



## 自由視点ARストリーミング技術



日本放送協会  
放送技術研究所  
伝送システム研究部  
研究員

ひるま のぶひろ  
蛭間 信博



日本放送協会  
放送技術研究所  
伝送システム研究部  
研究員

かわむら ゆうき  
河村 侑輝

### 1. はじめに

2次元映像の定量的なスペック向上による進化を果たした8K放送が開始され、スペックの向上とは異なる軸でのメディア技術の進化が期待されている。そこで筆者らは、8K放送の開始当時から高い関心を集めていたAR/VR (Augmented Reality/Virtual Reality) 技術を用いて、3次元コンテンツを放送に組み合わせる新しい視聴スタイルを提案し、それを実現する伝送技術の研究開発に取り組んでいる<sup>[1]</sup>。提案する伝送技術は、将来的に生放送番組にも対応させるため、3次元コンテンツを効率的にストリーミング伝送することを目指したものである。技研公開2022<sup>[2]</sup>では、提案する新しい視聴スタイルを体験することができるコンテンツを「NHKスペシャル 恐竜超世界」を題材に制作した。視聴者はリポーターと共にタイムカプセルに乗り込み、恐竜が存在していた時代にタイムスリップするという設定で、「恐竜超世界 特別編」を視聴しながら、番組に登場するスピノサウルスやアンモナイトなどを3DコンテンツとしてARで楽しめる。参考文献 [2] に体験動画が掲載されているので、ぜひご覧いただきたい。

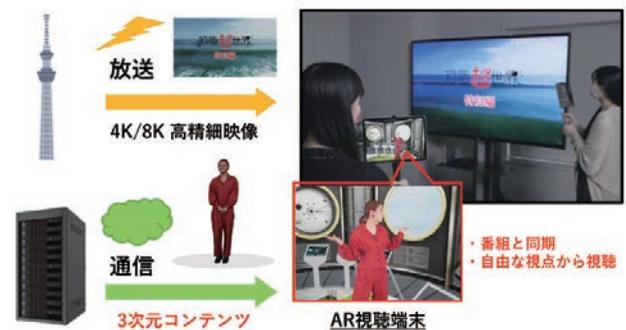
本稿ではまず、筆者らが提案してきた新しい視聴スタイルの概要と要求条件について説明する。そのあと、3次元コンテンツの効率的な伝送を目指して研究開発を進めている、オブジェクトベース伝送と、それを応用した伝送効率向上の仕組みについて述べる。最後に、技研公開2022で展示したコンテンツを例に、実際にオブジェクトベース伝送で配信するデータを作成し、1つのコンテンツとしてパッケージ化する具体的な手法を紹介する。

### 2. 提案する視聴スタイルの概要と要求条件

提案する新しい視聴スタイルの概要を図1に示す。テレビ放送による2次元の高精細映像に同期した、通信（インターネット）経由で伝送する3次元コンテンツをARで視聴端末

に提示する。テレビ放送番組と3次元コンテンツは同じ時間軸・ストーリーを共有しており、2つのコンテンツを連動させた新しい映像表現を演出することができる。例えば、技研公開2022のコンテンツにおいて、テレビ画面に映っていたリポーターがテレビ画面から消えると同時に、3DコンテンツとしてAR視聴端末に転送されてくる演出などである。また、3次元コンテンツは視聴者にとって自由な視点から視聴することができるため、従来の2次元映像にはなかった、新しい視聴体験の提供が可能となる。

上述したようなサービスに求められる要求条件として、3次元コンテンツをリアルタイムにストリーミング伝送することが挙げられる。放送に連動する3次元コンテンツを事前にダウンロードする場合と比較して、視聴者のインタラクションコストを下げることができ、生放送番組と同期させることが可能になるためである。一方で、3次元コンテンツは一般的に大容量のデータとなるため、効率的な伝送方法が求められる。そこで筆者らが研究開発を進めている伝送方式が、オブジェクトベース伝送である<sup>[3]</sup>。



■ 図1. 提案する視聴スタイル

### 3. オブジェクトベース伝送

提案するオブジェクトベース伝送は、コンテンツを構成する出演者や背景物をそれぞれ1つのオブジェクトとみなし、

パケットレベルで識別可能な状態で伝送することを特徴とした、UDP (User Datagram Protocol)/IP (Internet Protocol) ベースの独自の伝送プロトコルである。図2に、オブジェクトベース伝送のフレーム構造を示す。オブジェクトベース伝送では、後述する符号化した3次元データを、Packet IDという3次元オブジェクトを一意的に識別するための情報をヘッダ部に記載して伝送フレーム化する。そして、オブジェクトごとにフレーム化し、ペイロードヘッダ部に提示する時刻のタイムスタンプ (PTS: Presentation Time Stamp) などのメタデータを記載し、多重化する。これを用いることにより、複数の3次元モデルをまとめて伝送する場合と比較して、パケットレベルで各オブジェクトに対して柔軟な処理が可能となり、伝送効率の向上等に適用することができる。例えば、動きのない背景物などの3次元モデルは、動きのある出演者などの3次元モデルよりも低いフレームレートで伝送することができるため、圧縮したフレームレート分のデータ量を削減することができる。また、3次元モデルごとに優先度をつけて、3次元モデルそのものの品質を調整することで、合計の伝送データ量を調節するといったことも可能になる。

### 3.1 伝送する3次元モデル

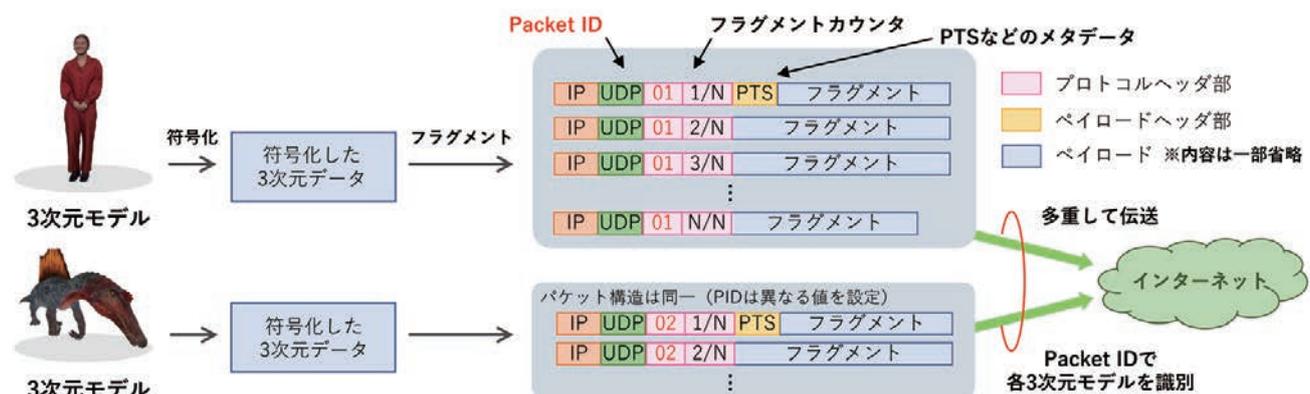
伝送する3次元モデルは、1フレームごとにメッシュジオメトリとテクスチャ画像で符号化されたものを想定する。なお、メッシュジオメトリの圧縮としてGoogle Draco<sup>[4]</sup>、テクスチャ画像圧縮としてJPEG、伝送フレーム形式としてglTF (GL Transmission Format) 2.0のバイナリ形式であるGLB File Format<sup>[5]</sup>をそれぞれ採用している。動きのある3次元モデルは、静止画のフレームを連続的に表示することで動画になる仕組みと同様に、1ファイルでは静止した3次元モデルを、30フレーム/秒などで伝送し、視聴端末で連続的

にレンダリングすることで表現する。一方で、動きのない背景物などの3次元モデルは、同一の静止画フレームを1フレーム/秒などで伝送する。なお、背景物であっても周期的に伝送する理由として、3次元コンテンツを途中から視聴を始める場合を想定している。

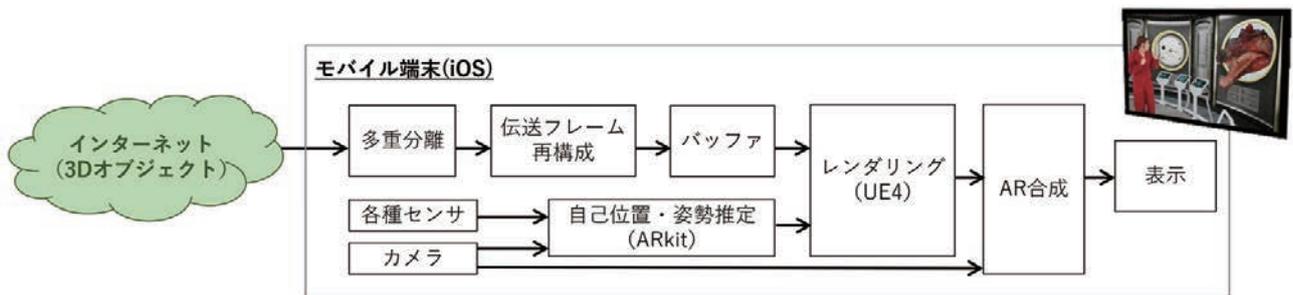
実際にオブジェクトベース伝送で配信する3次元モデルのデータは、便宜上、各オブジェクト名をフォルダ名とするフォルダに、伝送するフレーム数をファイル名とした連番のGLBデータとして管理する。例えば、120秒尺の技研公開のコンテンツにおいて、30フレーム/秒で伝送する1オブジェクトとみなしたりポーターの伝送データは、「reporter」というフォルダに、各フレームにおけるリポーターの3次元モデルの静止画フレームを上記の形式で符号化した、「00001~03600.glb」のような連番データとなる。

### 3.2 視聴アプリの試作

オブジェクトベース伝送により配信される3次元データを受信し、ARでレンダリングするアプリを、市販のモバイル端末 (iPad/Apple社) で動作するアプリとして実装した。図3にアプリのブロック構成図を示す。受信アプリは、まずオブジェクトベース伝送プロトコルでフレーム化されたIPパケットを受信し、多重化されている複数オブジェクトを分離する。次に、オブジェクトごとに伝送フレームを再構成し、パケットのヘッダ部に記載したタイムスタンプに従い、レンダリングするタイミングまでバッファする。そして、端末で撮影する画面に、3次元コンテンツをARで連続的にレンダリングして提示する。なお、レンダリングにはiOS/iPadOSのAR支援機能であるAR kitと、ゲームエンジン (Unreal Engine 4) を利用している。



■図2. オブジェクトベース伝送のフレーム構造



■図3. 視聴アプリのブロック構成

## 4. 伝送データ量の効率化

オブジェクトベース伝送を応用して、伝送するデータ量を効率化する仕組みの研究開発を進めている。ここでは、視聴者に応じて配信するデータを最適化することで、配信データ量の効率化を目指したオブジェクトフィルタ<sup>[6]</sup>と、配信する3次元データそのものを削減するテキストチャ間引き手法<sup>[7]</sup>について説明する。

### 4.1 オブジェクトフィルタ

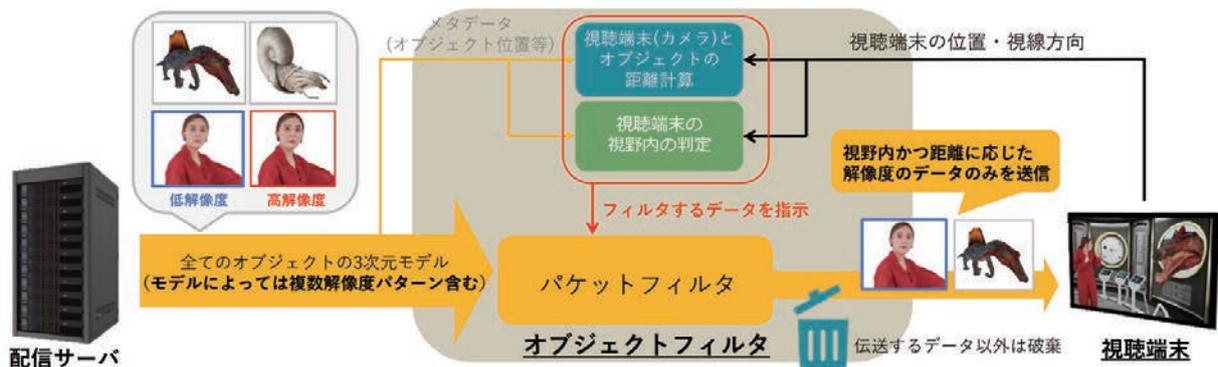
図4に、オブジェクトフィルタを含めたオブジェクトベース伝送のシステム構成図を示す。同図に示すように、オブジェクトフィルタは、配信サーバと視聴端末の伝送経路上に設置して、視聴者の位置・視線情報に応じて配信するデータを最適化する仕組みである。配信するデータは、視錐台カリリングと、解像度パターン選択の2つの手法で最適化される。視錐台カリリングは、一般的なCGレンダリングでも用いられる手法で、視聴端末の視野内にあるオブジェクトのみを伝送する。解像度パターン選択は、3次元コンテンツが視聴端末から遠い位置にある場合は小さく表示され、近い位置にある場合は大きく表示されることに着目し、複数解像度を用意したオブジェクトは、距離に応じた解像度のデータのみを伝送するという手法である。これにより、オ

ブジェクトフィルタと視聴端末の伝送区間において、常に高い解像度のデータを伝送する場合と比較して、効率的な伝送を行うことが可能となる。これらの処理は、Packet IDを識別子として用いることで、IPパケットレベルの簡易な処理として実装した。

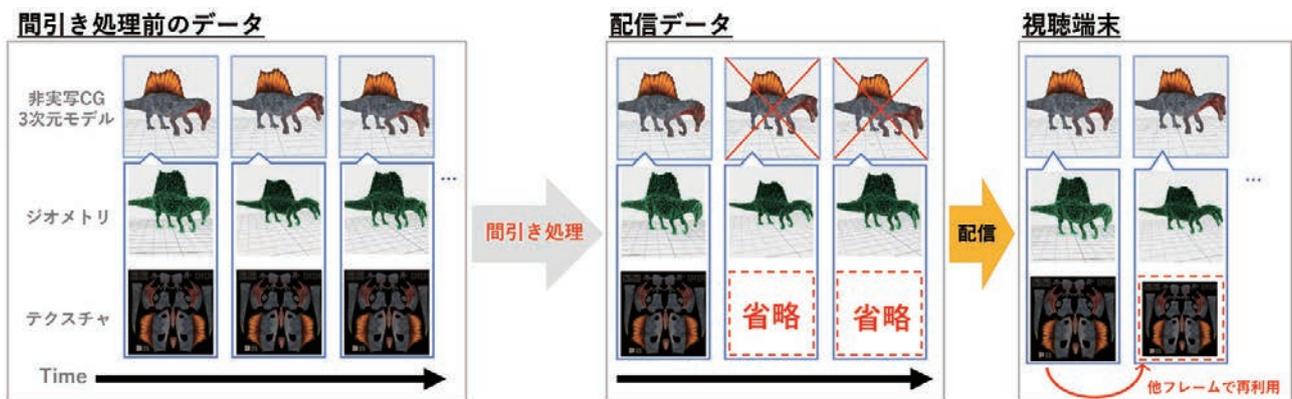
### 4.2 テキスチャの間引き

図5に、テキストチャ間引きの概念図を示す。CGモデリングツールなどで制作される、非実写のCGモデルは一般的にフレーム間でポリゴンの分割の仕方が変わらず、すべてのフレームで同一のテキストチャ画像が使われることに着目し、符号化された3次元モデルのうち情報量の多いテキストチャ画像を間引くというものである。テキストチャ画像を間引いたデータは、そのままでは3次元モデルとしてレンダリングすることはできないが、受信側でテキストチャ画像を含むデータを一度受信し、キャッシュして使いまわすことで、視聴端末では問題なくレンダリングすることができる。

技研公開2022コンテンツで使用したスピノサウルスを例に挙げてテキストチャ間引きの効果を紹介する。当該データは、頂点数が63479、テキストチャ解像度が1024×1024ピクセルであり、テキストチャを間引かない場合、伝送ビットレートは約109Mbpsになる。これに対して、30フレームに1度



■図4. オブジェクトフィルタのシステム構成



■図5. テクスチャ間引きの概念

だけテクスチャを含むように符号化されたデータを間引くことで、約34Mbpsまでデータ量を削減することができることを確認した。

## 5. コンテンツのパッケージ化

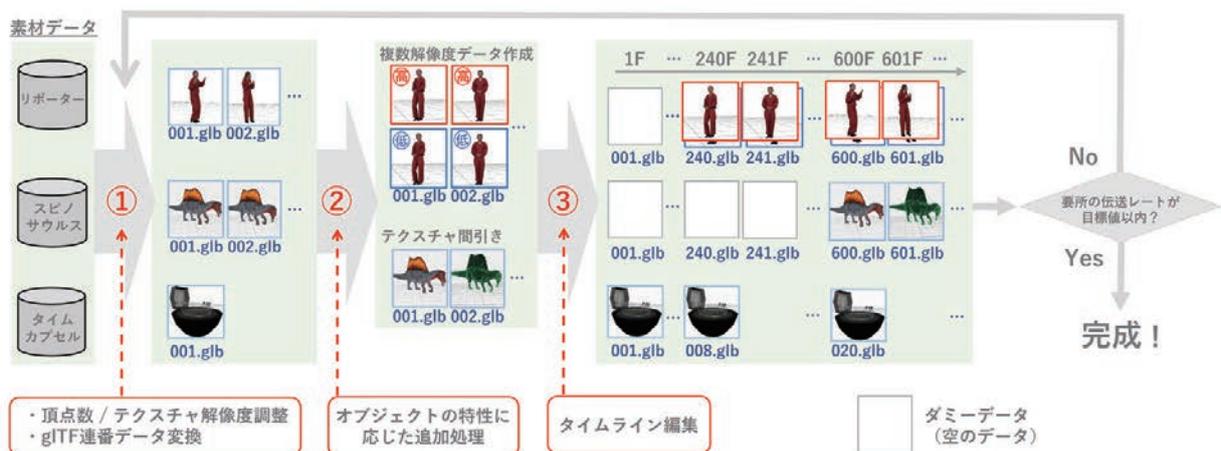
ここでは、技研公開2022で作成したコンテンツを例に、実際にオブジェクトベース伝送で配信するデータを作成し、コンテンツとしてパッケージ化するまでの流れを紹介する。

まずは、コンテンツにおける構成物のうち、静止部分と動きのある部分に分解できるかどうかを確認する。今回のコンテンツでは、タイムカプセルを、常に静止している床面部分と、陸上から海中への場面転換時に壁面と天井が粒子となって消える部分の2つに分解し、それぞれを別オブジェクトとして扱うことにした。

そして、実際に配信するデータを作成する。図6が、配信データ作成の流れを示したものである。同図では、データ作成において差異のある、ボリュメトリック映像データとしてリポーター、非実写CGデータとしてスピノサウルス、そ

して静止オブジェクトとしてタイムカプセル（床面部分）を例に示している。同図に示すように、作成方法は大きく3つの工程に分けられる。まず、①頂点数が多くテクスチャ解像度が高い状態の素材データを、頂点数とテクスチャ解像度を削減するとともに、オブジェクトごとにgltb形式の連番データに変換する。このとき、静止物オブジェクトはすべてのフレームで同一のため、作成するデータは1フレーム分となる。

次に、②オブジェクトごとに、オブジェクトの特性に応じてデータを作成・処理する。ここでは、ボリュメトリック映像はオブジェクトフィルタを前提とした、複数解像度パターンのデータを作成する。非実写CGオブジェクトは、4.2節で説明したテクスチャ間引き処理を適用する。なお、ボリュメトリック映像データ以外は、それぞれのデータ削減手法により、頂点数が多くテクスチャ解像度が高い高品質なデータとしても、安定したストリーミング伝送ができる伝送レートに収めることができたため、今回のコンテンツでは複数解像度は用意していない。



■図6. 配信データ作成手順



■図7. 技研公開2022のタイムライン

さらに、③作成した各オブジェクトのデータを、実際のコンテンツとして意図したタイミングで表示するためにタイムライン編集を行う。ここでは、ファイル名の編集作業になる。同図では、リポーターの素材データの1フレーム目が、実際のコンテンツでは240フレーム目に提示することが想定されているものとして、ファイル名を00001.glbから00240.glbに変更する例を示している。なお、オブジェクトが存在しないフレームではダミーのデータを挿入する。

最後に、伝送するオブジェクトによって変動する伝送ビットレートが、要所で安定したストリーミングを行える目安のビットレートに収まっているかを確認する。図7が、技研公開2022のコンテンツのタイムラインである。同図に示すように、特に複数のオブジェクトを同時に伝送するタイミングで、目標とするビットレートに収まっているかを確認し、収まっていない場合は①の工程に戻り、3次元モデルの頂点数、テクスチャ解像度を調整してデータを再度作成する。収まっている場合は、実際に伝送を行い、安定したストリーミング伝送を行えることを確認して、配信データのパッケージ化が完了する。

## 6. おわりに

本稿では、放送番組に3次元コンテンツを同期させた新しい視聴体験の提供を目指したサービスの概要について述べた後、それを実現するために研究開発を進めている伝送技術について説明した。その後、技研公開2022のコンテンツを例に、実際の配信データ作成手法について紹介した。

今後はサービスの具体化を目指し、配信データ作成にか

かるコストを削減するような手法の開発や、大規模配信を想定してCDN (Content Delivery Network) を利用することができるTCP(Transmission Control Protocol) /IPベースの伝送技術に加えて、視聴者のインタラクションコストが低いブラウザベースのアプリ開発・検証を進めていく予定である。

## 参考文献

- [1] 河村 侑輝, “テレビ映像と同期した自由視点ARコンテンツのリアルタイム伝送技術の開発,” 映像情報メディア学会誌, vol.75, no.1, p.131-138, Jan. 2021
- [2] NHK, “技研公開2022 展示5 自由視点ARストリーミング技術,” <https://www.nhk.or.jp/strl/open2022/tenji/5/index.html>
- [3] 河村 侑輝, 前田 恭孝, 久富 健介, 今村 浩一郎, “3次元コンテンツのオブジェクトベース伝送と複数提示方式による適応的視聴,” 電子情報通信学会技術報告, vol.121, no.73, SIS2021-6, p.32-36, June 2021
- [4] Frank Galligan, “Draco Bitstream Specification,” <https://google.github.io/draco/spec/>, Oct.2017.
- [5] Khronos Group, “The glTF 2.0 Specification,” <https://github.com/KhronosGroup/glTF/blob/main/specification/2.0/>
- [6] 河村 侑輝, 蛭間 信博, 今村 浩一郎, “3次元自由視点ARストリーミングの効率的配信に向けた可視オブジェクトフィルタターゲットウェイの実装,” 映情学技報, vol.45, no.35, p.35-40, Nov. 2021
- [7] 蛭間 信博, 河村 侑輝, 今村 浩一郎, “3次元空間コンテンツの効率的なストリーミング手法の一検討,” 映情学冬大, 22B-1, 2021