

放送・通信連携のための MMT の運用方法の検討

青木秀一^{†1} 大槻一博^{†1} 河村侑輝^{†1} 土田健一^{†1} 木村武史^{†1}

2016 年のスーパーハイビジョン (SHV) の試験放送開始を目指し、新たな放送システムの標準化が進められている。この放送システムの多重化方式として、筆者らが提案してきた MMT (MPEG Media Transport) が採用される見込みである。MMT は 2013 年 11 月に ISO/IEC 23008-1 として MPEG での標準化が完了した新たなメディアトランスポート方式であり、放送や通信など複数の伝送路でのコンテンツ配信に容易に対応できる特徴がある。MMT は伝送時のパケット形式やペイロード形式、制御メッセージなどを規定しているが、ISO/IEC 標準としてさまざまな用途に対応できるように、これらをどのように運用するかは規定しない。そのため、これらの運用はサービスに応じて検討する必要がある。そこで本稿では、スーパーハイビジョンの放送システムにおいて放送・通信連携を実現する観点から、MMT の運用方法について検討した結果について述べる。

A Study on MMT for Hybrid Delivery on Broadcast and Broadband

SHUICHI AOKI^{†1} KAZUHIRO OTSUKI^{†1} YUKI KAWAMURA^{†1}
 KENICHI TSUCHIDA^{†1} TAKESHI KIMURA^{†1}

The standardization of broadcasting system supporting Super Hi-Vision is in progress in order to launch its test broadcasting services in 2016. We have proposed MMT (MPEG Media Transport) as the appropriate multiplexing mechanism of the broadcasting system, and ARIB (Association of Radio Industries and Businesses), a standardization body for radio-communication systems in Japan would adopt it. MMT, ISO/IEC 23008-1, of which standardization has been finalized in MPEG in November 2013 is a new media transport protocol enabling hybrid delivery of content on various delivery channels such as broadcasting channels and broadband networks. MMT specifies MMT Protocol packet format, its payload format, and several signaling information assuming various use cases. However, the MMT specification does not specify the detailed operation of the MMT Protocol, signaling information and so on. The detailed operation needs to be specified in local SDO (Standard Development Organization) on the basis of its practical use cases. This report describes our study on detailed operation of MMT in order to achieve Super Hi-Vision broadcasting system from the viewpoint of hybrid delivery.

1. はじめに

2016 年のスーパーハイビジョン 4K/8K の試験放送開始を目指し、情報通信審議会 情報技術分科会 放送システム委員会並びに ARIB (電波産業会) において、技術方式の標準化が進められている[1]。超高精細度テレビジョン放送システムに関する中間報告 (多重化方式) [2]によれば、放送システムの多重化方式として MPEG Media Transport (MMT) が採用される見込みである。

MMT は、放送や通信など多様な伝送路を用いるコンテンツ配信のためのメディアトランスポート方式であり、ISO/IEC 23008-1[3]として標準化が進められ 2013 年 11 月に MPEG での標準化が完了した[4]。MMT は、ISO/IEC の標準技術としてさまざまなアプリケーションに対応できるように、伝送時のパケット形式やペイロード形式及び制御メッセージの形式などを規定している[5,6]が、それらをどのように用いるかはサービスに応じて検討する必要がある。そこで本稿では、スーパーハイビジョンの放送システムを実現するための MMT の運用方法について検討する。

2. MMT による放送システム

超高精細度テレビジョン放送サービスを実現する多重

化方式として検討されている MMT・TLV 方式の放送伝送路でのレイヤーモデルを図 1 に、参考として示されている

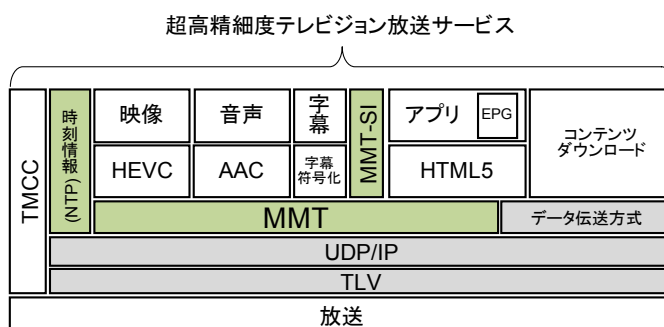


図 1 MMT・TLV 方式のレイヤーモデル (緑：新規に規定する部分, グレー：すでに規定されている部分)

Figure 1 Layer model of MMT and TLV for broadcast channel.

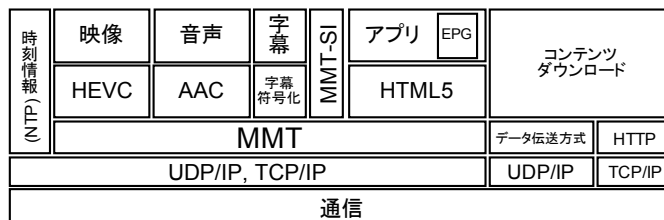


図 2 MMT・TLV 方式の通信回線でのレイヤーモデル

Figure 2 Layer model of MMT and TLV for broadband network.

^{†1} NHK 放送技術研究所
 NHK, Science and Technology Research Laboratories

通信回線でのレイヤーモデルを図2に示す[2]. MMT・TLV方式では、映像・音声や字幕の符号化信号はMMTP (MMT Protocol) パケット化する. MMTP パケットは、放送で伝送する場合にも通信で伝送する場合にも IP パケットで伝送する. このため、放送と通信でのサービスについて、両者のレイヤーモデルの差異を小さくすることができる. すなわち、MMT を用いることで、放送と通信のコンテンツ配信を共通化することが可能になる.

2.1 MMTP パケット化

映像符号化方式である HEVC (High Efficiency Video Coding) [7]とメディアトランスポート方式である MMT とのインターフェースは NAL (Network Abstraction Layer) ユニットとなる. HEVC 規格の NAL ユニットの例を表1に示す. NAL ユニットは、映像信号の動きベクトルや直交変換係数を符号化したデータである VCL (Video Coding Layer)

表 1 HEVC 規格の NAL ユニットの例

Table 1 Examples of NAL unit in HEVC specification.

(1) non-VCL

NAL ユニット名称	説明
VPS (Video Parameter Set)	階層符号化においてレイヤ間の共通情報や関係情報を示す.
SPS (Sequence Parameter Set)	ビデオシーケンスで共通のパラメータを示す.
PPS (Picture Parameter Set)	ビデオピクチャで共通のパラメータを示す.
Prefix-SEI (Supplemental Enhancement Information)	補足的な付加情報で VCL NAL ユニットの前に置かれる SEI.
Buffering period SEI	CPB リムーバルディレイや HRD の初期化のためのオフセット時間情報などを与える.
Picture timing SEI	CPB リムーバルディレイや DPB 出力ディレイの情報を与える.
Recovery point SEI	デコード開始位置の補助情報を与える.

(2) VCL

NAL ユニット名称	説明
CRA (Clean Random Access)	IRAP の一つである CRA スライスのデータを含む.
IDR (Instantaneous Decoding Refresh)	IRAP の一つである IDR スライスのデータを含む.
TRAIL	Trailing picture スライスのデータを含む.
TSA (Temporal Sub-layer Access)	時間方向階層符号化における復号レイヤ切り替えが可能な Trailing picture を含む.

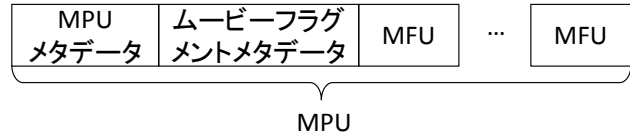


図 3 MPU の一般的な構成
 Figure 3 Generic structure of MPU.

NAL ユニットと、それらの付帯情報である non-VCL NAL ユニットとに分類される. MMT で HEVC を伝送する場合、non-VCL あるいは VCL を区別することなく NAL ユニットとして伝送することができる.

MMT では、メディアの符号を MPU (Media Processing Unit) と呼ぶ、復号処理が可能な単位として取り扱う. MPU の一般的な構成を図3に示す. MPU は、MPU 全体の構成に関する情報を含む MPU メタデータ、符号化したメディアデータの情報を含むムービーフラグメントメタデータ、さらに符号化したメディアデータである MFU (Media Fragment Unit) から構成される. HEVC の伝送では、NAL ユニットの MFU として扱うことができる.

HEVC エンコーダが出力する NAL ユニットから MPU を構成し、さらに MPU を MMTP パケット化する運用に関する提案の概要を図4に示す. MPU の先頭を IRAP (Intra Random Access Point) の AU とするため、図4では CRA の NAL ユニットとした. CRA の NAL ユニットに先立ち、このシーケンスを復号するために必要となる non-VCL の NAL ユニットである VPS, SPS, PPS や Prefix-SEI を配置する. CRA の NAL ユニットの後は、LP (Leading Picture) 及び TP (Trailing Picture) の NAL ユニットが引き続くこととなる.

MPU の大きさは、受信端末において、放送のチャンネル切り替え時に映像信号が出力されるまでの遅延に関する. そのため、映像符号化における IRAP の間隔で MPU を構成して運用することを提案する.

MMTP パケットは、MMTP パケットヘッダと MMTP ペイロード部とから構成される. MMTP ペイロードには、複数の NAL ユニットの格納することができる. また、NAL ユニットの大きさが伝送可能な IP パケットの大きさと比べ大きい場合は、NAL ユニットの分割してペイロードに格

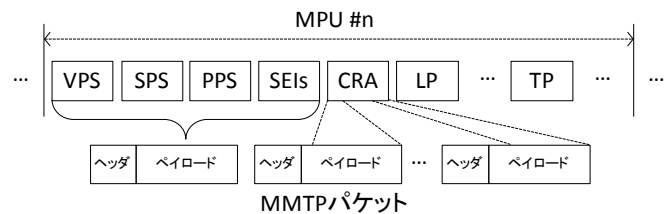


図 4 MPU を構成する NAL ユニットと MMTP パケット化の概要

Figure 4 MMTP packet generation from MPU consisting of NAL units.

納することとなる。図4には、non-VCLのNALユニットの大きさが一般にCRAのNALユニットよりも小さいため、複数のnon-VCLのNALユニットをまとめて一つのMMTPペイロードを構成し、これに対し、CRAのNALユニットは複数のMMTPペイロードに分割される様子を示した。

Prefix-SEIであるBuffering period SEIとPicture timing SEIを利用することで、ムービーフラグメントメタデータの送出手を省略したり、送出する場合にもMFUよりも後に送出したりするなどして、送出側がムービーフラグメントメタデータを構成するために必要となる遅延を抑制することができる。

なお、衛星放送では降雨時を除きエラーフリーの伝送が可能であることから、伝送時のIPパケットの大きさを通信回線よりも大きくし、伝送効率を向上することが期待できる。放送伝送路におけるIPパケットの多重化方式であるTLV多重化方式を記載するITU-R勧告BT.1869[8]ではIPパケットの大きさについて述べていないが、高度広帯域衛星デジタル放送でのIPパケットの伝送特性を評価した文献[9]では、1.5KB、4KB、6KBの大きさのIPパケットを用いた伝送性能の評価が行われている。受信端末は、上限サイズのIPパケットを受信するだけのバッファを備える必要があり、受信端末の実装の観点からは必要以上に大きなIPパケットを用いるのが困難であることを考慮すると、衛星放送伝送路でのIPパケットの上限サイズは4KB程度とすることが考えられる。

音声信号については、符号化方式としてMPEG-4 AACが、さらに、そのビットストリームの形式としてLATM/LOASストリーム形式が採用される見込みである[2]。LATM (Low Overhead MPEG-4 Audio Transport Multiplex) はオーディオデータのチャンネルコンフィグレーション情報を含むことに加え、オーディオデータの順序や連結などの多重機能を提供する。また、LOAS (Low Overhead Audio Stream) は同期機能を提供する。AACエンコーダが出力する音声信号は、1以上のオーディオフィームを含むAudioMuxElementの系列となる[10]。そのため音声信号を送信する際は、ムービーフラグメントメタデータを用いない場合、AudioMuxElementの単位でMFUとすることを提案する。

またMPEG-4 AACで符号化した音声信号は、いずれのオーディオフィームもランダムアクセスポイントとなることから、一つのMFUをMPUとすることも可能である。しかしこの場合、伝送制御信号の一つであるMP (MMT Package) テーブルをオーディオフィームごとに送出することになってしまいオーバーヘッドが増加する課題がある。そのため、映像信号のIRAPの間隔に相当する時間のオーディオフィームから、一つのMPUを構成することが妥当であると考えられる。

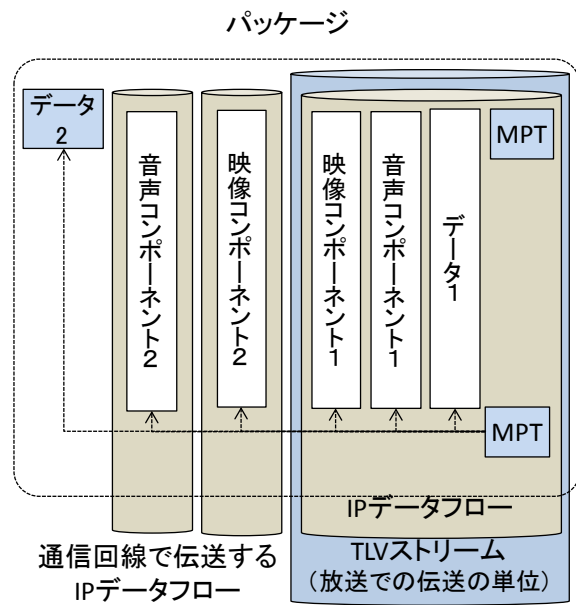


図5 MPTによるハイブリッド配信への対応

Figure 5 An MMT package through broadcast and broadband by MPT.

3. ハイブリッド配信への対応

放送伝送路と通信伝送路の両方を用いるサービスの構成を図5に示す。図5は、パッケージの構成を示す伝送制御信号であるMPテーブル (MPTと図示) 及び映像コンポーネント1、音声コンポーネント1、データ1が放送で伝送される様子を示している。MPTは、放送で伝送するこれらのコンポーネントに加え、通信回線で伝送する映像コンポーネント2、音声コンポーネント2をパッケージの構成要素として指定している。さらに、データ2についてはhttpで取得することも示している。

MPTは、パッケージを構成するアセットの種類、その取

ロケーションタイプ	パッケージ識別子				
ロケーションタイプ=0x00	16				
ロケーションタイプ=0x01	送信元IPv4アドレス 32	宛先IPv4アドレス 32	宛先ポート番号 16	パッケージ識別子 16	
ロケーションタイプ=0x02	送信元IPv6アドレス 128	宛先IPv6アドレス 128	宛先ポート番号 16	パッケージ識別子 16	
ロケーションタイプ=0x03	ネットワーク識別子 16	トランスポートストリーム識別子 16	将来予約 3	MPEG-2パッケージ識別子 13	
ロケーションタイプ=0x04	送信元IPv6アドレス 128	宛先IPv6アドレス 128	宛先ポート番号 16	将来予約 3	MPEG-2パッケージ識別子 13
ロケーションタイプ=0x05	URL長 8	URLバイト 8xN			

図6 ロケーション情報によるアセットの取得先 (数字はそれぞれのフィールドのビット数を示す)

Figure 6 Locations of asset identified by general location info structure.

得先、並びにアセットの詳細な情報を記述子として示す伝送制御信号である。アセットの取得先は図6に示す構造の `general_location_info` (ロケーション情報) で指定される。MPT と同一の packets フローに多重されたアセットを示す場合の他、任意の IP データフローの MMTP パケット、MPEG-2 TS の放送ネットワーク、MPEG-2 TS over IP の MPEG-2 TS パケットや URL を指定してアセットの取得先を指定することが可能である。

ロケーション情報を用いることで、コンテンツを構成するアセットの取得先をネットワーク横断的に指定できるため、MMT はハイブリッド配信に容易に対応することができる。

4. 時間方向階層符号化への MMT の対応

4.1 HEVC の時間方向階層符号化

フォーマットが 7680x4320/120/P あるいは 3840x2160/120/P の映像信号を送出した場合、それぞれ 7680x4320/60P あるいは 3840x2160/60P に対応した受信端末においても 60P の映像として復号・提示できることが望ましい。時間方向階層符号化をおこなった HEVC ストリームの 120P サブセットから 60P サブビットストリームを分離できる構造とし、受信端末が両者を識別できるよう伝送することで、この要件に対応することが可能である。時間

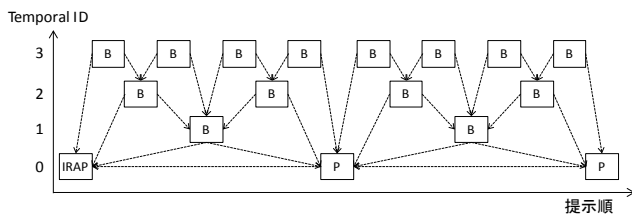


図 7 L4 の構造とした時間方向階層符号化の例

Figure 7 An example of four-level temporal layered coding.

方向階層符号化をおこなった HEVC ストリームの構造の例を図7に示す。この例では、Temporal ID が 0 から 2 までの NAL ユニットの 60P サブビットストリームとし、Temporal ID が 3 の NAL ユニットの 120P サブセットとする。

4.2 放送における 120P 対応

時間方向階層符号化 HEVC ビットストリームの伝送に関する提案の概要を図8に示す。パッケージを構成する際、60Hz 復号表示用の HEVC 時間方向サブビットストリームと 120Hz 復号表示用サブセットは、別々のアセットとする。図8では前者をアセット1、後者をアセット2として示した。別々のアセットであるため、アセット1とアセット2の AU は、別々のパケット ID の MMTP パケットで伝送されることとなる。

また、120Hz 復号表示用サブセットの AU を含む MPU は、それらの AU が属する GOP の 60Hz 復号表示用サブビットストリームの AU を含む MPU と同一の MPU シーケンス番号を付加することを提案する。時間的に関係のある MPU を同一の MPU シーケンス番号とすることで、受信端末はこれらのアセットが同一の GOP に属する AU を含むことを容易に識別することが可能である。

アセット2はアセット1に復号が依存することとなる。そのため、MP テーブルにおけるアセット2の情報を記述する記述子領域に依存関係記述子を挿入し、依存先としてアセット1のアセットIDを記述する。アセット1とアセット2が伝送されるネットワーク及びパケットIDは、MP テーブルのロケーション情報によりアセットごとに示すこととする。なお、アセット1とアセット2のいずれにもタイムスタンプ記述子を付加することとする。アセット1については、その先頭に位置する IRAP AU の提示時刻を示し、アセット2では、その先頭に位置する B ピクチャの提示時刻を示すこととする。

このように送付することで、120P に対応した受信端末は

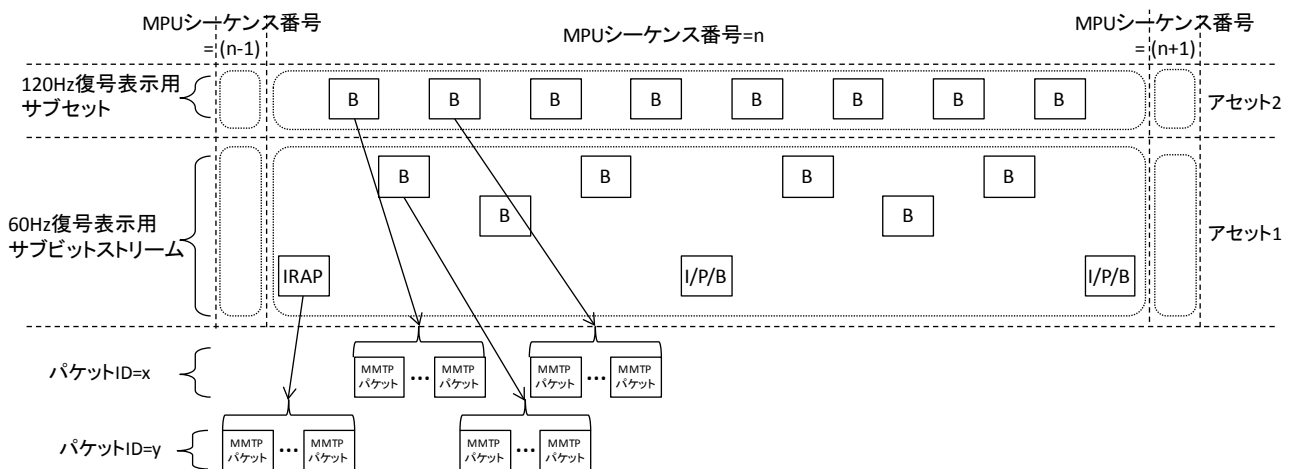


図 8 時間方向階層符号化 HEVC ビットストリームの伝送

Figure 8 Overview of transport of temporal layered HEVC code.

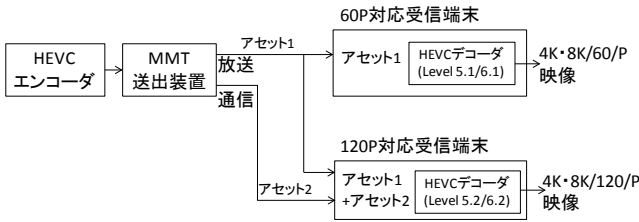


図 9 ハイブリッド配信における 120P 対応
 Figure 9 Support of 120P by hybrid delivery.

120P 対応の HEVC デコーダを用いて 120P の復号・提示を行うことができるだけでなく、60P に対応した受信端末でも 60P 対応の HEVC デコーダで 60P の復号・提示を行うことが可能となる。

4.3 ハイブリッド配信における 120P 対応

放送で 120P の映像を送出する場合の他、放送で 60P のためのサブビットストリームを送出し、120P のためのサブセットを通信回線で送することも想定される。3 節で述べたように MMT はハイブリッド配信に対応しているため、アセット 1 を放送で、アセット 2 を通信回線で送し、受信端末は必要に応じて通信回線からアセット 2 を受信することで、4.2 節と同様に 120P に対応することができる (図 9)。

5. シームレス切り替えのための MMT の運用方法

ある 1 つの放送サービスで映像信号のフォーマット (1920x1080/60/P, 3840x2160/60/P, 7680x4320/60/P など) を変更する際、受信端末においてシームレスな表示もしくは違和感の少ない表示が行えることが望ましい。

映像フォーマットを切り替える際の受信端末における映像表示の連続性について、MPEG-2 Video 並びに MPEG-4 AVC を用いる場合の MPEG2-TS の運用方法が ARIB 標準規格 STD-B32[11]に記載されている。そこで HEVC 及び MMT を用いる場合について、シームレスな切り替えを実現するための送出側及び受信端末の運用方法について検討する。

例として 4K 映像フォーマットの 3 個の番組から 8K 映像フォーマットの 1 個の番組に切り替える場合を記載する。2K 映像フォーマットの 3 個の番組から 4K 映像フォーマットの 1 個の番組に切り替える場合も同様である。

ある 1 つのサービス ID において映像フォーマットの変更時には、異なるフォーマットの映像ストリームのアセットの packet ID として、切り替えの前後で必ず異なる番号を使用することとする。

また、4K 映像を 3 番組から 8K 映像を 1 番組への切り替え、または 8K 映像を 1 番組から 4K 映像を 3 番組への切り替えにおいては、8K 映像の放送時にも 4K 映像の番組と同じサービス ID を持つ MPT を同じ数だけ送出することとする。これらに基づく例として、各番組のサービス ID、映像

アセットの packet ID を次のように例示する。

4K 番組 A : サービス ID=01, V_Asset_PID=101
4K 番組 B : サービス ID=02, V_Asset_PID=102
4K 番組 C : サービス ID=03, V_Asset_PID=103
↓
8K 番組 : サービス ID=01, V_Asset_PID=104
8K 番組 : サービス ID=02, V_Asset_PID=104
8K 番組 : サービス ID=03, V_Asset_PID=104

このとき、送出側では以下のように信号を送出することを提案する。

1. 4K/8K 番組の切り替え時刻を T1 とする。
2. 8K 映像の番組の MPT (映像アセットの packet ID として 104 を指定) は、通常の MPT の送出よりも早い時刻である、切り替え時刻 T1 より 1 秒前から送出する。切り替え時刻 T1 があらかじめ分かっているため、タイムスタンプ記述子に時刻 T1 を記述することが可能である。
3. 4K 映像ストリームの送出では、最後に EoS NAL ユニットの付加する。従って 4K 映像番組の最終の MPU は、EoS NAL ユニットの終了することとなる。
4. 切り替え時刻 T1 に、MMT 送出装置は 4K 映像のアセットの多重を停止し、8K 映像のアセットの多重を開始する。8K 映像のアセットの最初の MPU は、MPU シーケンス番号を 0 とし、先頭の AU は IRAP の AU とする。

これに対し受信端末は、以下のように動作することで、シームレス切り替えの表示に対応可能である。

1. 新たな MPT を取得する。
2. MPT のアセットの記述子領域に挿入されるコンポーネント記述子の内容により、4K 映像から 8K 映像への切り替えが発生することを判別する。この結果、MMT 多重分離装置は 4K と 8K の両方の映像アセットを同時にデコーダに入力する設定とする。ただし、両方のアセットは送出タイミングが異なるため重複してデコーダに入力されることはなく、4K 映像ストリームの受信終了後に 8K 映像ストリームがバッファに蓄積される。
3. 映像デコーダは、4K 映像アセットの EoS NAL ユニットの取得したら映像提示をフリーズする。
4. 8K 映像の VPS/SPS/PPS を取得したら、デコーダは 8K 映像のデコードを開始する。8K 映像を正常に出力可能な状態になったら映像提示のフリーズを解除し、8K の映像を出力する。
5. 8K 映像のデコードが開始されたことを確認したら、デコーダに入力するアセットを 8K のみとする。

このように送出し、また、受信端末が動作することで、映像フォーマットの切り替え時にもシームレスな切り替えが実現される。試験放送開始当初は 4K 映像の番組と 8K 映

像の番組とが混在することが予想されるため、こうした送出が必要になると言える。

6. おわりに

本稿では、スーパーハイビジョンの放送システムを実現するための MMT の運用方法として、MMTP パッケージ化の方法、映像の時間方向階層符号化への対応、シームレス切り替えの方法について検討した。こうした事項は ISO/IEC 標準として規定されないため、具体的なサービスに応じて検討し、各国の国内標準規格として規定する必要がある。

これらの検討を基に、今後、ARIB での標準規格策定並びに NexTV-F (次世代放送推進フォーラム) での運用規定策定の検討に寄与する予定である。

参考文献

- 1) 放送サービスの高度化に関する検討会 検討結果取りまとめ。
(http://www.soumu.go.jp/main_sosiki/kenkyu/bcservice/)
- 2) 情報通信審議会 情報通信技術分科会 放送システム委員会 超高精細度テレビジョン放送システム作業班配付資料。
(http://www.soumu.go.jp/main_sosiki/joho_tsusin/policyreports/joho_tsin/housou_system/02ryutsu08_03000151.html)

- 3) Text of ISO/IEC FDIS 23008-1 MPEG Media Transport, ISO/IEC JTC 1/SC 29/WG 11 N13982, (2013).
- 4) “MPEG Media Transport enables new services for the delivery of multimedia”, MPEG Press Release, ISO/IEC JTC 1/SC 29/WG 11 N13900, (2013).
- 5) Lim, Y., Park, K., Lee, J., Aoki, S. and Fernando, G., “MMT: An Emerging MPEG Standard for Multimedia Delivery over the Internet”, IEEE Multimedia Mag., Vol.20, No.1, pp.80-85, (2013).
- 6) Lim, Y., “MMT, new alternative to MPEG-2 TS and RTP”, IEEE International Symposium on Broadband Multimedia Systems and Broadcasting, (2013).
- 7) ITU-T Rec. H.265|ISO/IEC 23008-2, “Information technology - High efficiency coding and media delivery in heterogeneous environments - Part 2: High efficiency video coding”, (2013).
- 8) Recommendation ITU-R BT.1869, “Multiplexing scheme for variable-length packets in digital broadcasting systems”, (2010).
- 9) 青木他, “デジタル放送における IP パケット伝送方式の特性評価”, 映像情報メディア学会誌, Vol.64, No.7, pp.1020-1028, (2010).
- 10) ISO/IEC 14496-3, “Information technology - Coding of audio - visual objects - Part 3: Audio”, (2009).
- 11) ARIB STD-B32, “デジタル放送における映像符号化、音声符号化及び多重化方式”, (2012).