

インターネット画像通信

松本 修一

KDD 研究所 画像通信グループ

電子メールや ftp など一般に浸透してきたインターネットは、WWW(World Wide Web)の出現により爆発的に普及し、当初は文書や静止画が中心であったコンテンツも、圧縮技術の進歩とともに現在では動画像が数多く見られ、本格的なインターネットマルチメディア時代が到来したといえる。このインターネット画像通信を支える動画像圧縮、パケット多重化、伝送プロトコルの3つの基本技術について解説した上で、現状のインターネットビデオアプリケーションについて報告する。合わせて筆者らが開発した超低ビットレート画像圧縮 APC 方式について述べ、インターネットにおける画像通信技術の現状報告とする。

Internet Video Communication

Shuichi MATSUMOTO

Visual Communications Laboratory
KDD R&D Laboratories Inc.

Internet, initially known as E-mail and ftp, is dramatically coming into wide use due to WWW (World Wide Web) appearance. Its service contents, which were mainly text data and still pictures at the start, includes many motion pictures presently with a progress of bit-reduction video coding technologies. Hence, it can be said that the times with the internet multi-media have come to our actual life. In this paper, three essential technologies (bit reduction coding, packet multi plexy and transport protocol) are firstly explained as indispensable technologies for realizing the Internet video communication. And then, actual internet video applications are also introduced. In addition, a new low bit rate video coding system named APC which has been newly developed by KDD are also introduced.

1. まえがき

テレビ電話に端を発する低ビットレート動画画像圧縮技術は、インターネットの爆発的な普及に伴い、インターネットでの動画配信すなわちインターネットビデオに適用されるに至った。インターネットでの動画画像伝送はユーザ端末回線の制約から、伝送ビットレートとしてモデム 28.8kbps から ISDN (2B) 128kbps までを対象とするのが一般的であり、動画画像通信には十分な帯域とはいえず、しかも伝送帯域や遅延の時間的変動があるなど必ずしもリアルタイムの動画画像伝送に適している環境とはいえない。しかしながら、従来のインターネットコンテンツの中心であったテキストや静止画に対し、動画画像情報は映画やテレビといった形で人々の生活に深く浸透しており、品質的な面は別としてインターネット動画画像通信の存在意義はかなり大きなものとなっている。

本稿では、このインターネットビデオを支える動画画像圧縮、パケット多重化、伝送プロトコルの 3 つの基本技術についてまず解説する。また、インターネットビデオアプリケーションの代表例を紹介する。合わせて筆者らが開発した高品質超低ビットレート画像圧縮 APC 方式について述べ、インターネットにおけるビデオ通信技術の現状報告とする。

2. 要素技術

2.1 画像圧縮

動画画像圧縮では、圧縮処理と復号処理が対をなすいわゆる相互接続性が要求されるため、圧縮方式の統一化が必須であり、ITU や ISO は国際標準方式の確立にこれまで精力的に取り組んできた。図 1 にこれまでの国際標準化の流れを示す。超低レート動画画像圧縮の国際標準として、最初に定められたのが 1990 年の ITU 勧告の H.261¹⁾であり、64kbps の整数倍でのビットレートで主に画面の動きが少ないテレビ電話、会議を対象としていた。その後、1992 年に ISO は、CD-ROM などの蓄積媒体用の圧縮標準として H.261 の技術に基づいた MPEG-1²⁾を標準化した。ランダムア

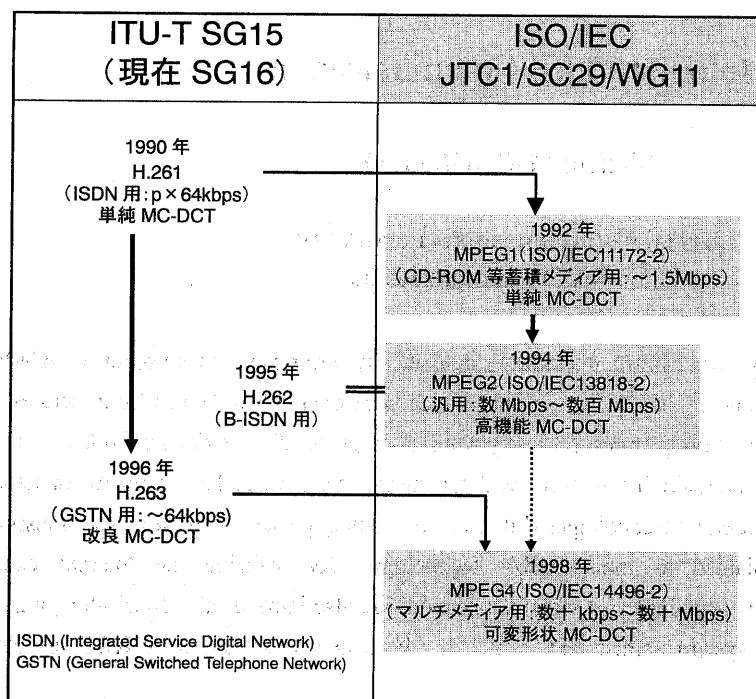


図 1 動画画像圧縮の国際標準化の流れ

クセス、高速サーチ、逆再生などに対応するもので、ビットレートの上限は 1.5Mbps と定められた。この MPEG-1 に続いて、1995 年に現行テレビから HDTV まで汎用的に圧縮が行なえる MPEG-2³⁾ (ITU の勧告番号は H.262) を標準化した。対象ビットレートは画像品質を重視し、数 Mbps $\sim$ 数百 Mbps でかなり高く設定されている。1996 年には超低ビットレート用として ITU から H.263⁴⁾ が勧告化された。これは先の H.261 をベースとして、動き検出精度の向上、圧縮情報に割り当てる可変長符号化の改良などにより、おおよそ

2 倍の圧縮率を実現させ、現在のインターネットビデオの配信アプリケーションやパソコンテレビ会議などの多くのサー

ビスに採用されている。

一方、ISO では MPEG-2 の標準化に引き続き、超低ビットレート用の符号化標準 MPEG-4⁵⁾の検討作業に入った。その結果、符号化対象とする画像や用いる技術に関して、H.263 の拡張に加えて、画面中の物体、背景毎の符号化や操作、自然画と合成画との統合、各種階層化、回線誤りに対する高度な耐性化といった圧縮標準の実現を目指し、標準化された。圧縮技術の立場から、強いて MPEG-4 の特長を一つ挙げるならば、その有意性は別として、従来矩形固定のブロック単位での動き検出や DCT を可変形状に拡張した点であろう。

その他の低レート用画像符号化方式として、フラクタル符号化が代表的である。これは、IFS (Iterated Function System) 符号化とも呼ばれ、ジョージア工科大学の Barnsley により発明された IFS 理論⁶⁾に基づく符号化方式である。IFS とは、画像の局所領域は同じ画像内から切り出されるある部分の縮小近似で表されるという理論である。同大学の Jacquin が IFS による画像符号化方式を発表している⁷⁾。圧縮効率 は JPEG 相当であり、縮小変換の繰り返しで復号されるため符号誤りに強いといった特徴がある。

特にインターネットを意識した超低ビットレート動画像符号化技術の究極の目標は、高画質、オールソフト、双方向通信、リアルタイム伝送となる。H.263 を始めとして、様々なソフトウェアがインターネット上で動画像配信、通信に用いられているものの、これら課題をすべて満たすものはいまのところ存在しない。すなわち、画像品質の面では、画質とフレーム数がトレードオフの関係にあり、両立はできない。たとえば、画質を優先させた場合、毎秒送るフレーム数が少なくなるいわゆるスキップフレームが発生する。さらに、これは伝送遅延時間を増大させる要因となり、双方向通信に支障をきたす。一方、画質を犠牲にして、フレーム数を増やすとパソコンに対する許容演算時間が短くなり、結果的に符号化処理が間に合わず、リアルタイム伝送が不可能となる。この様に、品質、遅延、演算処理の 3 つの課題を従来の圧縮技術は抱えていると言える。

2.2 パケット多重化

圧縮されたビデオビットストリームは、インターネット送出に適した形でパケット化され、音声信号等の他のメディアと同期多重される。マルチメディアシステムの標準方式でインターネットを対象としたものとしては、ITU-T H.323⁸⁾およびインターネットの標準化機関である IETF (Internet Engineering Task Force) の RTSP (Real Time Streaming Protocol)⁹⁾が挙げられる。

H.323 は IP ネットワーク上でのテレビ電話/会議システムを対象としており、オーディオ/ビデオデータの RTP (Real-time Transport Protocol)¹⁰⁾によるパケット同期多重、H.225.0¹¹⁾による呼制御、H.245¹²⁾による端末間制御、およびネットワークへのアクセス制御により端末間のリアルタイム通信制御を規定している。H.323 では、ビデオ符号化方式として H.261、H.263 を用いる。

一方、RTSP はサーバ/クライアント間のオーディオ/ビデオストリーミングを目的として 1998 年 4 月 IETF RFC 2326 として勧告化された。RTSP は HTTP¹³⁾をベースにした文字列形式となっており、クライアントからのストリームの再生/停止/早送り/巻戻し/指定位置からの再生といった機能を有し、いわばインターネット上でのリモート VTR 機能を提供するものである。ビデオ符号化方式は特に限定しておらず、またオーディオ/ビデオデータの配信については H.323 と同様 RTP が用いられる。

RTP は、IETF RFC1889 および ITU-T H.225.0 として勧告化されている。ここではメディアデータをパケット化し、メディアタイプやタイムスタンプ等のヘッダを付加した後、エラー訂正やその他の復帰制御のない”信頼性のない伝送路”すなわち UDP により伝送する。IETF では種々のビットストリームに対する RTP フォーマットを規定している。

2.3 伝送プロトコル

多重化等によって伝送準備の整ったパケットは、伝送プロトコルにより、実際にインターネット上に送出される。インターネットでは、TCP/IP、UDP/IP プロトコルが一般に用いられる。TCP (Transmission Control Protocol)¹⁴⁾はコネクション

型のプロトコルで、パケットは 20 バイトのヘッダとペイロードで構成されており、パケットの紛失、損傷、重複、遅延、順番の入れ違い等の起こりうる信頼性の低いネットワーク上で確実な通信を提供するものである。これに対し UDP (User Datagram Protocol)¹⁵⁾ はコネクションレス型のプロトコルで、相手からの応答や伝達の確認は行わないため TCP に比べて遅延が少なく、パケットヘッダは 4 バイトとオーバーヘッドの小さい伝送を提供する。

TCP ではパケットヘッダが大きく、さらにネットワークの状態によってはパケットの再送が行われることがあるため、インターネットでの動画配信については、伝送効率や遅延の面から UDP を用いるのが一般的である。しかし、企業内 LAN 等では、外部からの不正侵入防止やウィルスチェック等の目的でファイアーウォールを設置するのが一般的であり、この場合は Web アクセスのための HTTP のみがインターネット接続可能となる。HTTP は TCP 上のプロトコルであるため、ファイアーウォール透過の目的のためにリアルタイム性を多少犠牲にしながらも HTTP で動画配信を行う例もある。

UDP ではパケット損失等のエラーに対する処置が全く行われないため、これより上位の層すなわち画像コーデックによりエラーに対する処置を行う必要がある。これには定期的なリフレッシュフレームの挿入やイントラスライスによる段階的のリフレッシュ、エラー耐性 VLC コード、受信側からエラーのないフレーム番号をフィードバックしエラーのないフレームを参照画像として符号化する適応参照フレーム選択、FEC (Forward Error Correction) 等が用いられる。

3. インターネットビデオアプリケーション

ここでは、インターネットビデオ配信の代表的なアプリケーションである米国 RealNetworks 社の RealSystem、米国 Xing Technology 社の StreamWorks、イスラエル VDOnet 社の VDOLive を紹介する。

RealSystem は全世界で最も多くのサイトで用いられている事実上の標準方式となっている。システム構成は RealAudio を RealVideo の音声/画像オールソフトウェアエンコーダと RealServer サーバによりオンデマンド配信およびライブ配信を行い、RealPlayer クライアントソフトウェアで受信する形をとるが、これに加えて最新版 G2 からは RealText、RealPix の 2 つの新たなデータタイプが追加され、高機能テキストおよび高機能静止画像による超低ビットレートインターネットプレゼンテーションが可能となっている。伝送帯域の異なるクライアントからのアクセスに 1 本のビットストリームで対応するため、SmartStream と呼ばれる独自開発のスケラブルビットストリームも生成可能である。画像品質に関しては、アナログモデムユーザを想定した 20kbps 前後においては H.263TMN6 シミュレーションの品質よりやや高い。画像やテキストの特殊効果等を記述するマルチメディア記述言語 SMIL (Synchronized Multimedia Integration Language) を新たにサポートし、単なるオーディオビデオストリーミングから本格的なインターネットマルチメディアストリーミングシステムへと発展を続けている。コンテンツへのアクセスの利便性や GUI の完成度等の使い勝手も他に比べて大きくリードしており、現在最も進んだインターネットビデオシステムといえる。

StreamWorks は基本的には MPEG-1 システム (ISO/IEC 11172-1) に準拠しており、MPEG-1 ビデオ/オーディオを始めとして MPEG-2 オーディオレイヤ 2 およびレイヤ 3 もサポートする。ライブエンコーダ MPEGLive! Encoder はソフトウェアおよび MPEG 圧縮ハードウェアから構成され、インストールされた状態で PC ごと販売されるのが大きな特徴である。StreamWorks は MPEG-1 システムに準拠しているため、同社のエンコーダやオーサリングツール以外のアプリケーションにより作成された MPEG-1 コンテンツも配信可能である。動作ビットレートが高い、ファイアーウォール対応性が低い、MPEG-1 ビデオ/オーディオ配信にその機能を限定している。サーバの対応プラットフォームが幅広いといった点を考慮すると、インターネットビデオ配信というよりは企業内イントラネットもしくは衛星・ケーブルモデムによる広帯域配信にふさわしい構成になっていると思われる。

VDOLive は画像圧縮方式として Ver2.X では H.263 ベースの符号化方式を、Ver3.0 では VDOnet 社独自開発の Wavelet ベースの符号化方式を採用している。モデムアクセスを対象とした 20kbps 前後での画像品質は、H.263TMN6 シミュレーションと同程度である。無償配布のデコーダソフトウェアはオートマチックアップグレード機能を有しており、ソフトウェアのバージョンアップが行われると自動的に最新版をダウンロードしアップグレードが行われる。さらに Ver3.0

インターネットによる動画配信システムは、APC 用の独自開発のサーバ、エンコーダ、クライアントから構成されている。伝送ビットレートは24kbps～384kbpsで、APC符号化された動画は、サーバにおいてITU-T標準方式G.723.1 (5.3/6.3kbps デュアルレート音声符号化方式)で圧縮された音声データとともにオーバーヘッドの非常に小さい独自フォーマットの packets 多重化がなされる。配信形態はリアルタイムライブ形式とオンデマンド形式に対応し、UDP もしくは TCP 上の HTTP の各プロトコルによりインターネットへ配信される。HTTP を用いることにより、ファイアウォール環境下でのアクセスを可能としている。オンデマンドアクセス時には、クライアント側からの再生、一時停止、ランダムアクセス等の要求コマンドに対応してサーバが応答し、冗長度を極力小さくした独自のコマンド形態によりレスポンスの良いアクセスを実現している。

5. おわりに

急速に進歩を続けるこの分野も、動画圧縮の基本技術については完成された感が強く、今後は電子メールやマルチメディアデータベースといった他のメディアとの連携技術やサービス面での大きな進歩が予想され、その点において MPEG-4 をベースとした技術の発展が期待される。また、ユーザー要求がしっかり定義され、それに基づき規格が定められているデジタル TV 放送規格とは異なり、超低レート動画圧縮の対象となるインターネットや移動体による画像通信サービスでは伝送環境が不安定であるため、サービス品質の規定並びにその保証といったいわゆる QoS の問題が、今後大きくクローズアップされよう。

参考文献

- 1) ITU-T Recommendation H.261: Video codec for audiovisual services at p×64kbit/s (1993)
- 2) ISO/IEC 11172-2: Information Technology - Coding of moving pictures and associated audio for digital storage media at up to about 1.5Mbit/s - Part 2: Video (1993)
- 3) ISO/IEC 13818-2: Information Technology - Generic coding of moving picture and associated audio information - Part 2: Video (1995)
- 4) ITU-T Recommendation H.263: Video coding for low bit rate communication (Mar. 1996)
- 5) ISO/IEC CD 14496-2: Information Technology - Coding of audio-visual objects - Part 2: Visual (Mar. 1998)
- 6) M. F. Barnsley: "Fractals Everywhere," Academic Press (1988)
- 7) A. E. Jacquin: "Image coding based on a fractal theory of iterated contractive image transformations," IEEE Trans. Image Proc., 1, 1, pp.18-30 (Jan. 1992)
- 8) ITU-T Recommendation H.323: "Visual telephone systems and equipment for local area networks which provide a non-guaranteed quality of service" (Nov. 1996)
- 9) H. Schulzrinne, A. Rao and R. Lanphier: "Real time streaming protocol (RTSP)," RFC 2326, Internet Engineering Task Force (Apr. 1998)
- 10) H. Schulzrinne, S. Casner, R. Frederick and V. Jacobson: "RTP: A transport protocol for real-time applications," RFC 1889, Internet Engineering Task Force (Jan. 1996)
- 11) ITU-T Recommendation H.225.0: "Media stream packetization and synchronization for visual telephone system on non-guaranteed quality of service LANs" (Nov. 1996)
- 12) ITU-T Recommendation H.245: "Control protocol for multimedia communications" (Mar. 1996)
- 13) R. Fielding, J. Gettys, J. Mogul, H. Nielsen and T. Berners-Lee: "Hypertext transfer protocol - HTTP/1.1," RFC 2068, Internet Engineering Task Force (Jan. 1997)
- 14) J. Postel: "Transmission control protocol," RFC 761, USC/Information Sciences Institute (Jan. 1980)
- 15) J. Postel: "User datagram protocol," RFC 768, USC/Information Sciences Institute (Jan. 1980)
- 16) 宮地, 松本: "統合的時空間制御による超低ビットレート動画符号化方式," 映情学誌 51, 10, pp.1706-1714 (Oct. 1997)