

# XRコンテンツの撮影・編集・配信 向けスタジオ「NTT XR Studio」

移動機開発部

あべ	ももこ	ふじもと	ようへい
阿部	桃子†1	藤本	洋平†2
かとう	さだあつ	まとは	なおと
加藤	禎篤†2	的場	直人†2
よしやま	かずゆき	あさかわ	えいいち
吉山	和志†2	浅川	栄一†2

近年、XRサービスの本格的な普及に向けて、コンテンツの制作や編集、配信技術の確立が重要となっている。そこでドコモは2021年1月、東京台場のテレコムセンタービルにXR専用の撮影、編集、配信向けスタジオ「NTT XR Studio」を開設した。本稿ではこのスタジオ設備、およびスタジオ内に設置している各システムの技術や特長について解説するとともに、実際の活用事例について紹介する。

## 1. まえがき

近年、XRサービスの普及に向けて、コンテンツの制作や編集、配信技術の確立が重要となっている。特にアバターとして人物を立体的（3次元）に表現する技術や、人物の外見や動作を「あるがまま」の3次元動画として保存するVolumetric Video\*1の撮影技術は、XR空間を活用したサービス上で「自分」を含む「人物」を表現するために重要な技術となる。さらに、撮影したものや外部から取得したXR向けコ

ンテンツの編集から配信までを一元的に行う「ワンストップサービス」の実現により、XR向けコンテンツ制作・配信サービスの円滑な普及が可能になると考えられる。

ドコモは2021年1月、東京台場のテレコムセンタービルにXR専用の撮影、編集、配信向けスタジオ「NTT XR Studio」を開設した。スタジオ内ではVolumetric Videoの撮影や、撮影したコンテンツの編集、配信を行うための一連のシステムを配備している。また、人物の動きをリアルタイムにキャプ

©2022 NTT DOCOMO, INC.

本誌掲載記事の無断転載を禁じます。

本誌に掲載されている社名、製品およびソフトウェア、サービスなどの名称は、各社の商標または登録商標。

†1 現在、プロダクトデザイン部

†2 現在、スマートライフカンパニー スマートライフ戦略部

\*1 Volumetric Video：人物・物体の3D動画データのこと。対象物の形状データと表面のテクスチャ（\*2参照）データで構成されている。

チャして配信する「モーショキャプチャ」システムも構築している。本稿ではこれらのスタジオ内に設置している各システムの技術や特長について解説するとともに、これまでの活用事例について紹介する。

## 2. Volumetric Video撮影システム

NTT XR Studioには、多角的に撮影された被写体がデジタル空間上に立体（3次元）として再構成されたVolumetric Videoを生成するための最新機材を設置している。その中で特に特徴的なものが、TetaVi社（イスラエル）の「TetaVi Studio」という撮影機材で、通常はVolumetric Videoの撮影を行う際に必要なグリーンバック（緑色の単色背景）が不要である。これは、本機材が3次元のテクスチャ\*2を撮影する通常のカメラとは別に、対象物までの深度情報\*3を取得する手法をもつことで、撮影対象の人物の領域を抽出することが可能となるためである（図1）。

本撮影システムは、撮影対象が主に「人物」（1～2名）であり、直径2.9m、高さ1.8mの仮想円筒内の領域で撮影可能である。静止状態だけでなく動いている状態も撮影することができ、被写体が動いている場合は立体（3次元）のデータが時系列で記録される。撮影されたデータは360度どの方向からも視

聴することが可能であり、また視聴位置を前後に調整することによる拡大、縮小表示を実現できる。さらに本システムは、ライトによる明るさ制御が可能な屋内であれば移動して設置、撮影することも可能であり、背景の制限を受けることなく演劇用の舞台装置やスポーツ設備を背景に撮影することもできる。

本システムの活用事例としては、WebARを活用したVolumetric Videoの視聴が挙げられる（図2）。スマートフォンのブラウザを使い、外部カメラで撮影された背景映像上にVolumetric Videoの人物を重畳することで、例えば、手のひらの上に乗せるなど、自分の部屋の中にあたかも撮影された人物が存在するようにも見せることができる。視聴者がスマートフォンを任意の空間に向けると画面上にコンテンツが表示され、それに対してスマートフォンの角度や距離を変更することで、Volumetric Videoの人物の周りを移動したり、近づいたりするように感じられる。XRコンテンツならではの体験が可能である。

## 3. Volumetric Videoの配信システム・プレイヤー

NTT XR Studioは、前述したVolumetric Video撮影機材に加えて、マルチデバイスに対応したVolumetric Videoストリーミング配信システム・プレイヤーを備えており、撮影からコンテンツ配信まで



図1 TetaVi Studioの撮影例，データ作成例

\*2 テクスチャ：3D形式のデータにおいて、物体表面の質感を表現するために物体表面に貼り付ける画像のこと。

\*3 深度情報：本稿では、カメラ位置からの距離（奥行き）を表す。

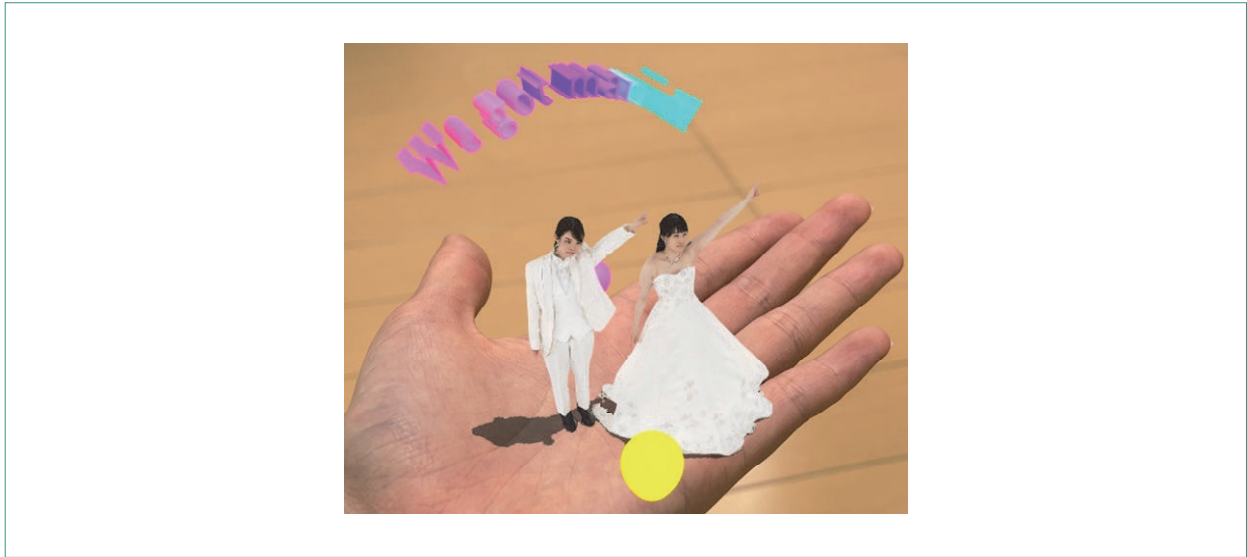


図2 TetaVi Studioで撮影したVolumetric Videoの活用事例

の一気通貫したサービスをワンストップで提供可能である。

### 3.1 Volumetric Video配信における課題

Volumetric Videoは、臨場感と存在感のある自由視点立体映像を表現することが可能であるが、2D動画に比べて情報量が多いため、ファイルサイズが比較的大きくなる。このため、Volumetric Videoの再生をダウンロード方式で行う場合、コンテンツを選択してから再生するまでにかかる時間が長くなり、ユーザ体験が損なわれることが懸念される。このような背景から、これまでのVolumetric Videoを用いたコンテンツ配信では、主にファイルサイズが小さい短尺コンテンツ配信が行われてきた。

ユーザ体験を損なうことなくファイルサイズが大きい長尺コンテンツを配信するためには、2D動画の配信方式で採用されているストリーミング配信方式を行うことが望ましい。ストリーミング配信は、コンテンツファイルを時系列に沿って細かく分割した上でサーバへ配置する。プレイヤーはコンテンツ冒頭のファイルのダウンロードを完了した時点で再生

を開始し、その後続く時系列データのダウンロードを継続しながらコンテンツの再生を行う。コンテンツファイルのすべてをダウンロードしてから再生を開始するダウンロード方式とは異なり、コンテンツ冒頭の小さなファイルをダウンロードした時点で再生を開始できるため、コンテンツ再生までにかかる時間を短縮することができる。しかし、Volumetric Videoのファイルフォーマットは2D動画とは異なるため、2D動画で用いられるストリーミング配信方式をVolumetric Videoへ単純に適用することはできず、Volumetric Video用の新しい配信方式が求められていた。

### 3.2 Volumetric Videoストリーミング配信

前述の課題を解決するため、ドコモはArcturus Studios Holdings, Inc. [1] の技術協力により、長尺コンテンツのストリーミング配信技術、およびコンテンツのクオリティを一定以上に保ちながら配信時のビットレートを最適化する技術を開発した(図3)。まずVolumetric Video撮影システムで被写体を立体のデータとして収録し、そのデータをストリーミン

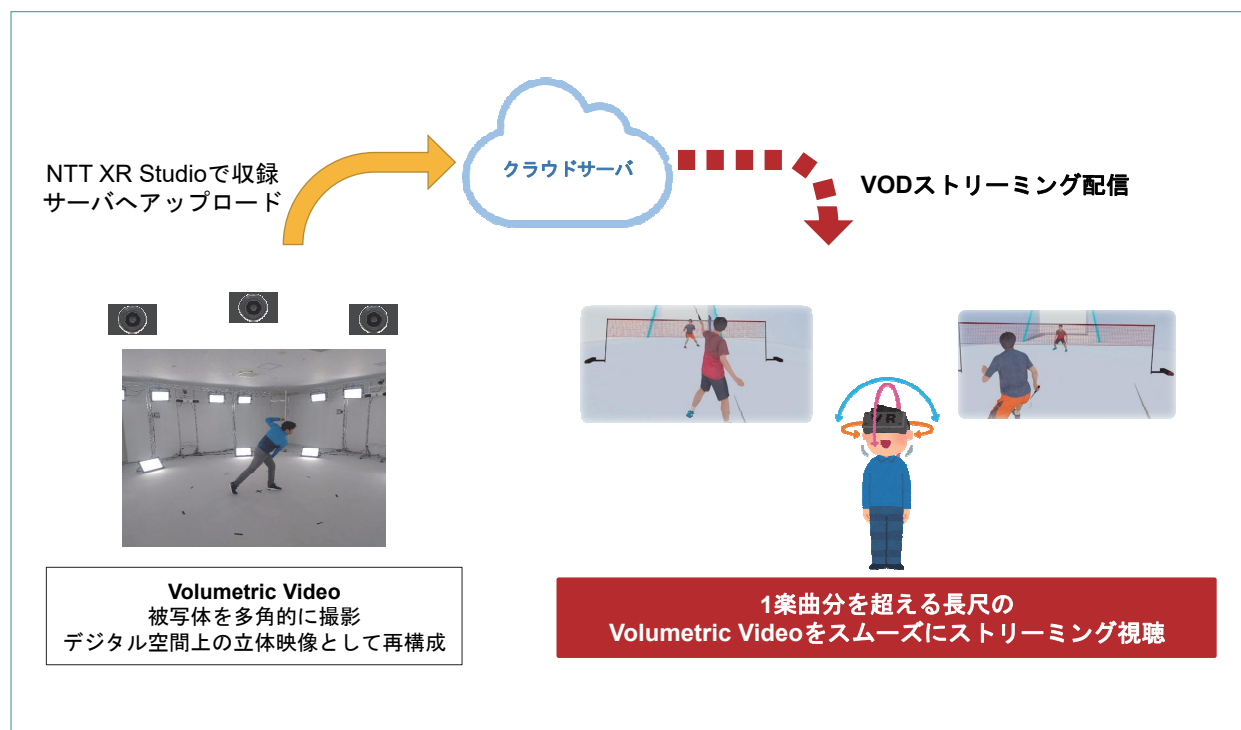


図3 Volumetric Videoストリーミング配信システム構成図

グ配信用データへ変換した後にクラウドサーバへアップロードする。ユーザはVolumetric Videoプレイヤーを備えたデバイスからそのサーバへアクセスすることで、コンテンツを視聴することができる。長尺で高品質なVolumetric Videoのストリーミング配信が可能となることで、ユーザの再生コストの低減や配信コンテンツの幅が広がることが期待される。本技術を利用したVolumetric Videoプレイヤーは、VR (Virtual Reality) ヘッドマウントディスプレイ\*4やMR (Mixed Reality)\*5デバイス、スマートフォン、PCといった多様なデバイスに対応しており、ユーザは場所や用途を問わずさまざまなシーンでVolumetric Videoコンテンツを視聴することが可能である。

また、Volumetric Videoは未だ統一されたファイルフォーマットが存在せず、各撮影機材によって異なるファイルフォーマットが採用されている。本技

術によるストリーミング配信では、代表的な複数のVolumetric Videoファイルフォーマットの入力に対応した専用の編集ツールを活用することで、撮影機材やファイルフォーマットに囚われないストリーミング配信が可能である。

#### 4. モーションキャプチャによる動き情報の取得とデータ配信技術

NTT XR Studioには、光学式モーションキャプチャ機材が導入されている。モーションキャプチャとは、現実のヒト・モノの動きをデジタルデータとして取得する技術であり、スポーツや医学的な動作解析といった用途や、映画やゲームにおけるコンピュータグラフィックス (CG) を用いた映像制作といったエンターテインメント用途などに幅広く活用されている。

\*4 VRヘッドマウントディスプレイ：頭部に装着するディスプレイ装置の総称。目のすぐ前における映像提示を可能とし、片目だけに映像を提示する単眼タイプと、両目に映像を提示する両眼タイプがある。メガネやゴーグルなどのレンズ部分に映像を提示するものが代表的である。

\*5 MR：現実世界を写した映像に、電子的な情報を重ねて、ユーザに提示する技術。ARと異なり、自由視点での表示など情報を実際にそこにあるかのように提示する。

#### 4.1 光学式モーションキャプチャ

光学式モーションキャプチャは、撮影対象である物体の動きを3次元データとして取得するシステムである。この方式では、まず撮影対象の各部位にマーカーを取り付け、そのマーカーの位置を複数台のカメラで撮影することでそれぞれのマーカーの3次元位置を推定する。対象物体がヒトである場合は、得られた各部位のマーカーの3次元位置を、あらかじめ定義された人体モデルと照らし合わせることで、撮影対象の関節位置や角度を推定することが可能となる。

光学式モーションキャプチャは、撮影のためにカメラを複数台設置する必要があり比較的広いスペースを必要とするが、他の手法に比べ検出精度が高く、撮影対象の動きが大きく制約されないことなどが特長として挙げられる。

#### 4.2 モーションキャプチャを用いたエンターテインメントコンテンツ制作・ライブ配信

前述のとおりモーションキャプチャは、映画やゲームのCG映像制作といったエンターテインメント用途で広く使われている。特に近年のエンターテインメント業界では、モーションキャプチャと3DCGを用いて制作した2D映像のライブ配信がさかんに行われている。このような映像は、モーションキャプチャにより取得した演者の動きを用いて3DCGのアバターキャラクターをリアルタイムに操作することで生み出される。クオリティの高い3DCGアバターをリアルタイムに操作することで、演者とユーザがライブならではのリアルタイムコミュニケーションを行えることが特長である。

この2D映像配信では、3DCGの描画は配信者側で行い、その描画結果を2D映像としてキャプチャしユーザへ配信することが一般的である。これは、ユーザがコンテンツ視聴に利用するデバイスにとって3DCG描画の処理負荷が大きいことや、3Dコンテン

ツ視聴に対応したデバイスがこれまで少なかったことが主な理由である。しかし、近年VRヘッドマウントディスプレイやMRデバイスなどの3Dコンテンツ視聴に特化したデバイスの普及が進み、それらデバイスの中には3DCG描画をリアルタイムに行うことができるものが存在する。これらデバイスに向けて2D映像よりも豊かな体験をユーザへ提供できる3Dコンテンツ配信の需要が高まっており、その対応が求められている。

#### 4.3 モーションキャプチャで取得したモーションデータのライブ配信技術

ドコモは、前述の課題に対して、モーションキャプチャを用いた3DCGコンテンツのユーザへの提供を目的とし、モーションキャプチャで取得したモーションデータを、音声データとともにユーザのデバイスへライブ配信するシステムを開発した(図4)。このシステムでは、取得したモーションデータと音声データの処理を配信PCで即座に行い、配信サーバへデータをアップロードする。また、ユーザのデバイスにインストールされたプレイヤーアプリケーションは、更新されていく配信サーバのデータを逐次的に受信する。プレイヤーアプリケーションには事前に3DCGアバターデータを送付しておき、デバイスが受信したモーションデータを基にアバターを描画することで、臨場感と存在感のある3Dコンテンツをユーザへ提供する。

本システムでは、モーションキャプチャデータの配信プロトコルとして2D映像配信で一般的に広く使われる方式を採用しており、既存の配信ファイル暗号化手法の実装やコンテンツ・デリバリー・ネットワーク<sup>\*6</sup>を用いた大規模配信へ容易に対応可能である。

また、本システムではモーションキャプチャデータに加えて、表情解析によりリアルタイムに取得した演者の表情データとマイクにより取得した音声

<sup>\*6</sup> コンテンツ・デリバリー・ネットワーク：コンテンツを高速、大規模配信するためのネットワークのこと。

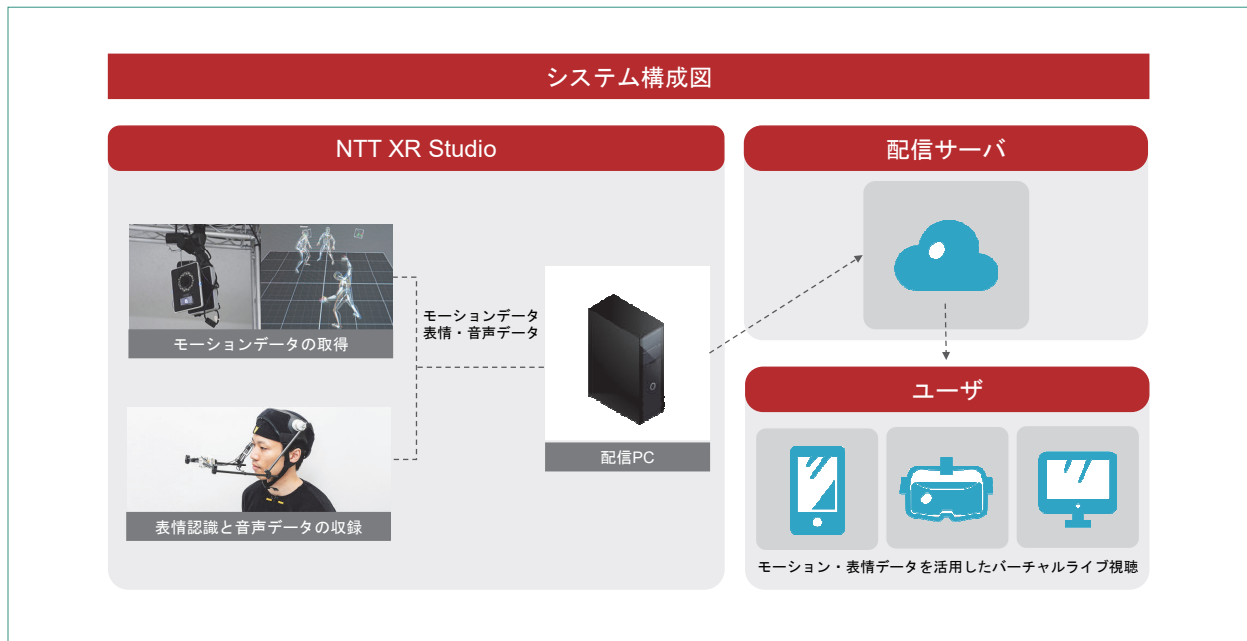


図4 モーションキャプチャデータライブ配信システム構成図

データを配信することが可能である。ユーザは自身のデバイスでこれらのデータを受信し、ユーザのデバイスにあらかじめ保存していた3DCGアバターデータに受信した動き、表情、音声を適用させることにより、ユーザは3DCGによる豊かなライブコンテンツを自由視点から鑑賞することができる。

## 5. あとがき

本稿では、「NTT XR Studio」の設備、およびスタジオ内に設置している撮影・編集・配信用の各シ

ステムの特長とその活用事例について解説した。ドコモは、今後も自社でXR Studioを構築することで撮影から配信までの一気通貫の環境を整備するとともに、この環境を外部のパートナーにも公開、開放して、社内外のコンテンツ開発者のフィードバックを得ながら協創することで、XR向けコンテンツの円滑な生成・普及に向けた活動を進めていく。

### 文献

- [1] ARCTURUS : "Powers volumetric video."  
<https://arcturus.studio/>