5章 360度VR映像の標準化動向

青木 秀一†

キーワード：360度VR映像，OMAF，AIAVシステム，ERP，イマーシブメディア，国際標準化

1. まえがき

360度映像に対応したカメラやヘッドマウントディスプレイの進化に伴い、360度VRコンテンツを体験する機会が増えてきた。さまざまな技術を用いた製品やサービスの市場への導入が先行したが、360度VRアプリケーションに関する技術の標準化も多くの団体で行われている。

本章では、360度VRアプリケーションに関する標準化について、MPEG、ITU、3GPPとVRインダストリーフォーラムの主な取り組みを概説する。

2. MPEGでのイマーシブメディアの標準化

ISO/IEC (International Organization for Standardization/ International Electrotechnical Commission：国際標準化機構/国際電気標準会議)において、マルチメディアの情報処理技術の標準化を行う作業部会であるMPEGは、ISO/IEC 23090 “Coded representation of immersive media” (MPEG-I)として、360度VR映像も含む、より没入感の高

い映像メディアに関する標準化を進めている。

2018年7月には、MPEG-Iパート1 “Immersive Media Architectures”と題するイマーシブメディアに関する技術レポートを発行した。この技術レポートには、イマーシブメディアに関して、3DoF (3 Degree of Freedom)、3DoF+、6DoFなどの用語を定義するとともに、それらのアプリケーションを実現する時期やアーキテクチャの例などが示されている(図1)。

MPEGは、最終的に6DoFのアプリケーションの実現を目指している。しかし、6DoFを実現する技術の検討には時間を要することから、6DoFの実現をフェーズ2とし、それに先行して3DoFのアプリケーションに限定したフェーズ1aと3DoF+にも対応したフェーズ1bを設け、標準化を進めることとした。フェーズ1aについては、3DoFの製品やサービスがすでに実現されているため、速やかな標準化が必要であった。そこで、可能な限り既存の映像符号化技術や音声符号化技術を応用する方向で検討が進められ、360度VRコンテンツのファイルフォーマットと伝送方法を

自由度	3DoF	3DoF+	Omnidirectional 6DoF	6DoF
イメージ図				
アプリケーション	固定点からの360度映像	いすに座った範囲の運動視差に対応する360度映像	6DoFのうち、移動が固定点から数歩程度に制限された範囲の視点での360度映像	空間を自由に移動した視点での360度映像
フェーズ	フェーズ 1a	フェーズ 1b	フェーズ 2	
想定時期	2017年末に OMAF バージョン1として標準化	2020年に標準化完了	2021~2022年に標準化完了	

図1 イマーシブメディアの技術レポートに示される各自由度の概要

† NHK 放送技術研究所
"Current Status of Standardization Activities on VR/360° Services" by Shuichi Aoki (Science and Technology Research Laboratories, NHK, Tokyo)

図2 単位球面と座標系

図4 ERPによる360度映像の矩形映像への射影変換の概要

規定する OMAF (Omnidirectional Media Format, ISO/IEC 23090-2) バージョン1として、2017年末に標準化を完了した。

OMAFは、主に以下の項目を規定している。

- ・X軸、Y軸、Z軸と単位球面から成る座標系(図2)
- ・360度映像を矩形映像に変換する射影変換と矩形映像の分割などの映像マッピングの方法
- ・360度映像・音声と、それらに必要なメタデータのファイルフォーマット
- ・MPEG-DASH (Dynamic Adaptive Streaming over HTTP, ISO/IEC 23009-1) と MMT (MPEG Media Transport, ISO/IEC 23008-1) を用いた伝送方法
- ・プレゼンテーションプロファイルとメディアプロファイル

OMAFにおける360度VRコンテンツの送信側と受信側

の処理フローを図3に示す。撮影した360度映像は、射影変換により矩形映像に変換される。OMAFバージョン1では射影変換として、ERP (Equirectangular Projection, 正距円筒図法) とキューブマップを選択可能である。ERPとキューブマップによる射影変換の概要を、それぞれ図4と図5に示す。ERPは、360度映像を地球に見立てたとき、緯線・経線が直角かつ等間隔に交差するように変換する射影変換である。

射影変換後の矩形映像をいくつかの領域に分割し、それらの領域の大きさや位置を変更するなどして別の矩形映像を生成するリージョンワイズパッキングを行うことも可能である(図6)。射影変換にERPを用いた360度映像では、赤道位置の情報量に対し両極位置の情報量が相対的に少なくなることから、リージョンワイズパッキングを適切に行うことで、同一ビットレートで圧縮符号化した際の画質を

図3 OMAFにおける360度VRコンテンツの処理フロー

図5 キューブマップによる360度映像の矩形映像への射影変換の概要

図6 リージョンワイズパッキングの例

向上することができる。また、矩形映像の両端にガードバンドと呼ばれるダミーの画素を加えることも可能である。これは、受信側で矩形映像の両端を接続して表示する際、符号化劣化が目立たないように、表示しない画素を加える処理である。

符号化した映像・音声および静止画は、360度映像に施した映像マッピングを特定するメタデータとともに、ISOベースメディアファイルフォーマットの形式でファイル化される。このようにして作られたファイルは、MPEG-DASHあるいはMMTにより受信側に伝送される。受信側では、メタデータを基に送信側と逆の処理を行い、360度映像のうちヘッドマウントディスプレイのセンサで特定された視線方向に対応した領域（ビューポート）の映像・静止画および音声がレンダリングされ表示・出力される。

また、コンテンツ制作者が、特定方向の視点での視聴を勧めるための初期視聴方向と推奨ビューポートに関する情報を時間付きメタデータとして付加することもできる。こうしたメタデータを用いることで、360度映像であっても特定方向の視聴に誘導することが可能である。

送信機器・受信機器の接続性を向上するため、OMAFは以下の2種類のプレゼンテーションプロファイルを規定している。

- ・ビューポートに非依存のベースラインプレゼンテーションプロファイル
- ・ビューポート依存のベースラインプレゼンテーションプロファイル

これらのプレゼンテーションプロファイルから参照するメディアに応じたプロファイルとして、映像について「HEVC (High Efficiency Video Coding) を用いるビューポート依存の映像プロファイル」など3種類、音声について「3Dオーディオベースラインプロファイル」など2種類、静止画について「HEVC静止画プロファイル」など2種類、テキストデータについて「IMSC1 (Internet Media Subtitles and Captions 1.0) プロファイル」など2種類のメディアプロファイルが規定されている。映像についてのメディアプロファイルには、リージョンワイズパッキングの有無などの制約も含まれ、アプリケーションに応じて適切なプロファイルを選択することで、送信機器と受信機器の接続性を向上することができる。

OMAFに対応したフォーマットを生成あるいは再生するソフトウェアとソースコードがGitHubで公開されていることもあり、今後、OMAFを用いて360度VRアプリケーションを実現する製品やサービスの増加が見込まれる。OMAFバージョン1の発行以降、MPEGでは、フェーズ1bとフェーズ2の実現に向け、OMAFの拡張と6DoFのアプリケーションのユースケースや要求条件、システムアーキテクチャなどの審議が行われている。

3. ITU-R SG6でのAIAVシステムの標準化

ITU-R (国際電気通信連合の無線通信部門) において、放送業務を所掌する研究委員会であるStudy Group 6は、放送局

図7 360度方向で7,680画素×4 (= 30K)の映像

においても360度VRコンテンツを制作する機会が増えたことを受け、2017年3月に、360度VR映像も対象とした、より没入感の高いAVシステムであるAIAV (Advanced Immersive Audio-Visual) システムの映像パラメータ、視聴環境、快適視聴の条件などを研究項目とする新たな研究課題を作成した。

2018年4月には、レポートBT.2420「放送におけるAIAVシステムの利用と現在の状況」*1を発行した。このレポートでは、日本を含む世界の放送局の360度VRコンテンツの制作や配信などの事例とスポーツや音楽ビデオといった360度VRコンテンツの分類、さらにMPEGやITU-Tなどの標準化団体の取り組みが簡潔にまとめられている。

また、3DoFの360度VR映像の制作・交換が容易に行えるよう、360度VR映像の制作・交換のためのフォーマットを規定する新勧告が、2018年10月に承認された。この新勧告は、360度映像から矩形映像への射影変換にERPを用いることや、矩形映像の空間解像度・フレーム周波数・カラリメトリ・ダイナミックレンジ (SDRおよびHDR) などの映像パラメータ値を規定している。空間解像度は、水平7,680画素の最適水平視覚が96度であることから、画素構造が検知されない360度方向の水平方向の画素数として30,720画素 (7,680画素×4) を規定している (図7)。

さらに、360度VRコンテンツを放送システムで伝送する方法についても検討されている。日本の4K・8K衛星放送システム (ISDB-S3) や北米の次世代地上放送 (ATSC 3.0) のようなIP (Internet Protocol) ベースの放送システムは、メディアトランスポートプロトコルにMMTあるいはMPEG-DASHを用いている。そのため、現在の放送システムのアプリケーションレイヤを拡張し、前述のOMAFを用いることで、360度VRコンテンツを容易に伝送することが可能

*1 レポートBT.2420 : <https://www.itu.int/pub/R-REP-BT.2420>

360度VRコンテンツ			
メタデータ	映像	音声	時間付きテキスト
OMAF			
MMT			
UDP/IP			
TLV多重化方式*1			
ISDB-S3 伝送路符号化・変調			

ISDB-S3での360度VRコンテンツの伝送

360度VRコンテンツ			
メタデータ	映像	音声	時間付き テキスト
OMAF			
DASH		MMT	
ROUTE*2			
UDP/IP			
ALP*3			
ATSC 3.0 伝送路符号化・変調			

ATSC 3.0での360度VRコンテンツの伝送

*1 TLV: Type Length Value

*2 ROUTE: Real-Time Object Delivery over Unidirectional Transport

*3 ALP: ATSC 3.0 Link layer Protocol

図8 360度VRコンテンツを伝送する放送システムの構成

である (図8)。放送システムは一对多の一方方向の伝送であることなどを考慮し、OMAFのビューポートに非依存のベースラインプレゼンテーションプロファイルを用いること、映像信号はHEVCを用いるビューポートに非依存の映像プロファイルで符号化すること、音声信号はMPEG-H 3D Audioを用いる3Dオーディオベースラインプロファイルで符号化することなどのガイドラインを示した文書が作成されている。2019年前半の承認を目指し、検討が進められている。

4. ITU-T SG12での360度VR映像の品質評価に関する標準化

ITU-T (国際電気通信連合の電気通信標準化部門) において、性能、サービス品質およびユーザ体感品質に関する研究を所掌する研究委員会であるStudy Group 12は、2017年1月に、VRアプリケーションにおけるユーザ体感品質測定のための主観評価方法と品質モデルに関する新勧告 (G.QoE-VR) に向けた検討を開始した。

現在、ライブ、VoD、ゲーム、ソーシャル、ショッピングなどVRサービスのユースケースと、視覚、聴覚、奥行き感、またコーデック、ビットレート、映像の解像度やフレームレートなど、さらにヘッドマウントディスプレイの視線トラッキングの遅延やディスプレイ解像度など、ユーザの体感品質に影響を与える要素の整理が行われている。これらの要素を基に、VRアプリケーションのユーザ体感

品質の評価方法をまとめ、新勧告を作成する計画である。

またG.QoE-VRとは別に、360度映像のユーザ体感品質の評価方法についても、P.360-VRとして検討が進められている。

5. 3GPPが発行した360度VRアプリケーションに関する技術レポートと技術仕様

地上の移動通信システムの仕様を検討・作成している3GPP (Third Generation Partnership Project) のSA (Service & Systems Aspects) WG4 (コーデック) では、主に移動通信システムでの360度VRアプリケーションの実現性について検討が行われた。

この結果2018年3月に、技術レポートTR 26.918「3GPP上のVRメディアサービス」*2を発行した。この技術レポートは、VRのユースケースや既存技術を精査することで、今後、標準化を必要とする項目を整理することを目的としており、3DoFの360度VRコンテンツの映像と音声それぞれについて、撮影・録音から圧縮符号化・復号、配信、レンダリングまでの処理例の詳細がまとめられている。また、VRストーリーミング、ソーシャルTVとVR、VRによるビデオカンファレンス、ユーザ制作型のVRなどさまざまなユースケースもまとめられている。さらに、映像品質の評価方法と音声品質の評価方法についても触れている。

この技術レポートの発行を受け、2018年9月には、VRのストーリーミングサービスのアーキテクチャや映像・音声のプロファイルを規定する技術仕様TS 26.118「ストーリーミングアプリケーションのための3GPP VRプロファイル (リリース15)」*3も発行された。

6. VRインダストリーフォーラムが発行したガイドライン

VRインダストリーフォーラムは、電機メーカーや半導体メーカー、欧州の放送連合など多くの分野の事業者から成るフォーラムである。ISO/IECやITUと異なり、標準規格を作成することはせず、コンテンツ利用者のため、標準規格を利用した高品質なVR体験の普及を促進することを目的としている。

VRインダストリーフォーラムは2018年3月に、利用者が高品質で快適なVR体験が得られるよう、360度VRコンテンツの制作から、圧縮符号化、蓄積・伝送、コンテンツ保護などさまざまな側面をカバーするガイドライン*4を発行した。

表1 映像信号のマスターフォーマットの仕様

項 目	内 容
座標系	OMAFの規定に従う
射影変換	ERP
解像度	モノスコピック：最小4,096 × 2,048 ステレオスコピック：右眼・左眼それぞれに 最小4,096 × 2,048
カラリメトリ	ITU-R勧告BT.709
ファイルフォーマット	MXF*1 Program Contribution, IMF*2 Application 2eなど

*1 MXF: Material eXchange Format

*2 IMF: Interoperable Master Format

360度VRコンテンツの制作に関するガイドラインでは、ビューポートの変更がいわゆるVR酔いの原因になり得るため、不必要なカメラの動きや移動速度の変化を避けるといった内容が示されている。また、子供がヘッドマウントディスプレイを使用する危険性についてもまとめられている。

360度VRコンテンツのマスターフォーマットに関するガイドラインでは、表1に示す内容が推奨されている。360度VRコンテンツの配信時のファイルフォーマットとしては、前記のOMAFを用いることが推奨されている。さらに、MPEG-DASHを用いるビューポート依存のストーリーミングとダウンロード配信のためのOMAF-DASHのクライアントモデルも示されている。

このガイドラインに沿って360度VRコンテンツを制作し配信することで、より多くの機器で受信・表示し、快適にコンテンツを視聴できることが期待される。

7. むすび

本章では、360度VRアプリケーションに関する標準化について、MPEG, ITU-R SG6, ITU-T SG12, 3GPP SA WG4とVRインダストリーフォーラムの取り組みを概説した。

360度VRアプリケーションとして、図1に示したように視聴者の自由度の観点から3DoF, 3DoF+, 6DoFなどに分類することができる。現在、標準化されている技術は、固定点からの360度映像の視聴が可能な3DoFのアプリケーションに関するものである。今後は、より自由度の高い3DoF+や6DoFのアプリケーションを実現するための検討が進められる見込みである。

(2018年9月28日受付)

青木 秀一 あおき しゅういち
2003年、東京大学大学院工学系研究科修士課程修了。2013年、同大学情報理工学系研究科博士課程修了。2003年にNHK入局。以来、放送技術研究所にて、IP技術を用いる放送システムの研究開発やITU-R, MPEGでの国際標準化に従事。2010年、本学会優秀研究発表賞、2018年、経済産業省国際標準化奨励者表彰(産業技術環境局長表彰)等を受賞。博士(情報理工学)。正会員。

*2 3GPP TR 26.918 : <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3053>

*3 3GPP TS 26.118 : <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3325>

*4 VRインダストリーフォーラムのガイドライン : <https://www.vr-if.org/guidelines/>