

放送業界におけるIP化

～IP番組制作とその関連技術の解説～

開講にあたって

編集幹事 倉掛卓也

近年の映像信号の高品質化や放送機器のIP化の進展は著しく、放送業界においては、これらに関するさまざまな規格や規格間の関係を網羅的に把握することが求められるようになっていきます。

会誌でも、2016年9号の技術解説「IPライブ制作」、2017年の講座「UHDを支える映像の高画質化技術」、2018年3月号の「知っておきたいキーワード SMPTE ST2022」および「知っておきたいキーワード 4K伝送のための12 G-SDIとVideo over IP」など、IP番組制作とその関連技術の話題をタイムリーに取り上げてきましたが、本講座では、それらの総括の意図も含め、全6回の講座を組むことにしました。講座での連載を通じて、有機的かつ一気通貫に理解を深める機会になれば、と期待しています。

本講座では、全体像からスタートし、UHDやHDRなどベースバンド信号の進化、IPパケットマッピング方式、IPライブ制作システムの制御、IPライブ制作技術の検証実験事例、JPEG軽量映像圧縮技術について、それぞれの分野の専門家に解説をお願いさせていただきました。ご多忙中、ご執筆を引き受けて頂いた著者の皆様には、この場を借りて深く感謝申し上げます。

なお、本講座は、春日康志編集幹事、蝶野慶一編集幹事と私が担当いたします。約1年間の講座となりますが、何卒よろしくお願いいたします。

予定目次

- (第1回) 放送業界におけるIP化の概要 ∞∞∞ 蝶野慶一氏 (NEC)
- (第2回) UHDやHDRから見たSDI信号の進化 ∞∞∞ 矢野浩二氏・鈴木則行氏 (リーダー電子)
- (第3回) 映像・音声信号のIPパケットマッピング方式 ∞∞∞ 川本潤一郎氏 (NHK)
- (第4回) IPライブ制作システムの制御技術 ∞∞∞ 小山智史氏 (NHK)
- (第5回) IPライブ制作技術の検証実験事例 ∞∞∞ 北島正司氏・小山智史氏 (NHK)
- (第6回) JPEG軽量映像圧縮技術の解説 ∞∞∞ 石川孝明氏 (早稲田大学)

放送業界におけるIP化の概要

正会員 蝶野慶一†

1. まえがき

放送業界におけるIP (Internet Protocol) 化とは、映像信号専用のSDI (Serial Digital Interface) ケーブルを汎用のEthernetケーブルに置き換える動きです。この動きは、Ethernetの速度向上により、SDI信号に含まれる非圧縮情報をIPパケット化して伝送することが可能になってきたことに由来します(図1)。

近年の映像信号の高品質化や放送機器のIP化の進展は著しく、放送業界においては、これらに関するさまざまな規格やそれらの関係を網羅的に把握することが求められています。

講座の第1回となる本稿では、映像信号のSDIの進化、映像インタフェースのIP化、IPの伝送帯域にてUHD (Ultra High Definition) 映像などを圧縮して伝送する軽量画像コーデック技術について概要を述べます。

2. 映像信号の高品質化とSDIの進化

映像のパラメータは、実現性ととも、視覚特性を考慮したシステム設計について規定されています。HD (High Definition) 映像の次世代のUHD映像は、高解像度化に加えて、広色域化、高ダイナミックレンジといった性能改善

が加わることによって、制作者には映像表現の拡大、視聴者には映像体験の向上をもたらすものです¹⁾。

広色域化に関するITU-R BT.2020²⁾と高ダイナミックレンジに関するITU-R BT.2100³⁾については、本講座の第2回に説明を委ね、本節では、高解像度化に伴うSDIの進化について概要を述べます。

2.1 HD-SDIから3G-SDIへ

SDIは、映像のデジタル信号を高速に伝送するインタフェースとして長らく利用されてきました。HDインタレース映像(1,920×1,080, 59.94i)には約1.5 Gbpsの伝送レートを有するHD-SDI⁴⁾が利用されており、HDインタレース映像の2倍の伝送レートとなるHDプログレッシブ映像(1,920×1,080, 59.94p)には、ケーブル2本で対応するDual-link-SDI⁵⁾やケーブル1本で対応する3G-SDI⁶⁾があります。3G-SDIのストリームは仮想的に二つの10bitの平行データストリームで構成され、Level-AとLevel-Bのマッピング方式があります。4:2:2の場合、Level-A方式では、一方の仮想的インタフェースに輝度、もう片方に色差がマッピングされます。Level-B方式では、Dual-LinkのペイロードのLink-AとLink-Bがそれぞれの仮想的インタフェースにマッピングされます。

2.2 ケーブル4本を用いたUHD機器接続

HDプログレッシブ映像の4倍の空間解像度となる4KのUHD映像(3840×2160, 59.94p)の伝送レートは約12 Gbpsとなるため、3G-SDIを使う場合、4本のケーブルを使って機器を接続します。

3G-SDIケーブル4本を使った接続では、SQD (Square Division) と2SI (Sample Interleave) の二つのマッピング方式があります。SQD方式では、4KのUHD映像を「田」の字に4分割し、それぞれの部分領域を3G-SDIケーブルで伝送します。一方、2SI方式は、映像の2画素サンプルを組みとして動作し、奇数ラインを1本目と2本目の3G-SDIケーブルから交互に伝送し、偶数ラインを3本目と4本目の3G-SDIケーブルから交互に伝送します。最終的に伝送される映像はSQD方式と2SI方式で同じですが、2SI方式は左上から順番に伝送が始まるため、遅延の観点で同方式の方が有利となります。

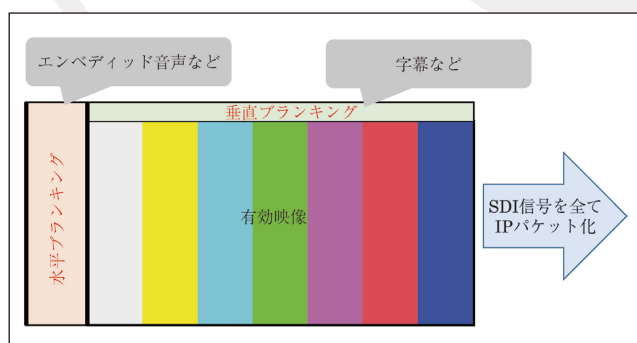


図1 SDI信号に含まれる映像・音声・アンシラリーのIPパケット化

† NEC バイオメトリクス研究所

"IP Adoption in Broadcast Industry (I): Brief Overview of IP Adoption in Broadcast Industry" by Keiichi Chono (Biometrics Research Laboratories, NEC Corp., Kawasaki)

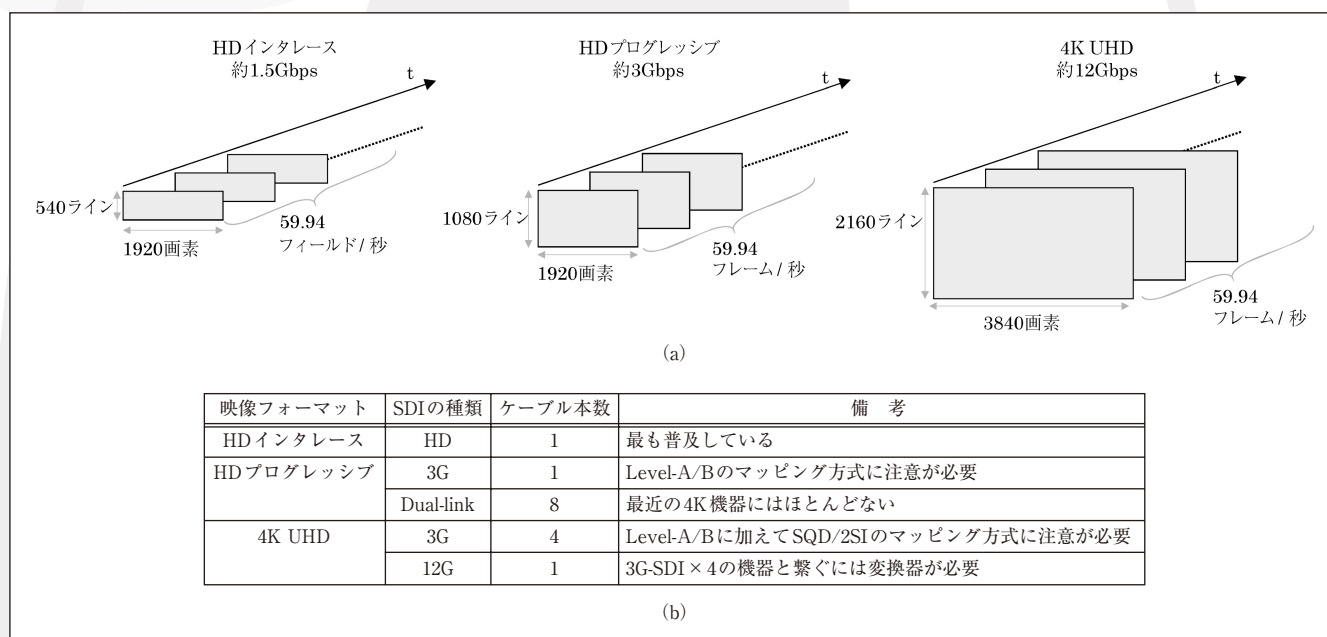


図2 (a) 各映像フォーマットと解像度の違い, (b) 機器接続にて各SDIが必要とするケーブル本数

2.3 12G-SDIを用いたUHD機器接続

4KのUHD映像のために複数のケーブルが要求されることは、特に機器間の接続を頻繁に変える運用では、受け入れがたい問題です。4K映像の登場に送れること数年、この問題を解決すべく従来のSDIの延長として現れたのが、12G-SDI⁷⁾です。しかしながら、12G-SDIの信号は高周波数信号のため、3G-SDIまでのケーブルでは伝送距離を確保できない課題がありました。この課題に対してケーブルメーカー各社から12G-SDI対応ケーブルが発売され、3G-SDIと同様の伝送距離も実現可能となっています⁸⁾。

2.4 SDIの進化と課題

図2は上述したSDI進化のサマリですが、インフラ運用の観点でみると、SDIベースでのシステムの4K対応には次に示す課題が見えます。① 3G-SDI × 4の機器をシステムに繋ぐには、同軸ケーブルの本数が4倍必要、② Level-A/BおよびSQD/2SIなどの複数のマッピング方式の考慮が必要、③ 3G-SDI × 4と12Gの混在への考慮が必要 (SEMTECH社のGearBoxを用いた変換器などが必要)。これらの課題に対して、次節で述べるIP化が一つの解として注目されています。

3. インタフェースのIP化

文献⁹⁾などでも紹介されているように、IP化のメリットは、汎用のEthernetを使って、複数の映像信号の多重化によるケーブル本数の削減や、さまざまな映像フォーマット対応への柔軟性です。

IPのネットワークに用いられるEthernetの速度は、対数オーダで向上しており、100 GbEに対応した機器も普及期を迎えようとしています¹⁰⁾。IPのネットワークはパケットネットワークであり、複数の映像信号を1本のケーブルに

多重して伝送することが容易です。例えば、100 GbEケーブル1本にて、1.5 GbpsのHDインタレース映像信号であれば約60のデータストリームを伝送できます。このようにケーブル本数を少なくできることはIP化のメリットです。

4KのUHD映像に対応した機器への更新期には、HDで制作する番組と4Kで制作する番組が共存することになります。IP化した番組制作システムでは、どのフォーマットの映像信号も伝送できます。このように映像フォーマットを意識しなくてもよいこともIP化のメリットとなります。

詳細については、本講座の第3回および第4回に説明を委ね、以下では、IP化のための要素技術である、機器同期、映像・音声・アンシラリのIPマッピング、伝送エラーに対する保護について概要を述べます。

3.1 機器同期

機器同期は、番組制作システムにおいて欠かせない機能であり、LAN (Local Area Network) 上で高精度に時刻を同期させるIEEE1588のPTP (Precision Time Protocol)¹¹⁾が重要な役割を果たしています。PTPはマスタとスレーブの間で、送信/受信時のタイムスタンプを挿入したPTPメッセージを定期的に交換し、このタイムスタンプを使って時刻を同期制御します。

PTPは汎用的な規格ですが、これだけでは番組制作システムに必要なタイミング情報がわかりません。このため、SMPTE (Society of Motion Picture and Television Engineers) は、SMPTE ST2059-1¹²⁾とSMPTE ST 2059-2¹³⁾にて、PTPタイムスタンプと映像位相信号の関係と番組制作システムに必要な同期精度を満たすためのプロファイルをそれぞれ規定しています。

3.2 IPマッピング

IPマッピングでは、文献¹⁴⁾で解説されているSMPTE

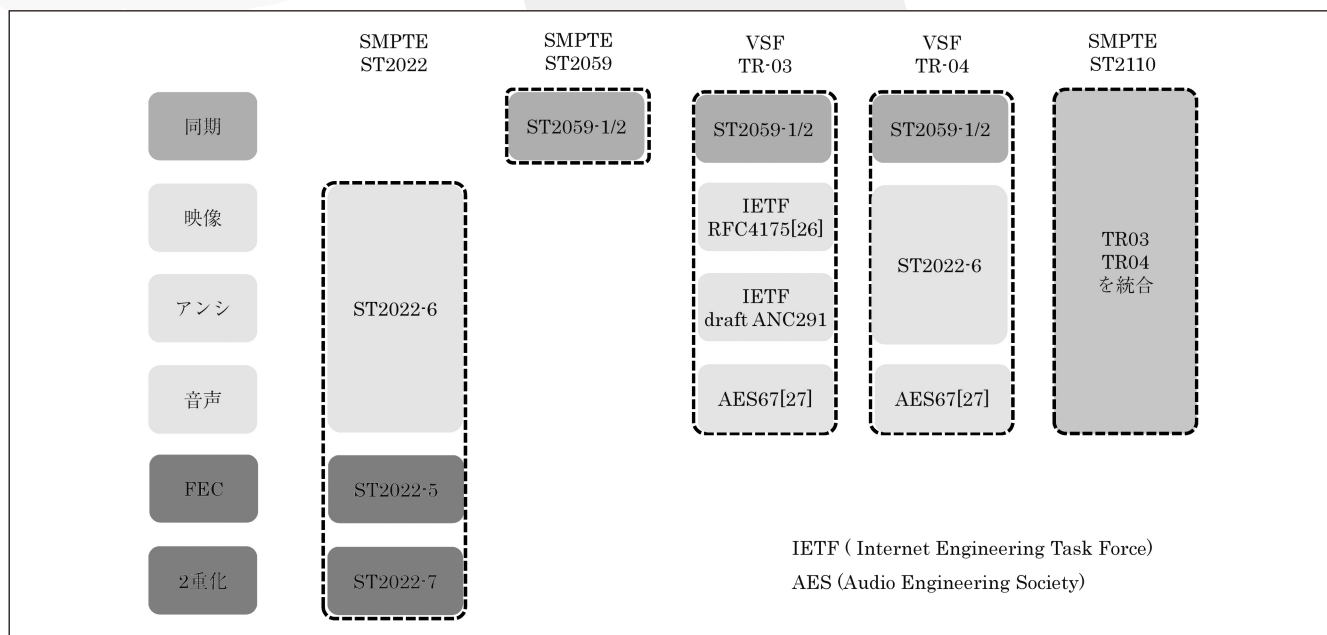


図3 各規格とIP伝送方式の対応図

ST2022シリーズが関係する規格となります。2007年に規格化されたSMPTE ST2022-2¹⁵⁾は、映像信号や音声信号の圧縮データなどで構成されるMPEG-2 TS (Transport Stream) のIPマッピングを規定しています。2012年に規格化されたSMPTE ST2022-6¹⁷⁾は、SDI信号で伝送していた映像、音声、アンシラリのIPマッピングを規定しています。10 GbEや40 GbEに対応した機器が普及期になり始めたことから、非圧縮データを低コストにIPネットワークで伝送可能になり、SMPTE ST2022-6は急速な広がりを見せています。

3.3 伝送エラー保護

リアルタイム性が重要となるライブ番組制作などにおいては、パケットを送りっぱなしのUDP (User Datagram Protocol) を利用するため、パケットロスが発生します。その保護技術としては、FEC (Forward Error Correction) や2重化があります。

FECは、パケットのパリティを送信して一定範囲のエラーを受信側で回復する技術です。FECに関連する規格には、MPEG-2 TSの圧縮データのIPパケットに対応するSMPTE ST2022-1¹⁵⁾や、SDIの非圧縮データのIPパケットに対応するSMPTE ST2022-5¹⁷⁾があります。

2重化は、送信側と受信側の間に冗長に二つの伝送経路を確保し、エラーからコンテンツを保護する技術です。2重化に関連する規格がSMPTE ST2022-7¹⁹⁾であり、冗長化伝送経路で流される同一コンテンツのストリームをシームレスに切替えるための受信器の要件が規定されています。

3.4 さまざまなIP伝送方式

SDIケーブルと同様に、上述したSMPTE ST2022-6でも、一つのIPパケットに映像、音声、アンシラリを含めて伝送できます。しかしながら、映像、音声、アンシラリを別々

のIPパケットにできる方が好ましいユースケースもあります。さまざまなユースケースに応じて、NMI, ASPEN, AIMSなど、IPマッピングと同期の考え方が異なる複数のIP伝送方式があります⁸⁾。

複数のIP伝送方式が混在するとデータ交換や相互接続が容易ではなくなるため、VSF (Video Services Forum) が開発した映像、音声、アンシラリの独立した取り扱いが可能なIPマッピング方式の勧告TR-03²⁰⁾やTR-04²¹⁾を統合する、共通に利用できるIP伝送方式として、SMPTE ST2110シリーズ^{22)～25)}が策定されています(図3)。

4. 軽量画像コーデック

番組制作システムなどに用いられる画像圧縮では、圧縮効率よりも、低レイテンシや視覚的に非圧縮と同等の画質(ビジュアルロスレス)が要件となります。加えて、IPネットワーク機器に組み込まれることが前提となるため、圧縮処理が軽量であることも重要視されます。

これらの要件を満たす軽量画像コーデックとしては、intoPIX社製TICOやソニー社製LLVCがあります。これらのビットストリーム、デコード処理、TICOビットストリームのIPマッピングなどは、SMPTEのRDD (Registered Disclosure Doc) で登録されています^{28) 29)}。

SMPTEとは別の標準化団体であるISO/IEC SC29 WG1 (通称JPEG) でも、これらの要件を満たすJPEG XS (ISO/IEC 21122)³⁰⁾の標準化作業を進めています。JPEG XSは、ビットストリームやデコード処理を定めるPart1、プロファイルなどを定めるPart2、配信や蓄積におけるフォーマットを定めるPart3、相互接続性を確保するためのコンフォーマンスを定めるPart4、リファレンスソフトウェアのPart5で構成されます。現時点ではPart1のIS化

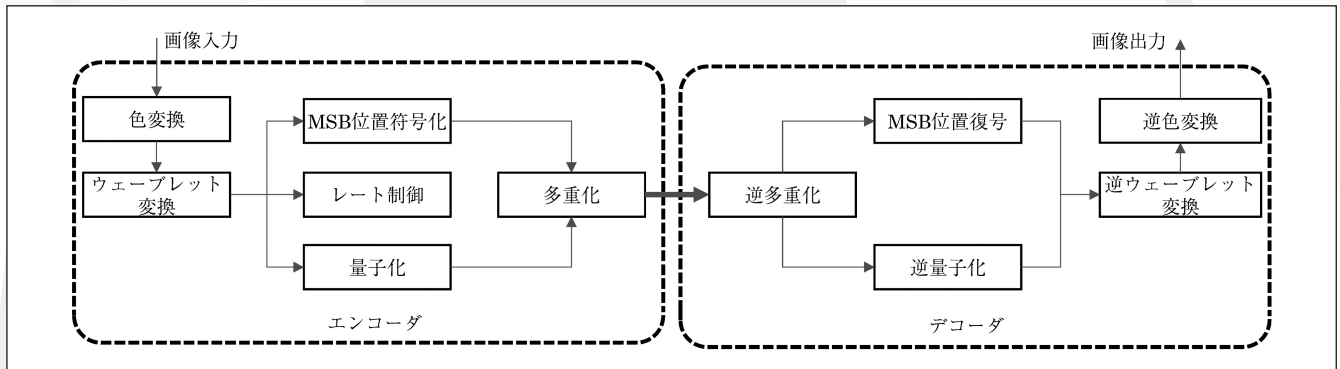


図4 JPEG-XSのエンコーダおよびデコーダのブロック図

までが進んでおり、2020年1月にすべてのPartが完成する予定となっています。

JPEG XSの基本となるJPEG2000やビジュアルロスレスの画質評価などの詳細については本講座の第6回に説明を委ね、図4を参照して簡単に概要を述べます。JPEG XSは、JPEG2000と同様にウェーブレット変換によるサブバンド分割に基づいた方式です。JPEG XSでは、低レイテンシと軽量処理の要件を満たすため、垂直方向のサブバンド分割を2段に制約します。さらに、エントロピ符号化処理を軽量にするため、係数ごとではなく、固定ビット数で表現された複数の係数をグループ化して処理するMSB (Most Significant Bit) 位置符号化を採用しています³¹⁾。JPEG2000よりも処理が簡略化されていますが、典型的な圧縮率としては1/2から1/6を達成します。

5. むすび

本稿では、講座「放送業界におけるIP化」の全体概要として、SDIの進化、映像伝送インタフェースのIP化、軽量画像コーデック技術について概要を述べました。

(2018年12月7日受付)

〔文 献〕

- 1) 西田：「UHD映像の概要」，映像学誌，71，2，pp.216-222（2017）
- 2) ITU-R Recommendation BT.2020-2: "Parameter values for ultra-high definition television systems for production and international programme exchange"（2015）
- 3) ITU-R Recommendation BT.2100-0: "Image parameter values for high dynamic range television for use in production and international programme exchange"（2016）
- 4) SMPTE ST 292-1: "1.5Gbps Signal/Data Serial Interface"
- 5) SMPTE ST 372: "Dual Link 1.5Gbps Digital Interface for 1920 × 1080 and 2048 × 1080 Picture Formats"
- 6) SMPTE ST 424: "3Gbps Signal/Data Serial Interface"
- 7) SMPTE 2082-1: "12Gbps Signal/Data Serial Interface - Electrical"
- 8) 和田：「4K伝送のための12G-SDIとVideo over IP」，映像学誌，72，2，pp.251-254（2018）
- 9) 児島：「IPライブ制作」，映像学誌，70，5，pp.783-787（2016）
- 10) <http://ethernetalliance.org/wp-content/uploads/2015/03/Ethernet-Roadmap-2sides-29Feb.pdf>
- 11) IEEE 1588-2008: "IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems"

- 12) SMPTE ST 2059-1:2015: "Generation and Alignment of Interface Signals to the SMPTE Epoch"
- 13) SMPTE ST 2059-2:2015: SMPTE Profile for Use of IEEE-1588 PTP in Professional Broadcast Applications"
- 14) 川本：「SMPTE ST 2022」，映像学誌，72，2，pp.247-250（2018）
- 15) SMPTE ST 2022-1: "Forward Error Correction for Real-Time Video/Audio Transport Over IP Networks"
- 16) SMPTE ST 2022-2: "Unidirectional Transport of Constant Bit Rate MPEG-2 Transport Streams on IP Networks"
- 17) SMPTE ST 2022-5: "Forward Error Correction for High Bit Rate Media Transport Over IP Networks"
- 18) SMPTE ST 2022-6: "Transport of High Bit Rate Media Signals over IP Networks (HBRMT)"
- 19) SMPTE ST 2022-7:2013: "Seamless Protection Switching of SMPTE ST 2022 IP Datagrams"
- 20) VSF TR-03: "Transport of Uncompressed Elementary Stream Media over IP"
- 21) VSF TR-04: "Utilization of ST-2022-6 Media Flows within a VSF TR-03 Environment"
- 22) SMPTE ST 2110-10: "Professional Media Over Managed IP Networks: System Timing and Definitions"
- 23) SMPTE ST 2110-20: "Professional Media Over Managed IP Networks: Uncompressed Active Video"
- 24) SMPTE ST 2110-30: "Professional Media Over Managed IP Networks: PCM Digital Audio"
- 25) SMPTE ST 2110-40: "Professional Media Over Managed IP Networks: SMPTE ST 291-1 Ancillary Data"
- 26) IETF RFC 4175: "RTP Payload Format for Uncompressed Video"
- 27) AES AES67: "AES standard for audio applications of networks - High-performance streaming audio-over-IP interoperability"
- 28) SMPTE RDD 34:2015: "Low Latency Video Codec for Network Transfer"
- 29) SMPTE RDD 35:2016: "TICO Lightweight Codec Used in IP Networked or in SDI Infrastructures"
- 30) <https://jpeg.org/jpegxs/>
- 31) T. Richter et al.: "Entropy Coding and Entropy Coding Improvements of JPEG XS", in Proc. of 2018 IEEE Data Compression Conference



蝶野 慶一（ちょうの けいいち）1999年，九州大学工学部電気系学科卒業。2001年，同大学大学院システム情報科学研究科修士課程修了。同年，NEC入社。2007年～2008年，スタンフォード大学客員研究員。入社より，映像符号化技術の研究，国際標準化，および製品化に従事。2014年，工業標準化事業表彰，国際標準化奨励者表彰，産業技術環境局長表彰など。現在，NECバイオメトリクス研究所主任研究員，正会員。

UHDやHDRから見たSDI信号の進化

矢野浩二[†], 鈴木則行[†]

1. まえがき

SDI信号は基本的なルールを変えることなく20年以上進化を続けているデジタルインタフェースです。SDTVの伝送に始まりHDTV、最近では4K、8Kの伝送もできるようになりました。SDI信号はこれまで放送業界で広く使用されてきた実績とさまざまなノウハウの共有により、強力なインタオペラビリティを確保し安定したシステム構築を実現しています。

SDI信号が進化してきた過程で変えられることなく守られてきた基本的ルールは、映像レベルの範囲と同期ワードTRS (Timing Reference Signal) です。このルールに基づいてSDI信号はビットレートを増やし、伝送できる映像情報を増やしてきました(図1)。HD-SDI信号による1080/59.94Iの伝送から3G-SDI信号による1080/59.94Pの伝送、そして12G-SDI信号によるUHD (4K) の伝送、12G-SDI信号を4本使用することでUHD (8K) の伝送が可能になりました。

今回は、このようなSDI信号の進化をUHD (Ultra High Definition) やHDR (High Dynamic Range), WCG (Wide Color Gamut) といった新技術の観点から見ていこうと思います。

まず、今後放送業界にIPの技術が浸透していくと思いますが、IPで伝送される映像、音声、補助データはSDI信号で伝送されている各データを念頭に規格化が進んでいます。今後IPを勉強していくうえでもこれらSDI信号を理解していることでIP規格の理解が早まります。

最初にSDI信号の進化をSMPTE規格、ITU-R規格の体系図で確認します。図2の左側に示すイメージフォーマットはHD (1,920 × 1,080), 4K (3,840 × 2,160), 8K (7,680 × 4,320) と下に行くにしたがって進化しています。図2の中央に示すように1つのイメージフォーマットの伝送に複数のインタフェースフォーマットが規定されている場合があります。

すべてのSDI信号で共通に使用されている規格は、図2の右側に示すようにSMPTE ST 291で規定されている補助データです。補助データは伝送する情報によって音声データ、タイムコード、ペイロードIDなどそれぞれ規格化されています。

UHDとともに注目されているHDRやWCGは、イメージフォーマットを拡張する形で主にITU-Rで規格化されています。

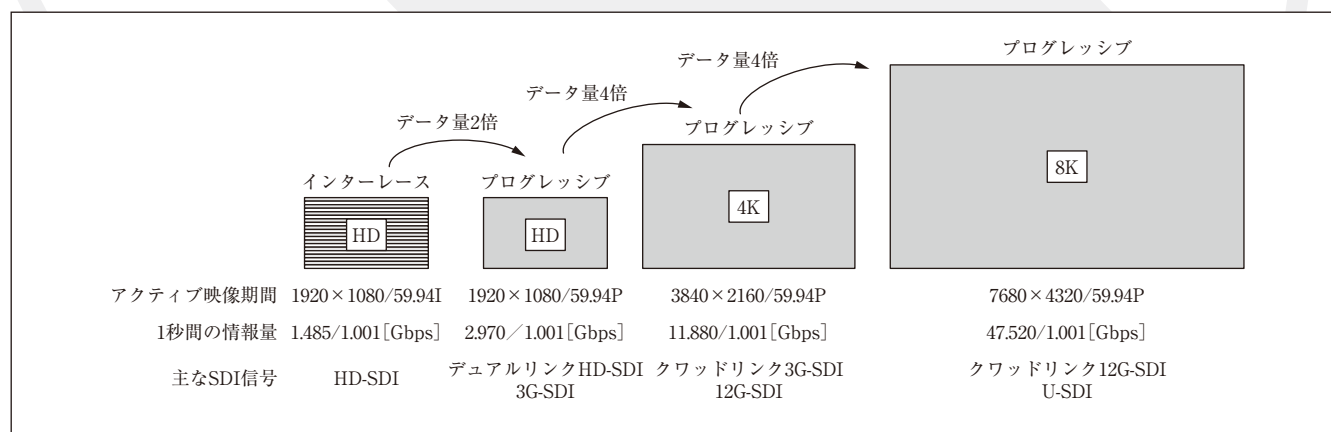


図1 SDI信号と伝送される情報量

[†] リーダー電子株式会社

"IP Adoption in Broadcast Industry (2): Transition of SDI Signal from the Viewpoint of UHD and HDR" by Koji Yano and Noriyuki Suzuki (Development Department, Leader Electronic Corporation, Yokohama)

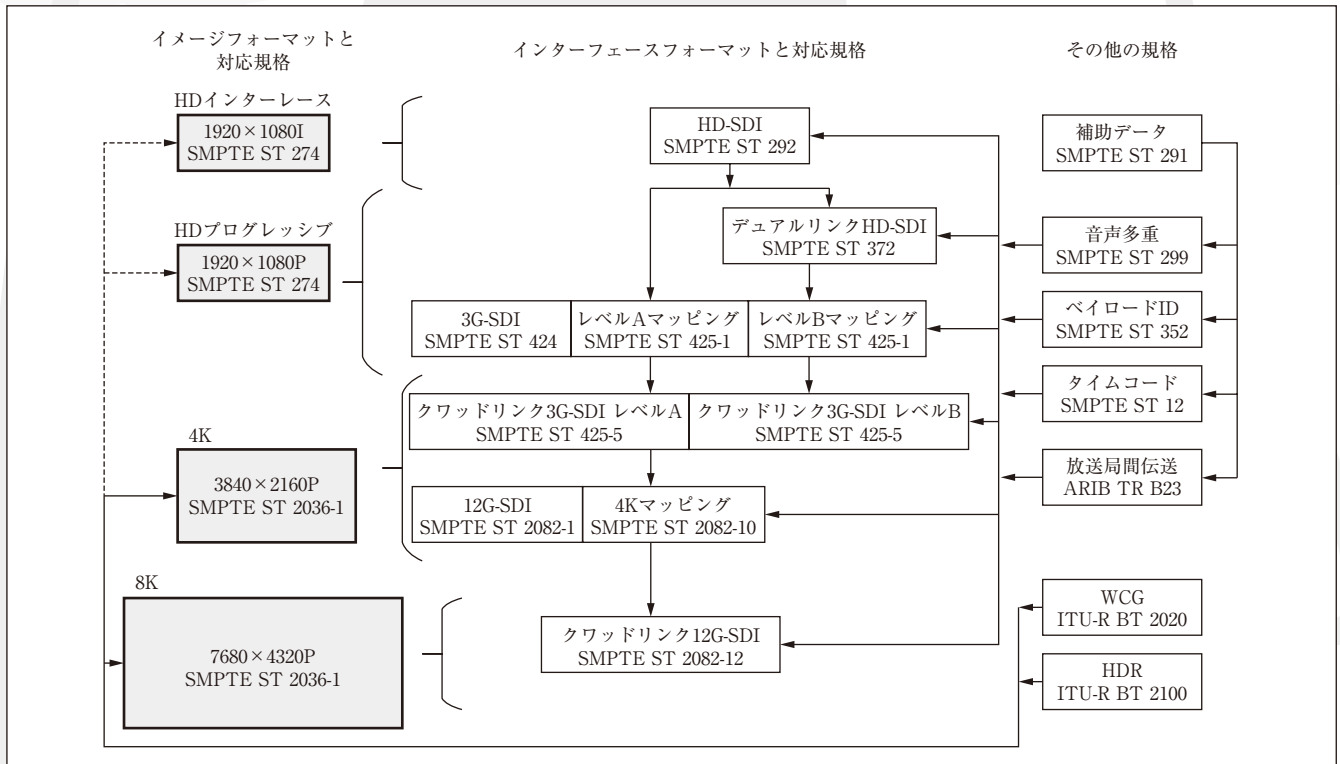


図2 SDI信号の規格体系図

2. HD-SDI信号の構造と特徴

4K、8KのSDI伝送を説明する前にHD-SDI信号をおさらいします。HD-SDI信号で伝送できるイメージフォーマットはSMPTE ST 274で規定されており、インターフェースフォーマットはSMPTE ST 292で規定されています。以下に1920×1080/59.94Iの伝送を例にHD-SDI信号の構造と特徴について述べます。

HD-SDI信号の1映像フレームは図3に示すように第1および第2の2フィールドからなるインターレース構造で1フレーム1125ラインで構成されます。1フレームのアクティブ映像期間のライン番号は第1フィールドで21～560、第2フィールドで584～1123です。第1、第2の両フィールドで1080ラインのアクティブ映像期間が構成されます。アクティブ映像期間以外のラインは垂直ブランキング期間とよびます。

HD-SDI信号の1水平ラインの合計ワード数はY系列、C系列それぞれ2200ワードです。そのうち映像データをマッピングするワードは水平アクティブ映像期間と呼び、それ以外のワードは水平ブランキング期間と呼びます。アクティブ映像期間はY系列、C系列ともに1920ワードで、Y系列には1920サンプル分の輝度データがマッピングされ、C系列には960サンプル分の色差データ（ C_B および C_R ）がマッピングされます（図4）。

水平ブランキング期間にはTRSとよばれる2種類の同期ワードが挿入され、水平アクティブ映像期間の始まりを示

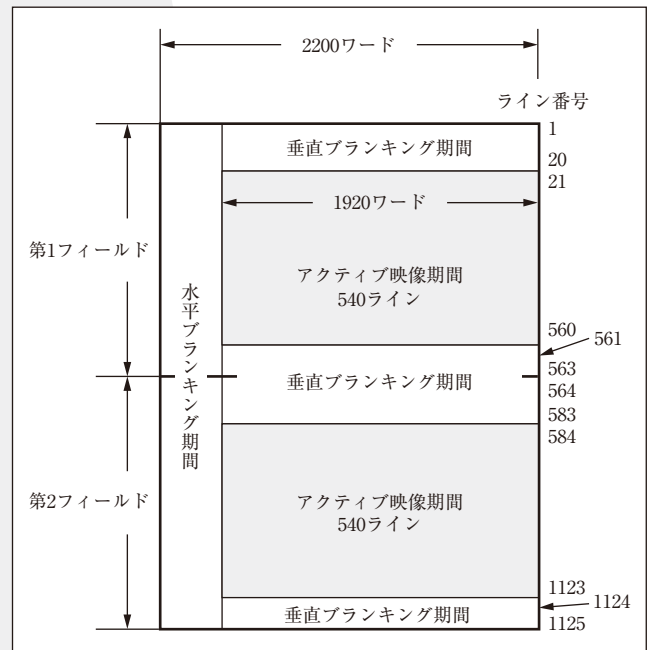


図3 1920×1080/59.94Iのフレーム構造

すSAV (Start of Active Video) とアクティブ映像期間の終わりを示すEAV (End of Active Video) があります。EAVのあとにはライン番号やエラー検出を行うCRCC (Cyclic Redundancy Check Codes) がY系列とC系列それぞれに挿入されます。

これら水平ブランキング期間に挿入されるTRSやライン番号、CRCCに関するルールは、3G-SDIや12G-SDIになっ

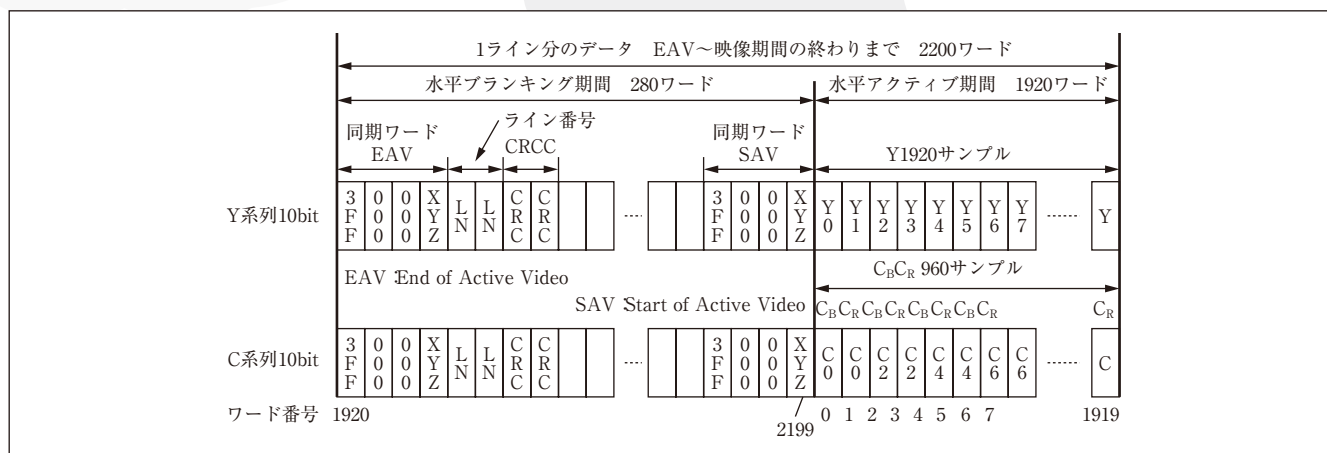


図4 HD-SDI信号の平行ビットデータ

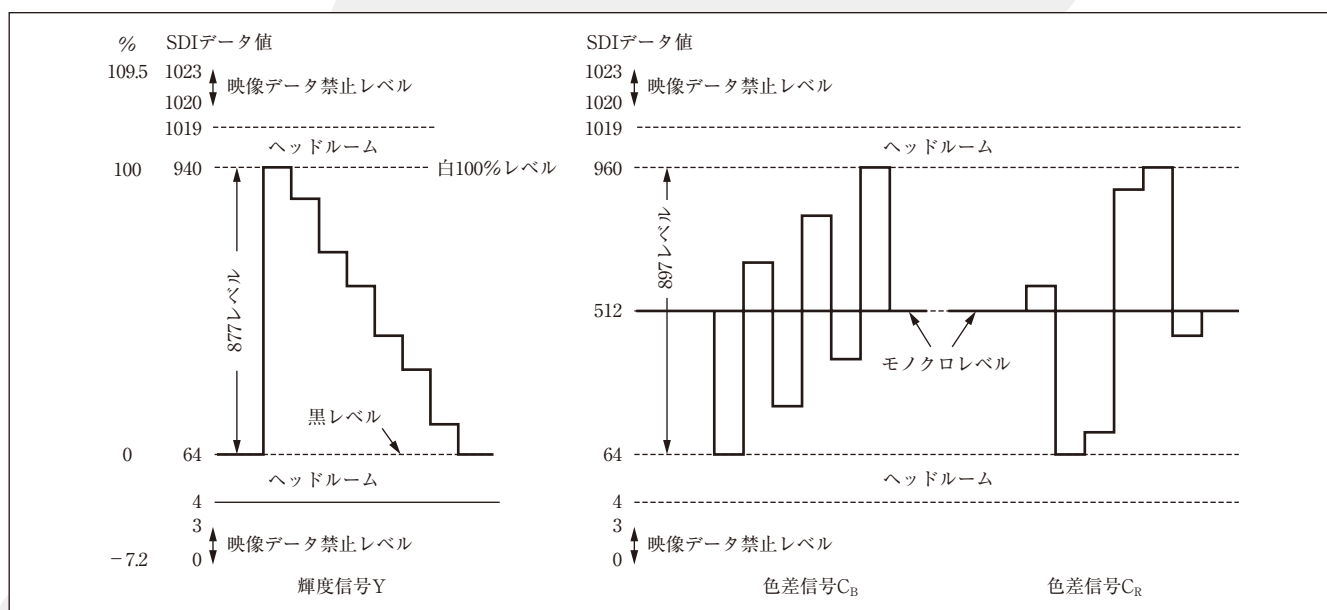


図5 HD-SDI信号の映像レベル

でも基本的には変わっていません。

図4に示すように、1水平ラインの開始点はEAV（ワード番号1920）で終了点はアクティブ映像の最終ワード（ワード番号1919）です*1。ライン番号はEAVの直後に2ワードで挿入され、ライン番号は1から1125になります。HD-SDIではこのSDI信号に挿入されている伝送ライン番号と映像ライン番号が一致しています。

HD-SDI信号のビット深度はY系列10bit、C系列10bitですが、図5に示すように映像レベルとして使用できる範囲は、輝度信号で64から940、色差信号で64から960となっています*2。この映像レベルの取り扱いは3G-SDI、12G-SDIになっても変更されていませんし、HDRやWCGも原則このレベルの範囲でマッピングされています。

映像レベルとして使用できない範囲のうち特に0から3と1020から1023はTRSや補助データのADFに割り当てられており、TRSやADF以外での使用は同期エラーを発生させる原因になるため厳しく制限されています。一方4から63までと941から1019まで（色差信号は961から1019まで）の値は、信号処理の過程で時折超えてしまうことが許容されている映像レベルでヘッドルームと呼ぶこともあります。ヘッドルームは映像レベルとして利用すると黒つぶれ、白飛び、色飽和の原因になりますのでできる限り避けるべきです。

(1) 補助データの多重

SDI信号は映像データの伝送以外にも、音声データ、タイムコード、字幕、パイロードID等の補助データを水平および垂直ブランキング期間に多重して伝送することができます。補助データは図6に示すようにパケット構造をしており、単独パケットとして伝送できるほか複数のパケットをつなぎ合わせて一連のデータとして伝送することもでき

*1 ワード番号0はアクティブ映像の最初のワード。

*2 リミテッドレンジの場合、ITU-R BT.2100には他にもフルレンジの記載があり、取り得る値の範囲はY<sub>C_BC_Rともに0から1023（ただし、伝送時には4から1020にクリップする）です。

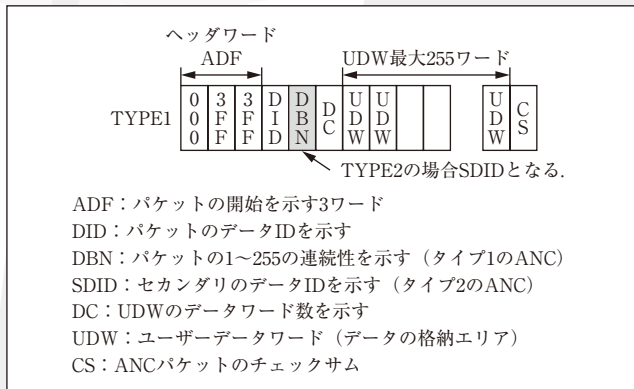


図6 補助データパケットの構造

ます。この補助データはADFの後に続くDIDの値によってデータの種類の識別でき、その値から大きく2つの形式に分類することができます。データの連続性をDBN (Data Block Number) によって確認できるタイプ1とデータの種類のSDIDによって細分化できるタイプ2があります。

SDI信号では音声データを多重する際、補助データとして多重し、C系列の水平ブランキング期間を独占的に使用します。補助データの形式はタイプ1を使用しており、DBNによって音声データパケットの連続性を確認できるようにしています*3。

音声データ以外の音声制御パケットやタイムコード、字幕、ペイロードID等の補助データはY系列に重畳することができます。

なお、補助データはSMPTE ST 291で規定されています。データの識別ができるDIDの値はSMPTEにて管理されており、DIDの一覧はSMPTEのWebサイト<https://smpte-ra.org/>から確認することができます。

(2) SDI信号における映像と音声の同期

音声は一定の周波数（通常48kHz）でサンプリングされデジタル音声データになります。しかし音声データは前述の通りSDI信号に補助データとして水平ブランキング期間に多重されますので、結果的にサンプリング周期通りに多重されず水平ブランキング期間に偏って多重されます。音声として再生を行うためには、補助データとして多重された音声データをSDI信号から分離して一度音声データのみをバッファに格納します。バッファに格納した音声データをサンプリング周波数で読み出すことで連続した音声データが再現できます。

このとき音声データの読み出しに使用するサンプリング周波数は、音声データパケット中の音声クロック位相情報を用いることで再生できます。この音声クロック位相情報が映像に同期していれば同期音声として、非同期であれば非同期音声として再生することが可能です。ただし、業務用映像機器においてはリップシンクの観点からも映像に同

期した音声を扱いますし、非同期音声を受け付けない機器もありますので、非同期音声の混入は注意が必要です。

1080/59.94Iの映像信号と音声データのサンプリング周波数48kHzが同期している場合、映像信号5フレームの周期と音声データ8008サンプルの周期が一致します。このとき映像信号5フレームをオーディオフレームシーケンスと呼んでいます。波形モニタにはオーディオフレームシーケンス内の音声データのサンプル数をカウントすることで同期音声のチェックを行っているモデルもあります。

音声クロック位相情報を使用せず、映像のサンプリング周波数から音声のサンプリング周波数を再生する機器もあります。このような機器は映像に非同期な音声を受け付けませんので、非同期音声を入力した際に定期的なノイズが発生してしまいます。規格上は非同期音声の取り扱いも許容されていますし非同期音声の検出は難しいので、音声データの同期/非同期は厳重に管理する必要があります。

(3) チェックフィールドパターン

SDI信号受信機の受信性能にストレスを与えるチェックフィールドパターンはSMPTE RP-198で規定されています。カラーバー信号で受信エラーが発生しなくとも、チェックフィールドパターンで受信エラーが発生するというケースがあります。チェックフィールドパターンは意図的に生成しない限り通常は発生しないストレスパターンですが、チェックフィールドパターンを使用することでシステムの受信エラーに対するマージンを確認することができます。

チェックフィールドパターンはHD-SDI、3G-SDI、12G-SDIでそれぞれパターンのデータが異なりますので、例えばHD-SDIのチェックフィールドパターンをアップコンバートして3G-SDIとしてシステムマージンチェックには使用できません。ダウンコンバートした場合も同様ですし、ゲインやオフセットなどデータを加工する信号処理を施してしまうとストレスパターンとしての効果がなくなってしまいます。チェックフィールドパターンでシステムのマージンチェックを実施する際は、専用の信号発生器を使用して、途中経路のデータ補正やデータ加工機能を停止した状態で行う必要があります。

3. 3G-SDI信号の構造と特徴

3G-SDI信号のシリアル伝送構造はSMPTE ST 424で規定されており、映像データのマッピング構造はSMPTE ST 425で規定されています。3G-SDI信号の特徴は映像データのマッピングの違いでレベルAとレベルBの2つの方式があることです*4。レベルAとレベルBでは映像デー

*3 機材によってはDBNを0固定で多重している場合もあるようです。この場合データの連続性を確認することができません。

*4 レベルBにはデュアルリンクHD-SDIをマッピングしたレベルB-DL (Dual Link) とHD-SDIを同時に2本マッピングすることができるレベルB-DS (Dual Stream) があります。本項ではレベルB-DLをレベルBとして説明しています。

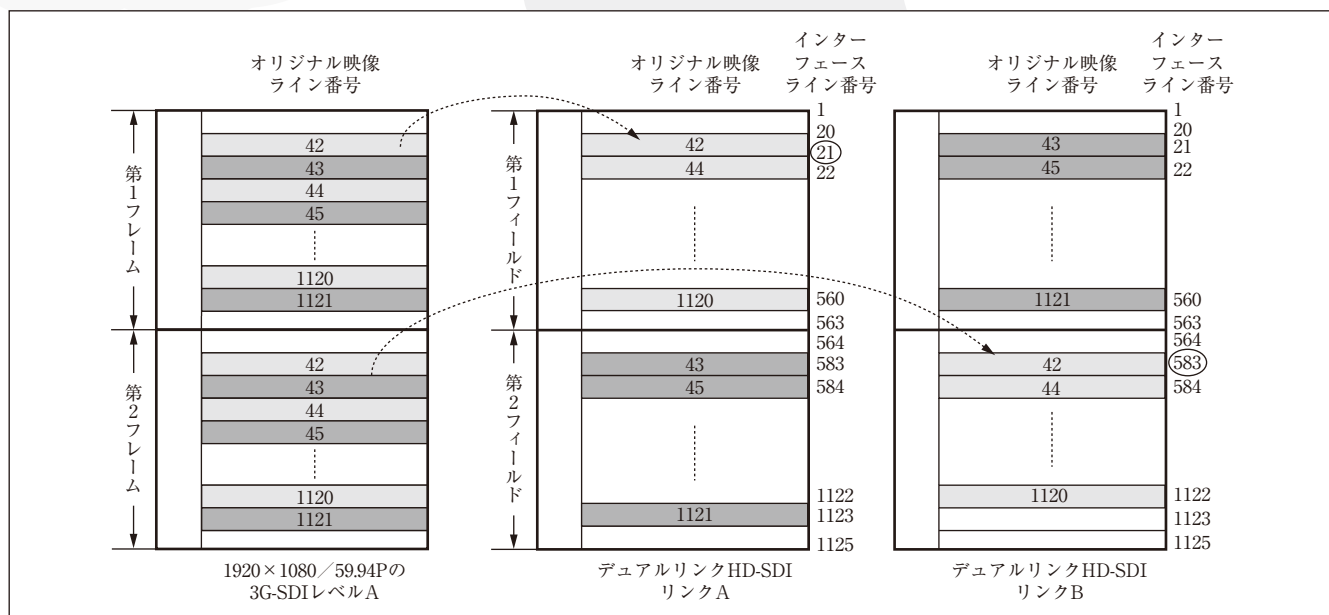


図8 1920×1080/59.94Pの3G-SDIレベルBマッピング

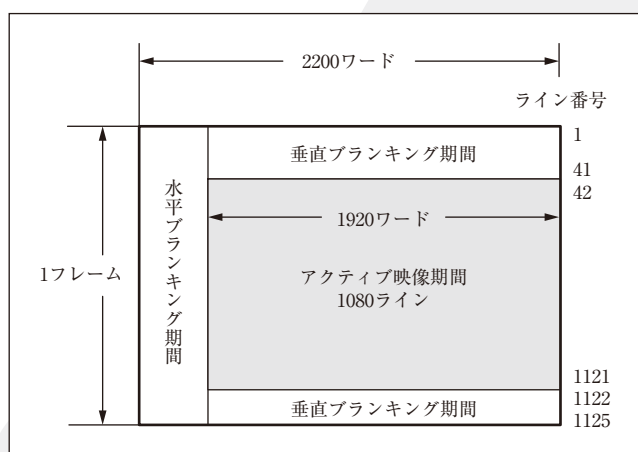


図7 1920×1080/59.94Pの3G-SDIレベルAフレーム構造

タのマッピング方式が異なりますがビットレートは同一なので、レベルA、レベルBどちらにも同じ映像フォーマットをマッピングすることができます。

図7に示す1920×1080/59.94Pの映像フォーマットを例にレベルA、レベルBそれぞれのマッピング方式について以下に述べます。

(1) レベルAの構造

3G-SDIレベルAの平行ビット構造は図4で示したHD-SDI信号と同じです。HD-SDIの各ワードの伝送レートを2倍にすることで水平周波数、フレーム周波数ともに2倍になり1920×1080/59.94Pの伝送を行うことができます。3G-SDIレベルAではY系列をデータストリーム1、C系列をデータストリーム2とも呼びます。

(2) レベルBの構造

3G-SDIレベルBは図8に示すように3G-SDIレベルAのデータのうちアクティブ映像期間のみを一度デュアルリンク

クHD-SDI*5にマッピングします。このデュアルリンクHD-SDIを再び1本に束ねた形をしているのが3G-SDIレベルBの構造です。

レベルBの特徴は、ストリーム1とストリーム2それぞれがHD-SDI信号の構造をしています。ストリーム1にデュアルリンクHD-SDIのリンクAをストリーム2にデュアルリンクHD-SDIのリンクBをマッピングしています。

デュアルリンクHD-SDIは図9に示すように1920×1080/59.94Pのアクティブ映像期間の最初のラインである42ラインの映像データは、連続する2フレームのうち最初のフレームはリンクAのインタフェースライン番号である21ラインにマッピングされ、次のフレームはリンクBのインタフェースライン番号である583ラインに分かれてマッピングされます。3G-SDIレベルBでは元の映像である1920×1080/59.94Pのオリジナル映像ライン番号と3G-SDI上でTRSに続く2ワードのインタフェースライン番号が異なります。

なお、オーディオ制御バケットやタイムコード、下記に示すペイロードIDなどの補助データの多重は、インタフェースライン番号に基づいて行われます。

4. ペイロードID

HD-SDIは1ラインのサンプル数や1フレームのライン数を計測することで映像フォーマットを自動で識別することができました。

しかし3G-SDIになってデータレートが倍になったことにより、YCbCr 4:2:2の映像信号だけでなくRGB 4:4:4を伝

*5 デュアルリンクHD-SDI信号は、HD-SDI信号で1920×1080/59.94Pを1フレーム伝送する時間に、リンクA、リンクBの2本を合わせて1920×1080/59.94P 2フレーム分の映像をマッピングできるものです。

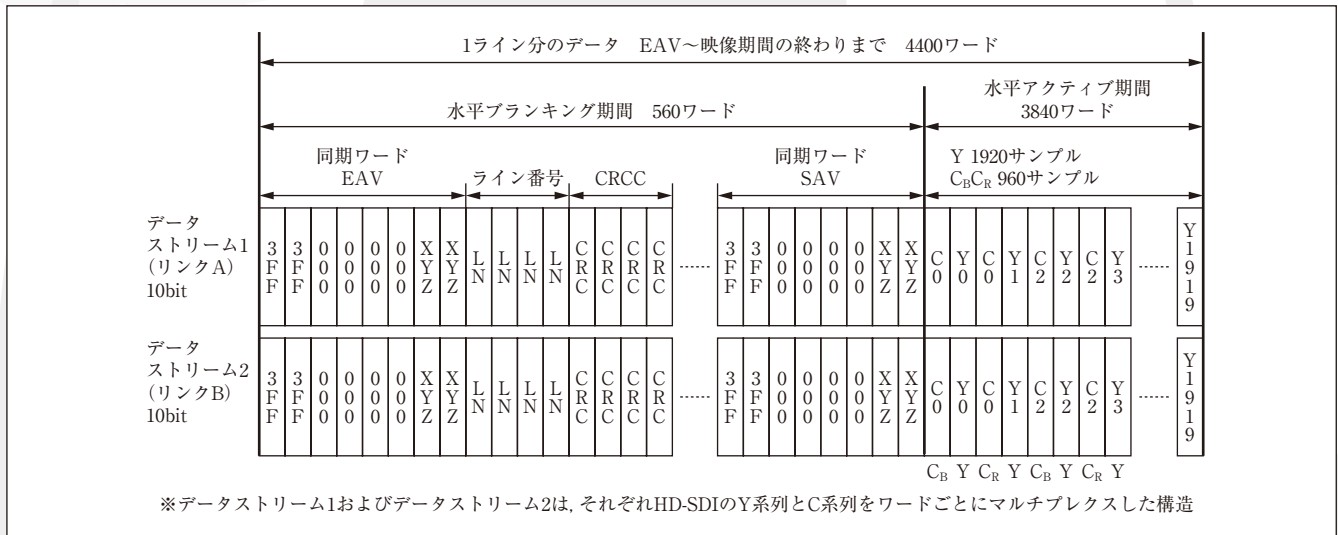


図9 3G-SDIレベルBのパラレルビットデータ

送することができるようになりましたし、量子化精度も10bitだけでなく12bitも伝送できるようになりました。その結果1ラインのサンプル数や1フレームのライン数を求めただけでは映像フォーマットの識別ができなくなったため、ペイロードIDという補助データを使用して映像フォーマットを自動識別できるように考案されました。

ペイロードIDは1フレームのライン数、1ラインのワード数、フレームレート、色域、ダイナミックレンジ等の識別が可能な4バイトデータです。SMPTE ST 352にその基本構造が規定されています。

しかしSMPTE ST 352では3G-SDI信号や12G-SDI信号の記載はなく、それぞれのインタフェースを規定しているSMPTE ST 425やSMPTE ST 2082に記載されています。またITU-Rにも同様にペイロードIDの規定がありBT.1614で基本構造を規定しています。HD-SDI信号、3G-SDI信号における各bitの規定はBT.1120に記載されています。

ここで注意すべき点は、SMPTE ST 425では3G-SDIのペイロードIDの多重は必須となっており、レベルAについてはデータストリーム1(Y系列)とデータストリーム2(C系列)両方に多重することになっています。これに対して、ITU-R BT.1120ではペイロードIDの多重はオプション扱いであり、多重する際はデータストリーム1(Yチャネル)のみに多重することになっています。

参照する規格によって3G-SDIのレベルAにおけるペイロードIDの扱いが異なっています。特にITU-Rのみに準拠している機材の場合、HD-SDIの原則同様C系列には音声パケットのみが多重される前提で設計されていることも考えられ、C系列のペイロードIDの多重を妨げてしまうこともあります。

ペイロードIDのバイト1の値は各インタフェースを識別する値となっており、その値は<https://smpte-ra.org/>で確認することができます。

5. UHD映像とSDIへのマッピング

SMPTE ST 2036-1で規定されているUHDTV1 4K映像3840×2160/59.94Pを例にSMPTE ST 425-5の3G-SDIレベルAクワッドリンクへのマッピングについて説明します。3G-SDIレベルBへのマッピングももちろん可能ですが、12G-SDI信号では、内部構造が3G-SDIレベルAクワッドリンクのみの規定ですので、3G-SDIレベルBへのマッピング方法は省略します。

図10に示すように4K映像にはブランキング期間は付加されていなく、有効映像期間のみ定義されています。4K映像の水平サンプル番号は0から3839、ライン番号は0から2159まで割り振られており、ライン番号も0から数えることが特徴です。4K映像の場合はライン番号というより垂直サンプル番号とよぶほうが適切かもしれません。

4K映像をその形のままSDI伝送する規定はなく、一度4つのHD映像に分割します。この4分割されたHD映像はサブイメージと呼びます。4K映像をサブイメージに分割する方法は2通りあります。1つは4K映像を上下左右に4分割するスクエアディビジョン方式、もう1つは4K映像を2画素ごとに抜き出して4分割する2サンプルインターリーブ方式です。ここでは、図10に示すように12G-SDI信号でも採用されている2サンプルインターリーブについて説明します。

まず4K映像の垂直サンプル番号0のラインについて、水平サンプル番号0から3839を2画素ごとに抜き出してそれぞれ1920個の画素列を2つ作ります。次に垂直サンプル番号1についても同様に2画素ごとに抜き出して1920個の画素列を2つ作ります。これを垂直サンプル番号2159まで繰り返します。偶数垂直サンプル番号と奇数垂直サンプル番号を集めて4つの1920×1080画素のサブイメージが作られます。

この分割された1920×1080/59.94Pのサブイメージを3G-

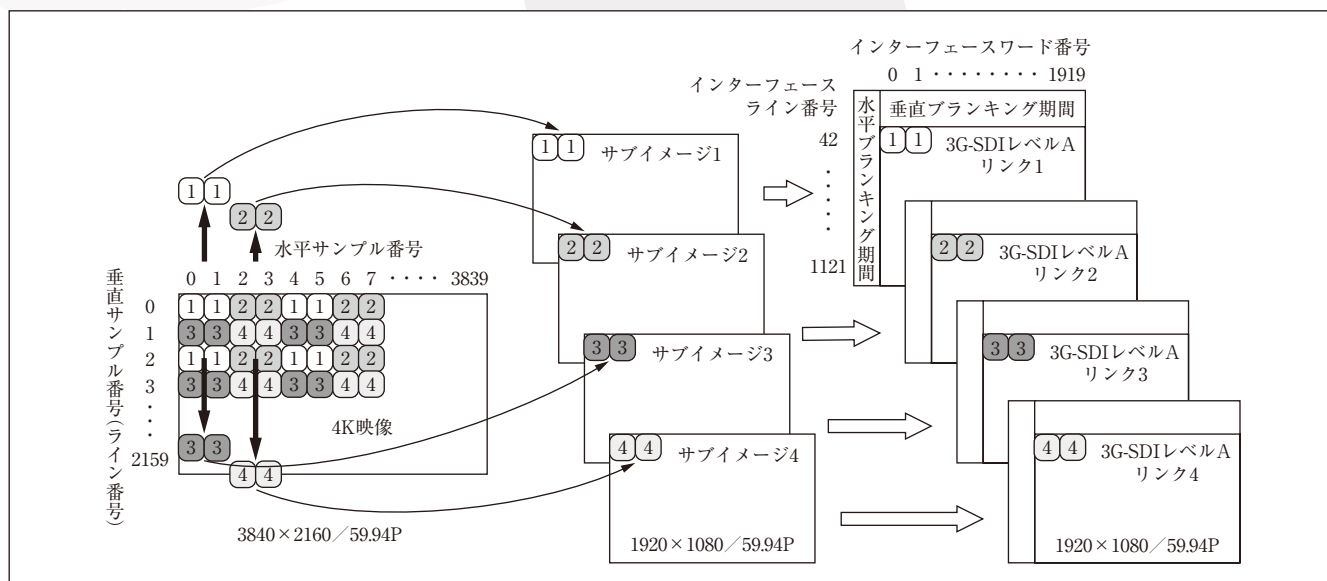


図10 4K映像と3G-SDIレベルAクワッドリンクへのマッピング

SDIレベルAのアクティブ映像期間にマッピングしていきます。音声データやタイムコード、4K字幕などの補助データはこの段階でできた3G-SDI信号のプランキング期間に多重されます。もちろん補助データの多重ラインはインタフェース上のライン番号である3G-SDI信号のライン番号に則って多重します。

波形モニタにSDI信号のデータを直接確認できるデータダンプ機能がありますが、このデータダンプ機能にも4K映像の水平サンプル番号と垂直サンプル番号でデータを確認する方法と3G-SDIレベルAの各リンク毎のインタフェース上のワード番号とライン番号でデータを確認する機能があります。映像データを確認するのか補助データなどインタフェース上のデータとして確認するのかによって使い分けができます。

6. 3G-SDIクワッドリンクの12G-SDIマッピング

4K映像3840×2160/59.94Pを1本の同軸ケーブルで伝送できるインタフェースとして12G-SDI信号があり、SMPTE ST 2082で規定されています。12G-SDI信号の内部構造は上記3G-SDIレベルAクワッドリンクを1本にまとめた形になっており、4K映像をやはり4つのサブイメージに分割して伝送します。

3G-SDIレベルAクワッドリンクの12G-SDI信号へのマッピング方法については図11に示します。3G-SDIレベルAの構造で説明した通りY系列であるデータストリーム1とC系列であるデータストリーム2がありますので、クワッドリンクで合計8つのデータストリームとして考えます。

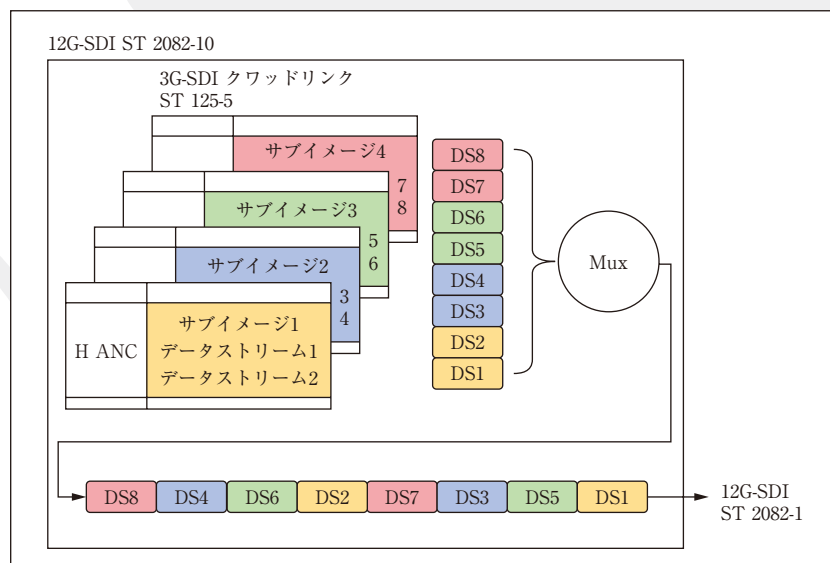


図11 3G-SDIクワッドリンクの12G-SDIマッピング

これらの8つのデータストリームは8→4→6→2→7→3→5→1の順番に画素単位でマルチプレクスして12G-SDI信号を生成します。

12G-SDI信号における4K映像の分割方式は2サンプルインターリーブのみであり、3G-SDI信号へのマッピングはレベルAのみとなっています。4K映像を3G-SDIレベルAにマッピングして12G-SDI信号を生成する方法をSMPTE ST 2082ではタイプ1とよんでいます。

補助データの多重については、5章で説明したようにサブイメージがマッピングされている3G-SDIレベルAでのインタフェース上のライン番号に則ります。

SMPTE ST 2036-1で規定されているUHDTV2 8K映像7680×4320/59.94PのSDI信号伝送はSMPTE ST 2082-12で規定されており、12G-SDIクワッドリンクで伝送できます。8K映像は12G-SDIクワッドリンクの他にARIB STD-B59で規定されていてU-SDIとよばれるインタフェースもありますが、これまでのSDI信号とは異なる部分が多いので省略します。

8K映像も4K映像と同様にサブイメージ分割を行いSDI伝送します。8K映像の場合は水平方向に2サンプルインターリーブで2画素ずつ抜き出し、垂直方向に1ラインおきに抜き出して一度4K映像を生成します。この4K映像を図10に示したようにさらに2サンプルインターリーブで1,920×1,080のサブイメージに分割し、3G-SDIレベルAにマッピングします。合計16のサブイメージへ分割され、サブイメージ1から4が12G-SDIのリンク1へマッピングが行われます。このマッピング過程は3G-SDIクワッドリンクの12G-SDIへのマッピングと同様です。他、サブイメージ5から8は12G-SDIのリンク2へ、サブイメージ9から12は12G-SDIのリンク3へ、サブイメージ13から16は12G-SDIのリンク4へマッピングされ、8K映像が12G-SDIクワッドリンクで伝送できます。

7. HDRのSDI信号マッピング

HDRとは映像信号レベル64 (0%) から940 (100%) の範囲を変えることなく、映像信号レベルに対する明るさ情報を変えることで、これまでのSDRと比較して数十倍から百倍程度の明るさ情報を扱えるようにしたものです。

映像信号レベルに対する明るさ情報は、非直線的なカーブを持った重みづけを行い広い明るさ情報を既存の映像信号レベルに押し込んでいます。このカーブはITU-R BT.2100で2種類の規定があり、それぞれHLG (Hybrid Log Gamma) とPQ (Perceptual Quantization) と呼ばれています。HLGはこれまでのSDRとの互換性を重視して作られたカーブで、PQは人の目の特性を実験で求めてカーブを再定義したものです。

HLGはカメラで撮影する際の光を電気信号に変換する特性を規定しており、SDRと比べてカメラに入射する光の量

を12倍から20倍まで扱えるように映像信号レベルを割り当てています。一方PQは映像信号レベルがそのままディスプレイ上の絶対輝度として規定しており、映像信号レベル100%がディスプレイ輝度値10,000cd/m²と定義しています。HLGとPQの相互変換も必要になることからITU-R BT.2100ではHLG、PQともにOOTF (Opto-Optical Transfer Function) という概念を導入して、カメラに入射する光とディスプレイで表示される明るさに一定の関係を持たせています。

HLGは従来の映像信号同様に100%映像信号レベルをディスプレイの最大輝度に割り当てています。そのため映像レベルによらずディスプレイ上で白飛びすることは原理的にはありませんが、最大輝度が異なるディスプレイに同じHLGのコンテンツを表示させると高輝度部分で明るさの差が出るのはもちろん、スキントーンなどの中間輝度部分にも大きな差が出てしまう恐れがあります。

そこでHLGはOOTF特性を変えることで対処しています。このOOTF特性の値をシステムガンマと呼び、システムガンマはディスプレイの最高輝度に応じて、ディスプレイ側で可変します。この可変カーブを適用することで、ディスプレイの最高輝度の差によって中間調の見え方に大きな差が出ないように工夫しています。ディスプレイがシステムガンマを持つことでHLGは制作の段階でディスプレイの最大輝度に応じた調整をする必要がありません。

一方PQはビデオレベルとディスプレイの明るさを一義的に決めているので、例えば、最大輝度1,000cd/m²のディスプレイに輝度値10,000cd/m²に相当する100%映像信号レベルを入力すると白飛びしてしまう恐れがあります*6。

HDRをSDI信号で伝送する際はSDRと比較して、まったく情報量に増減がありません。要するにSDRと同じSDIレートでHDRを伝送することができます。伝送系に依存しない技術なので従来のSDIシステムに最小限のコスト投資で高い映像表現効果が得られる技術です。

一方レシーバーが受信するSDI信号の構造はSDRと何ら変わらないことから、受信したSDI信号がSDRなのかHDRなのかを映像データのみから判別することができません。SMPTE ST 425で規定されているペイロードIDには2017年にSDRとHDR (HLG/PQ) の識別が追加され、各機器は自動でHDR/SDRを識別できるようになりました。しかし販売済みの多くの機器は、これらの識別子に対応していない場合も多いのでHDRとSDRの取り扱いには注意が必要です。

8. WCG (Wide Color Gamut) のSDI信号マッピング

より鮮やかな色表現が可能になったWCGもHDR同様に

*6 放送や配信、パッケージメディアではメタデータを使用することで白飛びを防ぐことができます。

映像信号レベル64(0%)から940(100%)(色差信号の場合64から960)の範囲を変えていません。WCGは映像信号レベルに対する色情報の割り当てを変えることで、色再現範囲である色域をこれまで以上に大きい範囲で扱えるようにしたものです。WCGはITU-R BT.2020で規定されています。HDで主に使用された従来の色域はITU-R BT.709で規定されていることから一般的に「709色域」と呼ばれ、WCGを「2020色域」と呼んでいます。ARIB TR-B36によるとBT.2020色域中のBT.709含有率は52.9%であるので、xy色度座標中の面積比で2倍弱の色表現範囲が広がっています。

色域の違いはHDR同様に映像データを見ただけで識別することはできません。SMPTE ST 425にて色域の識別ができるように改定されていますが、最新のペイロードIDに対応していない機材を用いる場合は、ARIB STD-B66やARIB STD-B72で規定されているWCGを含むカラーバーをコンテンツの頭に挿入するなど何らかの形で色域を管理する必要があります。

9. SDI信号の課題

このように進化を続けてきたSDI信号ですが、75Ωの同軸ケーブルで12G-SDIを超える伝送レートは物理的な特性からも難しいとされています。12G-SDIクワッドリンクで伝送することができる8K映像をケーブル1本で伝送するためには48G-SDIが必要ということになり、同軸ケーブルでの伝送は現時点では技術的なめどが立っていません。一方IPで伝送することを考えてみるとすでに100 Gbpsや400 Gbpsといった伝送レートは商用ベースで実現されており、48G-SDIで8Kを伝送するより技術的なハードルは低いといえます。

また現行のSDIシステムでは映像・音声を伝送するためのSDI信号のほか、システム同期のためのブラックバースト信号、機器を集中管理するための制御信号、タイムコード、インカムなど用途によってさまざまな信号を個別に使い分けしています。これらの信号はIP化することで1本のケーブルに集約することも技術的には可能で、システム構築が容易になるといえます。

今後IPが普及していくといっても、しばらくはSDIとIPの混在システムが続くかと思います。それぞれの特徴をよく理解したうえで、今何を選択すべきなのか答えを出していく必要があります。

(2019年1月31日受付)

【文 献】

- 1) SMPTE ST 12-2, Transmission of Time Code in the Ancillary Data Space

- 2) SMPTE ST 274 for Television 1920 x 1080 Image Sample Structure, Digital Representation and Digital Timing Reference Sequences for Multiple Picture Rates
- 3) SMPTE ST 291, Ancillary Data Packet and Space Formatting
- 4) SMPTE ST 299-1, 24-Bit Digital Audio Format for SMPTE 292 Bit-Serial Interface
- 5) SMPTE ST 299-2, Extension of the 24-Bit Digital Audio Format to 32 Channels for 3Gbps Bit-Serial Interfaces
- 6) SMPTE ST 352, Payload Identification Codes for Serial Digital Interfaces
- 7) SMPTE ST 372, Dual Link 1.5Gbps Digital Interface for 1920 × 1080 and 2048 × 1080 Picture Formats
- 8) SMPTE ST 424, 3Gbps Signal/Data Serial Interface
- 9) SMPTE ST 425-1, Source Image Format and Ancillary Data Mapping for the 3Gbps Serial Interface
- 10) SMPTE ST 425-5, Image Format and Ancillary Data Mapping for the Quad Link 3Gbps Serial Interface
- 11) SMPTE ST 2082-1, 12Gbps Signal/Data Serial Interface - Electrical
- 12) SMPTE ST 2082-10, 2160-line and 1080-line Source Image and Ancillary Data Mapping for 12G-SDI
- 13) SMPTE ST 2082-12, 4320-line and 2160-line Source Image and Ancillary Data Mapping for Quad-link 12G-SDI
- 14) SMPTE ST 2036-1, Ultra High Definition Television Image Parameter Values for Program Production
- 15) SMPTE RP 198, 1Bit-Serial Digital Checkfield for Use in High-Definition Interfaces
- 16) Rec. ITU-R BT.709-5, Parameter values for the HDTV standards for production and international programme exchange
- 17) Rec. ITU-R BT.1120-8, Digital interfaces for HDTV studio signals
- 18) Rec. ITU-R BT.1614-1, Payload identification data structure for digital television interfaces
- 19) Rec. ITU-R BT.2020, Parameter values for ultra-high definition television systems for production and international programme exchange
- 20) Rec. ITU-R BT.2100-0, Image parameter values for high dynamic range television for use in production and international programme exchange
- 21) ARIB STD-B58, 超高精細度テレビジョン信号スタジオ機器間インタフェース規格
- 22) ARIB STD-B66, UHDTVマルチフォーマット・カラーバー
- 23) ARIB STD-B72, Colour Bar Test Pattern for Hybrid Log-Gamma (HLG) High Dynamic Range Television (HDR-TV) System
- 24) ARIB TR-B23, 放送局間の情報伝送に使用する補助データ運用規定
- 25) ARIB TR-B36, 超高精細度テレビジョン番組制作用ディスプレイの色域包含率計算法



矢野 浩二 1994年、リーダー電子(株)に勤務。主に波形モニタの開発に従事。2018年、同社技術開発部長に就任し、現在に至る。



鈴木 則行 1997年、リーダー電子(株)に勤務。技術開発部にて、主にビデオ信号の測定器開発に従事。2018年、技術開発部グループマネージャに就任し、現在に至る。

映像・音声信号のIPトランスポート方式

正会員 川本潤一郎[†]

1. まえがき

近年、通信速度の高速化やビットコストの低下により、IP (Internet Protocol) ネットワークの導入が世界的に普及しています。放送局の番組制作現場においても、IP技術を導入した番組素材の信号伝送が盛んに行われるようになってきました。ここでは映像や音声、アンシラリデータなどの番組素材信号をIP信号にマッピングして伝送するIPトランスポート方式が必要となります。これまで、さまざまなメーカーが独自の優れたIP伝送方式を開発し、製品化・販売を競ってきました。一方で、機器間の相互接続性を担保することで、多くのメーカーが市場に参入でき、市場の活性化や導入コストの低下につながるため、標準化団体における早期の統一規格の策定も望まれていました。この流れを受けて、SMPTE (Society of Motion Picture and Television Engineers) では、番組素材信号のIPトランスポート規格である、ST 2022^{1)～7)}や、ST 2110^{8)～13)}を発行しています。これらの規格は、IPネットワークを介して番組素材信号を共有するための伝送に関する標準規格です。講座の第3回となる本稿では、放送局の番組制作におけるIP化に欠かせないこれら2種類の規格について、その概要を解説いたします。なお、本稿で扱うIPトランスポート方式は、非圧縮映像を対象としています。

2. 番組制作におけるIP化

2.1 IP化の背景

これまで放送局における番組制作では、映像・音声・アンシラリデータを1本のシリアル信号で伝送するSDI (Serial Digital Interface) が用いられてきました。SDIはシリアル信号に重畳された高精度の同期情報をもとに受信側でクロック再生を行うシステムのため、送信側と受信側を同軸ケーブルで直結すれば信号伝送が可能になり、SD (Standard-Definition) やHD (High-Definition) 映像の運用

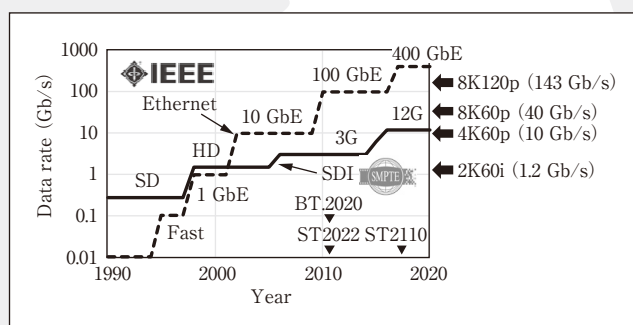


図1 伝送レートの変遷

では、システムが複雑にならずに実現可能です。他方、SDI信号は主には放送局の制作現場のみで用いられるプロプライエタリな信号形式のため、伝送機器の製造コストが高くなる場合があるということも言われてきました。

一方、世界的にはインターネット技術の普及によって、IP信号の高速化に伴う伝送コストの低下が進んでいます。図1に、IEEE (Institute of Electrical and Electronics Engineers) にて規格化されているイーサネット信号と SMPTE にて規格化されている SDI 信号の、年代ごとのデータレートを示します。HD放送が始まった2000年代初頭までは、2K60iを1本の同軸ケーブルで伝送できる1.5 GbpsのHD-SDI¹⁴⁾の伝送速度が、イーサネット信号の伝送速度 (ギガビットイーサネット, 1 GbE) よりも高速でした。その後、イーサネット信号規格は、2004年に10 GbE¹⁵⁾、2010年には100 GbE¹⁶⁾と高速化が続き、現在では400 GbE¹⁷⁾の規格化が完了するまでに至っています。一方、SDI信号は、2006年には3G-SDI¹⁸⁾、2016年には12G-SDI¹⁹⁾と規格化が進んでいるものの、イーサネット信号の伝送速度には遠く及ばない状況です。そのため、まずは伝送コストを低減することを目的に、より高速・汎用なオープンシステムのIP技術を番組制作システムに取り入れる検討が2000年代中頃から開始されました。その結果、2012年にはST 2022規格が発行され、今日に至るまで放送局の番組制作における伝送の部分で幅広く利用されています。

その後、IP技術の特長である双方向性や拡張性といったところにも関心が集まり、これらを用いてより柔軟な番組

[†] NHK放送技術研究所

"IP Adoption in Broadcast Industry (3): IP Packet Transport Method for Video and Audio Signals" by Junichiro Kawamoto (Science and Technology Research Laboratories, NHK, Tokyo)

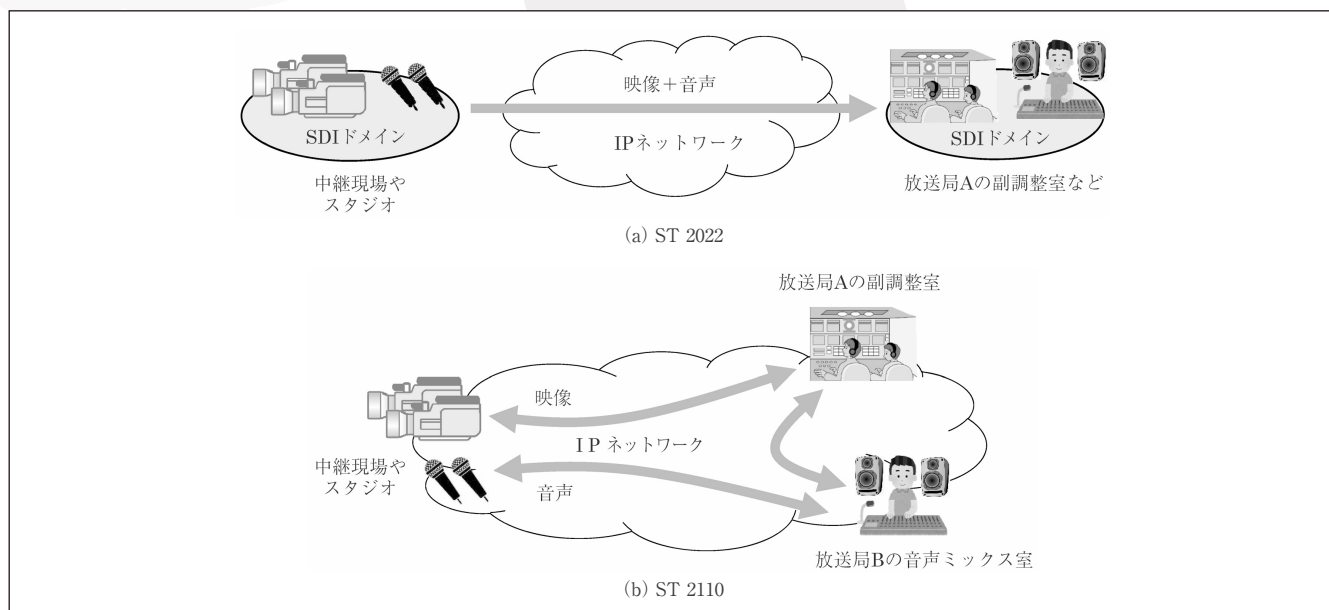


図2 ST 2022とST 2110の違い

制作システムの構築を目指す取り組みも始まりました。また、より高精細な4K/8K映像²⁰⁾をSDI信号で伝送する場合には、複数本のSDI信号を接続して伝送することが必要であり、システムが複雑になるという課題に対して、IP技術を用いて対応していこうといった検討も行われ始めました。これらを背景として、2017年にはST 2110が標準規格として発行されました。

2.2 ST 2022とST 2110の違い

ST 2022とST 2110の違いがイメージしやすいように、図2を使って説明します。ST 2022とST 2110の主な違いは、番組素材を構成する映像や音声、アンシラリデータ（これらをエッセンスと称する）を一つのIP信号の流れ（ストリーム）として伝送するのか、それとも別々のストリームとして伝送するのかにあります。前者がST 2022、後者がST 2110となります。ST 2022では、これまで放送現場で使われているSDI信号をインタフェースとして、伝送部分だけIPネットワークを介して伝送するイメージになります（図2(a)）。SDI信号の決まった場所に重畳されている各エッセンスのシリアル信号を順次IP化するため、中継現場から放送局へ、もしくは、スタジオから副調整室へといった、point to pointの伝送になります。ST 2022が規格化された2010年ごろは、IP技術を用いて伝送コストの低減を目指す流れがありました。その後、4K/8K映像の伝送や、IP技術を使ったより柔軟な番組制作システムの実現が望まれるようになった。そこで、ST 2110が規格化されました。ST 2110では、IP信号をインタフェースとして、各エッセンスを別々のストリームで伝送できるため、例えば、映像信号は放送局Aの副調整室へ、音声は放送局Bの音声ミックス室へ伝送してミキシング後に放送局Aへといったような、point to multi-pointの伝送も可能となり、ST 2022で

は実現が難しかった柔軟な番組制作システムが可能となります（図2(b)）。

2.3 各規格の対応図

ST 2022とST 2110はそれぞれ複数の規格文書から構成されます。各規格文書と対応する実現機能を図3に簡単にまとめます。非圧縮信号のST 2022を用いた伝送規格は、ST 2022-5、-6、-7になります。その中でもST 2022-6は、SDI信号をIPパケットへマッピングする方式や同期に必要な情報を定めるトランスポート規格です。先述した通り、SDI信号をIPパケットで伝送する方式のため、一つの規格文書で映像・音声・アンシラリデータの各エッセンスのマッピングを規定します。一方、ST 2110を用いた規格では、各エッセンスが別々のストリームで伝送されるため、各ストリームのマッピングに対応した規格がそれぞれ発行されています。具体的には、映像はST 2110-20/21、音声はST 2110-30/31、アンシラリデータはST 2110-40で規定されています。さらに、これらの各規格に共通のシステム

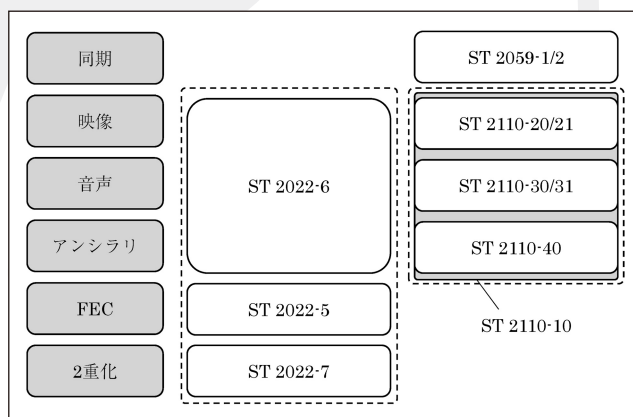


図3 各規格文書と対応する機能

全体のタイミングや基本的な制約を記述するST 2110-10からなります。すなわち、ST 2110はこれらの規格群から構成されるトランスポート規格です。関連して、エッセンス同期を規定するST 2059-1/2²¹⁾ 2²²⁾も発行されています。本章では、ST 2022とST 2110の各規格について説明します。

3. SMPTE ST 2022

SMPTE ST 2022は7つの規格から構成されています²³⁾が、本稿では、SDI信号をIPパケットで伝送する非圧縮映像のIPトランスポート方式であるST 2022-6規格についてもう少し詳細な中身について説明します。

3.1 パケットマッピング

RTP (Real-time Transport Protocol) パケットへのカプセル化構造を図4に、RTP HeaderおよびPayload Headerのフォーマットを図5 (a) および (b) に示します。RTP Headerの構成はRFC 3550²⁴⁾で規定されている構成となります。

SDI信号には、映像の他に音声やアンシラリデータがエンベデッドされていますが、この規格ではSDI信号を1376バイト毎に区切ってMedia Payloadとし、後述するPayload Headerを付加後、12バイトのRTP Headerをさら

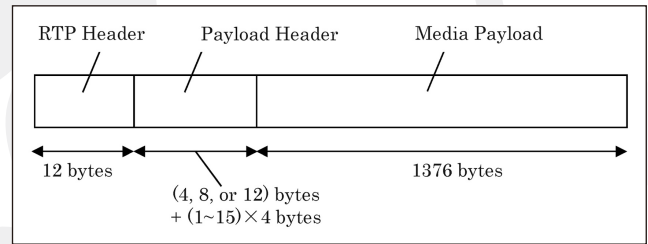
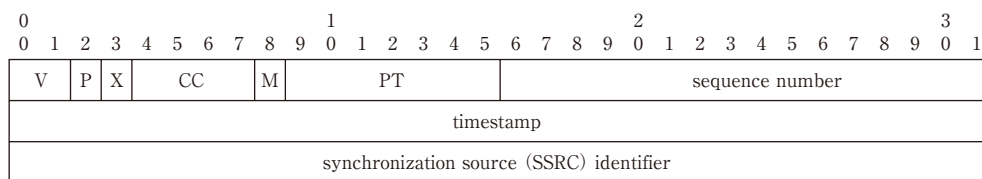


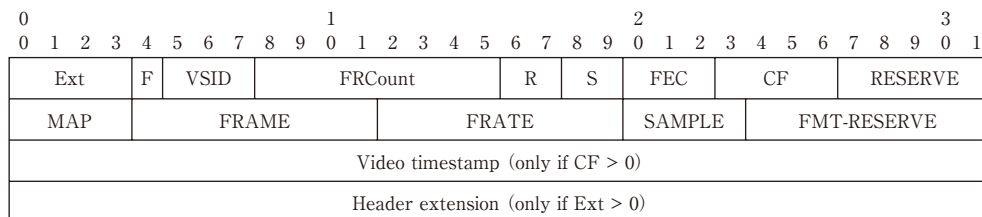
図4 RTPパケットの構成 (ST 2022)

に付加してRTPパケットにカプセル化します。

Media Payloadには、SDI信号に含まれる上述した映像、音声、アンシラリデータの他にそれらの位相関係を表すTRS (Timing Reference Sequence)、VANC (Vertical Ancillary)、HANC (Horizontal Ancillary) などすべてのデータが含まれます。図6にSDI信号のMedia Payloadのマッピングの参考例を示します。この例では、1映像フレームの1ライン目の先頭を表すEAV (End of Active Video) シーケンス (0xFF, 0x00, 0x00, 0x2DB) を、1バイトごとに区切ったシーケンス (0xFF, 0xC0, 0x00, 0x02, 0xD8) を先頭から順にMedia Payloadの1バイト目からマッピングしていきます。1映像フレームの最後の部分のSDI信号が1376バイトのMedia Payloadに満たすことができない



(a) RTP Header



(b) Payload Header

図5 Headerフォーマット (ST 2022)

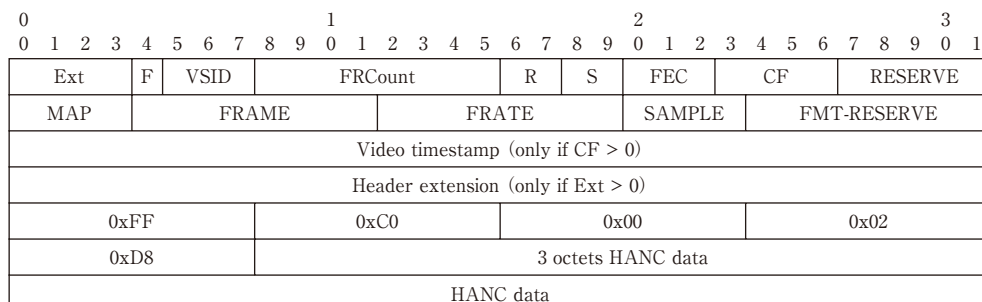


図6 マッピング例

場合には、ヌルバイトを追加して全長1376バイトのMedia Payloadを構成しなければなりません。この1映像フレームの最後のMedia Payloadに該当するパケットでは、当該RTP Header内のMビット(Markerビット)を1に設定し、それ以外のパケットはRFC 3497²⁵⁾に定義されているように0に設定します。

カプセル化されたRTPパケットは、さらにUDP (User Datagram Protocol) /IPヘッダを付加することで、IPパケットが生成されます。ST 2022-6では、1080pビデオを伝送する3G-SDIの非圧縮ストリームまでのIP伝送を規定しています。

3.2 同期

IP伝送では送受信間は通常非同期で運用されており、受信側でSDIストリームを安定して再生するためには、同期の復元(送信側の映像信号クロック周波数の再生)が必要となります。送受信間同期には、Video Timestamp情報を用いることが可能です。タイムスタンプ情報は、RTP HeaderとPayload Headerに含まれます。RTP Header内のタイムスタンプ情報を表す32ビットのtime stampは、RTPデータグラムの最初のオクテットの27 MHzクロックサンプリングの瞬間のタイムスタンプ値を反映します。このタイムスタンプ値は、固定長のパケット間隔でサンプリングされるため、送信側では単調増加します。そのため、受信側では、この値を監視することで、ネットワークのジッタを概算することができ、ジッタ吸収バッファの管理に使用することができます。一方、送受信間のSDIストリームの同期には、Payload Header内のタイムスタンプ情報を表す32ビットのVideo timestampを使用します。ここで用いるタイムスタンプ値は、送信側の入力SDIストリームに同期したクロックから算出された値が使用されます。先述した通り、各エッセンス間の位相情報はすべてSDIストリームの中に含まれているため、ST 2022では送受信間で当該Video timestamp情報からSDIストリームの再生周波数を同期すれば、各エッセンスの位相関係も復元することが可能です。

3.3 ストリーム情報の通知

伝送するパケットのMedia Payloadにどのような中身が含まれているかなどの情報を送受信間で共有することで、各エッセンスの正常な信号伝送が可能となります。これらの情報は、ST 2022では、主にPayload Headerに記されます。当該ヘッダ情報の例を表1に示します。映像フレームのカウント値を表すFrame CountやSDIストリームの種別を表すMAP、映像信号の画素数や映像フレーム周波数を表すFRAMEやFRATEなどが含まれます。またFEC usageでは、誤り訂正符号化の有無やその構成(列方向もしくは、行および列方向)を示します。

3.4 その他

ST 2022-6では誤り訂正符号化方式を、ST 2022-7では

表1 Payload Header内のデータ例(一部)

名 称	ビット数	機 能
Frame Count	8	映像フレーム毎にインクリメント
FEC usage	3	使用する誤り訂正符号化の種類
Clock Frequency	4	映像信号のクロック周波数
MAP	4	SDIストリームの種別
FRAME	8	映像信号の画素数など
FRATE	8	映像信号のフレーム周波数
SAMPLE	4	映像信号のサンプリング構造
Video Timestamp	32	送受信間の映像信号同期

シームレス切り替え方式が規定されています。これらの規格は、局外などの比較的長距離の素材伝送を想定して規格化されたST 2022の特徴的な規格の一つといえます。

4. SMPTE ST 2110

SMPTE ST 2110は、2019年3月現在、6つの規格から構成されています。本稿では、映像・音声・アンシラリデータの各エッセンスを別々のストリームで伝送する非圧縮映像のIPトランスポート方式であるST 2110規格について、もう少し詳細な中身について説明します。

4.1 パケットマッピング

4.1.1 映像エッセンス

ST 2110-20で規定する映像エッセンスのRTPパケットへのカプセル化構造を図7に、RTP HeaderおよびPayload Headerのフォーマットを図8(a)および(b)に示します。RTP Headerの構成は図5(a)で示したST 2022のヘッダ構成と同じとなり、RFC 3550で規定されている構成となります。RTP Header, Payload HeaderおよびMedia Payloadの合計のサイズは標準では1452バイト(UDPヘッダを含めて1460バイト)以下となる必要があります。また、拡張サイズも使用することができ、この場合UDPヘッダも含めて8960バイトまで使用することができます。RTPパケットのシーケンス番号はRTP HeaderのSequence Number(下位16ビット)とPayload HeaderのExtended Sequence Number(上位16ビット)を使用します。Media Payloadには、映像エッセンスのデータが含まれますが、そのデータに関する情報がPayload Header内のSRD(Sample Row Data)ヘッダで表されます。SRD LengthはMedia Payloadに含まれるバイト数を表し、この数字は後述するpgroup長の整数倍となります。SRD Row Numberは、Media Payloadに含

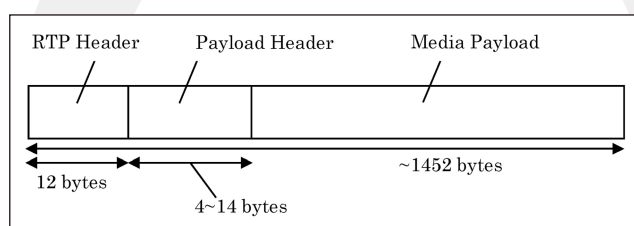
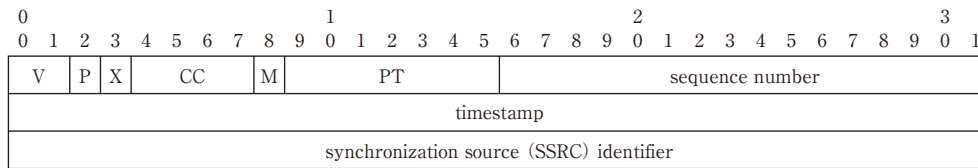
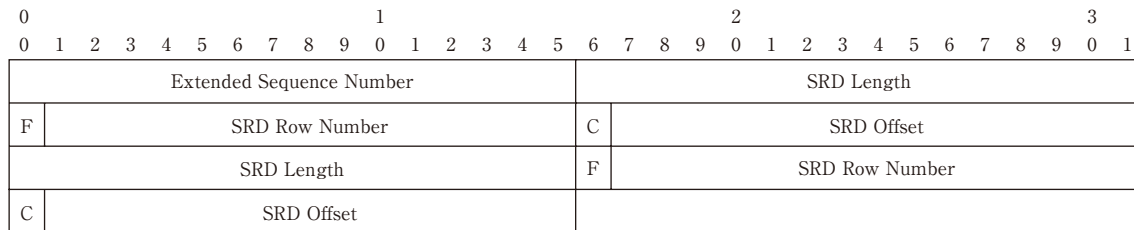


図7 RTPパケットの構成(ST 2110)



(a) RTP Header



(b) Payload Header

図8 Headerフォーマット (ST 2110)

まれる映像エッセンスのデータの行番号を表します。SRD Offsetは、当該データの最初のデータが位置する映像内の行方向のサンプリングの位置を示します。1映像フレームの左上を開始点として、各行ごとに水平方向の映像エッセンスをMedia Payloadにマッピングしていきます。SRD Row NumberおよびSRD Offsetはそれぞれ15ビットで構成されるため、映像サイズは32K×32Kまで対応できるようになります。一つのRTPパケットには三つ以上のSRDヘッダを含んではいけないことになっています。すなわち、三つ以上の行をまたぐような映像エッセンスデータの同一Media Payloadへのマッピングは禁止されています。

映像を構成する要素は、映像サンプリング構造やビット深度などがあり、同じ画素においても情報量が異なってきます。一方、マッピングにおいては8ビットを1バイトとするバイト単位での処理が通常行われます。そこで、映像画素をバイト単位で処理するためにpgroupという概念が導入されています。pgroupは上述した映像を構成する要素ごとに、バイト単位で処理できるような画素のグループを構成するパラメータとなります。参考に、4:2:2、10ビットの映像の例を図9に示します。この場合、2画素の情報を用いて10ビット×4の5バイトをpgroupとして(破線で囲んだ範囲)、これを処理単位としてMedia Payloadにマッピングしていくことになります。すなわち二つの異なるRTPパケットのMedia Payloadに、同一のpgroupのデータがマッピングされることを禁止しています。ST 2110-20では映像サンプリング構造が4:2:0、4:2:2、4:4:4、ビット深度が8、10、12、16ビットのpgroupを規定しています。

4.1.2 音声エッセンス

音声エッセンスのカプセル化はST 2110-30および-31で規定されています。ここでは非圧縮のPCM (Pulse Code Modulation) デジタル音声をRTPパケットにマッピング

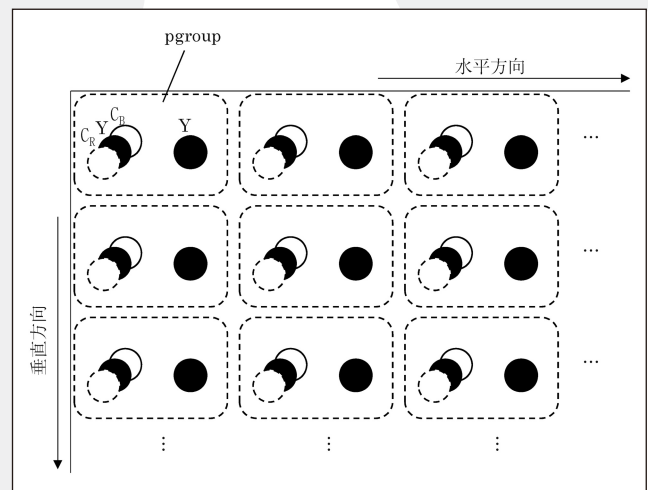


図9 pgroupの一例

するST 2110-30について概説します。この規格では、AES (Audio Engineering Society) 67²⁶⁾で規定される16ビット、24ビットのリニアPCMのデジタル音声ストリームを用います。PCMデジタル音声信号のサンプリング周波数は、48kHzを必須とし、96kHzはオプションとなります。その他のサンプリング周波数は使用しないことになっています。この規格の他にST 2110-31があります。これは、AES²⁷⁾準拠のデジタル音声信号のパケットマッピングを規定しています。

4.1.3 アンシラリデータエッセンス

アンシラリデータエッセンスのカプセル化はST 2110-40で規定されており、これは、ST 291-1²⁸⁾で規定されるAncillary Data PacketsをRTPパケットで伝送するための規格です。詳細なマッピングの内容はRFC8331²⁹⁾で文書化されています。図10に2個のAncillary Data PacketをRTPパケットにマッピングする例を示します。Payload Headerでは、Lengthにおいて、Media Payloadに含まれ

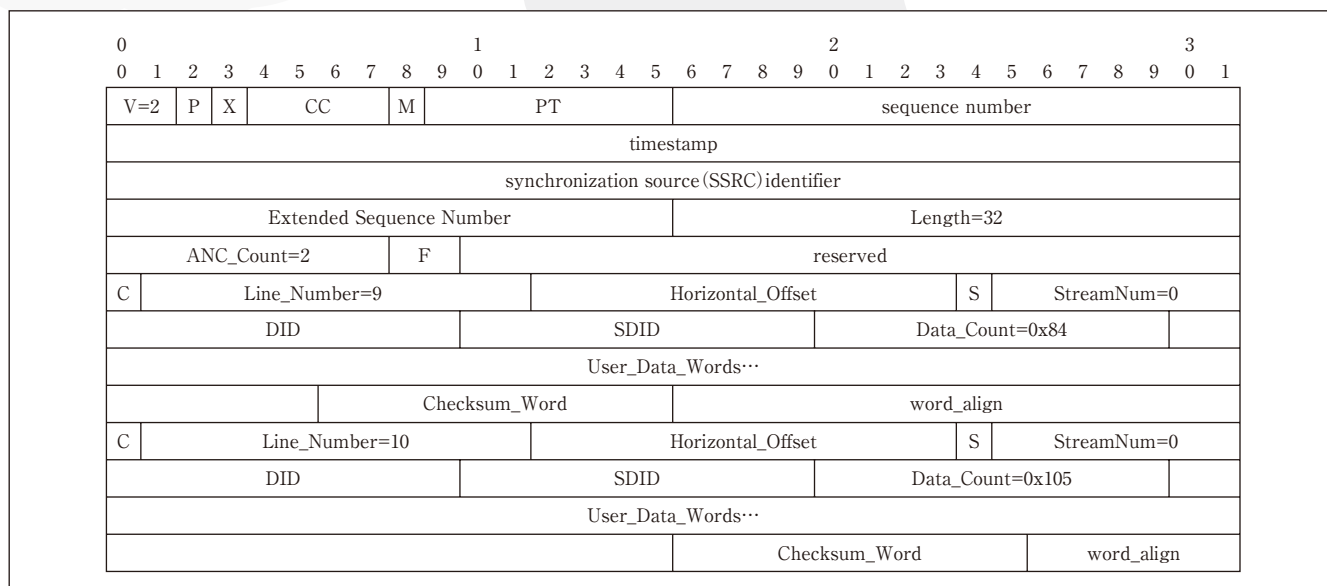


図10 マッピング例

る Ancillary Data Packetのサイズをバイト数で(この例では32バイト)示します。また、ANC_Countに Ancillary Data Packetの個数(この例では2個)を示します。Media Payloadには、フラグCビットを先頭として、Ancillary Data Packetがマッピングされます。

4.2 同期

ST 2022では、送受信間の同期には、RTPパケット内のタイムスタンプ値を用いていました。一方、ST 2110では、各エッセンスが別々のストリームで伝送されるため、それらの位相関係も再生するために、IEEE 1588-2008³⁰⁾のPTP (Precision Time Protocol)とST 2059-1および-2による正確な時刻配信を前提とした同期方式を規定しています。各エッセンスストリーム間の同期には各送受信機がPTPを介して同期した共通基準時計から生成されるRTPタイムスタンプ値を使用して実施します。RTPタイムスタンプは、共通基準時計から算出されたある瞬間のサンプリング値となります。連続した映像フレームのRTPタイムスタンプは、必要に応じて整数に切り捨てた映像フレームレートをベースに、単調増加します。同一映像フレームに含まれるRTPパケットのタイムスタンプ値は同じとなります。ST 2059-1および-2は、時刻情報から生成する各エッセンスの再生タイミングや、PTP通信に適用するパラメータ値を規定します。

4.3 ストリーム情報の通知

ST 2022では、ストリーム情報の通知に、当該伝送パケットのヘッダを用いて行っていました。一方、ST 2110ではRFC 4566³¹⁾で規定されたSDP (Session Description Protocol)を用いて通知します。SDPの例を図11に示します。この例では、ST 2110-20で規定されるソースIPアドレスが192.168.100.2、マルチキャストグループが239.100.9.10、UDPポート番号が50000の映像ストリーム(4:2:2, 10ビット, 720p, 59.94 Hz, 1系の場合)の情報や、同期について

```
v=0
o=- 123456 11 IN IP4 192.168.100.2
s=Example of a SMPTE ST2110-20 signal
i=this example is for 720p video at 59.94
t=0 0
a=recvonly
a=group:DUP primary secondary
m=video 50000 RTP/AVP 112
c=IN IP4 239.100.9.10/32
a=source-filter:incl IN IP4 239.100.9.10 192.168.100.2
a=rtptime:112 raw/90000
a=fmt:112 sampling=YCbCr-4:2:2; width=1280; height=720;
exactframerate=60000/10001; depth=10; TCS=SDR; colorimetry=BT709;
PM=2110GPM; SSN=ST2110-20:2017;
a=ts-refclk:ptp=IEEE1588-2008:39-A7-94-FF-FE-07-CB-D0:37
a=mediaclock:direct=0
a=mid:primary
m=video 50020 RTP/AVP 112
c=IN IP4 239.101.9.10/32
a=source-filter:incl IN IP4 239.101.9.10 192.168.101.2
a=rtptime:112 raw/90000
a=fmt:112 sampling=YCbCr-4:2:2; width=1280; height=720;
exactframerate=60000/10001; depth=10; TCS=SDR; colorimetry=BT709;
PM=2110GPM; SSN=ST2110-20:2017;
a=ts-refclk:ptp=IEEE1588-2008:39-A7-94-FF-FE-07-CB-D0:37
a=mediaclock:direct=0
a=mid:secondary
```

図11 SDP例

共通基準時計に使用するPTPマスタークロックのID(39-A7-94-FF-FE-07-CB-D0)などの情報が示されています。

4.4 その他

本稿で説明した以外に、圧縮映像の伝送規定であるST 2110-22や、広帯域の映像信号をマルチストリームに分割して伝送するためのRP 2110-23などが議論されています。

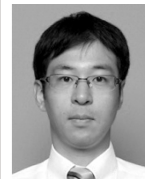
5. むすび

本稿では、IP化の技術的な背景やST 2022とST 2110の違い、およびこれらの概要について解説しました。

(2019年4月4日受付)

〔文 献〕

- 1) SMPTE ST 2022-1: "Forward Error Correction for Real-Time Video/Audio Transport Over IP Networks" (2007)
- 2) SMPTE ST 2022-2: "Unidirectional Transport of Constant Bit Rate MPEG-2 Transport Streams on IP Networks" (2007)
- 3) SMPTE ST 2022-3: "Unidirectional Transport of Variable Bit Rate MPEG-2 Transport Streams on IP Networks" (2010)
- 4) SMPTE ST 2022-4: "Unidirectional Transport of Non-Piecewise Constant Variable Bit Rate MPEG-2 Streams on IP Networks" (2011)
- 5) SMPTE ST 2022-5: "Forward Error Correction for High Bit Rate Media Transport Over IP Networks" (2012)
- 6) SMPTE ST 2022-6: "Transport of High Bit Rate Media Signals over IP Networks (HBRMT)" (2012)
- 7) SMPTE ST 2022-7: "Seamless Protection Switching of SMPTE ST 2022 IP Datagrams" (2013)
- 8) SMPTE ST 2110-10: "Professional Media Over Managed IP Networks: System Timing and Definitions" (2017)
- 9) SMPTE ST 2110-20: "Professional Media Over Managed IP Networks: Uncompressed Active Video" (2017)
- 10) SMPTE ST 2110-21: "Professional Media Over Managed IP Networks: Traffic Shaping and Delivery Timing for Video" (2017)
- 11) SMPTE ST 2110-30: "Professional Media Over Managed IP Networks: PCM Digital Audio" (2017)
- 12) SMPTE ST 2110-31: "Professional Media Over Managed IP Networks: AES3 Transparent Transport" (2018)
- 13) SMPTE ST 2110-40: "Professional Media Over Managed IP Networks: SMPTE ST 291-1 Ancillary Data" (2018)
- 14) SMPTE ST 292: "1.5Gbps Signal/Data Serial Interface" (1998)
- 15) IEEE Standard 802.3ae: "10Gbps Ethernet" (2002)
- 16) IEEE Standard 802.3ba: "40Gbps and 100Gbps Ethernet" (2010)
- 17) IEEE Standard 802.3bs: "Media Access Control Parameters, Physical Layers and Management Parameters for 200Gbps and 400Gbps Operation" (2017)
- 18) SMPTE ST 424: "3Gbps Signal/Data Serial Interface" (2006)
- 19) SMPTE ST 2082-1: "12Gbps Signal/Data Serial Interface - Electrical" (2015)
- 20) Recommendation ITU-R BT.2020-1: "Parameter values for ultra-high definition television systems for production and international programme exchange" (2012)
- 21) SMPTE ST 2059-1: "Generation and Alignment of Interface Signals to the SMPTE Epoch" (2015)
- 22) SMPTE ST 2059-2: "SMPTE Profile for Use of IEEE-1588 Precision Time Protocol in Professional Broadcast Applications" (2015)
- 23) 川本： "SMPTE ST 2022", 映情学誌, 72, 2, pp.247-250 (2018)
- 24) RFC 3550: "RTP: A Transport Protocol for Real-Time Applications" (2003)
- 25) RFC 3497: "RTP Payload Format for Society of Motion Picture and Television Engineers (SMPTE) 292M Video" (2003)
- 26) AES67: "AES standard for audio applications of networks - High-performance streaming audio-over-IP interoperability" (2013)
- 27) AES3-3: "AES standard for digital audio - Digital input-output interfacing - Serial transmission format for two-channel linearly represented digital audio data" (2009)
- 28) SMPTE ST 291-1: "Ancillary Data Packet and Space Formatting" (2011)
- 29) RFC 8331: "RTP Payload for Society of Motion Picture and Television Engineers (SMPTE) ST 291-1 Ancillary Data" (2018)
- 30) IEEE Standard 1588-2008: "A Precision Clock Synchronization Protocol for Networked Measurement and Control Systems" (2008)
- 31) RFC 4566: "SDP: Session Description Protocol" (2006)



かわもとじゅんいちろう

川本潤一郎

2002年、筑波大学大学院工学研究科修士課程相当修了。2008年、NHK入局。2011年より、NHK放送技術研究所にて、番組素材の光伝送技術に関する研究に従事。2017年、筑波大学大学院システム情報工学研究科博士課程修了。博士(工学)。正会員。

IP番組制作システムの制御方式

正会員 小山 智史[†]

1. まえがき

番組制作システムをIP (Internet Protocol) 化する動きが始まった当初は、「これまで利用していたSDI (Serial Digital Interface) によって実現できていた機能をどのようにIPネットワーク上で実現するか」ということが検討されていました。そしてこれまでにIPネットワークを利用して機器間の同期を実現するためのSMPTE ST 2059-1/2^{1) 2)}や、映像信号をIPパケットにマッピングするためのSMPTE ST 2110-20³⁾などの標準規格が発行されました。しかし、SDIの機能をIPネットワーク上に移植しただけでは、高速なネットワークを安価に利用できるという、IP化によってもたらされるメリットの一部を利用しているだけにすぎません。現在では、IP化の特徴を生かした新しい機能の実現に向けた制御技術の検討と標準化が進められています。

本稿では、IP番組制作システムにおける制御技術が果たす役割や、制御技術の標準規格として広まりを見せているAMWA NMOSの内容について紹介します。

2. 制御方式の役割

番組制作システムのIP化と聞いて、「IPアドレスを各機器に設定するのが面倒」、「パケットロスが起きないか心配」といった印象を持ったことはないでしょうか？ ケーブルを接続すれば信号が流れるSDIシステムとは異なり、IPシステムでは運用の際に番組制作機器のIPアドレスの設定やネットワークスイッチのルーティング設定などさまざまな設定・制御が必要となります。これまでも素材伝送などでIP回線を利用できる機器が一部に用いられてきましたが、システム全体をIP化すると設定が必要な機器の数は膨大になり、手動での設定がとても大変になるだろうと想像できます。

ところで、IPネットワークは映像・音声以外にもさまざまなデータを伝送できるネットワークであり、制御情報も

伝送できます。そこでこの特長を生かして機器同士で制御情報を交換するようにすると、設定の自動化が可能になります。さらに、受信ステータスやエラーなどを集約して監視し、障害対応を自動化すれば、従来よりも効率よく番組制作システムを運用できるようになるかもしれません。また、IP化によって実現が期待されているリモートプロダクションにおいても、離れた場所で運用する機器の制御がIPネットワーク越しにできることが必須条件となります。

このように、番組制作システムのIP化にとってIPネットワーク経由での機器の設定は欠かせないものであり、現在ではさまざまな機能を実現するための提案が行われるようになりました。制御情報をIPネットワーク経由で通信することで実現が期待される新しい機能として以下のような機能が考えられています。

- ・機器のIPアドレスやその他コンフィグの自動設定
- ・ネットワーク内の他の機器の情報の取得
- ・リモートプロダクション
- ・ネットワークの帯域確保などのQoS制御
- ・機器監視の自動化
- ・機器共有
- ・セキュリティ強化

従来、機器制御機能はベンダー独自の方式が用いられてきました。IP経由の制御方式にもEmber+⁴⁾やNDCP⁵⁾などベンダー発の技術が存在しています。一方、オープンな標準規格としての制御方式の議論はSMPTEでは行われていなかったのですが、AMWA (Advanced Media Workflow Association)⁶⁾において議論が進められ、2016年にはネットワーク内の機器探索と登録についての制御方式の規格が発行されました。AMWAの制御方式は、番組制作機器ベンダーと放送局が協力して作り上げており、制作設備のIP化を推進するベンダーのアライアンスであるAIMS⁷⁾が推奨する制御方式として採用したり、EBUがベンダーに対してAMWAの制御方式への対応を求めることを表明⁸⁾したりするなど、広がりを見せています。

そこで、次章ではこのAMWAの制御方式の概要と機能について紹介します。

[†] NHK 放送技術研究所

"IP Adoption in Broadcast Industry (4): Control Method for IP-based Program Production System" by Tomofumi Koyama (Science and Technology Research Laboratories, NHK, Tokyo)

3. AMWA NMOSについて

3.1 AMWA NMOSとは

IP化された番組制作システムで利用する機器の制御方式が機器ごとに異なっていると、各機器に応じた制御システムが必要となってしまいます。IPネットワークによって機器同士が互いに通信して連携ができるようになったにも関わらず、制御方式の違いによってソフトウェアレベルで分断されているのはIP化のメリットが生かせません。そこで、IP化された番組制作システムに対応した機器で必要となる標準的な機能を提供するためのAPI (Application Program Interface) について、2015年にAMWAで検討が始まりました。定義されたAPIはNMOS (Networked Media Open Specifications) という名前がつけられ、github上で公開されています⁹⁾。

本記事を執筆している2019年5月の時点では、AMWA NMOSとしてIS-04¹⁰⁾、IS-05¹¹⁾、IS-06¹²⁾、IS-07¹³⁾、IS-08¹⁴⁾の5つのIS (Interface Specification) 文書が発行されています。また、IS-09¹⁵⁾、IS-10¹⁶⁾が作業中です。規格一覧を表1にまとめます。その他、複数の機器の関係についての記述方法やセキュリティを実現するための運用方法について記したBCP (Best Current Practice) もNMOSの一部として同様に公開されています。本章では各APIの目的や使い方、特徴について概要を解説します。詳細な実装方法や機能についてはgithub⁹⁾を参照してください。

3.2 IS-04 Registration and Discovery

IS-04はRegistration and Discoveryというタイトルがつ

けられており、その名の通り番組制作システム上の機器を自動的に発見し、登録して運用するための機能を提供します。本節では執筆時点の最新バージョンであるバージョン1.2.1 (stable) について、その内容を紹介します。

IS-04では「Registration API」、「Query API」、「Node API」の3つが定義されています。図1に、各APIの役割を示します。Registration APIは、RDS (Registration & Discovery System) と呼ばれる機器情報を管理するサーバが提供するAPIで、制作システムに接続した機器が自分自身の情報を登録する際に利用します。Query APIも同様にRDSが提供するAPIで、監視装置や制御装置など機器情報を利用するアプリケーションがシステム内の機器情報を取得するために用います。一方、Node APIは各機器が提供するAPIです。RDSがシステムに存在しない場合にもNode APIを利用して各機器から直接情報を取得することができます。

まず、機器はネットワーク内のRDSのIPアドレスをDNS-SD (Domain Name System - Service Discovery)¹⁷⁾によって取得します。DNS-SDはインターネットで用いられているDNSと同じような仕組みで、サービス名 (Registration APIでは“_nmos-registration._tcp”) からIPアドレスを取得することができます。RDSがシステム上に存在してそのIPアドレスが取得できた場合、機器はRegistration APIにより自分の情報を登録します。制御装置なども同様にDNS-SDからQuery APIが提供されているIPアドレスを検索し、ネットワークに接続されている機器の一覧を取得します。また、Query APIではWebSocketにも対応しており、

表1 AMWA NMOS ISシリーズ一覧

規格番号	タイトル	バージョン	概 要
IS-04	Discovery and Registration	v1.2.1 (stable)	機器の発見と登録
IS-05	Device Connection Management	v1.0.1 (stable)	機器の接続制御
IS-06	Network Control	v1.0	ネットワークの管理
IS-07	Event & Tally	v1.0	機器ステータスの配信
IS-08	Audio Channel Mapping	v1.0	音声チャンネルの入れ替え
IS-09	System	作業中	制作システムのパラメータ管理
IS-10	Authorization	作業中	認可情報の管理

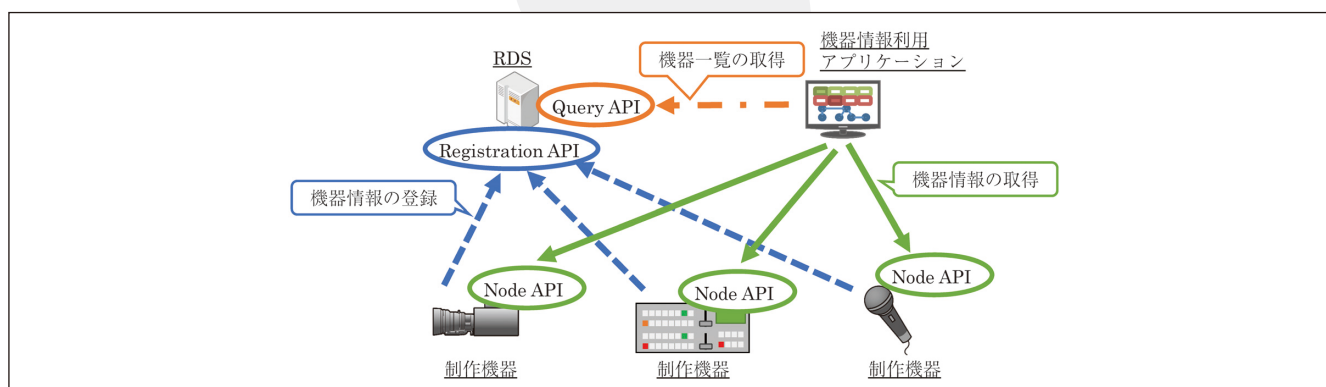


図1 IS-04各APIの概要

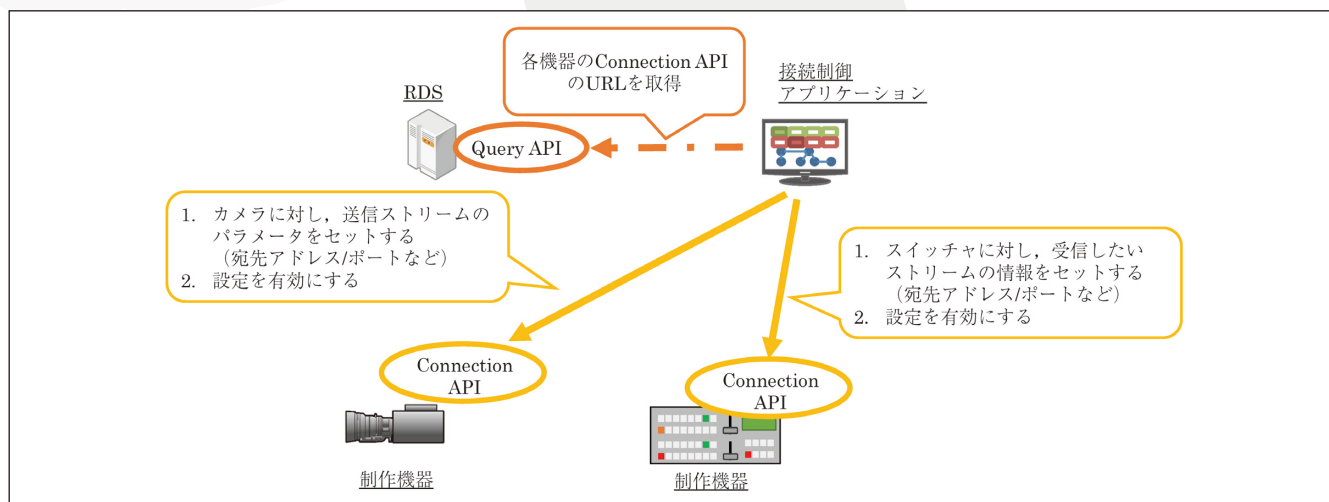


図3 IS-05 Connection APIの動作概要

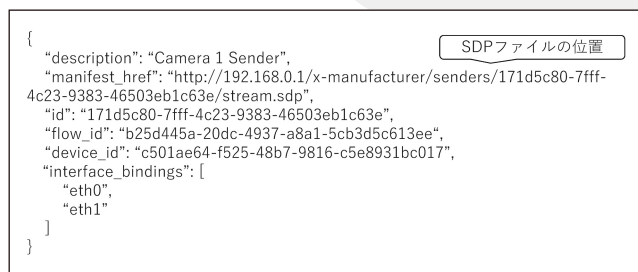


図2 登録パラメータの一部（送信機能）

RDSに登録された情報に変更があった場合にすぐに通知を受けることができます。

IS-04ではJSON (JavaScript Object Notation)¹⁸⁾で情報を記述し、HTTPによってAPIにアクセスします。これらの技術はインターネットで広く用いられており、開発や検証するための敷居が低いことも特徴です。図2に送信機能の情報を登録する際のJSONによる記述の一例を示します。JSONでは“key”と“value”のペアによりデータを表現しており、人も読みやすい記述となっています。例えば、“manifest_href”というkeyには、SMPTE ST2110で定められている送信データのパラメータを記述したSDPファイルの取得URLがvalueとして記述されていることがわかります。

3.3 IS-05 Device Connection Management

IS-05 Connection APIはカメラとモニタといった、送信機器と受信機器を接続するための機能を提供します。執筆時点の最新バージョンは1.0.1 (stable)です。SMPTE ST2110ファミリーでは、送信フローの宛先IPアドレスや映像パラメータなどをSDP (Service Description Protocol)¹⁹⁾ファイルに記述しますが、その受け渡し方法については定義されていません。IS-05を用いることで、機器間でSDPファイルの情報を交換することができます。

NMOSが検討され始めた当初は、IS-04にも同様の機能がありましたが、時間指定切り替えなどの高度な接続機能

を追加する際に、分離して独立した規格となりました。なお、IS-04の機器接続機能はバージョン2以降で廃止予定となっています。

図3に動作イメージを示します。まず、接続を制御するアプリケーションはIS-04のQuery APIなどによって制御対象の機器が提供するConnection APIのアドレスを取得します。次に、送信機器のパラメータ（送信フローの宛先IPアドレスやFECの有無など）を送信機器のConnection APIを利用して取得・設定します。その後、受信機器のConnection APIを利用して、受信フローのパラメータを設定し受信を開始します。

IS-05でもJSONを用いて設定情報を表現し、HTTPによって伝送します。

3.4 IS-06 Network Control

IS-04/05は、番組制作機器を対象としたAPIでしたが、IS-06はネットワークを対象としたAPIについて定義しており、ネットワークポロジの取得、送信するフローの帯域保証、登録されたフロー以外は転送しないことによるセキュリティの強化などの機能を提供します。執筆時点での最新バージョンは1.0です。

図4に適用するシステムの概要を示します。IS-06はネットワーク機器を直接制御するのではなく、Network Controllerに対して用いるAPIです。このようにすることで実際の機器の制御方法は抽象化され、番組制作システムの制御システム（図中のBroadcast Controller）は、どのベンダーのネットワーク機器でも同じように制御できるようになります。

まずBroadcast ControllerはIS-04 Query APIなどを利用して得たシステム内の機器の情報を、Network Controllerに登録します。この情報と、ネットワークスイッチからLLDP (Link Layer Discovery Protocol)²⁰⁾によって得られた情報を統合し、Network Controllerではネットワークポロジを計算します。また、登録された機器以外からのフローをスイッチで遮断することで不正なトラフィックが

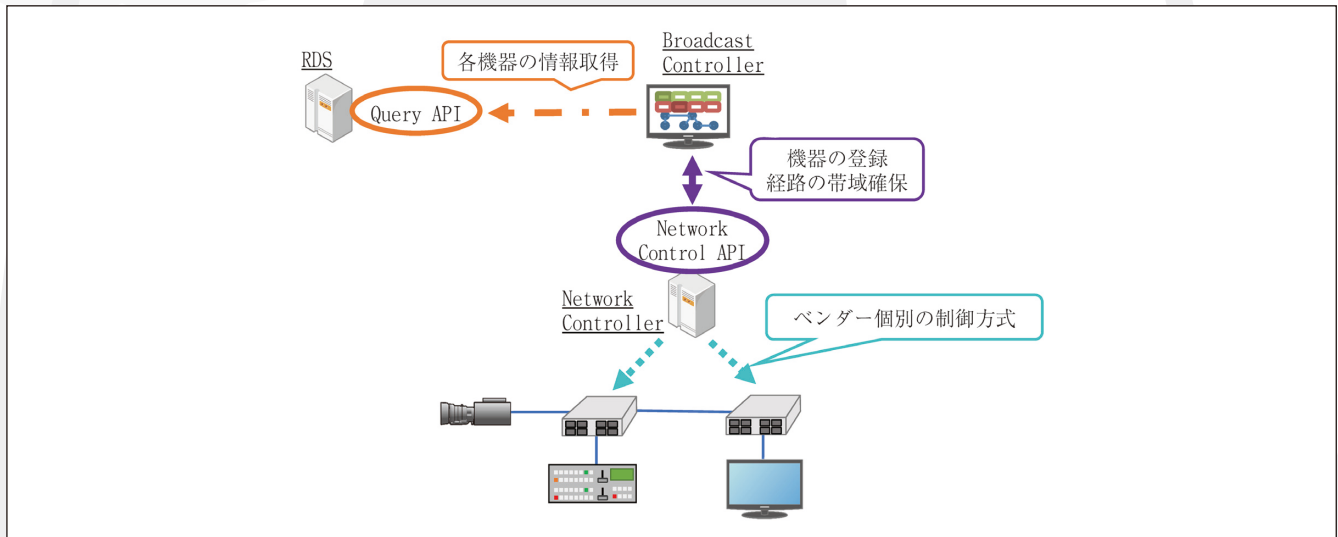


図4 IS-06 Network Control APIの概要

流れることによるサービスへの影響を抑えることができます。次に、IS-05などによって番組制作機器の間でフローを伝送する際に、IS-06のAPIを利用してネットワーク内の転送ルートを作成を行います。Network Controllerではネットワーク内の帯域を管理しており、新しく伝送を開始した場合に既存のフローへ影響を与えるかどうかを確認してからネットワーク機器の制御を行います。IS-05だけを用いてフローの送受信切り替えを行うと輻輳が発生する可能性があります。IS-06を組み合わせることで安全に送受信設定ができるようになります。

IS-04と同様に、情報はJSONで表現し、HTTPによって伝送します。

3.5 IS-07 Event & Tally

IS-07はタリーなどの機器の状態を伝送するための規格です。執筆時点の最新バージョンは1.0です。

情報を伝送するトランスポートとしてWebSocketに加えてMQTT (Message Queue Telemetry Transport)²¹⁾を採用している点が特徴です。MQTTではメッセージブローカーと呼ばれるサーバが送信機器と受信機器の間に設置さ

れ、メッセージの保持・配信を行います(図5)。このため、一対一でコネクションを確立して情報をやりとりするWebSocketと比べて少ない負荷で送信機器が不特定多数の受信機器に対してメッセージを配信することが可能になります。

利用例として、タリーの他に機器温度のモニタリングや音声ミキサーのVUメータ値の配信などが考えられています。

3.6 IS-08 Audio Channel Mapping

IS-08は送信機器や受信機器で音声チャンネルの割り当てを変更する機能を提供します。例えばSMPTE ST2110-30では、一つのIPフローで複数の音声チャンネルを送信することができます。受信装置でこの音声チャンネルの割り当てを変更することで、再生するチャンネルを自由に選択できるようになります。従来、このような制御は機器ベンダーが独自に実装していましたが、共通化の要望が強くNMOSのAPIとして標準化されました。他の多くのNMOSシリーズと同じように、JSONを用いて設定情報を表現し、HTTPによって伝送します。執筆時点の最新バージョンは1.0となっています。

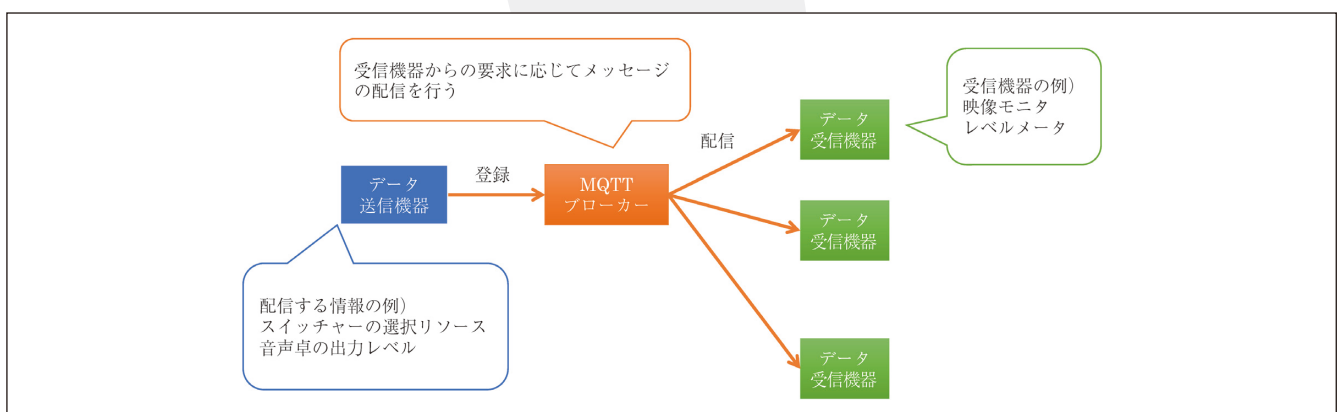


図5 IS-07 Events APIの概要

3.7 IS-09 System, IS-10 Authorization

執筆時点において、IS-09とIS-10が作業中となっています。IS-09は制作システム全体で共有する情報を取得する方法について扱っています。例えば、システム毎にIDを設定しておき、機器が電源を入れた際にシステムIDを再度取得して前回と同じIDであれば保存された設定のまますぐに動作を開始することで起動時間を短縮することなどが考えられています。IS-10は制御アプリケーションがどのようなAPIを実行することができるかという権限を管理するための仕組みを検討しています。この仕組みによって、別の番組制作に利用している機器を誤って制御してしまうといったことを防ぐことができますようになります。

4. 運用ガイドラインJT-NM TR-1001-1について

2018年12月に、JT-NM (Joint Taskforce on Networked Media: AMWA, EBU, SMPTE, VSF)からなる制作設備のIP化を検討する団体²²⁾は機器を接続して最小限の設定のみで運用できるIP制作システムを実現するための「構成」や「機器の動作」などについて記載したTR-1001-1を発行しました²³⁾。制御方式の規格ではありませんが、NMOSを運用するための推奨環境・動作が規定されていますので、参考に紹介します。

図6はTR-1001-1で想定するネットワーク構成です。信号伝送を2重化してシステムの安定性を向上させるSMPTE ST 2022-7²⁴⁾に対応したメディア用ネットワーク2つ (AmberとBlue) と、制御用のネットワーク一つから構成されます。制御用ネットワークには、IS-04のRDSの他、DHCPサーバやDNSサーバ、Broadcast Controllerが存在します。

ネットワークに番組制作機器を接続すると、機器はDHCPにより各IPインタフェースにIPアドレスを設定します。IPアドレスの設定後、機器はDHCPサーバから取得したDNSサーバのIPアドレスに対し、IS-04 RDSのIPアドレスの解決を行います。こうして得られたIS-04 RDSのRegistration APIを使い機器情報を登録すると、Broadcast ControllerにはQuery APIを通じて機器の情報が伝えられ、監視や制御が可能となります。

このような環境を構築することで、機器をネットワークに接続するだけで利用が可能になり、システムを運用する負担を軽くすることができます。

5. むすび

本稿では、IP制作システムにおける制御方式の役割と、その標準規格であるAMWA NMOSについて紹介しました。NMOSによって、機器の管理をIPネットワーク経由で行う手段 (API) が整備されつつあります。今後は、このAPIを活用して、例えば、障害発生時の原因特定や一時対

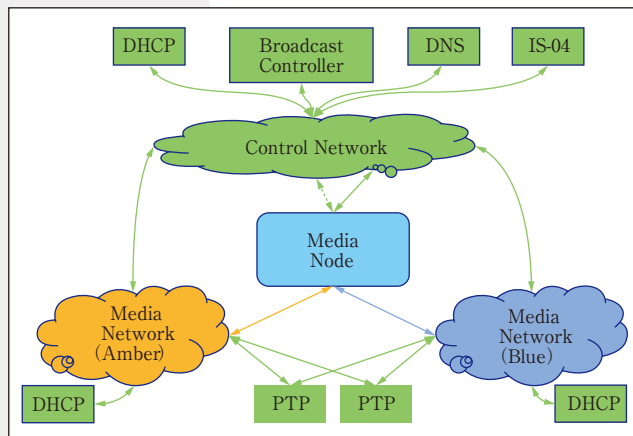


図6 JT-NM TR 1001-1で前提となるネットワーク環境

応までを自動的に行うことができるような高度な機能をもったアプリケーションの開発が進むことが期待されます。

NMOSでは新しい機能が次々と提案されていますが、その内容はgithubで公開されていますので是非チェックしてみてください。
(2019年5月31日受付)

〔文 献〕

- 1) SMPTE ST 2059-1: "Generation and Alignment of Interface Signals to the SMPTE Epoch" (2015)
- 2) SMPTE ST 2059-2: "SMPTE Profile for Use of IEEE-1588 Precision Time Protocol in Professional Broadcast Applications" (2015)
- 3) SMPTE ST 2110-20: "Professional Media Over Managed IP Networks: Uncompressed Active Video" (2017)
- 4) <https://github.com/Lawo/ember-plus/wiki>
- 5) SMPTE RDD38: "Networked Device Control Protocol" (2016)
- 6) <https://www.amwa.tv/>
- 7) <https://aimsalliance.org/>
- 8) EBU R152: "STRATEGY to ACCELERATE ADOPTION of AN OPEN DISCOVERY & CONNECTION PROTOCOL (NMOS) " (2019)
- 9) <https://amwa-tv.github.io/nmos/>
- 10) <https://amwa-tv.github.io/nmos-discovery-registration/>
- 11) <https://amwa-tv.github.io/nmos-device-connection-management/>
- 12) <https://amwa-tv.github.io/nmos-network-control/>
- 13) <https://amwa-tv.github.io/nmos-event-tally/>
- 14) <https://amwa-tv.github.io/nmos-audio-channel-mapping/>
- 15) <https://amwa-tv.github.io/nmos-system/>
- 16) <https://amwa-tv.github.io/nmos-authorization/>
- 17) RFC6763: "DNS-Based Service Discovery" (2013)
- 18) RFC8259: "The JavaScript Object Notation (JSON) Data Interchange Format" (2017)
- 19) RFC4566: "SDP: Session Description Protocol", 4566 (2006)
- 20) IEEE 802.1AB: "Station and Media Access Control Connectivity Discovery (LLDP) " (2016)
- 21) <http://mqtt.org/>
- 22) <http://jt-nm.org/>
- 23) JT-NM TR-1001-1:2018 v1.0 (2018)
- 24) SMPTE ST 2022-7: "Seamless Protection Switching of RTP Datagrams" (2019)



小山 智史 2005年、東京工業大学大学院修士課程修了。同年、NHK入局。2009年より、放送技術研究所に勤務。現在、番組制作システムのIP化の研究・開発に従事。正会員。

IP番組制作システムの検証事例

正会員 北島正司[†], 正会員 河原木政宏^{††}, 正会員 小山智史^{††}

1. まえがき

IPネットワーク上で映像や音声を伝送するためのパケットフォーマットの標準規格であるSMPTE ST 2110スイート^{1)~4)}が2017年に発行され、展示会でも特設コーナーが設けられるなどライブ番組制作を行うための番組システムのIP化について放送業界の関心が高まっています。

SDIで構築されている番組制作システムをIPネットワークで実現するため、前述したSMPTE ST 2110の規格化の他、同期を伝送するPTP (Precision Time Protocol) を番組制作用途で用いる際のパラメータを定めるSMPTE ST 2059^{5) 6)}や機器の制御方式の標準規格であるAMWA NMOS⁷⁾などの規格化が進められています。実際に番組制作を行うためには、要素技術の開発の他に、機器の相互接続性検証やIPネットワーク構築手法の確立など検討すべき事項が多くあります。

本稿では、IP番組制作システムの実現へ向けて筆者らが実施した「IPライブプロダクション相互接続実験」と「2018 NHK杯フィギュア音声リモートプロダクション検証」の2つの事例について紹介します。これらの事例が、今後の機器の開発やIP制作システムを構築する際の参考になれば幸いです。

2. IPライブプロダクション相互接続実験

筆者らは、多くのベンダにご協力いただき、SMPTE ST 2110スイートやAMWA NMOSなどのオープンな規格に対応した番組制作機器の相互接続実験を2018年6月、12月、2019年7月と3度実施してきました。本実験の主な目的は以下の3点です。

- (1) 規格に対応した機器の相互接続性確認
- (2) IPライブプロダクションに適したネットワーク構造の検討

[†] NHK 放送技術局

^{††} NHK 放送技術研究所

"IP Adoption in Broadcast Industry (5): Experiments of IP-based Program Production System" by Shoji Kitajima (NHK Broadcast Engineering Department, Tokyo), Masahiro Kawaragi and Tomofumi Koyama (NHK Science and Technology Research Laboratories, Tokyo)

(3) 監視手法の確認

実験ではリモート制作を想定し、複数の現場やスタジオをNHK放送技術研究所(以下技研)内の実験スペースに構築してそれらをバックボーンネットワークを介して接続する形としました。実験で構築した制作システムの概要を図1に示します。番組制作機器やネットワーク機器、PTP GM (PTP Grand Master Clock)などはすべて複数ベンダの機器を用いています。また、パケットキャプチャ装置やネットワーク監視装置も利用して、制作システムの監視や障害対応時の調査方法についても検証しました。

2.1 機器の相互接続検証

従来の制作システムではSDIという共通のインターフェースを利用することで、異なるベンダの機器であっても信号のやりとりが可能となっています。同様にIPライブプロダクションシステムにおいても機器の相互接続を可能とするための規格が発行されています。相互接続実験では、映像伝送にSMPTE ST 2110-20、音声伝送にSMPTE ST 2110-30、同期伝送にPTPを用いて異なるベンダの機器間で正常に信号の受け渡しができるかどうかを検証しました。また、機器制御の標準方式であるAMWA NMOSについても各機器の対応状況を検証しました。

(1) 映像

SMPTE ST 2110-20に対応した機器間では、多くの組み合わせで正常に映像伝送ができることが確認できました。4K映像の伝送に対応している機器には、SDIでも12G-SDIによる伝送と3G-SDIを4本用いた伝送の2方式が存在するように、1つの4K ST 2110-20フローとして伝送する機器と、4つの2K ST 2110-20フローとして伝送する機器が存在しています。これらの機器間では相互接続ができないため注意が必要です。一方で、4つの2K ST 2110-20フローとして伝送する機器が出力する各フローは、独立した映像信号として扱うことが可能なため、2Kに対応したマルチビューアーなどで受信して表示することができ、監視に有効でした。

(2) 音声

音声制作の分野では、映像よりも先行してIPネットワークの利用が始まっており、SMPTE ST 2110-30が発行される以前からDante⁸⁾やAES67⁹⁾などの規格が発行されてい

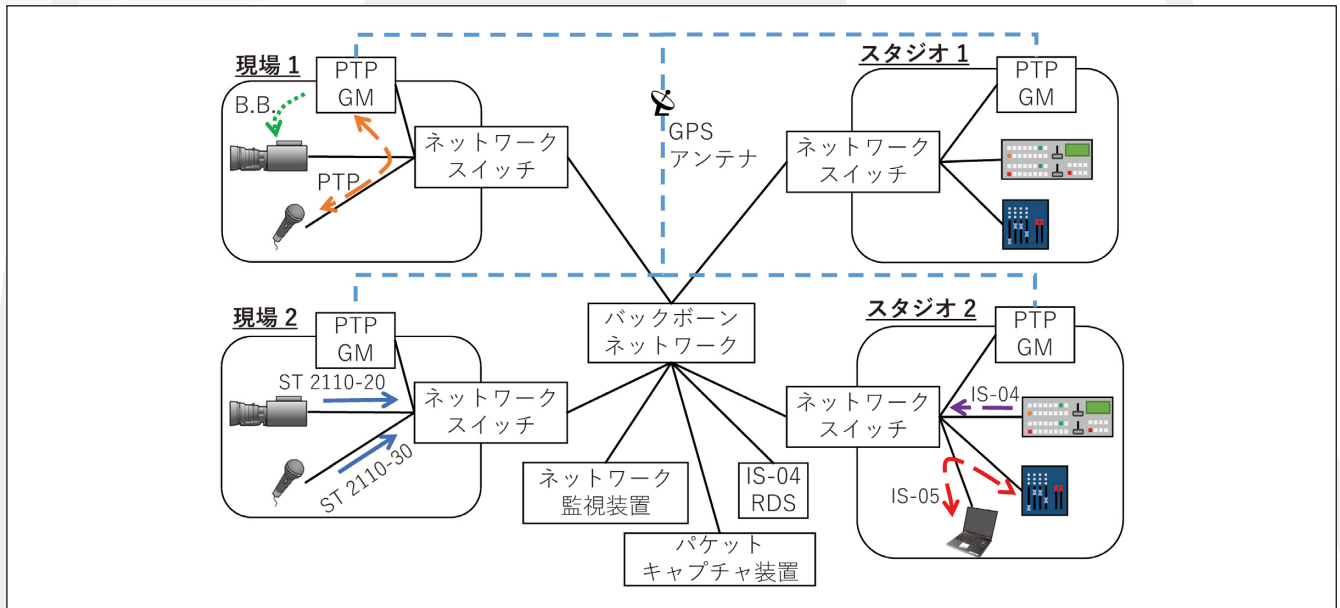


図1 実験用制作システム概要

ました。SMPTE ST 2110-30はパケットフォーマットがAES67と共通であることから、実験ではAES67対応機器も対象としました。また、広く利用されているDante規格もAES67モードを備えているため、Dante対応機器との接続実験も実施しました。

SMPTE ST 2110-30/AES67/Dante AES67モードのいずれの組み合わせでも音声の再生を確認することができました。ただし、受信のために入力するSDP (Session Description Protocol) ファイル¹⁰⁾ (音声のチャネル情報やIPフローのアドレス情報などを含んだファイル) の記述内容にベンダ独自のフィールドを含むことがあり、各機器に応じた内容に修正しないと受信できない場合もありました。

(3) 同期

リモート制作を想定し、各ベニューにGPSにロックしたPTP GMを準備しました。制作機器のうち、PTPに対応した機器はPTP SlaveとしてGMに同期し、PTPに対応していない機器については、PTPの時刻情報からSMPTE ST 2059-1に基づいて生成したBlack Burst (以下BB) 信号を入力することですべての機器が同じ同期源にロックした環境を実現し、フレームシンクロナイザを用いずにベニュー間の信号受け渡しができるかどうかを検証しました。

各ベニューに設置したPTP GMの位相差は1μs以下であり、フレームシンクロナイザ (FS) を使うことなくベニュー間を跨いで信号を伝送してもフレームスキップなどなく運用できることが確認できました。また、各ベニューのネットワークスイッチをBC (Boundary Clock) またはTC (Transparent Clock) として運用し、正常に動作することも確認しました。

機器をネットワークに接続した際にPTPの同期がうまく

いかないケースもありました。機器がSMPTE ST 2059-2で指定されているパラメータに対応していなかったり、マルチキャストで送信されてくるPTPパケットに対してIGMP (Internet Group Management Protocol)¹¹⁾ による受信要求を行っていなかったりすることが原因でした。実運用の際には対応状況を確認する必要があります。

(4) 制御

IP制作システムの大きな特徴として、IPネットワーク経由で機器の制御ができることが挙げられます。これまでの実験では、番組制作機器をネットワーク上の管理サーバへ登録する機能を提供するAMWA NMOS IS-04¹²⁾ と、カメラとモニタといった送信機器と受信機器を接続するための機能を提供するAMWA NMOS IS-05¹³⁾ について検証しました。

IS-04は、AMWA NMOSシリーズで最初に発行された規格であり、実験を重ねるたびに対応機器数が増加しています。登録された機器情報から、機器のIPアドレスや死活情報、実装している機能について知ることができます。課題としてはルータを超えた管理サーバの発見への対応があります。IS-04では、DNS-SD¹⁴⁾ という仕組み用いて各機器が管理サーバを探索して自身の情報を登録します。IS-04の初期バージョンでは、マルチキャストDNS¹⁵⁾ によって管理サーバの情報を探索していましたが、この方法ではルータ越しの管理サーバを探索することができません。現在ではルータを超えた探索が可能なユニキャストDNSを用いた方法が推奨されていますが、制作機器の対応はまだ進んでいないため今後の実装が待たれます。

IS-05もIS-04ほどではありませんが、徐々に対応機器数が増えてきています。IPネットワーク上で伝送される映像

や音声を受信するためには、IPアドレスや映像・音声のパラメータを設定しなければなりません。手でその設定を行うと、設定ミスが発生しますし、システムの規模が大きくなると作業量が増え対応できなくなってしまう。実験では、受信確認のために頻繁に受信するフローを切り替えましたが、IS-05に対応した機器では1クリックで設定を変更することができ、とても効率的でした。迅速に設定変更を行うためにも、今後対応することが必須の規格と言えるでしょう。

2.2 ネットワーク構造

番組制作用のIPネットワークを構築する上で、どのような構造のネットワークを用いるかは重要なポイントとなります。筆者らはシステム全体を1つのL2ネットワークとして構築することから検証を始めました(図2)。L2ネットワークでは、IPパケットのルーティング設定を行う必要がなく、比較的容易にネットワークを構築することができます。一方で、マルチキャストパケットはブロードキャストパケットと同様にネットワーク全体へ転送されてしまうため、輻輳が発生しやすいというデメリットがあります。不必要なマルチキャストの転送を抑制するためIGMP snoopingという機能を用いましたが、ネットワークベンダ毎にデフォルトパラメータが異なったり、各ベニューでマルチキャストトラフィックが発生するという番組制作システムの特徴により、L2ネットワークによる運用ではトラフィックの制御が困難で、実験中に輻輳が発生し映像や音声の伝送に影響が出ることが多くありました。

そこでベニュー内はL2ネットワークで、ベニュー間はL3ネットワークで接続する構成を検証しました(図3)。L3スイッチでは受信要求があったマルチキャストフローのみを転送するため、意図しない輻輳の発生を防ぐことができ

ます。ただし、L3ネットワークではベニュー間を跨いでパケットを転送するために、ユニキャストとマルチキャストのそれぞれでルーティングの設定が必要となります。ルーティングプロトコルによる自動設定を行うため、ユニキャストルーティングにOSPF¹⁶⁾、マルチキャストルーティングにPIM¹⁷⁾を用いました。この変更により、必要とするマルチキャストトラフィックのみがベニュー間を流れるようになり、輻輳の発生をかなり抑制することができました。フローの管理を容易にし、輻輳による障害を防ぐことができるため、番組制作システムで用いるIPネットワークはL3ネットワークを用いて設計する必要があると言えます。

実験で用いた制作機器で利用できるマルチキャストルーティングのPIMには、PIM-SMとPIM-SSM¹⁸⁾の2つの方式がありました。PIM-SMでは、マルチキャストを受信する際にパケットの宛先マルチキャストグループのみを指定するのに対し、PIM-SSMでは宛先マルチキャストグループと送信元IPアドレスを指定します。PIM-SMを利用する場合、マルチキャストパケットはランデブーポイント(RP)と呼ばれるルータへいったん転送されます。実験ではPIM-SSMに対応していない機器があったため、すべての機器をPIM-SMで動作させたところ非圧縮映像のような高レートフローがRPに集中し、輻輳が発生してしまいました。PIM-SSMではマルチキャストパケットはRPを経由せずに送信元から受信機器へ転送されるため、このような問題は発生しません。番組制作機器はPIM-SSMへの対応が必要です。

2.3 監視手法

IP番組制作システムを実現するために取り組む必要のある課題として、IPネットワークの監視があります。IPネットワークでは、ケーブルを接続するだけでなく前述したと

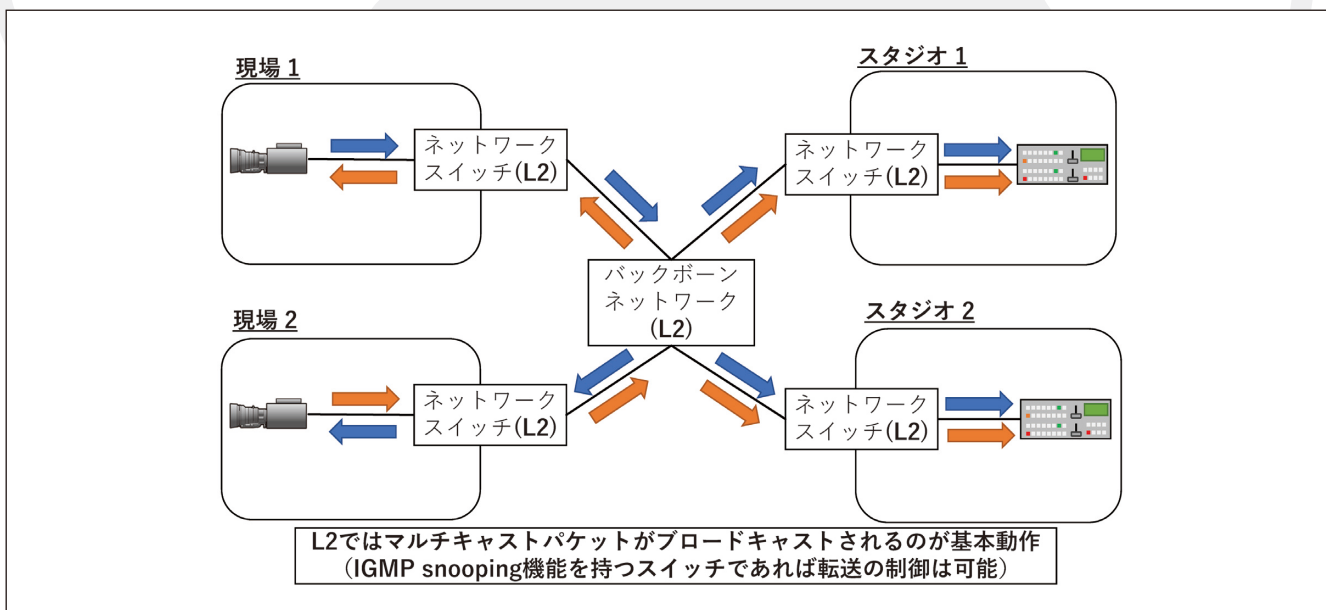


図2 L2ネットワークでのマルチキャスト転送

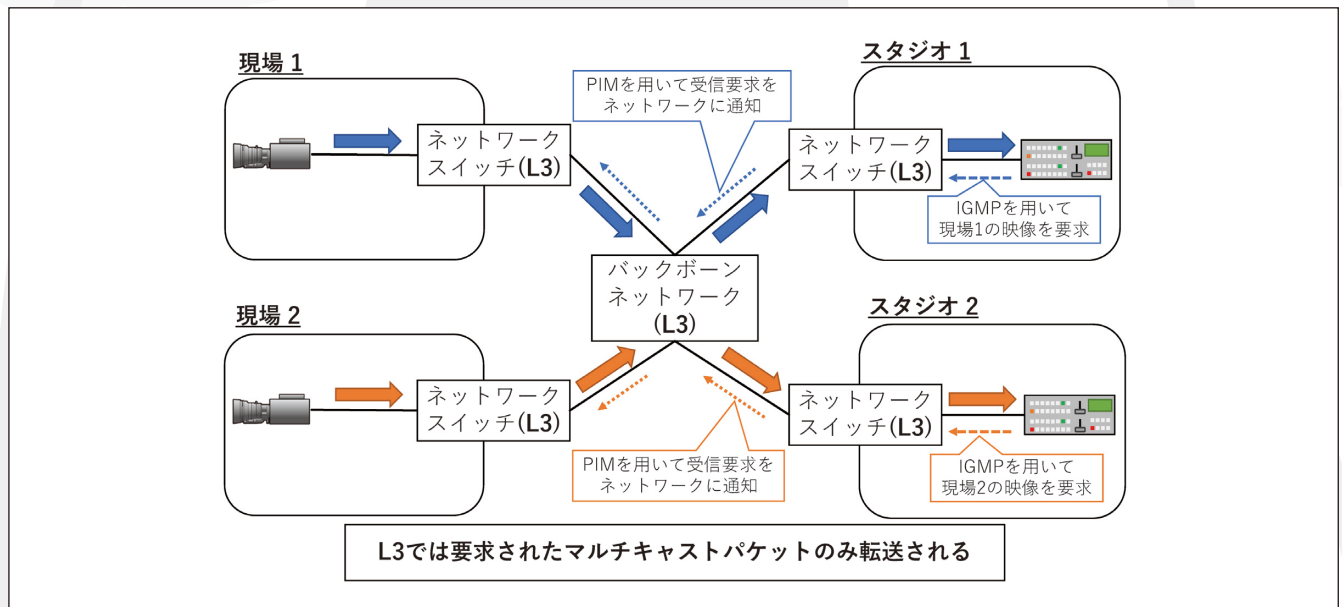


図3 L3ネットワークでのマルチキャスト転送

おり、ネットワーク機器のルーティング設定や制作機器で受信したいフローのアドレス設定などの作業を正しく行わないと信号を送受信することができません。また、正しく設定を行ったとしても、伝送路上で輻輳が発生し、パケットが消失することもあります。映像や音声再生できない場合、「ネットワークからパケットが到達していないため」、「機器内の処理で不具合があるため」、「同期がずれているため」などさまざまな原因が考えられ、原因を迅速に絞り込むためのネットワークの可視化ツールがとても重要となります。ネットワークスイッチの各ポートの利用帯域情報は取得可能ですが、個々の映像・音声フローの監視は難しく、また制作機器ではデータの受信状況や同期状態などを詳細に取得できる機器はあまりありませんでした。機器の状態を取得するための機能を充実させていくことが望まれます。

2.4 その他の課題

制作機器のIPアドレス設定やネットワークスイッチに機器を接続するポートの間違いなど、基本的だと考えられる箇所でも設定ミスは多く発生しました。DHCP (Dynamic Host Configuration Protocol)¹⁹⁾やL2認証²⁰⁾など、すでに通信業界では当たり前のように入用いられている機能を使って動作を自動化していく取り組みが必要であると考えられます。

また、IPネットワークを利用することで、意図しないアクセスが機器に対して発生する可能性があり、今後はセキュリティへの対応も考えなくてはならないでしょう。機器上で動作するアプリケーションで使用するライブラリーを最新版にアップデートすることや、不要なサービスは実行しないといった、一般的な通信機器が行う対策の他、アプリケーション自体も認証や通信の暗号化などの機能に

対応することが求められます。

3. 2018 NHK杯国際フィギュアスケート音声リモート制作

2018年11月9日から11日に広島市の広島県立総合体育館グリーンアリーナで開催された「2018 NHK杯国際フィギュアスケート競技大会」の8K制作において、22.2ch音声ミックスをNHK広島拠点放送局（以下広島局）でIPリモート制作しました²¹⁾。リモート制作で制作した音声は全国5ヵ所のPV会場で利用された他、新4K8K衛星放送で番組として放送されました。これは、NHK内でSMPTE ST 2110スイートを用いて制作した番組が放送された初の例となりました。本章では、その技術的な概要と得られた効果について紹介します。

3.1 システム概要

今回のリモート制作では、8K番組制作のうち、22.2ch音声制作をIPリモート制作で行い、映像制作は従来通り現場中継車で実施しました。音声制作のリモート制作システムを組むにあたり、考慮した要件は以下の通りです。

- (1) 現場音声素材256chをST 2110-30で伝送する。
- (2) 広島局T1スタジオに仮設する設備で22.2chの音声リモートミックスを行い、ミックス後の音声を現場に戻して本線音声として利用する。
- (3) 遠隔地からの最低限のバックアップの検証として、東京の技研でも5.1chの音声リモートミックスを同時に実施する。
- (4) 高品質で低遅延な8K映像をモニタしながら音声ミックスができる環境を準備する。
- (5) 現場と広島局間の伝送は、ST 2022-7²²⁾により冗長化されたシステムで行う。

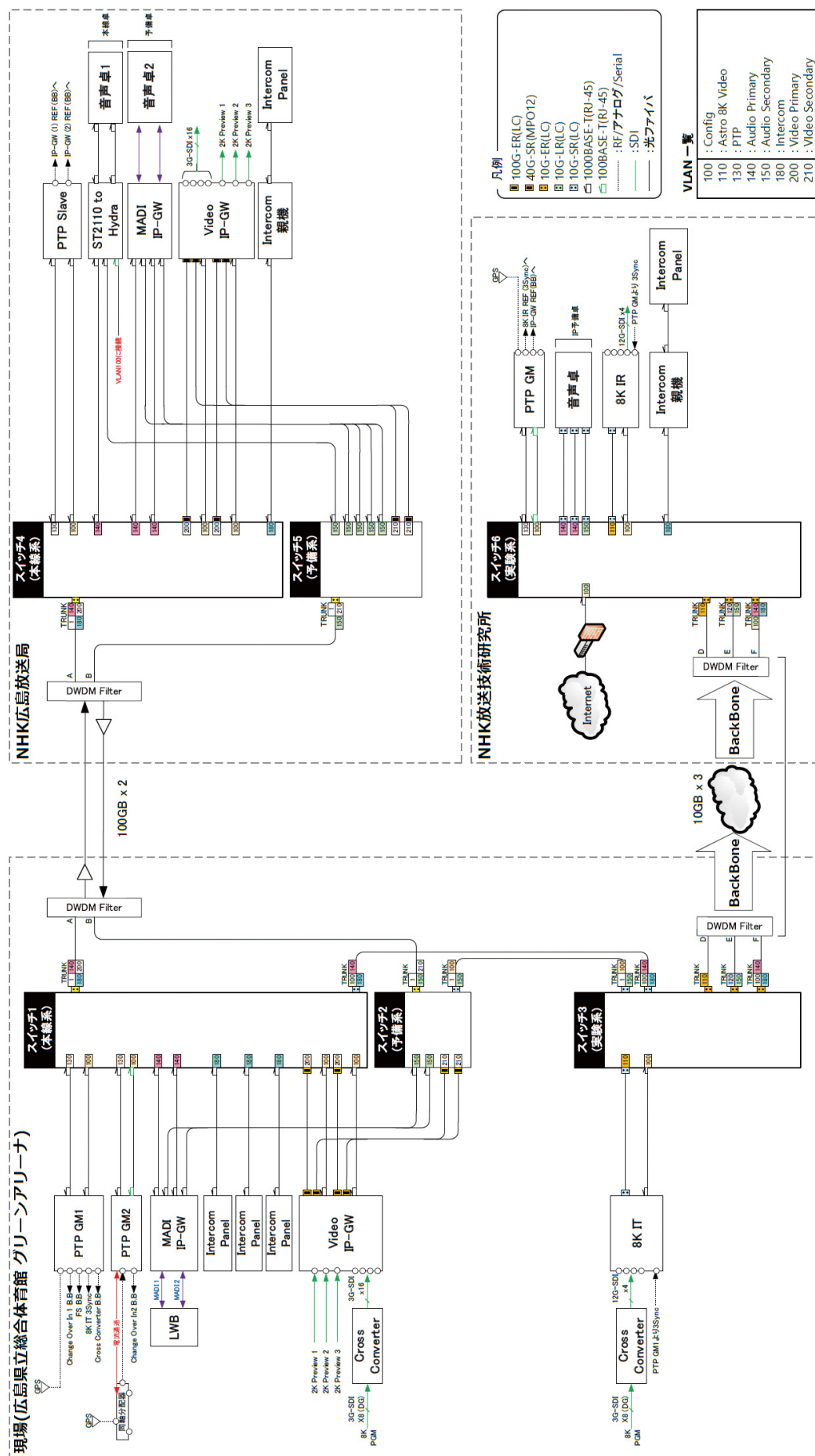


図4 音声ミックスのためのIPリモート制作システム図

上記要件を元に、図4に示す系統を設計しました。

ネットワークは現場を中心として広島局と技研までそれぞれ回線を伸ばす構成としました。利用可能なダークファイバや長距離回線、機材を考慮し、現場と広島局間は100 GbEを2系統、現場と技研間は10 GbEを3系統構築しました。全体をL2のネットワークとし、マルチキャストトラフィックの不必要な転送を抑制するIGMP snoopingを有効にしました。また、映像や音声、インカムといったメディア種別毎にVLAN (Virtual Local Area Network) でネットワークを分離し、障害時の影響範囲を限定できるようにしました。

同期方式はST 2059-1/2を用いてすべての機器をGPSに同期できるような構成としました。音声系はIP制作システム上で作業を行う一方で、映像系は現場中継車でSDIシステムによる制作を行っており、IPとSDIが混在したシステムに同期信号を提供する必要がありました。そのため、BB信号の出力が可能なPTP GMを現場に設置し、そのBB出力を中継車を駆動する同期信号として供給しました。広島局の仮設設備ではGPS信号の受信系統を構築することが難しかったため、現場のPTP GMを参照するようにネットワークを設定しました。一方、技研ではGPSにロックしたPTP GMを設置しました。現場と技研で、制作機器が参照するPTP GMは異なりますが、どちらのPTP GMもGPSという共通の時刻源にロックしているため、FSやサンプルレートコンバータを利用せずに信号の受け渡しが可能になっています。PTP GMの障害を考慮し、現場のPTP GMは2重化構成としました。

音声伝送にはST 2110-30を利用しましたが、現場のシステムや映像中継車との接続はMADI (Multichannel Audio Digital Interface) を利用していたため、MADIとST 2110-30を相互変換できるIP-GWを用いて音声信号をIP化して広島局や技研へ伝送しました。音声ミックスにはST 2110-

30を直接受信できる音声卓を利用しました。図5は広島局に仮設した音声ミックス環境です。音声の中継車と比較して広さに余裕があり、スピーカを規格通りの位置に配置できたり、スピーカとオペレータの間に機器によって塞がれたりしないため、理想的な制作環境を構築することができました。

リモート制作において、現場で制作している場合と変わらない制作環境を提示することは重要です。そのため、リモート制作時にモニタする8K映像として広島局へST 2110-20による非圧縮の8K映像を伝送しました。技研への回線は8K非圧縮映像を送るだけの帯域がないため、技研で開発した軽圧縮伝送装置²³⁾を用いて8 Gbpsに圧縮して伝送しました。

3.2 ネットワーク監視

ネットワークスイッチを監視するため、ネットワーク業界で広く使用されているSNMP (Simple Network Management Protocol)²⁴⁾を用いて1分ごとにスイッチに流れているトラフィック量や温度を集計し、グラフ化して監視できるようにしました。SNMPコレクタにはTelegraf²⁵⁾、取得したデータを保存するデータベースにはinfluxDB²⁶⁾、データのグラフ化にはGrafana²⁷⁾とオープンソースソフトウェアを組み合わせて構築しました。

また、一部のスイッチについては、ベンダが開発する専用の監視システムも併用しました。これにより、PTPの同期状態などSNMPではわからないスイッチの状態をリアルタイムに把握できるようになりました。

3.3 IPリモート制作の効果

現場から広島局への伝送遅延は数ミリ秒であり、中継現場の素材を高品質なまリアルタイムに共有することができました。実際にミックスを行った担当者からは「音声制作車での制作と比較し、理想のモニタ環境でミックスを行うことができた。特に、空間が広いことで音のつながり聞



(a) 22.2chスピーカの配置



(b) オペレータ視点の環境

図5 広島局に仮設した音声ミックス環境

の確認がしやすく、音声制作車の制作では気がつかない細かい部分のバランスをとることができた。」という感想があり、IPリモート制作によって番組品質を向上できたと考えています。

今回は広島局の設備を仮設しましたが、将来IP制作に対応した常設スタジオでの制作が可能になれば準備にかかる時間の短縮も見込めます。

現場から遠く離れた技研でのミックス作業も通常の作業と変わりなく行うことができました。音声ストリームが現場から技研に伝送されて、ミックス後の音声広島局に届くまでの遅延時間は50ms程度でした。この時間は8K映像中継車のスイッチャなどで発生する遅延よりも小さく、もし技研でリモート制作した音声信号を本線音声としても、映像に遅延を加える必要がありません。このように、遠隔地にある別拠点からのバックアップも可能であることがわかりました。

また、IPリモート制作では素材の共有が容易であることが人材育成において大きなメリットであることもわかりました。ネットワークとIP対応の制作スタジオがあれば、現場に行けなくても現場同様に本番で使用する素材を用いた作業ができます。そのため、番組制作を行っている裏で、各地では若手がスキルアップのためにミックスを行い、番組用にチーフがミックスした音声と聞き比べるといったことができるようになります。これまでよりも簡単に研修の場を持てるようになることは番組品質の向上にも役立つと言えます。

4. むすび

本稿ではIP制作システムを構築する際の相互接続性やネットワーク構築手法などについて実験した「IPライブプロダクション相互接続実験」と、実際に22.2chの音声ミックスをIPリモート制作によって行った「2018 NHK 杯国際フィギュアスケート音声リモート制作」について紹介しました。

IP制作システムによる番組制作は、ネットワーク監視や機器接続の自動化対応など改善すべき点がありますが、SMPTE ST 2110スイートにより基本的な相互接続性は実現されており、用途によってはすでに実用できる段階になっていると言えます。今後ますますIP制作システムによる番組制作の事例を目にすることが多くなるでしょう。その際には利用している標準規格に加え、番組制作システムの運用を支える機器の制御方式やネットワーク監視方式、さらには運用により得られた効果などにも注目するとより理解が深まると思います。

最後に、本稿で紹介した実験に協力いただいたベンダ、回線事業者の皆様にご礼を申し上げます。

(2019年7月31日受付)

〔文 献〕

- 1) SMPTE ST 2110-10: "Professional Media Over Managed IP Networks: System Timing and Definitions" (2017)
- 2) SMPTE ST 2110-20: "Professional Media Over Managed IP Networks: Uncompressed Active Video" (2017)
- 3) SMPTE ST 2110-30: "Professional Media Over Managed IP Networks: PCM Digital Audio" (2017)
- 4) SMPTE ST 2110-40: "Professional Media Over Managed IP Networks: SMPTE ST 291-1 Ancillary Data" (2018)
- 5) SMPTE ST 2059-1: "Generation and Alignment of Interface Signals to the SMPTE Epoch" (2015)
- 6) SMPTE ST 2059-2: "SMPTE Profile for Use of IEEE-1588 Precision Time Protocol in Professional Broadcast Applications" (2015)
- 7) <https://www.amwa.tv/>
- 8) <https://www.audinate.com/solutions/dante-overview>
- 9) AES 67: "AES standard for audio applications of networks - High-performance streaming audio-over-IP interoperability" (2015)
- 10) RFC 4566: "SDP: Session Description Protocol" (2006)
- 11) RFC 3376: "Internet Group Management Protocol, Version 3" (2002)
- 12) <https://amwa-tv.github.io/nmos-discovery-registration/>
- 13) <https://amwa-tv.github.io/nmos-device-connection-management/>
- 14) RFC 6763: "DNS-Based Service Discovery" (2013)
- 15) RFC 6762: "Multicast DNS" (2013)
- 16) RFC 2328: "OSPF Version 2" (1998)
- 17) RFC 7761: "Protocol Independent Multicast - Sparse Mode (PIM-SM) : Protocol Specification (Revised) " (2016)
- 18) RFC 4607: "Source-Specific Multicast for IP" (2006)
- 19) RFC 2131: "Dynamic Host Configuration Protocol" (1997)
- 20) IEEE 802.1X-2010: "IEEE Standard for Local and metropolitan area networks-Port-Based Network Access Control" (2010)
- 21) 河原木ほか: "SMPTE ST 2110を用いた8K中継番組のIPリモート制作", 放送技術, 72, 2, pp.138-142 (2019)
- 22) SMPTE ST 2022-7: "Seamless Protection Switching of RTP Datagrams" (2019)
- 23) 川本ほか: "リモートプロダクションを実現する軽圧縮8K IP伝送装置の開発", 放送技術, 71, 8, pp.107-111 (2018)
- 24) RFC 3584: "Best Current Practice- Coexistence between Version 1, Version 2 and Version 3 of the Internet-standard Network Management Framework" (2003)
- 25) <https://github.com/influxdata/telegraf>
- 26) <https://www.influxdata.com/>
- 27) <https://grafana.com/>



きたじま しょうじ
北島 正司 1994年、NHK入局。2016年頃より、IPリモートプロダクションの調査・検証を実施。現在、番組制作システムのIP化に取り組んでいる。正会員。



かわら せいひろ
河原木 政宏 2014年、豊橋技術科学大学大学院修士課程修了。同年、NHK入局。2017年より、放送技術研究所に勤務。現在、ネットワーク監視技術の研究・開発に従事。正会員。



こやま ともひさ
小山 智史 2005年、東京工業大学大学院修士課程修了。同年、NHK入局。2009年より、放送技術研究所に勤務。現在、番組制作システムのIP化の研究・開発に従事。正会員。

JPEG 軽量映像圧縮技術

石川 孝明[†]

1. まえがき

ISO/IEC JTC 1/SC 29/WG 1 (JPEG) では静止画像符号化技術の国際規格を審議しており、フレーム間の予測を用いない圧縮方式については、各フレームを独立した静止画像とみなして、動画像符号化の規格も策定している。JPEG XS (ISO/IEC 21122シリーズ) は、そのような規格のうち最新の規格であり、低遅延かつ軽負荷の映像圧縮を実現している。

JPEG XSの符号化方式は、JPEG 2000符号化方式にも使われているウェーブレット変換を利用して、主に水平方向の空間的な冗長性を除去する。また、垂直方向に対する処理を簡略化することで、フレーム単位ではなく数ライン単位での処理を可能とし、符号化・復号処理の低遅延化を実現している。このような特徴から、JPEG XSは、低ビットレートに対する圧縮率の向上よりも、視覚的なロスレスの画像品質を維持しながら、4:4:4および4:2:2の画像シーケンスを1/2から1/10程度の圧縮率で符号化する方式である。また、コーデックの複雑度が低く、フレームバッファを必要としないため、さまざまなハードウェアにおいて低コストでの実装が可能である。

本稿では、JPEG XSの規格化経緯に加えて、JPEG XSの符号化方式と評価方式を解説し、さらに、超高解像度映像の映像制作への応用事例について述べる。

2. 規格化の経緯と審議状況

2.1 規格化経緯

JPEG XSの規格化が開始された契機は、2015年2月開催の第68回WG1シドニー会合で審議したVideo Electronics Standards Association (VESA) のリエゾンレターと考えられる。WG1は、それまでもVESAとの連携を続けてきたが、同レターは特にVESAのDSC規格 (Display Stream Compression) を拡張する技術提案への応答を促す内容で

あった。しかし当時のWG1にはDSC規格の拡張仕様を満たす規格は存在せず、また提案募集に対応するための時間を十分に確保することが困難であった。一方でWG1は、高品質映像のIP伝送等、プロフェッショナル用途での低遅延・軽負荷なコーデックへのニーズがあることをすでに把握していたため、DSC規格の拡張仕様に限定せず、より広い応用領域を対象とする新たな規格策定に向けた議論を開始した。

第69回会合 (2015年6月) において、新たな規格の想定利用シーンを記述するユースケースと、それを実現するための要求仕様を定めた文書の初版が承認され、同時に正式なプロジェクトとして承認を求める新規提案も作成された。JPEG XSの要求仕様は、許容される最大遅延量、レート制御性能、ハードウェア実装を前提とする複雑度、必要とするメモリー量などについて、数会合に渡り議論され更新された。さらに、WG1以外の専門家や産業界からの意見も反映する目的から、第71回会合 (2016年2月) にてワークショップを開催し、それから得られたさまざまな意見を要求仕様に反映した¹⁾。第69回会合で申請した新規提案は、2016年1月締切の投票にて正式なプロジェクトとして承認され、規格番号としてISO/IEC 21122が付与された。

WG1は、第71会合での合意内容を踏まえて、2016年2月に提案募集をリリースした²⁾。提案募集に示された要求仕様のうち一部を抜粋したものを表1に示す。低遅延については最大で32ラインの範囲に収めることが条件とされ、軽負荷については、リアルタイムで処理可能な負荷量として、例えば、ソフトウェア実装であればIntel Core i7プロセッサを用いて、あるいはハードウェア実装であれば、Xilinx Artix 7のようなFPGAの約半分の性能において、4K映像 (4:4:4, 8ビット, 60 fps) をリアルタイムに符号化または復号処理することが目安として示された。

第72回会合 (2016年6月) にて、提案募集の応答として、のべ8機関 (企業6件、大学2件) から7つの方式が入力された。入力された提案方式には、JPEG 2000を改良した方式、ウェーブレット変換をラインベースで適用する方式、整数DCTやアダマール変換を用いる方式、さらには、H.265/HEVCをベースとする提案まで、さまざまな方式が

[†] 早稲田大学 国際情報通信研究センター

"IP Adoption in Broadcast Industry (Final study): Lightweight Image Coding Technology in JPEG Standard" by Takaaki Ishikawa (Global Information and Telecommunication Institute, Waseda University, Tokyo)

表1 提案募集に示された要求仕様(抜粋)

入力信号の仕様	
解像度	VGA (640×480)～8K (8192×4320 pel)
サブサンプリング	4:2:2, 4:4:4
コンポーネント	RGB, YCbCr
ビット深度	8から12ビット
フレームレート	24～120 fps
ビットストリームの仕様	
画像品質	視覚的ロスレス
圧縮率	1/2～1/6
レート制御	VBR, CBR
レート指定	フレーム単位の最大符号量指定
独立性	フレーム単位の符号化・復号処理
反復圧縮	7回以上の反復(画質低下耐性)
設計に対する仕様	
低遅延	最大32ライン以下
軽負荷	4K映像のリアルタイム処理(目安)
実装環境	FPGA/ASIC/GPU/CPU
実装の拡張性	スループットに応じた規模を維持
並列処理	エンコーダ・デコーダ双方でサポート

含まれた。それらの方式は、基準となるアンカー方式とともに、計7機関が参加する客観評価および主観評価試験により比較された。審議の結果、2016年10月の第73回会合にて、エントロピー符号化の構造が類似している2方式に絞り込まれ、詳細な比較をコア実験の結果で確認するためのテストモデルの作成に着手した。続く第73回会合では、基本となる符号化システムの草案が作成された。

第74回会合にて、基本アルゴリズムの規定からバッファモデルの規定を分離する目的から新パートの策定申請が出され承認された。現在のJPEG XS (ISO/IEC 21122シリーズ)は、5つのパートで構成されている。パート1 (2019年5月IS発行)は、Core coding systemであり、JPEG XSの復号方式を規定し、パート2 (2019年7月、IS発行)は異なる送受信環境およびアプリケーションに対応するバッファモデルやプロファイル・レベルを規定する。さらに、パート3 (2019年8月、IS発行)は、ファイルフォーマットを規定し、続くパート4およびパート5では、それぞれ適合性試験と参照ソフトウェアを規定している。

2.2 審議状況

表2に、JPEG XSの各パートの名称と作業計画を示す。表中の略語は、作業原案(WD)、委員会原案(CD)、国際規格案(DIS)、最終国際規格案(FDIS)、国際規格(IS)に対応している。2020年以降は、すでに審議が開始されている追補の開発作業に注力する予定である。現時点で追補は2件あり、パート1に対する追補1としてExtended capabilities for JPEG XSと、パート2に対する追補1としてJPEG XS profiles extensionsがある。いずれも2021年1月の審議完了を目標としている。これらの追補は、4:2:0信号とセンサ圧縮をサポートするための拡張である。

表2 JPEG XS (ISO/IEC 21122シリーズ)の作業計画

Part	Title	WD	CD	DIS	FDIS	IS
1	Core coding system	17/07	17/10	18/02	—	18/10
2	Profiles and buffer models	17/07	18/04	18/07	—	19/01
3	Transport and container formats	17/04	18/07	18/10	19/03	19/11
4	Conformance testing	18/07	18/10	19/01	19/07	20/11
5	Reference software	18/07	19/03	19/07	—	20/01

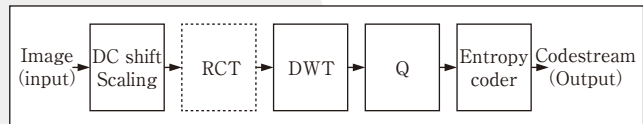


図1 JPEG XSの符号化処理ブロック

3. JPEG XS符号化システム

3.1 符号化方式の概要

JPEG XSの符号化処理ブロックを図1に示す。まず入力信号をビット深度に応じてスケールし、DCレベルシフトにより0を中心とする信号を生成する。そして、入力信号の色コンポーネントがRGBである場合は、必要に応じて輝度信号と色差信号にコンポーネント変換する。同変換は、JPEG 2000のロスレス変換で使われている色変換であるReversible Component Transform (RCT)と同一である。次に、5-3ウェーブレット変換により信号の冗長性を除去し、得られたウェーブレット係数を量子化して情報量を削減する。最後にエントロピー符号化を施して出力となるコードストリームを得る。これらの一連の処理を、画像全体を読み込むことなく処理することで、必要なバッファ量を減らし、要求仕様に規定された範囲の遅延量で符号化・復号処理を完了する。

3.2 コードストリームの構造

JPEG XSのコードストリームは、入力画像全体をピクチャとして捉え、その内部をスライス、プレシント、サブバンド、パケットの順で区分する。コードストリーム構成の概要を図2に示す。パケットはサブバンドの各ラインのウェーブレット係数をエントロピー符号化したものであり、ピクチャから得られるすべてのパケットを集約することで、JPEG XSのコードストリームが構築される。なお、図2のLL成分を含むLL_{5,2}からHL_{2,2}のサブバンドは、これらを集めたものを1つのパケットとして扱う。図2に示した区分の最小単位はパケットであるが、エントロピー符号化では、パケットをサブパケットに分類しながら処理する。各サブパケットの生成では、パケットに含まれる4つのウェーブレット係数をまとめて1つの処理単位としており、これをコードグループと呼ぶ。エントロピー符号化の詳細は、第3.5節で述べる。

JPEG XSコードストリームのピクチャヘッダには、プロファイルやレベルを表すフィールドに加えて、コードスト

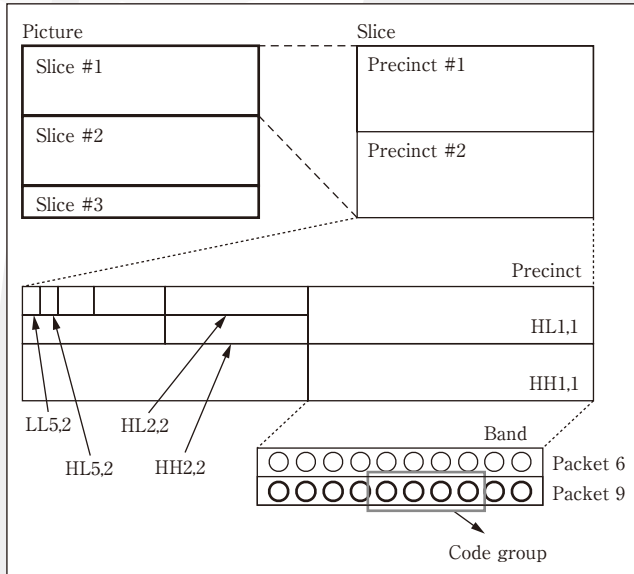


図2 ピクチャ、スライス、プレシント

リームの長さ、スライスの高さ、色変換の種別、ウェーブレット変換の水平・垂直分割数、量子化種別などが記録される。ウェーブレット変換、係数量子化の詳細は、第3.3節および第3.4節で述べる。なお、JPEG 2000では複数のプログRESSIONオーダを選択できるが、JPEG XSではResolution-Line-Band-ComponentのみがデフォルトのプログRESSIONオーダとして規定されている。

コードストリームは、必要に応じてファイルフォーマット (JXS形式) に格納することで、一般的なファイルとして利用できる。この他にも、ISO/BMFFにコードストリームを格納するMotion JPEG XS、さらには、HEIF画像フォーマットに格納する方式も策定されている。これらはJPEG XSパート3に規定されている。また、コードストリームに含まれる各スライスをRTPプロトコルのペイロードに格納すれば、リアルタイム伝送に応用可能である。同様に、MPEG-2 TSへの格納もISO/IEC 13818-1:2019/Amd.1として規格化されており、さらに、SMPTE ST 2110スイートでの利用も検討されている。

3.3 ウェーブレット変換

JPEG XSのウェーブレット変換には、JPEG 2000符号化方式でも採用された5-3ウェーブレット変換が用いられる。よりパフォーマンスの良い13-7ウェーブレット変換の利用も検討されたが、遅延量の増加に対してPSNR値の改善に有意性がみられないことから採用が見送られている。

$$\begin{cases} y(2n+1) = x(2n+1) - \left\lfloor \frac{x(2n) + x(2n+2)}{2} \right\rfloor \\ y(2n) = x(2n) + \left\lfloor \frac{y(2n-1) + y(2n+1) + 2}{4} \right\rfloor \end{cases} \quad (1)$$

5-3ウェーブレット変換は整数変換であり、Le Gallのリフティング構造で実装できる。式(1)のように、シフト演算を

用いて簡単に表現できるため、軽量の処理で画像の空間的な冗長性を除去できる。\$x\$が入力信号、\$y(2n)\$がローパスバンドのウェーブレット係数、\$y(2n+1)\$がハイパスバンドのウェーブレット係数を表す。復号処理で用いられる逆ウェーブレット変換も同様のリフティング構造で表現される。

上記のウェーブレット変換は、1次元信号に対する処理である。2次元信号である入力画像を垂直および水平方向の2方向について変換することで、2次元のサブバンドが得られる。また、得られたLLバンドを再帰的にウェーブレット変換することで、空間的なスケーラビリティが得られると同時に符号化効率の向上が期待される。

しかし、垂直方向の分割数を増やすためには、係数の算出に必要な信号をバッファに読み込まなければならず、処理に必要なメモリー量の増加と遅延量の増加につながる。そのためJPEG XSでは、垂直方向の分割数を水平方向の分割数に応じて制約している。規定では水平方向の分割数\$N_{L,x}\$は、1～8回まで指定できるが、垂直方向の分割数\$N_{L,y}\$は、\$0 \sim \min(N_{L,x}, 6)\$で表される範囲に限定される。実際に使われる分割数は、プロファイルとレベルに応じて異なるが、垂直方向の分割数は最大でも2回が目安である。プロファイルとレベルについては、第4章で述べる。

3.4 係数の量子化

ウェーブレット変換により得られる係数群は、デッドゾーン付き量子化あるいは一様量子化を用いて量子化インデックスで表現される。どちらの量子化方法を用いるかは、ピクチャヘッダで指定できる。デッドゾーン付き量子化は、ビットプレーン展開した係数の下位を打ち切る処理に相当する。一方、一様量子化は、量子化幅\$\Delta\$が式(2)で表されるスカラ量子化に相当する。

$$\Delta = \frac{2^{M+1}}{2^{M+1-T} - 1} \quad (2)$$

$$= 2^T \frac{1}{1 - \frac{1}{2^{M+1-T}}} = 2^T \sum_{k=0}^{\infty} (2^{M+1-T})^{-k} \quad (3)$$

\$M\$はビットプレーンカウントであり、コードグループごとに一意に定まる有意ビットのあるビットプレーンを表す。また、\$T\$は打ち切り点を表し、各ウェーブレット係数の表現精度とビットレート制御に影響を与える。\$M\$と\$T\$の詳細は、第3.5節で述べる。

符号化器における式(2)による除算処理の実装は、シフト演算を用いて容易に表現できる。しかし復号器での積算処理は、分母に\$-1\$が含まれるため、そのままではシフト演算に落とし込めない。この問題は、式(2)を式(3)のようにノイマン級数で展開し、所定の精度まで加算処理を繰り返すように実装することで解決できる³⁾。

デッドゾーン付き量子化は、スケーリング処理が不要となり回路規模を抑制できるメリットがある。その一方で、

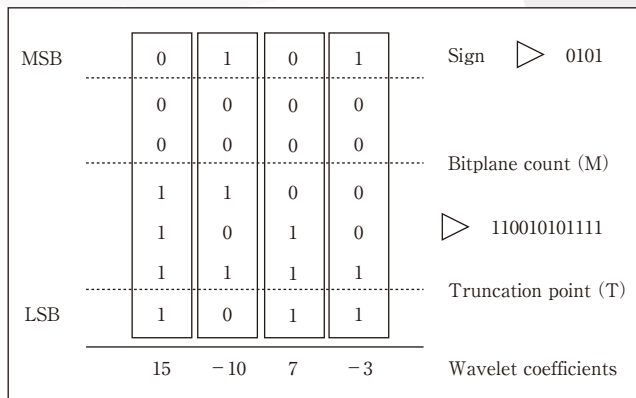


図3 ビットプレーン展開 (コードグループ)

一様量子化よりもPSNR値が若干低下することがコア実験により確認されている。また量子化パラメータ(ビットプレーンの打ち切り点)は、プレシント単位で制御されるが、サブバンド単位に重み付けを変えることで、より細かく画質を制御することもできる。

3.5 エントロピー符号化

ウェーブレット変換後の係数は、4係数ごとにコードグループとしてまとめられ、エントロピー符号化される。第3.4節で述べたように、ウェーブレット変換後の係数は、ビットプレーン展開におけるビットプレーンカウント M と打ち切り点 T により量子化される。打ち切り点 T は、プレシント単位の量子化パラメータ Q と、プレシントおよびサブバンド単位の補正情報を用いて記録される。したがって、コードグループ単位での符号化が必要となるのは、ビットプレーンカウント M 、係数の大きさ v 、係数の符号 s となる。これらを以下に述べる4種類の符号化方法を用いて効率的に表現する。

図3に示すように、係数の大きさ v と係数の符号 s は、ビットプレーン展開によりすでにバイナリー表現されている。そのため、実質的なエントロピー符号化は、Bit-plane count codingにおけるビットプレーンカウント M に適用する可変長符号化のみとなる。この点からも、JPEG XSが軽量の符号化方式であることがわかる。

(1) Significance coding

8つのコードグループをまとめたsignificance groupについて、有意ビットが存在するかをフラグで表現する。Bit-plane count codingの予測モードにrawモードを用いる場合は、Significance codingを使用できない。Significance codingは、オプションであり、サブバンド単位で切り替え可能である。

(2) Bit-plane count coding

ビットプレーンカウント M を符号化する。ビットプレーンカウント M は、コードグループに含まれる4係数をビットプレーン展開し、MSBから検査して最初に有意ビットが現れるビットプレーンのインデックスである。インデックス

を可変長符号化により圧縮するが、いくつかの予測モードを選択できる。予測モードについては、第3.6節で述べる。

(3) Data coding

ウェーブレット係数の大きさを符号化する。係数の大きさ v は、ビットプレーンカウント M と打ち切り点 T に挟まれたプレーンのみを記録すればよく、MSB側から順番に記録する。

(4) Sign coding

ウェーブレット係数の符号を符号化する。ただし符号情報 s をSign codingで扱わずに、Data codingで符号化することもできる。ピクチャ単位で切り替え可能である。

3.6 予測モード

ビットプレーンカウント M を符号化するBit-plane count codingでは、符号化効率を向上させるために、サブバンドごとに予測モードを適切に切り替える。予測モードには、ビットプレーンカウント M をそのまま出力するrawモード、現在のラインの直上にあるラインから予測するvertical predictionモード、そして、打ち切り点 T から予測するnon-predictionモードがある。

vertical predictionモードまたはnon-predictionモードを用いる場合は、予測誤差値を可変長符号化する。可変長符号化には、打ち切り点 T を考慮したユニナリー符号が使われる。例えば、 $\{0, -1, +1, -2, +2\}$ に対して、 $\{0, 10, 110, 1110, 11110\}$ が割り当てられる。予測値が打ち切り点 T と一致するnon-predictionモードの場合、予測誤差値が非負になるため、ビットプレーンカウント M が増加するにつれて、ユニナリー符号の符号長が単純に増加する。vertical predictionモードは、常に直上のラインの情報を必要とするため、画像の第一ラインを含むサブバンドには適用できない。同様に、スライスを独立して復号可能にするために、スライス境界を越える予測となる場合はvertical predictionモードを使用できない。

予測モードがrawモードの場合、ビットプレーンカウント M は4ビットの固定長で符号化されるため、必要な符号量を事前に知ることができる。一方、vertical predictionモードまたはnon-predictionモードでは、予測が外れる場合にrawモードよりも長い符号長になる可能性がある。そこで、rawモードで計算できるデータ長 L_{cnt}^{raw} を上限として、予測誤差のデータ長が L_{cnt}^{raw} を超える場合にはrawモードを選択する。これにより、指定したビットレートを越えてしまうことを防ぐことが可能になる。符号長の上限は、サブバンド単位で制御される。

4. プロファイルとバッファモデル

JPEG XSの要求仕様は、幅広いアプリケーションを想定している。要求仕様のうち、許容される複雑度、遅延量、圧縮率などのパラメータは、各アプリケーションで対応すべき範囲が異なる。JPEG XSパート2は、それらの違いに

表3 プロファイル(抜粋)

(Main profile)	422.10	444.12
サンプリング	4:0:0, 4:2:2	4:0:0, 4:2:2, 4:4:4
ビット深度	8, 10	8, 10, 12
垂直分割数	0, 1	0, 1
水平分割数	[1-5]	[1-5]
量子化	DZQ, UNQ	DZQ, UNQ
スライスの高さ	16	16
(Light profile)	422.10	Subline 422.10
サンプリング	4:0:0, 4:2:2	4:0:0, 4:2:2
ビット深度	8, 10	8, 10
垂直分割数	0, 1	0
水平分割数	[1-5]	[1-5]
量子化	DZQ	DZQ, UNQ
スライスの高さ	16	16
(High profile)	444.12	4444.12
サンプリング	4:0:0, 4:2:2, 4:4:4	4:0:0, 4:2:2, 4:4:4, 4:2:2:4, 4:4:4:4
ビット深度	8, 10, 12	8, 10, 12
垂直分割数	[0-2]	[0-2]
水平分割数	[1-5]	[1-5]
量子化	DZQ, UNQ	DZQ, UNQ
スライスの高さ	16	16

対応するためにプロファイルとレベルを規定している。また、バッファオーバーフローやアンダーフローを生じることなく、復号器が適切にコードストリームを処理できるようにバッファモデルを規定している。

JPEG XSパート2には、ビット深度やサブサンプリング精度で分類されるプロファイル(Main, Light, High)と、解像度とビットレートで分類される8種類のレベルが規定されている。表3は、JPEG XSパート2に規定されたプロファイルの一部を抜粋したものである。表中のDZQはデッドゾーン付き量子化、UNQは一様量子化を表す。

5. 符号化性能の評価基準

JPEG XSの符号化方式は、PSNR, MSEによる客観品質、視覚的ロスレスを満たす主観品質、実装規模を考慮した複雑度、低消費電力性、要求メモリーサイズ、符号化・復号の反復処理への耐性など、複数の基準を用いて符号化性能を総合的に評価している。ただし、視覚的なロスレスを厳密に定義するのは困難である。一般に、PSNRによる客観評価において約40 dBを上回る場合に、人は視覚的な違いを感じる事が困難であるため、これをもって視覚的なロスレスと述べる場合が多い。

これに対して、JPEG XS符号化方式の規格化では、ISO/IEC 29170-2 (Evaluation procedure for nearly lossless coding)のAnnex B: Forced choice paradigm with interleaved images test protocolに規定されたフリッキング検査法を用いて、視覚的ロスレスを満たすかを評価している。フリッキング検査法の検査手順は、①モニタの表示領域を左右に区分する、②一方に元画像、他方に復号後の

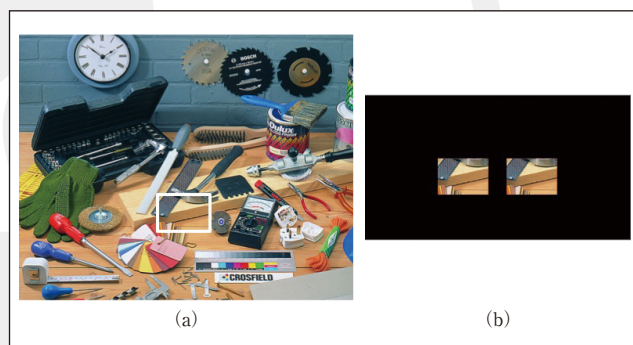


図4 試験画像の切り出しと配置例

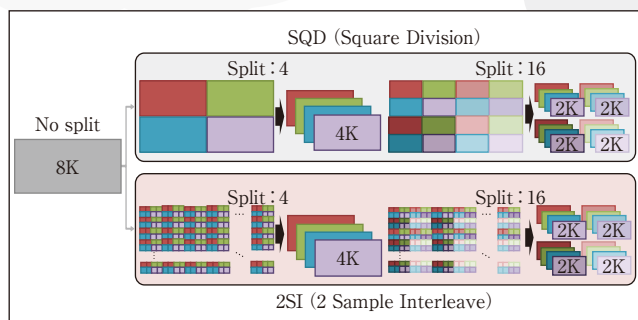
画像を表示する、③復号後の画像が表示された側を所定の間隔で交互に元画像と切り替える、④切り替えの結果生じる変化が観測されるモニタ領域(左右)を被験者が選択する、④切り替えの結果生じる変化が観測できない場合は変化なしとする、以上の手順が1つの試行になる。

図4(a)に試験画像の一例(TOOLS)を示す。モニタに表示する画像は、試験画像のうち一部を切り出した画像であり、拡大縮小やアンチエイリアシングなどの処理を加えてはならない。また、試験画像で画面全体を覆うような表示ではなく、図4(b)のように画面中央に少ない視線移動で確認できる範囲に元画像および復号画像が表示される。復号画像を表示する領域は、試行のたびに左右をランダムに切り替える。以上の検査から得られる結果を統計的に分析し、視覚的なロスレスを満たす圧縮率を導き出している。なお、ISO/IEC 29170-2には、本試験の前に実施するトレーニングセッション、被験者の人数や試験環境に関する規定も詳細に記述されている。

6. 応用事例

本章では、JPEG XS符号化方式のフルスペック8K制作での利用を想定した応用事例を示す⁴⁾。フルスペック8Kは、7,680 × 4,320画素の空間解像度、フレームレートが120 fps、4:4:4サンプル、ビット深度が12ビットの超高品質な映像であり、非圧縮映像の伝送レートは144 Gbpsに及ぶ。そのような広帯域を必要とする映像を既存の物理インターフェースを用いて伝送するには、視覚的ロスレスを満たすデータ圧縮と、適切な画面分割による並列伝送が必要となる。画面分割の手段は複数あるが、8K映像の伝送については、図5にある田の字分割方式(SQD)と2画素毎サンプル方式(2SI)が考えられる。SQDは空間的な信号の相関性を十分に活かすことができるため高い圧縮性能が見込まれるが、フレームバッファを要し、遅延量が増加する。一方、2SI方式は、圧縮性能の低下が予想されるが、フレームバッファが不要であり、低遅延で済むという特徴がある。

比較に用いた画像シーケンスを図6に示す。画面の分割数は、4および16としている。40 GbEをIP伝送のインターフェースに用いる場合、乱れを生じることなく映像を切り

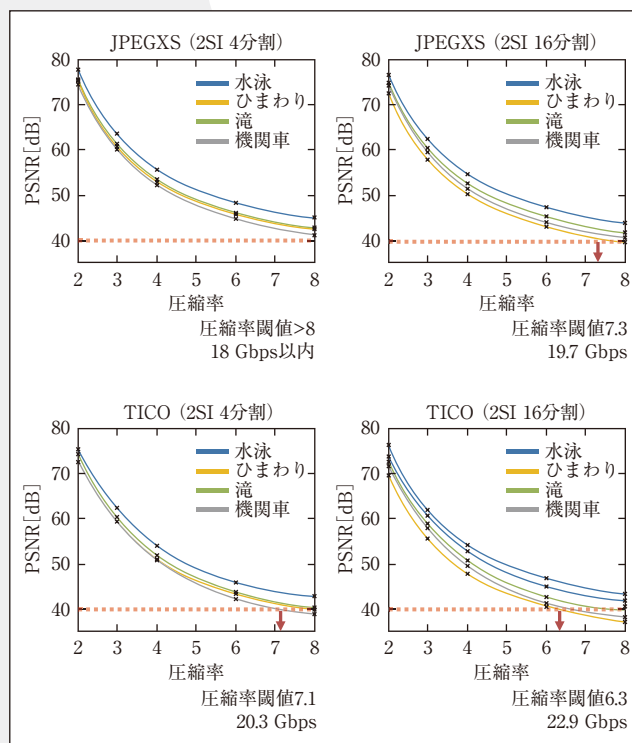
図5 画像分割 (SQD, 2SI) (文献⁴⁾より引用)図6 実験に用いた8K映像 (文献⁴⁾より引用)

替えるためには、20 Gbps 以下にレート制御する必要がある。したがって、圧縮率1/7.2が条件を満たす閾値となる。図7は、前述の2つの画面分割方法について、JPEG XS符号化とJPEG XSのベースとなったTICOを用いて圧縮率1/2から1/8の範囲で符号化した場合の実験結果である。

各実験映像をJPEG XSを用いて符号化すると、すべての場合において40 dB以上の品質を保ちながら20 Gbps以下に圧縮できることが確認されている。一方、TICOを用いた同じ条件の符号化実験の結果では、2SIの場合に20 Gbps以下に収まらず、PSNR値も40 dBを下回った。また同実験では、TICOと比較してJPEG XSの符号化効率が3 dB程度優位であった。このように、JPEG XSは、従来の方式と比較して同等の品質をより低いビットレートで達成しており、フルスペック8Kの映像伝送にも充分に応用可能なことがわかる。

7. むすび

本稿では、低遅延・軽負荷な画像符号化方式であるJPEG XSの規格化経緯、符号化方式の構成、評価基準、応用事例について述べた。JPEG XSは、他の最新規格と比較して圧縮率の低い方式ではあるが、視覚的ロスレスの品質を保ちながら、超高解像度映像の低遅延伝送が可能である。

図7 実験結果 (JPEG XS (2SI), TICO (2SI)) (文献⁴⁾より引用)

また、処理の複雑度も低く、さまざまなプラットフォームへの実装が容易であることから、高品質映像の低遅延伝送が要求されるライブ配信やAR/MR/VR分野等を中心として、今後の幅広い普及が期待される。

謝辞 本稿を執筆するにあたり、フルスペック8K映像の制作に向けた軽量コーデックの応用に関する情報をご提供いただいた、NHK放送技術研究所の林田哲哉氏、青木秀一氏に謝意を表します。

(2019年10月21日受付)

〔文 献〕

- 1) "JPEG XS Workshop proceedings - Use cases for a low-latency lightweight image coding system", https://jpeg.org/items/20160229_program_xs_workshop_proceedings.html (2016)
- 2) "Call for Proposals for a low-latency lightweight image coding system", https://jpeg.org/items/20160229_program_xs_workshop_proceedings.html (2016)
- 3) T. Richter, J. Keinert, A. Descampe and G. Rouvroy: "Entropy Coding and Entropy Coding Improvements of JPEG XS", 2018 Data Compression Conference, p.8796 (Mar. 2018)
- 4) 林田哲哉, 青木秀一, 三浦菊佳, 中村友洋, 米内淳, 瀧口吉郎: "軽圧縮コーデックとフルスペック8K制作への応用", 第47回画像電子学会年次大会予稿集, pp.OS1-3 (June 2019)



石川 孝明 2005年、早稲田大学大学院国際情報通信研究科修士課程修了。2007年、同大学国際情報通信研究センター助手。2010年、同センター招聘研究員となり、現在に至る。画像符号化の研究と応用開発に従事。2014年より、ISO/IEC JTC 1/SC 29/WG 1国内小委員会主査。