに使用します。このオプションを選択すると、ダウンロード中、画像として低解像度プロキシが最初に表示されます。このコーデックは、圧縮効率を若干低下させるため、アーカイブ目的にはあまり適していません。

PNG は常に可逆ですが、Filter モードがエンコード時間と最終的なファイルサイズ間のトレードオフを制御します。圧縮効率を最大にするには、「Best」オプションを使用します。一定の圧縮効率を得ながら、速度を最大にするには、「Up」オプションを使用します。

Microcosm

Microcosm コーデックは、もう 1 つの可逆 RGB インターフレーム専用コーデックです。このコーデックの PNG に対する大きな利点は、16 ビット / チャンネルをサポートしていることです。したがって、After Effects など、16 ビット / チャンネルをサポートするツールとともに使用すると、このコーデックはコンテンツの可逆アーカイブを実行できます。この 16 ビットサポートが優れた効果を実際に発揮するのは、フィルムや放送用に配信する場合、特に Cineon や 1 チャンネルあたり 8 ビットを超えるソースで作業する場合です。Microcosm コーデックは、アルファチャンネルを伝送することもできます。

Microcosm コーデックの無料デコーダはダウンロードにより入手できますが、エンコーダは購入する必要があります。

None16

None16 は、Microcosm の開発者による無料コーデックです。このコーデックは、16 ビット / チャンネル RGB とアルファチャンネルを提供します。ただし、完全に非圧縮であるため、ファイルサイズが非常に大きくなります。

Photo-JPEG

Photo-JPEG は、標準の JPEG 静止画コーデックを実装したものです。このコーデックは、8 ビット / チャンネルの 4:2:0 Y'CbCr またはグレースケールをサポートする不可逆のコーデックです。ただし、ほとんどのコンテンツにおいて、品質設定が 85 以上の場合、見た目には可逆です (Photo-JPEG には明示的なデータレートコントロールがありません)。

Photo-JPEG は、スライドショー用のコンテンツ配信には適している可能性があります。ほとんどのビデオアーカイブや転送の用途には、Motion JPEG の方が適しています。Photo-JPEG ではフィールドが認識されないため、プログレッシブコンテンツに最適です。

モーション JPEG

モーション JPEG は、JPEG の 4:2:2 バリエーションであり、フィールドをサポートしています。モーション JPEG は、多くの NLE システムのネイティブ形式になっています。NLE システムの多くは、独自のネイティブ形式から汎用 QuickTime モーション JPEG コーデックへの可逆トランスコードを実行できます。

モーション JPEG コーデックには、A と B の 2 種類があります。もともとこれらには、初期の NLE カードでサポートされていた別々のチップセットに対応する意味がありました。これらは機能的には同等であるため、A を使用するのが最も簡単です。オプションダイアログボックスで、正しいフィールドモードとフィールド優先度を手動で設定する必要があります。

Component

Component コーデックは、非圧縮 Y'CbCr 8 ビット / チャンネル 4:2:2 コーデックです。このコーデックは非圧縮のため極めて高速ですが、データレート要件が高くなる可能性があります。このコーデックは、ディスク容量が大量にあるユーザ用の中間形式として主に使用されます。Component コーデックは、アルファチャンネルをサポートしていません。

None

None コーデックは、非圧縮 RGB コーデックです。このコーデックは、16 ビット / チャンネルモードを除く、アルファチャンネルも含めたすべての RGB QuickTime 色深度を提供します。QuickTime の可逆 RGB コーデックの多様性を考えると、純粋な速度が必要でない限り、None コーデックを使用する必要はめったにありません。

QuickTime オーディオコーデック

QuickTime オーディオコーデックのメインユーザインタフェイスには、特定のデータレートコントロールがありません。その代わりに、オプションダイアログボックスでデータレートを指定します (制御可能な場合)。

MPEG-4

現在、QuickTime で最高の汎用オーディオコーデックは、MPEG-4 AAC-LC です。デフォルトでは、エンコーディングは Mac OS にしか含まれません。Windows ユーザは、ほとんどのアプリケーションでこのコーデックを利用できません。MPEG-4 コーデックでは、AAC-LC (汎用コーデック) と低ビットレートスピーチに使用される Code Excited Linear Prediction (CELP) のオプションがリストされます。実際には、CELP は QuickTime 6.5 以降では使用できません。

MPEG-4 ビデオと同様に、MPEG-4 オーディオはオープンな MPEG-4 規格に基づいています。サードパーティの AAC-LC エンコーダを使用して、QuickTime にインポートし、ビデオトラックとして使用できるファイルを作成できます。MPEG-4 には、ネイティブパケタイザがあるため、リアルタイムストリーミングとファイルベース配信の両方に適しています。

AAC-LC は非常に効率的な最新のコーデックです。128Kbps を超えるデータレートは、48KHz のステレオの場合であっても、ほとんどのコンテンツに対して過剰です。音楽には、48Kbps、44.1KHz のモノラルで十分です。場合によっては、これらを下回る値でも十分です。

なお、このコーデックでは、Audio ダイアログボックスで指定したサンプルレートから、Codec ダイアログボックスで指定したレートに自動的に変換されます。44.1 または 48 で AAC-LC にエンコードする場合、これと一致するように Audio ダイアログボックスでサンプルレートを設定します。低サンプルレートを使用する場合、ソースのサンプルレート (通常 48KHz) と一致するように Source ダイアログボックスでサンプルレートを設定します。2 つのサンプルレート変換を行うことは避けてください。

「Encoder Quality」オプションは、コーデックについて品質と速度のトレードオフを設定します。16 ビットソースの場合、Better が適切なデフォルトです。Best は 16 ビットを越えるソースにのみ有効であり、Better より低速になります。

QDesign Music 2

QDesign Music 2 (QDM2) は、QuickTime 6 および MPEG-4 の前は主要なストリーミングオーディオコーデックでした。このコーデックは現在でも、QuickTime for Windows でデフォルトで使用できる最高のコーデックとなっています。QDM2 コーデックにはネイティブパケタイザがあるため、ストリーミングに適したオプションです。

QDM2 は、インストルメンタル音楽には非常に適していますが、高データレートであっても、スピーチのエンコード時には不快な圧縮ノイズが発生する可能性があります。QDM2 によって圧縮ノイズが発生した場合、ステレオからモノラルにチャンネル数を減らしてから、サンプルレートを下げてください。混合コンテンツを含む 44.1 のステレオを圧縮するには、Pro バージョンと 48Kbps を超えるデータレートが必要になります。QDM2 の無料バージョンは、最大のデータレートが 48Kbps に制限されています。

QDesign Music 2 Pro

QDM2 の Pro バージョンは、より高い圧縮効率と、はるかに高速なエンコーディング速度、48Kbps を超えるデータレートを提供します。QDM2 の基本バージョンでは、8 ~ 48Kbps がサポートされます。Pro バージョンでは、128Kbps まで圧縮できます。

このコーデックには、サンプルレートとチャンネルの推奨される組み合わせが含まれています。これらの推奨事項は、インストルメンタル音楽には適切ですが、スピーチを収録したコンテンツには過剰な設定です。このコーデックは、Quality モードか Speed モードで使用できます。Pro バージョンは基本バージョンよりもはるかに高速であるため、Quality モードを使用してください。

ほとんどのコンテンツで、Attack と Spread のデフォルト設定である 50/50 を使用できます。圧縮したコンテンツに多数の圧縮ノイズが含まれている場合、試行錯誤に基づいてこれらの設定を調整することで、より良い結果が得られることがあります。

Qualcomm PureVoice

Qualcomm PureVoice は、Code Division Multiple Access (CDMA) の携帯電話で使用されるコーデックであり、音質は携帯電話とよく似ています。このコーデックは非常に明瞭なスピーチを実現しますが、他の種類のコンテンツでは問題があります。PureVoice コーデックでは、モノラルオーディオしかサポートされません。

このコーデックは、特にリアルタイムストリーミングの場合に、8KHz で最適に機能します。このコーデックを 22KHz 以上で実行するのに十分なデータレートがある場合は、おそらく別のコーデックの方が全体として優れた鑑賞品質を提供できます。

PureVoice は、ネイティブパケタイザがあるため、優れたストリーミングコーデックです。このパケタイザは、8KHz で最適に機能するため、ストリーミングプロジェクトにはこのサンプルレートを使用してください。

Qualcomm PureVoice コーデックには 2 つのモードがあります。Full Rate モードはパルス符号変調 (PCM) 方式で 9:1 圧縮を提供するため、8KHz で約 15Kbps になります。品質は、接続状態の良い携帯電話と同等です。Half Rate モードは PCM 方式で 19:1 圧縮を提供するため、8KHz で約 7.5Kbps になります。スピーチは明瞭ですが、品質は Full Rate モードよりも大幅に低下します。この品質は、接続状態が悪い携帯電話と同等です。

「Optimize Compression For Streaming」オプションは、ストリーミング時に各オーディオブロックのブランクデータを送信しないようにヒントに指示します。このオプションを使用すると実際のデータレートを若干削減できますが、この削減は予測できません。このオプションは RTSP に対してのみ有効にしてください。

IMA

IMA は、Interactive Multimedia Association が開発した古い CD-ROM コーデックです。このコーデックは、None モードで 4:1 圧縮を提供します。品質は 44.1kHz で良好ですが、これより低いレベルでは低下します。IMA はデコーダが非常に高速であり、幅広い互換性がある (QuickTime 2.1 との下位互換性があり、AVI でも機能します) ため有効です。このコーデックは、再生パフォーマンスの方が圧縮効率よりも重要な CD-ROM やキオスクの用途に適したオプションです。

QuickTime は、Microsoft 社が開発した IMA Adaptive Differential Pulse Code Modulation (ADPCM) のバリエーションもサポートしています。このコーデックは実質的に同じコーデックですが、古い AVI プレーヤとの互換性がより高くなっています。

MP3

QuickTime には独自の MP3 エンコーダは搭載されていませんが、非常に高機能な MP3 デコーダが搭載されているため、様々なツールでエンコードされた MP3 ファイルを再生できます。エンコーディングツールによっては、オーディオトラックとして MP3 のエンコーディングをネイティブサポートするものもあります。

優れた MP3 エンコーダは、MPEG-4 以外のほかのどの QuickTime コーデックよりも混合コンテンツについて高いパフォーマンスを実現します。品質は、44.1 のステレオの場合は 128Kbps で非常に良好となり、160Kbps でほとんどのリスナーは圧縮されていることを感知できなくなります。

QuickTime 4.1 以降では VBR MP3 を問題なく処理できます。ただし、VBR MP3 は QuickTime 4.0 をクラッシュさせるため、QuickTime 4.0 との互換性が必要な場合は、CBR のみを使用する必要があります。

MP3 用の公式 RTSP モードが存在しますが、QuickTime はこのモードをサポートしていません。RTSP ストリーミングには、MP3 は使用しないでください。

None

None コーデックは非圧縮の PCM オーディオトラックです。このコーデックが配信コーデックに適することはまれですが、中間ファイルとしての用途や転送用としては最適のオプションです。None コーデックは 8 ビットと 16 ビットをサポートしていますが、16 ビットだけを使用してください。

24 ビット / 32 ビット 整数型

QuickTime は、24 ビットおよび 32 ビットの高いサンプル深度でも PCM をサポートしています。これらのコーデックは、こうした高いサンプルレートをサポートする 2 つのツール間での転送に使用されます。

浮動小数点型

浮動小数点型コーデックは、オーディオを一部のハイエンドオーディオツールや信号処理ツールに書き出すために使用されることがあります。通常のワークフローでは、このコーデックを使用することはありません。

MPEG-1 形式への書き出し

MPEG-1 は最初のオープンなマルチメディア標準でした。当初、コンピュータ用の MPEG-1 に関しては、ユーザが自分のコンピュータ用にアドオン MPEG-1 デコーダボードを購入して再生を行うという形態が想定されていました。しかし、この使い方が一般に広がることはなく、デコーディングをすべてソフトウェアで実行するために十分な速度をコンピュータが獲得するまで、MPEG-1 再生が広く普及することはありませんでした。

DVD の先駆けとなった VideoCD でも MPEG-1 を使用していました。VideoCD は、世界のほかの地域ではそれほど使用されませんでした。アジアでは広く普及しました。

MP3 は、MPEG-1 の関連技術でもあります。

MPEG-1 を使用したプログレッシブダウンロード

MPEG-1 の圧縮効率は、現在の最先端の Web コーデックの半分程度に過ぎませんが、この技術はプログレッシブダウンロードにおいて一定の役割を果たしています。この分野では、幅広い互換性の方が圧縮効率よりも重要だからです。

MPEG-1 のデータレートは他の形式よりも高くなります。良好なオプションは、336Kbps による 256 × 192 のビデオと 64Kbps、32kHz のモノラルオーディオとの組み合わせです。このデータレートは、優れた 2 パス VBR MPEG-1 エンコーダと併用した場合に、多くのコンテンツで適切な品質を実現できます。Web 用の MPEG-1 は常に正方形ピクセルになります。

MPEG-1 を使用したリアルタイムストリーミング

MPEG-1 は、リアルタイムストリーミングでも一定の役割を果たし、Apple 社による無料の Darwin Streaming Server や、Cisco 社の様々な製品でサポートされています。なお、MPEG-1 の圧縮効率は、他の最新のコーデックほど高くありません。

MPEG-1を使用したCD-ROM 配信

MPEG-1は、CD-ROM 作品でよく使用されます。この分野では、デコードパフォーマンスと互換性が圧縮効率よりも重要になります。QuickTimeをインストールできない、クロスプラットフォームのCD-ROMプロジェクトの場合は特にそうです。MPEG-1はPowerPoint用の唯一の優れたクロスプラットフォームファイル形式になっています。Windows上でQuickTimeファイルを使用したり、Mac OS上でWindows Mediaを使用することは容易でないためです。

CD-ROM用にエンコードするMPEG-1は、正方形ピクセルでエンコードする必要があります。これは、Microsoft DirectShowの古いバージョンが非正方形ピクセルを正しく処理できないためです。

MPEG-1を使用したVideoCD 配信

VideoCD形式は、主にアジアで普及していますが、ほとんどのパソコンでソフトウェアを追加せずに再生できます。また、DVDプレーヤでも再生できます。

なお、VideoCDは、DVDと同様、PALまたはNTSCの形式になります。コンピュータでは両方の形式を再生できますが、家庭用のDVDプレーヤやVCDプレーヤでは、独自形式のディスクしか再生できない場合があります。

VideoCDのパラメータセットは非常に厳格に制限されています。解像度は、NTSCで352×240、PALで352×288になります。また、データレートも固定されています。PALでは常に25fps、NTSCでは、ビデオ用の29.97fpsかフィルムソース用の23.976fpsのいずれかになります。このデータレートは、オーディオCDのデータレートであるCBRレートに厳密に固定されています。このため、CDに保持できるオーディオ1分あたり1分のVideoCDビデオを保持できます。

なお、CD-RでVideoCDを作成できますが、CD-RのVideoCDは、レーザーとの互換性がない一部の初期モデルのDVDプレーヤでは使用できません。工場生産のCD-ROMメディアは問題なく使用できます。

古いディスクや、安価なリアルタイムエンコーダやソフトウェアエンコーダでエンコードされたディスクなど、多くのVideoCDが使用されています。最新のプロフェッショナルグレードのエンコーダは非常に優れています。このため、解像度は低くても、さまざまなコンテンツで画質と音質が向上する可能性があります。

MPEG-1 ビデオ / オーディオコーデック

MPEG-1でサポートされるビデオコーデックは1つで、これもMPEG-1と呼ばれます。また、このコーデックには、Layer I、Layer IIおよびLayer IIIという3つのオーディオモードも含まれています。

MPEG-1 (ビデオ用)

MPEG-1は当時としては驚くべき圧縮効率を実現しましたが、最新のエンコーダでは、この形式の開発者が可能と考えたデータレートをはるかに下回るデータレートで良好に動作します。MPEG-1の最大の障害は後処理が欠けていることです。したがって、低データレートではソフトさがなくなり濃淡のむらが発生します。このため、良好な結果を得るために必要な最低ビットレートが相対的に高くなります。2パスVBRはMPEG-1に非常に適しており、ほとんどの実装で使用できます。

MPEG-1 (オーディオ用)

実際には、Layer IIオーディオモードしか使用されていません。Layer Iの圧縮効率は、より低いものでした。Layer IIIは、デコーダの一部としての標準にはなりません。ただし、Layer IIIはMP3という名前を使用し続けています。Layer IIの圧縮効率は、MP3より約50%低くなります。VideoCDは224Kbpsを使用しますが、一般に音質は非常に良好です。さらに低いKbps値でも、Webコンテンツとしては使用できます。

Layer IIはジョイントステレオをサポートしており、これにより、特に低データレートで圧縮効率が向上します。大多数のLayer IIエンコーダは、デフォルトではジョイントステレオが有効になっておらず、その場合はこの設定を手動で行う必要があります。Dolby Pro Logicサラウンドエンコーディングを維持する必要がない限り、ジョイントステレオを使用します。

DVD/MPEG-2 形式への書き出し

視聴者がビデオを見る時間で測定すると、MPEG-2は間違いなく世界的に主要なデジタルビデオ形式です。この形式は、DVDだけでなく、ほぼすべてのデジタル放送システム（デジタルケーブル、衛星および放送）で使用されています。

MPEG-2の大きな強みは、そのプログレッシブコンテンツとインターレースコンテンツの処理方法にあります。ストリーム全体をインターレースまたはプログレッシブにできます。インターレースストリームにはプログレッシブフレームを含めることができ、インターレースフレームにはプログレッシブブロックを含めることができます。なお、コンピュータで再生する場合、プレーヤによっては、インターレースMPEG-2ファイルをインターレース解除できる場合とできない場合があります。

DVD 形式の MPEG-2 サブセットは、実際はプログレッシブストリームをサポートしていません。NTSC DVD 上のテレシネフィルムは、通常の 3:2 プルダウンパターンを使用して格納されます。しかし、プログレッシブのフレームとインターレースのフレームをマークし、重複したフィールドを毎秒 12 個示す「フィールド繰り返し」タグを使用することによって、プログレッシブ DVD プレーヤを使用してリアルタイム逆テレシネを実行し、ビデオを 23.976fps に戻すことができます。プログレッシブ DVD プレーヤをプログレッシブディスプレイに接続すると、ビデオがネイティブフレームレートで再生できるため、画質と滑らかさが向上します。

優れた MPEG-2 エンコーダは自動的に 24P (または 23.976fps) のソースを使用し、そのソースから正しい「プログレッシブ」DVD を作成します。

MPEG-2 を使用した DVD 配信

DVD は、MPEG-2 ビデオコーデックを使用しますが、MPEG-2 ファイル形式は使用しません。代わりに VOB (Video Object) ファイルを使用します。このファイルは、他のオーディオコーデックを使用し、字幕など、他の種類の情報が含まれます。

多くの圧縮ツールでは VOB ファイルを直接書き出しできません。この制限を回避するには、(エレメンタリと呼ばれる) MPEG-2 ビデオファイルのストリーム (一般に M2V) と、スタンドアロンのオーディオファイル (WAV または AC3) を生成した後、制作ツールを使用して、これらを 1 つの VOB ファイルに結合することができます。

MPEG-2 を使用した CD-ROM/DVD-ROM 配信

MPEG-2 は実用上のディスクベース再生の標準にはなっていません。これは、デコーダがすべてのパソコンにバンドルされているわけではなく、相応のライセンス料を支払わなければデコーダを合法的に配布できないためです。

キオスクアプリケーションの場合、MPEG-2 は有効に機能します。MPEG-2 では、他の多くの形式に比べて 1 ピクセルあたりのプロセッサ要件が低いため、ディスク容量が問題にならない場合は、MPEG-2 を使用して高解像度、高品質のビデオを極めて良好に再生できます。

DVD の GOP の処理

DVD の場合、PAL およびテレシネフィルムコンテンツの標準的な GOP (Group Of Pictures) サイズは 12 です。NTSC では 15 を使用します。標準では、PAL の場合に最大 15 フレーム、NTSC の場合に最大 18 フレームが許可されています。

GOP は、オープンまたはクローズにすることができます。Open GOP では、GOP の最後の B フレームが次の I フレームを参照用に使用できます。次の I フレームを使用できることで、圧縮効率が若干向上しますが、各 GOP は完全な自己完結型ではないため、ランダムアクセスが少し複雑になります。

DVD 制作の場合、マルチアングルのタイトルを除いて Open GOP が適しています。マルチアングルのタイトルでは、Closed GOP が必要になり、アングル間でスムーズに切り替えられるようにするには、すべての I フレームをアングル間に整列させる必要があります。

MPEG-2 オーディオコーデック

デフォルトの MPEG-2 オーディオコーデックが存在しますが、ほとんどの実装では他のコーデックを使用します。

Layer II オーディオ

MPEG-1 の場合と同様に、スタンドアロンの MPEG-2 ファイルで使用される理論上の標準オーディオコーデックは Layer II です。MPEG-4 が使用する AAC オーディオコーデックも MPEG-2 コーデックとして開発されたものですが、MPEG-2 ファイルでは事実上使用されていません。コンピュータベースの再生に使用される MPG または MP2 ファイルでは、Layer II が唯一の使用されるオーディオコーデックです。

ただし、DVD は Layer II を使用しません。DVD は、公式の MPEG-2 仕様に含まれていない他のオーディオコーデックを使用します。

PCM

消費者向け DVD 制作パッケージは、非圧縮 PCM オーディオを使用します。品質は極めて良好ですが、データレートが高くなります。1 時間のホームビデオの場合、品質は問題になりませんが、Dolby Digital の圧縮効率を利用した方が、より高品質のビデオを 1 枚のディスクに追加できます。

PCM を高いサンプルレートおよびチャンネルに使用できますが、ほとんどの場合、PCM は 48KHz、16 ビットのステレオでしか使用されません。

Dolby Digital

Dolby Digital (AC-3 と呼ばれます) は、プロフェッショナル DVD オーディオの標準です。AC-3 エンコーダを搭載していることが、プロフェッショナル DVD 制作ツールの特徴です。

従来の標準的な AC-3 データレートは 384Kbps ですが、DVD 標準では 448Kbps が許可されています。DVD 標準は、データレートが最重要視されない限り、ほとんどのマルチチャンネルプロジェクトに最適です。ステレオエンコーディングでは、これよりも低いデータレートが使用できます。一般に、192Kbps で十分です。

DTS

ハイエンドタイトルによっては、Digital Theater System (DTS) オーディオ形式を含むものもあります。DTS では、はるかに高いデータレートが可能ですが (ただし、こうした高いレートの一因は圧縮効率が低いことです。つまり、DTS では音響心理学を利用しません)。一般に、DTS は A 級のハリウッド映画にしか使用されませんが、これにはいくつかの理由があります。まず、マスタリングツールが非常に高価であり、圧縮ツールに付属しないことがあります。第 2 に、DTS デコーディングが DVD 規格の一部として指定されていないため、DTS タイトルには Dolby Digital トラックも含まれています。DTS トラックと Dolby Digital トラックの両方を収録することで、使用可能なディスクスペースが減少します。また、オーディオ品質は、448Kbps AC-3 と比較して少ししか改善されません。

MPEG-4 形式への書き出し

MPEG-4 は、MPEG-1 および MPEG-2 の後継となる ISO 規格です。インターネットが一般に普及する前に開発され、オーディオやビデオの再生に厳密に焦点を当てていたこれらの形式とは異なり、MPEG-4 ははるかに広範で拡張可能な形式です。この形式の知名度が高まったのは最近のことです。

プロファイルとレベル「MPEG-4 互換性」という単純なものはありません。MPEG-4 標準の対象は極めて広範です。特定の製品や業界のサポートニーズに対応するため、MPEG-4 ではプロファイルとレベルという概念を使用します。プロファイルとは、使用できる機能 (B フレームなど) を定義するものです。レベルとは、これらの機能で使用できる最大パラメータ (最大フレームレート、最大フレームサイズ、最大データレートなど) を定義するものです。

ビデオ圧縮を処理するプロファイルで、最も多く見られるものと、少なくとも今後の実装について検討されているものを以下に示します。ここに記載されていないものは、リッチメディアなど、より高度なプロファイルで、圧縮よりも制作でより重要となるプロファイルです。

- **Short Header** このプロファイルは、ベースライン H.263 と同じもので、Simple Visual プロファイルのサブセットです。これまでのところ、Short Header モードは一部の携帯電話でしか必要とされていません。標準的な PDA でも、より強力なプロファイルを使用できます。それでも、Short Header は、こうしたデバイスに搭載されている小さな画面上に低いデータレートで適切な品質を提供できます。
- **Simple Visual** このプロファイルは、コンピュータ再生の現在のベースラインであり、QuickTime がサポートする最高のビデオプロファイルです。Simple Visual の圧縮効率は H.263 より高くなりますが、現代のコーデックに対する優位性はありません。
- **Advanced Simple** このプロファイルは、Simple Visual プロファイルの拡張バージョンで、(高レベルの) インターレースビデオ、B フレームなど、圧縮効率を向上できる要素のサポートを追加しています。QuickTime を除くほとんどのコンピュータプレーヤで、Advanced Simple プロファイルがサポートされています。
- **Simple Scalable** このプロファイルは、MPEG-4 のリアルタイムストリーミング用スケーラビリティの最も基本的な形態です。Simple Scalable には、エンハンスメントレイヤーが 1 つしかないため、可能なデータレートは 2 つになります。Simple Scalable は PacketVideo 社のツールスイートに採用されてきましたが、それ以上の実績はありません。
- **Fine Grain Scalability** このプロファイルは、MPEG-4 に強力なスケーラビリティを提供します。Fine Grain Scalability は、複数のエンハンスメントレイヤーと、特定のレイヤーの一部だけを使用する機能とを組み合わせ、使用可能なデータレートで非常に厳密に自己最適化を可能にする非常に柔軟なスケーリングを実現します。

残念ながら、このプロファイルを使用する製品はまだ存在しません。Fine Grain Scalability または同等の Master Boot Record (MBR) 技術が実装されるまで、MPEG-4 には、Windows Media や RealMedia の形式と比較してインターネット上でのリアルタイムストリーミングに関する優位性はないでしょう。

- **AVC (Advanced Video Coding)** このプロファイルは、MPEG-4 の世界で広く知られたプロファイルです。H.264、JVT、H.26L および MPEG-4 Part 10 とも呼ばれるこのプロファイルは、H.263 をベースとしない全く新しいコーデックです。AVC は、MPEG-2 のように最低 10 年は存続することを期待して開発されたため、パフォーマンスよりも圧縮効率を重視した設計になっています。

NAB (National Association of Broadcasters) 2003 では、開発者が最初に AVC エンコーダのデモを行いました。現在、AVC エンコーダは初期段階で、製品が登場したのは 2004 年です。

DIVX

最近、最も話題を呼んでいる AVI コーデックは DivX です。DivX は、.avi 拡張子を使用せず、代わりに .divx を使用します。DivX はもともと Microsoft MPEG-4 バージョン 3 コーデックの不正なハッキングに基づいていましたが、バージョン 4 以降は、標準準拠のビデオ用 MPEG-4 (V4 で Simple、V5 で Advanced Simple に) に基づいています。

オーディオについては、DivX は通常 MP3 形式を使用します。他の種類の .avi ファイルとの混同を避けるため、DivX のファイル形式自体はほとんどプレーン AVI であっても、拡張子に .divx を採用しています。

DivX 形式がプレーン AVI と見なされない理由は、多くの場合、DivX ファイルには VBR MP3 オーディオが含まれるためです。AVI は従来 VBR オーディオをサポートしてきませんでした。このサポートが原因で、一部のプレーヤ、特に QuickTime で再生の欠損が発生します。

- **Studio I** フレームのみの MPEG-2 や、モーション JPEG、多くのベンダーのプロプライエタリなキャプチャコーデックと同様に、このプロファイルはスタジオ用です。MPEG-4 ビデオ技術を使用して、Studio プロファイルは現在の形式よりも低いデータレートでマスタリング品質を実現します。このプロファイルには、10 ビット / チャンネルモード、4:2:2 および 4:4:4 など、ハイエンドビデオ制作に望まれる機能が含まれています。

Studio プロファイルの製品はまだ存在しません。また、近い将来に予想される製品もありません。このプロファイルが MPEG-4 の重要な用途になるかどうかは、現時点では不明です。

- **N-Bit** このプロファイルは、医療画像のアーカイブなど、ハイエンド用途に使用されます。MPEG-4 は、16 ビット / チャンネルなどのモードでのエンコードを可能にする N-Bit プロファイルを提供します。N-Bit プロファイルは独自の領域では有効ですが、圧縮ビデオ配信やビデオ制作でも使用されることはまずありません。
- **3GPP プロファイル** このプロファイルは Third Generation Partnership Project の略語で、携帯電話用に設計されています。3GPP には 2 種類あり、オーディオコーデックがそれぞれ異なります。ただし、ほとんどの通信事業者が通信事業者固有の種類の 3GPP を指定しています。これまでのところ、MPEG-4 を再生できるすべての携帯電話で機能する単一のファイルを作成することは不可能です。携帯電話がより強力になるにつれて、この制限は変化する可能性があります。
- **ISMA プロファイル** Web ベースのコンピュータ再生用 MPEG-4 に取り組む最も重要な団体が Internet Streaming Media Alliance (ISMA) です。ISMA は、2001 年後半に ISMA 1.0 の仕様を発表しました。この仕様は、2 つの基本プロファイル（紛らわしいことに、MPEG-4 とは異なり、MPEG-4 プロファイルとレベルの組み合わせになっています）を定義するものです。
 - ▶ **ISMA Profile 0** この基本プロファイルは、携帯電話など、非常に低電力のモバイル機器向けに設計されています。これまでのところ、こうした種類の機器では、3GPP か通信事業者固有の形式を使用しています。
 - ▶ **ISMA Profile 1** この基本プロファイルは、もともとはコンピュータ再生向けでしたが、最新の XScale ベースの PDA もこのプロファイルをサポートできるほど高速になっています。Profile 1 は、最も一般的にサポートが検討されるプロファイルです。残念ながら、多くのプレーヤーは仕様全体をサポートしているわけではありません。CELP サポートを使用できるプレーヤーは少数です。

MPEG-4 ライセンス問題

これまでのところ、MPEG-4 の普及を妨げる最大の障害はライセンス問題でした。ライセンス条件は禁止を目的としたものではありませんが、長期間にわたって決着が付かなかったため、あいまいさに関する懸念が発生しました。

ほとんどの形式で、エンコーダおよびデコーダを作成するために一定のライセンス料が必要で、MPEG-4 でも違いはありません。製品ベンダーはこうしたコストを支払っていますが、これらは、圧縮を実行するユーザや配信を行う立場の人の問題ではありません。しかし、MPEG-4 では、MPEG-4 形式でコンテンツを配信するための料金が含まれています。ディスクベースのコンテンツの場合、プレス業者が、DVD を作成するための MPEG-4 料金を長年回収してきました。インターネットでコンテンツを配信するための料金は、MPEG-4 にとって新しいものです。

こうした料金は想像するほど威圧的なものではありません。ほとんどのディストリビュータは一切支払をする必要はありません。まず、料金は営利目的のコンテンツにのみ適用され、マーケティング、広告および非営利コンテンツには適用されません。営利目的のサイト（主に購読コンテンツや広告費によるコンテンツを提供するサイト）の場合、1 法人あたり最初の 50,000 人までは無料です。50,000 を超えると、ディストリビュータは、年間 1 加入者あたり 0.25 ドルを支払うか、コンテンツのストリーミング 1 時間あたり 0.02 ドルを支払うかを選択できます。こうした数字はストリーミングプロジェクトの予算のごく一部にすぎません。それでも、使用状況の追跡という単純な問題があるため、場合によっては採用が制限されています。

MPEG-4 オーディオコーデック

MPEG-4 のオーディオプロファイルは、可能なコーデックセットを組み合わせたものです。ほとんどの実装では、プロファイル全体ではなく特定のコーデックをサポートしているため、一般的には、プロファイルではなくオーディオコーデックについて説明した方が有益です。

AAC-LC

これまでのところ、ほとんどの MPEG-4 実装で仕様されているオーディオコーデックは、Advanced Audio Coding-Low Complexity (AAC-LC) です。「Low Complexity」の部分は非常に高品質のコーデックです。他の種類の AAC も存在しますが、リスニングテストで AAC-LC を大きく上回る圧縮効率を示すものは 1 つだけです。

ISMA 標準に AAC-LC が含まれており、すべてのコンピュータベースの MPEG-4 プレーヤでも AAC-LC をサポートしています。

HE-AAC

AACの新しいバージョンは、HE-AAC (HEはHigh Efficiencyの略) です。HE-AACは、Spectral Band Replicationという技術を使用するもので、本質的に非常に効率的なエンハンスメントレイヤーです。

現在、HE-AACはかなりの差をつけて最適なオーディオコーデックとなっています。リスニングテストによって、48KbpsでCDオーディオとほぼ同じ品質を持つこと、32KbpsではFMラジオと同等のエンターテインメント品質を持つことが示されました。HE-AACはまだ主流にはなっていませんが、業界の専門家は広範に採用されることを予想しています。

CELP

CELPは低ビットレートのスピーチコーデックです。ライセンス問題が原因で、CELPはクリティカルマスに達するだけのプレーヤに展開されていないため、普及していません。他の類似のコーデックと同様、CELPは低データレートで明瞭なスピーチを実現します。MPEG-4には、他にもスピーチコーデックがいくつか存在しますが、CELPが最もよく使用されるコーデックです。AAC-LCで低ビットレートのスピーチが良好に処理されるため、MPEG-4ではスピーチコーデックがそれほど必要とされていません。

AMR

MPEG-4でそれ以外に重要な唯一のコーデックがAdaptive MultiRate (AMR) です。CELP同様、AMRは低ビットレートのスピーチコーデックです。CELPとは異なり、AMRはリアルタイムスケラビリティをサポートしています。AMRは、主に携帯電話向け3GPPプロファイルで使用されます。

AVI形式への書き出し

Microsoft Audio Video Interleave形式は、ごく初期のCD-ROM形式でしたが、今でも様々なアプリケーションで広く使用されています。

AVIを使用したWeb配信

AVIは、リアルタイムストリーミングをサポートしていないため、Web形式には適していません。また、すべてのコンピュータでデフォルトで使用できるコーデックのセットは、比較的限られています。Web上のAVIファイルは、ほとんど常に、完全なファイルとして配信されます。

AVIを使用したディスクベース配信

AVIは今なおディスクベース再生に広く使用されています。AVIは、Adobe® Directorなどのアプリケーションでよくサポートされており、コーデックをディスクとともに配布できるため、インストール環境についての問題ははるかに少なくなります。

AVIビデオコーデック

AVIでは、数多くのコーデックが使用できます。その大多数では、手動のインストールが必要になります。最も一般的で便利なコーデックのいくつかを以下に示します。

Cinepak

古いAVIコーデックはCinepakでした。QuickTimeの節で述べたように、Cinepakで提供される品質は限られたものですが、再生の互換性が高く、デコーダーパフォーマンスが非常に高くなります。Cinepakは、すべてのAVIプレーヤで使用可能と想定できる唯一のビデオコーデックです。

Indeo

IndeoコーデックはIntel社によって開発されましたが、現在はLigos社の所有になっています。Indeo 3、4、5は、CD-ROM全盛の時期には非常に広く使用されていました。しかし、何年にもわたって大きな更新が行われていません。従来は優れたデフォルトのコーデックでしたが、Indeo再生はWindows XPにプレインストールされておらず、Mac OS Xでも使用できません。特に、4×4ブロックごとにクロマ値サンプルが1つしか存在しないYUV-9色空間を使用していたため、Indeo 3、4、5の品質は最新のコーデックよりも低いものでした。

VP3

QuickTimeの節で述べたように、VP3は、もともとOn2社が開発したオープンソースのコーデックです。さまざまなプラットフォームでVP3の無料エンコーダとデコーダを使用できます。ただし、WindowsやMac OSにプレインストールはされていません (VP3ファイルのAVI再生にはQuickTimeを使用します。QuickTimeの自動コンポーネントダウンロードプログラムにはVP3が含まれています)。

VP3コーデックの現在のバージョンには、QuickTimeバージョンでは使用できない機能がいくつか含まれています。Lowest Allowed Qualityコントロールでは、品質について最低しきい値を設定します。データレートを維持するには、最小値63を使用する必要があります。Sharpnessコントロールでは、高精細でも斑点が含まれる可能性のあるフレームか、ソフトになるが斑点は含まれないフレームか、その中間のフレームかの目標を設定します。最適な選択は、ソースとその方法に固有なものとなります。

VP5、VP6

これらの新しいコーデックも On2 社によるもので、VP3 に基づいていますが、特に VP6 で圧縮効率が大幅に向上しています。VP6 は、個人用途であれば www.on2.com で無償でダウンロードして入手できます。

DivX

DivX コーデックは、当初問題がありましたが、AVI インフラストラクチャの重要な部分への進化しています。現在の DivX 5 は、DivX 5 ビデオコーデックの優れた実装であり、Simple プロファイルまたは Advanced Simple プロファイルをサポートしています。

DivX コーデックも、コーデック自体に多数の前処理機能が含まれているため、これらの機能がないツールで書き出しを行う場合に便利です。

AVI コーデックの DivX のユニークな機能の 1 つは、複数のパスで圧縮を連続的に改善できる N パスエンコーディングをサポートしていることです。DivX は、ラウンドアバウト方式で N パスエンコーディングを実行するため、最初のパスで複雑なデータを格納したダミー AVI ファイルを作成し、次にそのダミーファイルを 2 番目のパスに使用します。このプロセスを自動化するツールもあれば、手動で実行する必要があるツールもあります。4 つ以上のパスで圧縮が改善され続けることはめったにありません。2 パスと 3 パスとの違いもごくわずかなものです。

DivX には特殊な心理視覚モードがあります。このモードでは、人間の視覚システムにとってより重要な画像の部分が内部モデルによって選択されます。次に、コーデックがそうした部分にビットの分配を試みます。Light 設定を使用すると非常に安全です。しかし、Extreme 設定を使用すると、奇妙な結果が生じる可能性があります。それでも、非常に低いビットレートでは、このモードを使用することで、一部のコンテンツを大幅に改善できます。

DivX を使用すると、新しいナチュラルキーフレームをトリガするためのしきい値を手動で設定できます。デフォルトの 50 は、ほとんどのコンテンツに対して最適な出発点です。また、DivX では、品質とエンコード速度のトレードオフに 5 つのレベルが許可されています。時間が許すのであれば、常に最も遅いモードを使用してください。

AVI オーディオコーデック

AVI のオーディオコーデックは、ビデオコーデックほど注目を集めていませんが、オーディオコーデックも同様に重要であり、様々なコーデックが存在します。最も重要なものを以下にいくつか示します。

None

最も基本的なオーディオコーデックは None で、QuickTime None コーデックのような非圧縮モードです。None は配信用ではなく、主に中間ファイルに使用されます。

IMA ADPCM

この IMA コーデックは、QuickTime の IMA コーデックと同じもので、適切な品質で None コーデックに対する 4:1 圧縮を実現します。IMA ADPCM は、AVI 用の最も汎用的なオーディオコーデックであり、デコーダは非常に軽量です。

MS ADPCM

MS ADPCM は、IMA をわずかに変更したものです。このコーデックには、IMA に対する利点はなく、互換性も若干低くなるため、制作にこのコーデックを使用する必要はありません。

MP3

MP3 コーデックは、AVI で広く使用されています。しかし、VBR MP3 オーディオは CBR MP3 ほど広くはサポートされていません。Windows コンピュータには、最大で 56Kbps の非常に制限が大きい MP3 エンコーダが付属します。しかし、数多くの優れたサードパーティ製エンコーダが使用可能です。

詳細情報

アドビ製品の包括的な内容については、アドビシステムズ株式会社の Web サイト (www.adobe.co.jp) を参照してください。



アドビシステムズ株式会社
〒141-0032 東京都品川区大崎 1-11-2 ゲートシティ大崎 イーストタワー
www.adobe.co.jp

Adobe Systems Incorporated
345 Park Avenue, San Jose, CA 95110-2704 USA
www.adobe.com

Adobe、Adobe ロゴ、Acrobat、Adobe Audition、Adobe Premiere、After Effects、Illustrator および Photoshop は、Adobe Systems Incorporated (アドビシステムズ社) の米国ならびに他の国における商標または登録商標です。Mac および Macintosh は、米国およびその他の国々で登録された Apple Computer, Inc. の商標です。Intel は、米国 Intel Corporation の米国ならびに他の国における商標または登録商標です。Microsoft および Windows は、米国 Microsoft Corporation の米国ならびに他の国における商標または登録商標です。そのほかすべての商標は、それぞれの権利帰属者の所有物です。

© 2006 Adobe Systems Incorporated. All rights reserved.