

映像音声符号化技術と国際標準化

浅井 光太郎 (三菱電機株式会社) 守谷 健弘 (日本電信電話株式会社)

概要 映像音声符号化技術は家庭用 AV 機器の市場を背景にした国際標準化活動に牽引され、急速に発達した。特に ISO/IEC JTC 1/SC 29/WG 11 (MPEG) と ITU-T/SG16 によって策定された標準方式はデジタル放送や DVD などのパッケージ、携帯音楽プレーヤなどの様々な製品やサービスに適用され、世界的に普及している。過去およそ 20 年余りに渡り、映像音声符号化技術は圧縮性能や幅広いビットレートへの対応、高品質化、誤り耐性の強化、スケールビリティやマルチチャンネル化など、性能と機能を向上させてきた。本稿では、これらの進歩を実現した技術の概要について説明する。標準は最小限の互換性を保証し、標準の適用には競争的要素を残している。標準の策定は技術開発の終止符ではなく、現在も新たな用途や市場に向けた標準化活動が技術開発を促進している。

1. はじめに

今日、映像や音声デジタルメディア情報として流通している技術的基盤の一つは映像音声符号化技術であり、同技術がもたらす最大の利点は圧縮性能である。メディア情報の圧縮により、長時間記録や帯域制限下の通信・放送が可能になる。互換性や相互運用性への要求から、同技術は 1980 年代以降、国際標準化活動と連動して発達してきた。同技術の標準化を担ってきた代表的な国際標準化組織として ISO (International Organization for Standardization, 国際標準化機構) と IEC (International Electrotechnical Commission, 国際電気標準会議) の傘下にある JTC 1 (Joint Technical Committee, 合同技術委員会) SC 29 (Sub Committee, 分科委員会) WG 11 (Working Group) 通称 MPEG (Moving Picture Experts Group) と、ITU-T (International Telecommunication Union Telecommunication Standardization Sector, 国際電気通信連合電気通信標準化部門) SG16 (Study Group, 研究委員会) 傘下のグループが挙げられる。特に MPEG (ISO/IEC JTC 1/SC 29/WG 11) は 1988 年の発足以来、映像や音声、音響などメディア情報の符号化、多重化・伝送・蓄積の各種フォーマットなど広範な分野において、世界各国の参加による技術提案と改善を通じ、高度な国際標準を開発してきた。その結果、符号化技術は急速に進展し、デジタルメディアを活用する機器やソフトウェア、コンテンツが今や年間 40 兆円もの巨大な市場を形成している。次章より、MPEG のビデオとオーディオのそれぞれについて、MPEG 標準がどのような特質を持ち、どのような用途に活用されているか、標準方式の基本となる技術内容とそれらの改善技術について概説する。続けて、標準化活動による技術の発展や標準方式を活用するための課題について述べる。

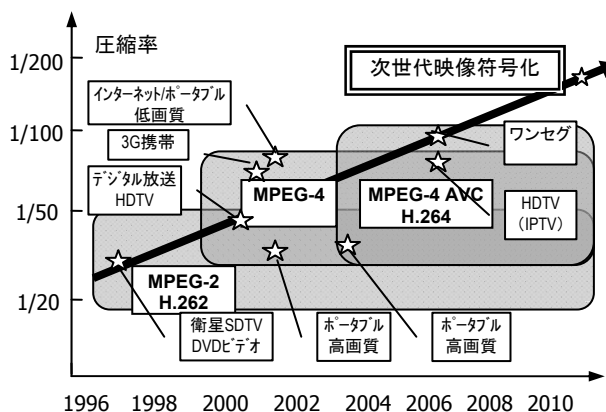


図 1 MPEG 映像符号化標準の進展

2. 映像符号化

2.1 MPEG 映像符号化標準の概要

図 1 は MPEG 映像符号化標準を圧縮率やアプリケーションとともに時系列で示している。MPEG が開発した映像符号化標準は CD-ROM を媒体に想定して開発した MPEG-1 に始まり、MPEG-2、MPEG-4 と進化した、符号化性能を向上させている。また、図 2 は MPEG 映像符号化標準と ITU-T 勧告との関係を示している。MPEG-2 および MPEG-4 AVC (Advanced Video Coding) は ITU-T/SG16 との協力関係の産物である。H.26X は ITU-T 勧告番号のシリーズである。1994 年に標準化された MPEG-2 | H.262 は SDTV、HDTV とともに現行のテレビジョン信号の性質を利用した符号化方式を提供し、世界各国のデジタル放送や DVD ビデオに採用されている。日本でも、2000 年の BS 放送、2003 年の地デジで HDTV 放送に適用され、薄型テレビや DVD レコーダの中心的技術となった。1999 年に標準化された MPEG-4 は低レートかつ低品質の回線に対するロバスト性を重視して策定され、第 3 世代携帯

電話の動画機能やインターネットストリーミング、ポータブル機器などに採用されている。2003年に標準化された MPEG-4 AVC | H.264 はきめ細かな改善要素を盛り込んで圧縮率を大きく向上し、ワンセグ放送やブルーレイ、

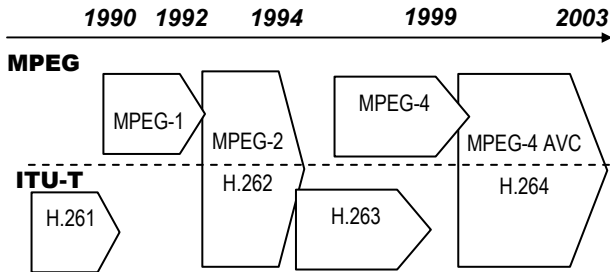


図2 映像符号化に関する MPEG 標準と ITU-T 勧告

IPTV、ポータブル機器に適用され、現在、市場に浸透しつつある。

2.2 映像符号化技術の概要

2.2.1 映像符号化の基本ブロック図

図3は MPEG 映像符号化の基本ブロック図である。MPEG-1 から MPEG-4 AVC の最新方式まで、MPEG 映像符号化標準は共通の基本構造を踏襲しており、大きな技術要素として、動き補償予測、変換と量子化、エントロピー符号化とに分けられる。動き補償予測とは、画素のブロック（例えば 16×16 ）を単位として、既に符号化済みの画像上で、符号化対象のブロックに最も良い予測を与える変位を決定し、予測誤差を符号化対象とすることで符号化効率を向上させる技術である。変換と量子化は、画素ブロック（例えば 8×8 ）単位で DCT (Discrete Cosine Transform) を行い、信号電力が局在化された変換係数を量子化することで圧縮効率を得る技術

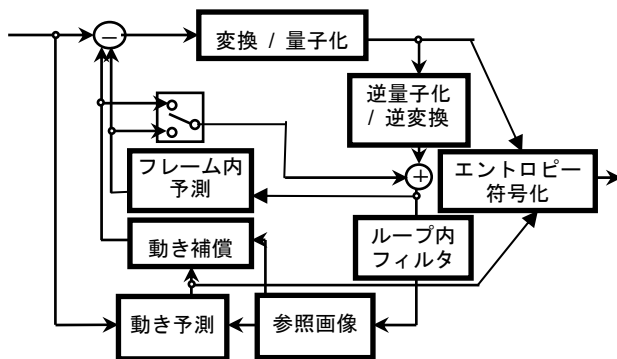


図3 映像符号化器の基本ブロック図

である。エントロピー符号化は動き補償予測の変位情報や変換係数の量子化情報から統計的冗長度を除く可逆な符号化技術である。この基本構造は MPEG 以前に標準化された ITU-T (当時は CCITT) の H.261 が原型である。

2.2.2 MPEG-2 | H.262

先行する H.261 に対して MPEG-1 が追加した主要要素は、映像フレームの並べ替えを行い、本来の時間軸では過去と未来に相当する両方向からの予測を導入したこと、動き補償予測の変位に整数画素単位でなく、 $1/2$ 精度を導入したことである。両方向予測は MPEG-1 が CD-ROM などパッケージを対象としたため、ランダムアクセスや特殊再生を目的として導入されたが、同時に符号化効率を向上させる効果をもたらした。MPEG-2 ではさらに現行テレビジョン信号のインターレース構造に対応し、映像フレームを構成する奇数走査線および偶数走査線の二つのフィールドを個別あるいは組み合わせて動き補償予測や変換を行う適応手法を導入し、それぞれ予測と変換の効率を向上させた。

2.2.3 MPEG-4

MPEG-4 では第3世代携帯電話を想定し、低速低品質の回線でもロバストな符号化、低解像度の映像に対する最適性を重視した検討が行われた。その結果、動き補償

予測の変位情報は $1/4$ 画素精度、予測単位となるブロックサイズに従来の 16×16 から 8×8 も使用可能とし、ブロックの一部が画面をはみ出す画面外参照を許容するなど、動き補償予測の強化が行われた。変換と量子化では、変換係数の画面内予測や、

ゼロでない変換係数の符号化イベントを低レート用に変更するなどの追加が行われた。また、エントロピー符号化では、回線のビットエラーによる同期はずれに耐性を持つ再同期マーカ、双方向からの一意復号性を保証する符号語の活用なども追加された。もう一つ、大きな要素として、MPEG-4 では映像をオブジェクトに分解し、たとえば人物と背景を別のオブジェクトとして、形状とテクスチャに分けて符号化を行う方式も定義されている。ただし、不定形オブジェクトの自動分割には広範なアプリケーションに適した実用的手法が確立されていないため、MPEG-4 は矩形オブジェクト（しばしば画面全体）を対象とした符号化として用いられることがもっぱらである。

**映像符号化標準
方式の基本は動き
補償予測と DCT**

2.2.4 MPEG-4 AVC | H.264 とその拡張

MPEG-4 AVC | H.264 は ITU-T で検討されていた H.26L (L は Long term) を原型に, MPEG と ITU-T の VCEG (Video Coding Experts Group) とが Joint して開発した標準である。動き補償予測の変位情報は 1/4 画素精度であるが, 端数位置の画素計算にはより高次の演算を行うようにした。予測単位となるブロックのサイズと形状は 16×16 , 8×8 , 16×8 など 7 つのパターンを使えるようにした。また, 参照画像の数は従来方式が両方向予測のために 2 面までであったが, さらに多く (例えば 5 面) の参照画像を使えるようにした。これらが予測効率の向上につながっている。変換は従来の 8×8 DCT から 4×4 の整数変換に変更 (後に 8×8 の変換モードを追加), 変換前の画素に画面内で多方向からの空間的予測を行う手法を追加した。エントロピー符号化では, 符号化履歴をコンテキストとして符号の適応化を行う CAVLC (Context Adaptive Variable Length Coding), 算術符号化 CABAC (Context Adaptive Binary Arithmetic Coding) が導入され, さらに圧縮効率を向上させた。また, 符号化による歪みを低減するためのデブロッキングフィルタを追加して, 歪みが符号化ループの中で伝播しないようにした。

MPEG-4 AVC は 2003 年に標準化された後, 2005 年には HDTV の符号化性能改善に顕著な効果を持つ 8×8 変換の追加などを含む High Profile を策定し, これがブルーレイなどの HDTV 符号化に用いられている。さらに 2007 年には輝度成分と色差成分の画素密度を同等にする 4:4:4 サンプリングを含む Professional Profiles を定義するなど, 高品質対応の拡張を行ってきた。機能面の拡張では, 時間解像度, 空間解像度, 品質などを階層的に扱うスケーラビリティを実現しつつ, 階層分割による損失を 10%程度に抑えた SVC (Scalable Video Coding), 複数のカメラで捕えた映像を時間方向の予測に加え, 視点間の予測を用いて 20%程度の符号量削減効果を得る MVC (Multi-view Video Coding) がある。

上記のように, MPEG-4 AVC では動き補償予測+変換と量子化+エントロピー符号化という基本構造を踏襲しつつ, それぞれの要素の効率を向上させるための適応化や高度な処理を追加して積分的な改善効果を実現している。これらはいずれも演算負荷の高い要素であり, すべての要素をフルに活用して符号化を行うには, MPEG-4 に比べて 20 倍以上の演算量と 4 倍のメモリーアクセスが

必要とも言われ, 演算量の削減が課題である。もう一つの課題として, 同じ基本構造で改善を重ねているため, 全く異なるアプローチの対抗案が出にくいということがある。異なるアーキテクチャを持つ符号化器は MPEG-4 AVC と異なる最適化要素を必要とするため, 標準化の場で同列に比較すると, 言わば「熟成度が足りない」方式はどうしても符号化性能で見劣りすることになる。このため, MPEG ではワークショップやアドホックグループの形で, 異なるアプローチを検討する機会を設定している。

2.3 MPEG 映像符号化標準化の現状

MPEG は現在, HVC (High-performance Video Coding) の検討を行っており, 2010 年 1 月の会合では, ITU-T/VCEG と連携して次世代の映像符号化標準を策定することが合意された。HVC では HDTV を越える超高精細映像, たとえば $4K \times 2K$, $8K \times 4K$ を符号化対象に含めるため, 従来にも増して圧縮率を高めることへの要求がある。MPEG や VCEG の会合でこれまで報告されてい

る手法のほとんどは MPEG-4 AVC | H.264 の構造を用い, さらに改善しようとするアプローチをとっている。例えば動き補償予測では, 適応的な補間フィルタを用いて予測画素を生成する手法, 1/8 画素精度の変位で動きを表す手法, 動き補償を行うブロックのサイズと形状を多様化する

**高品質・高機能化を
追求して新たな
標準化課題に挑む**

手法, 変換ではブロックサイズの拡大や適応化, 画面内の空間的な予測方向と連動した変換基底の適応化, Wiener フィルタによる局所復号画像の品質改善などが報告されている。一方, 超高精細映像とは別方向となるモバイル向けアプリケーションのための要求条件として, 処理負荷の軽減, 省電力化がある。標準を実装する者にとって, 符号化性能と演算負荷のトレードオフを調整できる標準であることが, 新しい要求条件になっている。

3. 音声・音響符号化

3.1 MPEG オーディオの概要

MPEG オーディオの標準は図 4 に示すようにデジタル放送や CD-ROM 等への蓄積媒体の符号化として広く使われている。MPEG-1/2 レイヤー I は DCC (Digital Compact Cassette) に使われ, レイヤー II はビデオ CD, DVD, 欧州のデジタル放送に広く使われている。レイヤー III はインターネットや携帯プレーヤで MP3 として幅

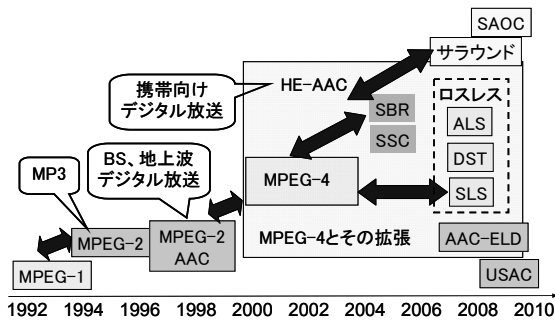


図4 MPEG オーディオの進展

広く使われて 1997 年に新たに制定された MPEG-2 の AAC (Advanced Audio Coding) は日本のデジタル放送や世界中の携帯プレーヤで広く使われている。これらの符号化技術によって元の音声や音響信号の品質をほとんど損なわないままで、その伝送(記録)情報量を 10 分の 1 程度まで圧縮できている。このような符号化技術がなければ放送や音楽の鑑賞形態はまったく成立しなかったと思われる。

AAC に引き続き, MPEG-4 およびその拡張規格も制定され, 今後の普及が期待されている。特に SBR (Spectral Band Replication)によるスペクトル帯域拡張と SSC (Sinusoidal Coding)のパラメトリックステレオ符号化を低レートの AAC を組み合わせた HE-AAC (High efficiency AAC) はワンセグ放送で採用されている。このほか ALS (Audio Lossless coding), SLS (Scalable Lossless coding), DST (Direct Stream Transfer)などの高品質向けの歪のないロスレス符号化 やマルチチャネルの効率よい MPEG サラウンド符号化, 遅延の小さい AAC-ELD (Enhanced Low

MP3 から AAC へ
さらに多様化へ

Delay)が制定されている。さらに音声と音楽を両立できる低レート符号化 USAC (Unified Speech and Audio) の制定作業が続けられている。

3.2 オーディオ技術の概要

3.2.1 MP3

MPEG-1, -2 の規格はレイヤー(階層)という名称で区別される 3 種のアルゴリズムが存在するが, いずれも周波数領域の係数を適応的に量子化する際に聴覚心理モデルに基づく雑音の制御で聴感歪を軽減している。言い換えれば, 信号の周波数領域での偏りがあること (時間領域ではサンプル間の相関があること) と人間の聴覚特性を利用して圧縮符号化が実現されている。レイヤーIII もレイヤーI と II との共通化で 32 サブバンド分割を行い, さらに入力窓長が 36 点の時間領域での重なりをもつ MDCT(modified Discrete Cosine Transform)で周波数領域 (周波数領域の独立な係数は 18 点) に変換する。また過渡的な信号の場合には変換長が 1/3 の 6 点の変換を 3 回

行う。結果的に図 5 のような AAC とも共通な変換符号化となる。MDCT 係数は高域での値が小さいので, 高域から 0 が連続する領域でランレングス符号化, 絶対値が 1 を越えない領域では 2 次元, 1 を越える領域 4 次元の可変長符号化を行う。このような可変長符号化による情報量変動を吸収するため, 複数フレーム間で量子化ビットをやりくりするビット貯蔵(bit reservoir)手法も使う。

3.2.2 AAC

MPEG-2 は, 後方互換性を保ったまま, マルチチャネル対応と, 低サンプリング周波数対応の拡張が定義された。後方互換のための構成と最高ビットレートの制約のために十分な品質が得られないという問題が生じた。特にマルチチャネルの高品質化の要請により AAC の策定が開始され, 1997 年に制定された。基本構成は MDCT 係数を不均一な帯域に分割して適応的な正規化と量子化を行い, 可変長符号化するものである。フレーム長の適応化, フレームをまたがるビットの貯蔵, ノイズシェイピングを初めとする数多くの要素技術が統合されている。聴感歪を最小とするようエンコーダを制御する機能が柔軟に盛り込まれているため, MP3 より品質が高く, 128 kbit/s で, ほぼ聴覚的な劣化のない CD 帯域のステレオ信号の符号化が可能になった。

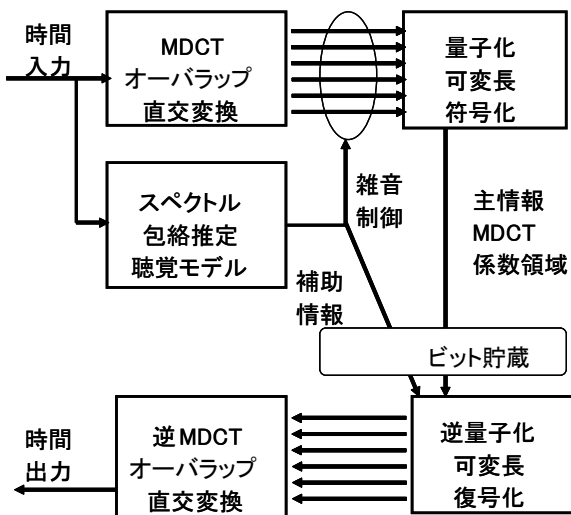


図5 MP3 と AAC

3.2.3 MPEG-4 と拡張

MPEG-2 に続き、2000 年に MPEG-4 オーディオの規格が完成した。幅広いビットレート、帯域、そのスケールビリティ、符号誤り耐性などに柔軟に対応し、合成音声、音楽合成（構造化音響符号化）も含まれている。ストリーミング、双方向通信、無線など広範囲の用途を想定しているが、2010 年時点では市場への展開は限定的で、詳細は割愛する。

2002 年以降の MPEG-4 の新たな拡張として、下記の規格が重要である。

- SBR (Spectral Band Replication) 広帯域の入力信号の低域部分のみを AAC で符号化し、高域部分はサブバンド領域で表現したスペクトル包絡情報だけから擬似波形を再構成する技術である。結果的に AAC より 30-50% 少ない情報量で広帯域信号の符号化ができ、HE-AAC として使われている。低域は従来の復号器と互換性があり、新たな拡張復号器では補助情報から高域の再生が可能である。また低演算のローパワーモードもあり、低ビットの音楽配信用途として有効である。
- SSC (Sinusoidal Coding) 正弦波と雑音で楽音を表現する。低ビットで、ピッチや再生スピードの変更が容易である。この中の一つの機能である PS (parametric stereo) はチャンネル間の相関パラメータを追加するだけで効率よくステレオ信号が再構成でき、HE-AAC にも使われている。
- MPEG surround: 多チャンネル信号をステレオと同等の低ビットで符号化する技術で、例えばサラウンドのスピーカ配置用に作られた 5 チャンネル音源をステレオに混合して符号化（例えば AAC, PCM も可能）する。同時に 5 チャンネル信号のもつ空間情報（レベル差、位相差、相関）のステレオ信号に対する対応をサブバンド領域で抽出して低レートの補助情報とする。従来の復号器でもステレオ信号の再生が可能であり、拡張復号器では補助情報とステレオ信号から 5 チャンネル信号を再構成できる。さらに音源情報も加えて機能拡張した SAOC (Spatial Audio Object Coding) の標準化も進められている。
- ALS (Audio Lossless coding) 時間領域の予測と可変長符号に基づくロスレス符号化である。浮動小数点も

含む幅広い入力信号に対応し、用途に応じて演算量や圧縮率を柔軟に選択できる。このためオーディオ信号の蓄積、編集、伝送、長期間のアーカイブへの適用が想定されている。また音楽以外の生体や地震などの時系列信号にも応用可能である。

- SLS (Scalable Lossless coding) 周波数領域ロスレス符号化で、AAC などの周波数領域の符号化の誤差信号を階層的に符号化し最高レートで無歪となるスケラブル符号化である。全体のビット列を使って完全な波形を保存し、AAC の部分だけを取り出すことで、携帯プレーヤなどの保存容量に応じた品質の音を利用できる。
- DST (Direct Stream Transfer) SACD (Super Audio CD) の 1 ビットオーバーサンプル量子化 (DSD: direct stream digital) を対象としたロスレス符号化である。+1 または -1 の値を線形予測し、予測誤差系列を可変長符号化する。これまで SACD で使われている圧縮と互換性を持ちつつ機能拡張され、編集や蓄積など汎用的に利用可能である。

4. 標準化に関する競争と課題

4.1 標準策定後の競争

4.1.1 標準方式の実装に関する競争

標準は通常、実装方法を規定しないため、実装の規模、コスト、消費電力は目標とする動作速度や符号化性能を考慮した上で、設計者の思想に委ねられる。このため、実装方法や採用するデバイスの選択や開発は小さく安く、

など製品の競争力に直結する。特にエンコーダは実装に依存する。MPEG 標準ではビットストリームのシンタックス（構文規則）とデコーダ処理を規定し、エンコーダの処理を規定しない。一方、エンコーダ側の負荷

は特に映像符号化の場合、デコーダ側の負荷の何倍も重い。動き補償予測のために最適な画素変位を探索する処理、最適な符号化条件を決定する処理はエンコーダ側のみ必要な動作であり、規定はなく自由度が大きい。最小の実装規模と要求条件をみたす符号化性能を可能にするには、符号化方式を熟知することが必要である。

4.1.2 標準方式による符号化性能の最適化競争

メディア情報の圧縮符号化において、再生品質はエン

標準の成立は競争の終結でも
技術開発の終止符でもなく
次の技術開発を促進してきた

コーデ側の処理でほとんど決定されるため、符号化性能には競争的要素が多分にある。4.1.1のようにエンコーダの処理には任意性があるため、標準を採用というだけで性能が保証されない一方、同じ標準を用いた装置の中で他より優れた性能を発揮することができる。エンコーダの性能を左右する大きな要素は符号化制御である。符号化制御とは、標準方式が定義する多くの選択可能な適応アルゴリズムやパラメータをいかに決定し、符号量の多寡をいかに配分するかを制御することである。

映像符号化の場合、制御の基本は圧縮率と SNR (Signal to Noise Ratio) のトレードオフを最適化することである。演算負荷や演算時間を考慮しない計算機シミュレーションでは、適切な条件設定と簡略化条件の下に最適なアルゴリズム選択やパラメータ決定を求めることが可能である。従来はシミュレーションに留まる最適化手法であったが、デバイスやプロセッサの高速化により、擬似的に同様の最適化手法が用いられるようになってきている。SNR の高さは当該符号化方式の入力画像に対するロバスト性を示す一方、SNR と主観的な画像品質の高さは必ずしも一致しないことが知られている。SNR の最適化制御を基本にし、視覚特性を加味した符号化制御が、映像符号化標準を最大限に活用するノウハウとなる。

オーディオ符号化でも標準仕様が決まったあとでもその再生品質は圧縮（エンコーダ）側の処理に大きく依存する。また品質評価においては、SNR は頼りにならず、評価はほとんど主観品質評価が必須である。映像の場合と同様に、聴覚特性を駆使した符号化制御が、オーディオ符号化標準を最大限に活用するノウハウとなる。

これらの競争が標準方式成立後も標準および標準を採用した製品を進化させ、市場拡大と標準の普及を後押ししてきた。今後とも同じ傾向が続くであろう。

4.2 標準活用の課題

4.2.1 オーディオ標準化の普及

オーディオの符号化規格は多岐にわたっており、標準規格であっても必ずしも世界には普及していない。MP3 は制定当初は機器や放送に使われる見通しが明らかであったレイヤーI やレイヤーII に比較して処理が複雑で、参照ソフトも不備が多く、実用的なアプリケーションに採用される見通しはなかった。品質評価結果で低ビットでの品質の高さがアピールされ、同時にインターネット環境で

使える試用ソフトが配布され、草の根のように利用者を獲得した。これに対し、AAC は品質評価結果で MP3 などより有意に品質が高く、機能も充実していたので、日本のデジタル放送や携帯プレーヤでトップダウン的に採用が決まった。いずれにしても世界的な普及に必須であるのは品質などの技術的競争力である。逆に MPEG-4 オーディオではスケーラブル符号化や低ビットの符号化をはじめとする多くの新機能をもつ符号化が制定されたが、普及が進んでいない。この原因として制御の複雑さ、すでに普及している方式とからみた利点が小さいことが考えられる。また主流となった IP 伝送では遅延が少ないと IP ヘッダの相対伝送量が多く、低ビット化やスケーラビリティの意義が薄れてしまったことにもよると考えられる。

4.2.2 IPR コスト

IPR (Intellectual property right) すなわち特許権は、標準化を提案する立場では競争力のある技術を開発し開示する原動力となる。一方、技術を利用する立場からは技術の普及の障害となることがある。映像符号化の場合、技術を開示する機関と利用する機関がともに多く、両者が重複する場合もしばしばあるため、特許プールが有効に機能する条件が整っている。オーディオ符号化では技術を開示する機関と利用する機関が異なる場合が多く、利害の戦いとなる。ロイヤリティーフリーの符号化の開発も活発になりつつあるが、オーディオ符号化では主流ではない。標準化では公知の技術だけでは競争力の高い規格とはなりえないし、オープンソースでは互換性の管理や保守性の問題がある。またいずれにしても真に抵触する特許がないことを確認するのは難しく、性能の優れた技術に対する妥当な額の特許料は必要コストとして社会が負担していくことが望ましい。

4.2.3 MPEG と ITU-T

映像の符号化では長く MPEG と ITU-T が共同歩調をとっている。映像では MPEG-2 ビデオ (ITU-T 勧告 H.262) と MPEG-4 AVC (ITU-T 勧告 H.264) が共同作業の大きな成果であり、両標準化組織が策定した一つの標準であることが、これらの標準の世界的に採用される根拠になったとすることができる。これは映像圧縮が多く

の技術要素の集積によって実現されること、すなわちチームを組んで方式開発を行う利点が多いこと、エンコーダやデコーダの規模が比較的

**音声音響符号化標準は
MPEG と ITU-T ですみわけ**

大きく、通信・放送・蓄積媒体で共通に使える単一標準のニーズが高いこと、特に放送と蓄積など片方向の市場が大きいことなどによるものと考えられる。

一方、オーディオや音声符号化は共同で国際標準を制定したことがない。

オーディオや音声符号化では、表1のように多くの観点で独立に標準規格が作られている。音声は双方向の電話という巨大な市場が存在し、遅延がITU-Tの標準化にとって重要な要件となる。これに対してMPEGオーディオは入力段階で符号化対象とする信号の帯域が広く、放送や蓄積媒体が大きな市場であり、双方向の市場は小さいことから、これまで共同作業による標準策定の機会がなかったと考えられる。

しかしながら、MPEG-4オーディオでは双方向通信を想定した符号化も制定され、例えばAAC-ELD (Enhanced Low Delay)は30ms程度の遅延でのオーディオ符号化である。2008年からはMPEGでUSAC (Unified Speech and Audio)の標準化を進めている。遅延は100ms程度まで許すので片方向の用途となるが、モノラルで12 kbit/s ステレオで16 kbit/sの低レートで、音声も音楽もできるだけ広帯域（ラジオ以上）で高品質に符号化が可能となる見込みである。これには従来のオーディオ符号化と電話音声符号化技術が結集される見込みであり、さらにこれが低遅延化されれば万能の符号化となり、MPEGとITU-Tの技術の境界が完全に消滅することになる。

表1 ITU-TとMPEGの相違

機関	ITU-T	MPEG
典型入力	8 kHz サンプル モノラル	48 kHz サンプル ステレオ
用途	電話音声 双方向, 低遅延	オーディオ 片方向
方向性	広帯域化, 高品質化	低ビット化, 多チャンネル化
規定範囲	符号器復号器すべてを規定	ビット列と復号器のみの規定
実用対応	遅延, 演算量の制約を規定	用途に合わせてプロファイル
符号復号	符号復号器はほぼ同数, 同居	符号器は少量 復号器は多量
品質	品質が保証される	品質改良の余地あり

5. おわりに

過去20年以上に渡り、MPEGおよびITU-Tによる国際標準化は映像音声符号化技術の発展を牽引してきた。共同作業が技術の急速な進化を促し、標準が内包する競争的要素が標準の進化と普及を後押しした。標準方式の決定は技術開発の終止符とならず、新たな市場を目指して、標準方式の開発が続けられてきた。今後とも、高品質化や3D化などデジタルメディアの領域そのものが拡大し、標準化が技術開発を促進する要素であることを期待する。

謝辞 国際標準化の活動に貢献しておられる皆様に深謝いたします。

参考文献

- 1) ISO/IEC 11172-2, (MPEG-1 Video)
- 2) ISO/IEC 13818-2, (MPEG-2 Video)
- 3) ISO/IEC 14496-2, (MPEG-4 Visual)
- 4) ISO/IEC 14496-10, (MPEG-4 AVC)
- 5) 安田, “MPEG/マルチメディア符号化の国際標準”, 丸善 1994年9月
- 6) 浅井, “符号化圧縮技術の現状”, 映情学誌 2009年12月
- 7) M. Bosi and R. E. Goldberg, "Introduction to digital audio coding and standards," Kluwer Academic Publisher, 2003.
- 8) ISO /IEC 11172-3, (MPEG-1 Audio)
- 9) ISO /IEC 13818-3, (MPEG-2 Audio)
- 10) ISO/IEC 13818-7, (MPEG-2 AAC)
- 11) ISO/IEC 14496-3, (MPEG-4 Audio)
- 12) 守谷, “音声・音響符号化技術と標準化動向”, 信学技法 2009年5月
- 13) 守谷, “オーディオ符号化技術の現状と最新動向”, 電子情報通信学会誌, vol. 93 No. 2, pp. 133 - 138, 2010.

浅井 光太郎 (正会員)

E-mail asai.kotaro@db.mitsubishielectric.co.jp

著者略歴：1981年東大・工・電気工学科卒業。同年三菱電機株式会社入社。以来、映像信号の符号化伝送の研究開発，標準化に従事。1992-3年コロンビア大学客員研究員。2006年よりISO/IEC JTC 1/SC 29議長。情報処理学会，電子情報通信学会，映像情報メディア学会，画像電子学会，IEEE各会員

守谷 健弘 (正会員)

E-mail moriya.takehiro@lab.ntt.co.jp

著者略歴：1980年東大院・工・計数修士課程了。同年日本電信電話公社(現 NTT)入社。以来、音声音響信号の符号化の研究開発，標準化に従事。1989年工博。同年 AT&T ベル研究所客員研究員。現在情報処理学会規格調査会 SC 29 専門委員長，NTT コミュニケーション科学基礎研究所特別研究室長，NTT フェロー。情報処理学会，電子情報通信学会，日本音響学会，IEEE (Fellow)各会員

投稿受付：2010年1月7日

採録決定：2010年2月8日

メンタ：大蒔和仁 (東洋大学)