

3

ソフトウェアコーデックによる 高品質 HD 映像の IP 伝送

熊谷誠治

(株) 電通国際情報サービス

今回の日韓ハイビジョン・テレビ中継実験は JGN2plus を利用し、雪まつり会場の映像を韓国、京仁テレビ (OBS) まで伝送し、韓国国内でテレビの生放送を行うというものである (写真 1)。今回の使用可能帯域は 1Gbps であり、リアルタイムでの映像圧縮技術の導入が不可欠である。また、汎用性を考慮して PC をベースとするため、ソフトウェアコーデック方式を採用することとした。今回我々は 2007 年から開発を続けてきた PC を用いた伝送システム (ProRes 422 HD 伝送システム) を用い生放送のコンテンツの伝送に成功した。本稿では、今回開発した ProRes 422 HD IP 伝送システムの特徴、動作特性と実際の長距離伝送 (実際には広帯域高遅延条件での伝送となる) における問題点の解決に関して検討した結果を報告する。なお、放送現場の状況は、1140 ページのコラムで報告されておりあわせて参照されたい。

PC による IP 伝送システム

我々は、2000 年ごろから PC を利用したデジタル・ビデオ (Digital Video - DV) の IP 伝送に取り組み、2004 年には多地点の DV 映像をソフトウェアで合成して各地点へ送り返すテレビ会議システムを開発し、遠隔授業など向けに提供してきた。

ここに来て、地上波テレビ放送のデジタル化に伴う HD (High Definition) テレビ放送受信機の普及に合わせる形で、映像伝送システムにおいても HD 化の要求が高まってきた。そこで、我々の映像伝送システムも HD 化の検討を始めた。

市場では HDD レコーダや Blu-ray レコーダが録画時間を増加させるために安価で圧縮率の高いコーデックが必要となり、ワンチップの H.264 コーデックが開発され、製品に組み込まれている。テレビ放送受信機も「アクトビラ」のようなインターネット経由の映像配信に対応するため、H.264 デコーダを搭載するようになっており、民生機器分野での高圧縮技術はどんどん進んでいる。業務用映像機器分野においても、H.264 などの高圧縮



写真 1 京仁テレビ (OBS) に設置した受信用 Apple Xserve

技術を採用した製品が増加している。しかし、H.264 は計算量が多いため CPU では計算が追い付かず、専用のハードウェアを使用するのが一般的である。ニッチ分野で特別なハードウェアを付加することは、製品価格の上昇を招き、ハードウェアの固定は日々発達する技術への対応を難しくする。

そこで、放送局や映像プロダクションが映像編集用に使用している PC を利用し、簡単なしくみで放送局でも使用できる高画質 HD 映像の IP 伝送を実現するための技術検討を始めた。2007 年のことである。

DVCPRO HD 伝送

ソフトウェア・コーデックを調べていくと、Mac 上で利用されている Final Cut Studio という映像編集ソフトが、DVCPRO HD のソフトウェア・コーデックを持っていることが分かった。DVCPRO HD は、Panasonic が開発したコーデックで、ポータブル・ビデオカメラのコーデックとして使用されている。DVCPRO HD は、放送局でのニュース取材用や屋外撮影でも使われており、その画質には定評がある。画素数 1280 × 1080、クロマレ

コーデック	画素数(最高)	クロマサンプリング	カラー深度(最高)	フレーム間圧縮	伝送遅延
DVCPRO HD	1280 × 1080	4:2:2	8bit	なし	約 300m sec.
ProRes 422 Normal	1920 × 1080	4:2:2	10bit	なし	約 200m sec.
ProRes 422 HQ	1920 × 1080	4:2:2	10bit	なし	約 200m sec.
HDV	1440 × 1080	4:2:0	8bit	あり	約 1sec.

表-1 各種コーデックの機能比較

ベル 4:2:2, カラー深度 8bit である。

DVCPRO HD のコーデックをプログラムから利用できるかどうかを調べたところ, QuickTime の API から利用できることが分かった。また, DVCPRO HD は, DV を HD 化しているため, データフォーマットも DV と同様の 80 バイトの DIF ブロックを基本とした構造となっている。

このため, これまで DV の伝送に用いてきた DIF ブロックを RTP のペイロードに載せる RFC3189¹⁾ と同じ方式で, DIF ブロック数を増やすだけでそのまま DVCPRO HD の IP 伝送に適用することが可能である。ただし, ソフトウェア・コーデックであることから, HD-SDI と呼ばれる放送局用のカメラ映像入力, エンコード, ネットワーク出力を 1 台のマシンで実行するため, 4 コア搭載のマシンでは, CPU 負荷が心配であった。とりあえず手慣れた RFC3189 をベースに Mac 上に DVCPRO HD 伝送システムを実装したところ, 2.66GHz, 4 コアでも問題なく送信または受信できることを確認できた。

HD-SDI は, 同軸ケーブルを利用して信号を伝送する方式で, 放送局内の映像伝送に利用されている。映像と音声を同一のケーブルに乗せるエンベデッドが可能である。

幕張で開催された放送技術展 (InterBEE) に出展したところ, 編集用に使っている Mac を使ってライブ映像伝送ができるということから, 注目を集めることができた。しかし, 地上デジタル放送が 1440 × 1080, BS デジタル放送が 1920 × 1080 を採用しているという現状では, 画質的に劣るといふ指摘を受けることがあった。

ProRes 422 コーデック

その後, 新バージョンとして Final Cut Studio 2 がリリースされ, これには Apple 社による新開発のビデオコーデック ProRes 422²⁾ が搭載された。ProRes 422 は, 編集映像の保管用に開発されたソフトウェアによる圧縮方式で, フル HD (1920 × 1080) に対応し, DVCPRO HD や HDV に比較して高画質である。また, 非可逆圧縮であるため圧縮・伸長を行うと画質の劣化を伴うが, 圧縮・伸長を繰り返しても画質の劣化が少ないのが特徴という。VBR (Variable Bit Rate) エンコーディングに対

応し, Normal, High Quality という 2 つのビットレートを選択できる。これらの圧縮方式の比較を表-1 にまとめた。

多くの映像システムでは, 映像記録時に輝度信号と色信号に分解して記録している。このときに人間の眼は色情報に鈍感であるということを利用してデータ量を減らしている。クロマサンプリング 4:2:2 とは, 輝度信号 Y, 色差信号 U (Y-B), 色差信号 V (Y-R) の比率を表しており, 輝度信号に比べて 2 つの色差信号が半分になっていることを表す。4:2:0 は, 垂直方向で 1 走査線ごとに U と V を交互にサンプリングするため, 情報量を減らすことが可能であるが, その分画質が低下する。

ProRes 422 の技術詳細は非公開となっているが, 4 コア CPU を用いればリアルタイムで HD-SDI から入力し, エンコードしてディスクに書き込めるため, 我々が考える高画質ネットワーク伝送に適していると判断した。

ProRes 422 の HQ モードは, 非圧縮の標準画質 (SD - Standard Definition) よりも少ないディスク保管容量を目指して開発されており, 伝送帯域においても非圧縮 SD 程度に収まるため, 伝送路の確保という面でも優れている。

IP 伝送システム

このような状況のもと, ProRes 422 伝送システムの開発に着手した。ターゲット・マシンとして, 3GHz 8 コアの Xserve (Mac OS X Server 10.5.4) を 2 台と, HD-SDI 入出力ボード AJA Kona 3 を 2 枚使用した。

ProRes 422 においては, DVCPRO HD で用いた RFC3189 のような標準がないため, 単純に圧縮後のデータをパケットのサイズに分割して, 順次送信する方法を採用した。音声サンプルも同様にパケットのサイズに分割し, 順次送信する。映像/音声の区別は RTP ヘッダの直後に 1 バイトのタグを付加することで実現した。なお, 60i と呼ぶ 59.4fps のインターレース映像にのみ対応することとして, 必要最小限の開発で動作確認することを優先させた。

Apple ProRes 422 (Normal, HQ とも) でエンコードされた映像や音声 (16bit PCM, 2ch) や FEC のパリティパケ

バイト位置	0 ... 11						12 ...
内容	RTP ヘッダ						データ
	バージョン	マーカ	ペイロード タイプ	シーケンス番号	タイムスタンプ	SSRC	
ビット位置	0 ... 7	8	9 ... 15	16 ... 31	32 ... 63	64 ... 95	96 ...

表-2 パケット・フォーマット

ペイロードタイプ	意味	
0	NONDV	NONDV パケタイザでパケット化されている Apple ProRes 422 の場合はこの値になる
1	DVCPRO HD	DVCPRO HD パケタイザでパケット化されている
2	RAW YUV	RAW パケタイザでパケット化されている 非圧縮送信の場合
3	FEC	FEC のパリティパケット

表-3 ペイロードタイプ・ビットの意味

バイト位置	0 ... 11 (12 バイト)	12	13 (パケット末尾まで)
内容	RTP ヘッダ	タグ	メディアデータ

表-4 NONDV タイプのパケット・フォーマット

タグの値	意味	メディアデータの内容
0	映像	パケットごとに分割されたエンコード済み映像データの断片
1	音声	パケットごとに分割された 2ch 分の PCM 音声データの断片
2	テスト	テスト用データ
3	ヘッダ	使っているコーデックや映像のサイズなどの付加情報

表-5 メディアデータを示すタグの意味

バイト位置	0 ... 11 (12 バイト)	12 (タグ)	13 (パケット末尾まで)
内容	RTP ヘッダ	0	分割された映像データの断片

表-6 映像パケットの構造

ットは、表-2のようなフォーマットの RTP パケットとして送信する。なお各フィールドの値はネットワークバイトオーダー(Big Endian)にて格納している。

このとき、RTP のペイロードタイプとして表-3の値をいれることで、データ部の内容を区別できるようにする。ここではペイロードタイプ 0 (NONDV) および 3 (FEC) のパケット・フォーマットを記述する。NONDV ペイロードタイプのパケットは、表-4のようにタグを付加することで、その内容(映像・音声など)を区別する。Apple ProRes 422 のデータはこのフォーマットで送信される。

この「タグ」は表-5のような意味を持ち、バイト位置 13 (14 バイト目) 以降のメディアデータの内容を表示している。映像データは 1 フレーム分をパケットのサイズ

に収まるように分割し、RTP のシーケンス番号の若いパケットから順に詰める(表-6)。1 フレームの最後のパケットは RTP ヘッダ中のマーカビット(ビット位置 8)を立て、フレームの最後のパケットであることが区別できるようにする。受信側では、シーケンス番号の若いほうから順に並べることでフレームのデータを構成し、マーカビットの立っているパケットが来た時点で 1 フレーム分が届いたと見なしてデコードする。

音声データは、ほぼ 1 フレーム分のデータを、映像と同じくシーケンス番号の若いパケットから順に詰める(表-7)。このとき音声サンプルは L, R, L, R... の順に並べ、パケット末尾は必ず R チャンネルのサンプルで終える。ただし音声の場合は 1 フレーム分のデータ量に規定はなく、「1 フレームの映像データを用意している間にキャ

バイト位置	0...11 (12 バイト)	12 (タグ)	13... (パケット末尾まで)
内容	RTP ヘッダ	1	分割された音声データの断片 (L, R, L, R...)

表-7 音声パケットの構造

バイト位置	0... 11	12... 21	22... 25 または 22... 29	26... または 30...
内容	RTP ヘッダ	FEC ヘッダ	FEC レベルヘッダ	パリティ (XOR)

表-8 FEC ペイロードタイプのパケット・フォーマット

コーデック	スペック	計算値	実測値 (8bit)	実測値 (10bit)
Apple ProRes 422 Normal	145Mbps	151Mbps	139 ~ 160Mbps	143 ~ 164Mbps
Apple ProRes 422 HQ	220Mbps	228Mbps	224 ~ 246Mbps	239 ~ 243Mbps
DVCPRO HD(60i)	109Mbps	113Mbps	114Mbps	—

表-9 コーデックごとの伝送帯域の比較

プチャされた音声データ」が 1 フレーム分の音声データということになる。また、将来は音声パケット用タグを追加し、3 チャンネル目以降の音声チャンネルの伝送に対応する予定である。その場合は既存の音声用タグ (1) に加えて 4, 5, 6 を音声用に用いる。

Forward Error Correction

RTP ストリーム用の FEC (Forward Error Correction) として、RFC5109³⁾ に準拠した方式を採用し、パケットロストからの回復機能を実現した。RFC5109 方式の FEC は以下のような特徴を持つ。

1. XOR 演算 (パリティ) によるシンプルな方式であるので、リアルタイムなアプリケーションに向く
2. FEC に対応していない受信者は、単純にパリティパケットを無視することで、通常通りの受信が可能
3. マルチキャスト送信の際に、FEC 対応/非対応ホストを混在させることが可能
4. パリティパケットの量を増減することで、データの信頼性と使用帯域幅の調整が可能
5. 個々のメディアパケット (保護の対象となるデータパケットをメディアパケットと呼ぶ) の重要度に応じて、動的にパリティパケットの量を変更可能

今回の実装では、復元の対象となるデータが Apple ProRes 422 であり、個別のメディアパケットごとの重要度の違いが不明である。そのため、5 番目の「メディアパケットの重要度に応じて、動的にパリティパケットの量を変更」させる機能については実装していない。パリティパケットの付与率を変更する方法としては、コマ

ンドラインオプションにおいて指定する方法を用意した。ここで使用する FEC ペイロードタイプのパケットは表-8 のようになっている。なお、RFC5109 に準拠しているが、マルチレベルパリティは使用していない。FEC ヘッダおよび FEC レベルヘッダには (12 バイト以降) がパケット保護のための基本情報が、パリティ部分 (26 または 30 バイト以降) には保護対象のパケット (RTP ヘッダ除く) の XOR をとった値が入る。

使用帯域と遅延

Apple ProRes 422 の使用帯域は表-9 のようになる。ここでいうスペックとはコーデックの仕様上の数値²⁾、計算値は MTU=1500 として RTP ストリーム化した場合のネットワーク上の帯域、コーデックから計算したものの (音声込み) である。

比較のため DVCPRO HD (60i) を使用した場合の帯域も併記する。送信した映像データは動きがあるため、VBR である Apple ProRes 422 では帯域が変動する。計算値と実測値を記載しているが、Apple ProRes 422 HQ の場合を除き、おおむね計算と一致している。Apple ProRes 422 HQ の帯域が高めな理由は不明であるが、スペックどおりのサイズに圧縮できていないのではないかと推測される。

本実装では、FEC によっても使用帯域が増大する。FEC の仕様上、パリティパケットは最小で 48 個のメディアパケットにつき 1 つ、また最大でメディアパケット 1 つにつき 1 つの割合で付加することができる。その結果 FEC つきの Apple ProRes 422 のストリーム (1920 ×

	FEC なし	48 につき 1	10 につき 1	1 につき 1
ProRes 422 Normal	145Mbps	148Mbps	160Mbps	290Mbps
ProRes 422 HQ	240Mbps	245Mbps	264Mbps	480Mbps

表 -10 FEC を使用した場合の伝送帯域

時間 スキップ量	1 分後のバッファ内 フレーム数	3 分後のバッファ内 フレーム数	5 分後のバッファ内 フレーム数	5 分後の遅延フレーム数
なし	6	24	28	36
1 フレーム	1	1	1	6 か 7
4 フレーム	1	1 か 2	1 か 2	6 から 8

表 -11 映像のスキップ量と遅延

時間 スキップ率	1 分後のバッファ内 フレーム数	3 分後のバッファ内 フレーム数	5 分後のバッファ内 フレーム数	5 分後の遅延秒数
なし	16	25	32	1.2 秒程度
1	7	8 から 10	8 から 10	0.3 秒程度
4	4 から 5	4 から 5	4 から 5	微小

表 -12 音声のスキップ率と遅延

1080, 60i)は表-10 に示す帯域を占める。

遅延量の計測は、カメラが発生するタイムコードを受信側のモニタに表示する方法で実施した。その結果、遅延は5から6フレーム(約170から200ms)という結果が得られた(8bit, 220Mbps モード時)。この遅延量は、DVCPRO HD 伝送の遅延量(約300ms)を大幅に下回っている。総務省の「IP ネットワーク技術に関する研究会」報告書⁴⁾によると、固定電話並みのクラスAは、100msec. 以下、携帯電話並みのクラスBは150msec. 以下、最低限の通話ができるクラスCを400msec. 以下と規定している。この数値に照らすと、本システムの約200msec. という遅延は、十分に双方向での掛け合いが可能な範囲内にある。

遅延の増大とその対策

長時間の送信においては、内部バッファに滞留するデータが増えることで、遅延が増大することがある。この原因としては、以下のような理由が考えられる。

1. 送受信のクロックずれにより、送信側が生成するデータ量が受信側の消費するデータ量より多いケース
2. 受信側で、再生デバイスへのデータの送出手が滞ることで、データが余るケース

結果的に発生する遅延の増大の対策として、受信側の内部バッファにあるデータを捨ててしまう(スキップすること)で、遅延を減少させる機能を実装した。

まず、映像の遅延であるが、表-11のような結果が得られた。

「スキップ量」は1フレームの再生あたり最大何フレームスキップするかを表す。つまり、この値が1であれば1フレーム再生するたびに1フレームスキップする可能性がある。また、4であれば1フレーム再生するたびに最大4フレームまでスキップする可能性がある。ただし、内部バッファにフレームがたまっていない場合は当然スキップしない。

「X分後のバッファ内フレーム数」はその時点で内部バッファにある映像フレームの数「5分後の遅延フレーム数」は5分後の時点で、計測した遅延フレーム数である。5分後で28フレーム分のバッファ内データという値は通常あり得ないほど極端な遅延の増大であるが、この場合でも最大1フレームのスキップで十分な効果が得られることが分かる。

次に音声の遅延については表-12のような結果が得られた。音声における「スキップ率」は、内部バッファ内の音声データの量と現在の音声レベルからスキップすべき音声サンプルの数を計算するためのパラメータである。バッファ内の音声データ量が多い、音声レベルが低い、「スキップ率」が大きい場合に、多くの音声サンプルがスキップされる。逆にバッファ内の音声データ量が少ない場合や、音声レベルが高く「スキップ率」が小さい場合には音声サンプルはスキップされず、通常通り再生される。

「X分後のバッファ内フレーム数」はその時点で内部バ



写真2 大通会場に設置した岡山向けの送信用 Xserve

ッファにある音声サンプルのフレームの数、「5分後の遅延秒数」は5分後の時点で、ストップウォッチで計測した音声の遅延秒数である。こちらの場合はスキップ率4が遅延増大の抑止のためには望ましいことが分かる。

これらの結果から分かるようにバッファ内データのスキップは遅延の抑制に有効であるが、反面本来出力すべきデータをスキップするため、コマ落ちや音声落ちが発生するという問題がある。そのため、実際の使用の際にはコマ落ちや遅延、それぞれの許容量を比較しつつ、パラメータの調節を行うべきである。また、現在のスキップ機能は映像と音声を同期させるリップシンクを考慮していない。これについては今後の検討課題である。

今回の実証実験では、本システムを日韓中継だけでなく、エリアワンセグの映像素材伝送用として、札幌→岡山間でも使用した。こちらは、IPv6 マルチキャストで送信し、岡山だけでなく東京の品川でも同時受信し、仮想化ネットワーク上において240Mbps以上の広帯域であっても、マルチキャスト可能なことを確認した(写真2)。

本システムは、放送局内での利用を意識して、外部映像機器と同期をとる GenLock にも対応している一方で、民生機器に使われている HDMI インタフェースからの入出力にも対応しており、安価なビデオカメラやカラーモニター、カラーテレビを接続することも可能である。

また、Backmagic Design 社製の HD-SDI 入出力ボードを合わせて搭載することで、8 コアの CPU を用いれば双方向伝送が可能とも確認した。同一メーカー製の HD-SDI ボードを2枚挿入することはできなかったが、メーカーが違えば2枚挿入が可能で、それぞれを入力と出力に使うことで、双方向での伝送が可能となることを確認した。

今後は、受信映像の Final Cut Studio ファイルへの書き出し、Final Cut Studio ファイルのネットワークへの送信対応なども検討している。

謝辞 本実験、システム開発ならびに実用化には、(独)情報通信研究機構 JGN2plus の東京→大阪間折り返し1000Kmの長距離遅延テストベッドの利用が不可欠であったことを付記し、関係各位に謝意を表す。

参考文献

- 1) RTP Payload Format for DV (IEC 61834) Video, <http://www.rfc-editor.org/rfc/rfc3189.txt>
- 2) Apple ProRes 422 White Paper, http://images.apple.com/jp/finalcutstudio/docs/ProRes422_WP_070518.pdf
- 3) RTP Payload Format for Generic Forward Error Correction, <http://www.rfc-editor.org/rfc/rfc5109.txt>
- 4) IP 電話通話品質評価法の標準化動向, http://www.oki.com/jp/Home/JIS/Books/KENKAI/n202/pdf/202_R28.pdf

(平成21年9月10日受付)

熊谷誠治 | kuma@isid.co.jp

京都工芸繊維大学工学部電気工学科卒業。現在は、(株)電通国際情報サービス 事業推進本部 開発技術センター 主幹研究員。(財)インターネット協会 企画運営会議メンバ、同 Net-Cast 研究部会 会長。著書に日経 BP 社刊「誰も教えてくれなかったインターネットのしくみ」など多数。

