

5G が開く VR

Immersive VR Experience Realized by 5G Network

福井啓允 大塚裕太 野中敬介 内藤 整



モバイル映像サービスとして VR 視聴への期待が高まっている。VR 視聴をモバイルサービスとして実現する上では、画質や遅延といった体感品質の向上が不可欠であり、5G の積極的な利用が想定されている。日本国内での VR を使用した 5G トライアルとしては、新宿での au 5G 体感トライアル（報道機関向け、2017 年 4 月）、上野駅での「南三陸さんさん商店街へ瞬間移動」（一般向け、2018 年 1 月）が開催された。本稿では VR に関わる技術動向を導入した上で、上述の 5G トライアルの紹介を行い、最後に将来の VR 技術として自由視点 VR の技術解説を行う。

キーワード：ロケーション型 VR, VR ライブ中継, 5G 体験イベント, 自由視点 VR

1. ま え が き

モバイル映像サービスとして VR コンテンツの視聴に対する期待が高まっている。VR 視聴をモバイルデバイスにより実現する上では、端末のサイズ・消費電力を抑える一方で、画質や遅延といった体感品質の向上が不可欠であり、5G の積極的な利用が期待されている。日本国内での VR を使用した 5G トライアルとしては、2017 年 4 月に新宿で au 5G 体感トライアルを報道機関向けに実施した。更に一般向けには、2018 年 1 月に上野駅で実施される地域再発見プロジェクト「宮城産直市」において「南三陸さんさん商店街へ瞬間移動」という、地方都市を都会で感じられる体験イベントが実施された。

本稿では、2. で VR に関わるグローバルな技術動向を解説した上で、3. で上述の VR 伝送トライアルについて紹介する。最後に 4. では将来の VR 技術として期待される自由視点 VR に着目し、技術解説を行う。

2. 技 術 動 向

本章では VR 伝送を取り巻く技術動向について解説する。

2.1 伝送方式

VR 伝送方式の一つとして、全天球映像を配信し、クライアントで視野映像を合成する方式がある。全天球映像はユーザから全外界方向（360 度）を見渡したときの映像である。MPEG のシステム規格である OMAF (Omnidirectional Media Application Format) においては、最も一般的な Equirectangular 方式をはじめ Cubic や Cylinder など様々な投影形式が提案され、代表的な伝送方式が ISO/IEC23000-20 として規格化された。

また、その他の伝送方式として Facebook 社の Pyramid 方式に代表される Region-wise Packing 方式⁽¹⁾などがある。正面方向の画質や解像度を高くするように方向ごとの複数の映像をあらかじめ用意し、ユーザの向きに応じて伝送対象を切り換える。全天球方式に比べて品質向上（帯域削減）ができる反面、ストレージや符号化回数の点で効率が悪くなるという問題がある。

そこで、新しく Tile-Based Streaming 方式が提案されている⁽²⁾。全天球映像を HEVC の Tile に分割して符号化する。このとき動きベクトルが他の Tile を参照しないように制限を加えて符号化する。このような符号化映像を複数の画質や解像度分用意する。そして、伝送時

福井啓允 正員 (株)KDDI 総合研究所超臨場感通信グループ
大塚裕太 KDDI 株式会社次世代ネットワーク開発部
野中敬介 正員 (株)KDDI 総合研究所超臨場感通信グループ
内藤 整 正員：シニア会員 (株)KDDI 総合研究所超臨場感通信グループ
Hiromasa FUKUI, Keisuke NONAKA, Members, Sei NAITO, Senior Member
(Ultra-Realistic Communications Laboratory, KDDI Research, Inc., Fujimino-shi, 356-8502 Japan), and Yuta OTSUKA, Nonmember (Next Generation Network Development Department, KDDI Corporation, Tokyo, 102-8460 Japan).
電子情報通信学会誌 Vol.101 No.11 pp.1085-1090 2018 年 11 月
©電子情報通信学会 2018

はユーザの向きに応じて Tile 単位で映像を合成する。例えば、ユーザの正面方向の Tile は高解像度、逆方向は低解像度となるように一つの符号化映像に結合する。この結合操作を復号映像に対してではなく、符号化データに対して直接行うことで処理負荷の低減を追求する手法が提案されている。Region-wise Packing 方式に比べ、ストレージや符号化回数の効率が良いという利点がある。更に視野 (FOV と呼ばれる) の範囲が異なるデバイスにも柔軟に対応できる。

一方で、このような方式を実環境で利用する場合、伝送遅延が問題となる。ユーザが頭の向きを変えると、それがサーバに送信され、必要な映像が動的に選択・生成される。したがって、遅延が大きくなると、向きを変えた先でしばらく低解像度の再生が続くことになり、VR 視聴の体感品質が低下してしまう。このような問題に対して、低遅延コーデックの開発及び RAP や配信プロトコルの調整など、各社が多面的に改善を図っている。

2.2 映像符号化

ITU-T VCEG と ISO/IEC MPEG の連携組織である JVET (Joint Video Exploration Team) では、次世代符号化方式として、全天球映像に特化した符号化ツールが検討されており、射影方式を考慮した方式提案及び評価が行われている。

2.3 ヘッドマウントディスプレイ

Oculus 社の Oculus Rift や HTC 社の VIVE に代表される PC 形、Oculus Go やレノボ社の Mirage Solo などのスタンドアロン形に加えて、スマートフォンと専用ゴーグルを組み合わせる Samsung 社の GearVR や Google 社の DayDream などのモバイル形が広く登場した。

スペック面では、Pimax 8K のように高解像度化や視野の広角化が進み、また Mirage Solo のように単体で 6DoF に対応するものが登場し、VR コンテンツの表現の幅が拡大している。

2.4 360 度カメラ

Go Pro Fusion をはじめコンシューマ製品でも解像度やステッチング精度が向上し、SNS の普及もあいまって 360 度映像のトラヒックが増大すると見込まれる。ハイエンドモデルとしては Insta 360 pro が 8K 撮影に対応し、4K であれば単体でのリアルタイム配信も可能となった。

上述のように VR 伝送は複数の技術分野において進化が続いている。ユーザの向きに応じた伝送では低遅延性が求められ、8K 解像度の映像配信では高速・大容量が必要といったように、今後 5G は VR 伝送にとって必要不可欠となると考えられる。

3. 5G を使った VR 伝送トライアル

本章では KDDI 株式会社が行った実証実験について紹介する。

3.1 ロケーション型 VR

2017 年 4 月に KDDI 新宿ビル周辺に 28 GHz 帯を使った 5G の試験エリアを構築し、人や車が往来する実環境で動作するアプリケーションを発表した。基地局側のアンテナは約 10 m の高さの柱に固定、受信機側のアンテナはバスの上部に設置し、5G 通信を行った。

「超高速」と「低遅延」を訴求するデモの一つとして、バス移動と連動した VR 視聴を車内で体験できるシステムを構築した。図 1 にシステム構成を示す。デモシナリオとしては月面探査機に搭載したカメラからの視点を VR ゴーグル越しに体感するというものである。バスが左折すると月面探査機も左折するといったように、実際の動きの感覚に視聴映像が連動することで、あたかも月面を移動しているかのような高い没入感を実現した。

より具体的には、まずバスに搭載した GPS と加速度センサの情報を 5G 経由でサーバに送信する。サーバでは位置情報に応じた視点映像を 4K 解像度で CG データから合成する。そして、低遅延配信向けにチューニングされたコーデック及び配信プロトコルを用いて、符号化映像を 5G 経由で VR ゴーグルに送信する。VR ゴーグルでは、受信した符号化映像を復号し、ユーザの頭の向

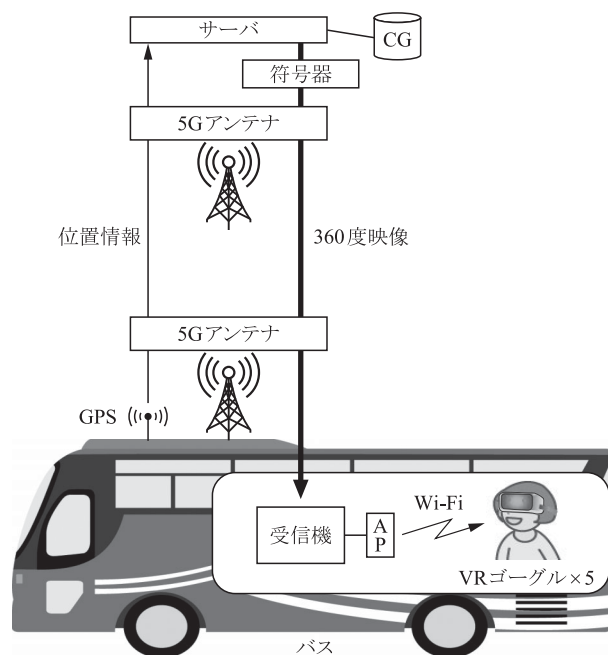


図1 システム構成 サーバとバスを5Gで接続。サーバはバスの位置を受信し、対応する視点画像を合成・符号化して送信する。VR ゴーグルは符号化映像を復号し、ユーザの向きに応じて必要な視野映像をレンダリングする。

きに応じて必要な視野映像をレンダリングする。

バスの位置情報の取得から VR ゴーグルでの映像表示までに要する遅延は 100 ms 程度であり、本デモのような形態では違和感のない体験が得られることを確認した。なお、VR ゴーグルとしては Samsung 社の GalaxyS7 と GearVR を用いた。商用モデルであり、5G の電波を直接受信することはできないため、本デモではバス内の評価用受信機が受信したデータを Wi-Fi 経由で中継した。将来的にはこれらの機器群が一つのモバイルデバイスに統合される見通しである。

このように高負荷なレンダリングをクラウド側で行い、また VR 対応のモバイルデバイスが普及すれば、一般のモバイルユーザにおいても手軽に上述の視聴体験を享受できる。

3.2 5G と VR によるお買物体験イベント

本節では、5G と VR でリアルタイムにコミュニケーションが可能な現地体感イベントを国内で初めて実現した事例を述べる。本イベントは東日本旅客鉄道株式会社の協力により、KDDI 株式会社が 2018 年 1 月に上野駅

で実施したものである。

本イベントは、5G と VR を利用し、上野駅と南三陸さんさん商店街をリアルタイムでつなぐという新しい現地体感イベントである。図 2 のとおり、上野駅の特設会場にて装着する専用の VR ゴーグルと、南三陸さんさん商店街の現地ガイドが持つ 4K 解像度の 360 度カメラ（以下、VR カメラ）をリアルタイムに接続する。これにより、上野駅での参加者がガイドを通じて南三陸さんさん商店街でのショッピングを擬似体験でき、宮城の魅力を存分に体感可能である。

本イベントで用いた 5G による VR 映像伝送システムの構成を図 3 に示す。なお、本イベントにおいては音声による双方向コミュニケーションも併せて実現しているが、本稿では音声システムのシステムについては割愛する。

本イベントで用いた VR 映像伝送システムは、アプリケーション部、5G システム部、ネットワーク部に分けられる。アプリケーション部は VR 映像の撮影、撮影した映像の符号化や復号、IP パケットの生成や送受信を行う。5G システム部は、IP パケットの無線伝送を行う。ネットワーク部は、IP パケットの有線伝送を行う。

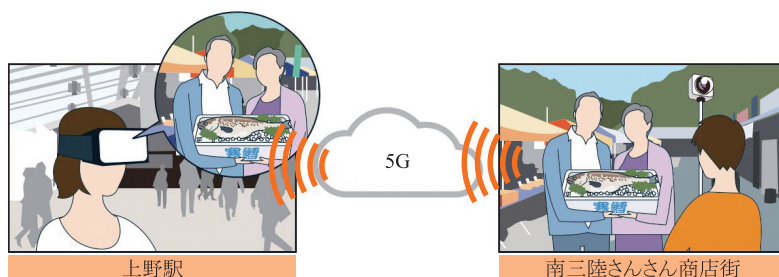


図 2 現地体感イベント概要 南三陸さんさん商店街の様様を VR カメラで撮影し、5G ネットワークを介して上野駅の体験ブースにライブ中継する。

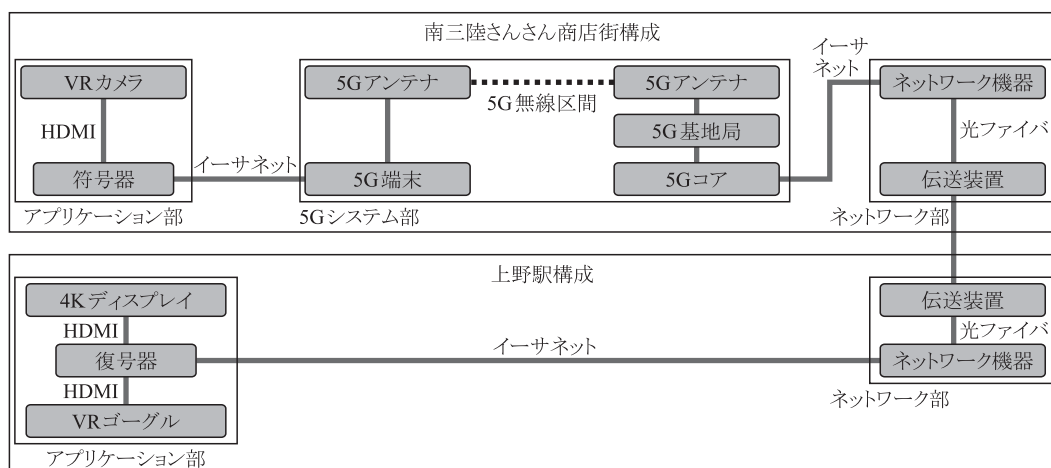


図 3 システム構成 南三陸さんさん商店街において VR カメラで撮影した映像を 5G 回線によって基地局までライブ中継する。5G のカバーエリアは商店街のストリート部分である。



(a) 基地局側 5G アンテナ



(b) 端末側 5G アンテナ

図 4 5G アンテナの外観

VR カメラで撮影された映像は、HDMI ケーブルにより符号器に伝送される。符号器にて処理された IP パケットは、5G システム、ネットワーク機器、伝送用回線を経て、上野駅に伝送される。上野駅に設置した復号器にて VR 映像を復号し、ゴーグルにてその映像を視聴できる。これにより、参加者は南三陸さんさん商店街の様子を上野駅にいながら体感できる。

本イベントでは、南三陸商店街敷地内に設置した基地局から同敷地内の端末間を 5G 無線で伝送している。図 4 に 5G アンテナの外観を示す。5G エリアとしては、商店街のストリート部分をカバーした。これにより、5G 端末は 5G 基地局との通信品質を維持した状態で、イベントに協力頂いた各商店街店舗の軒先を移動することができた。

本イベントにより、5G による VR 映像伝送システム及び双方向の音声通話を併用することで、遠隔地との臨場感あふれる双方向コミュニケーションを実証できた。またイベント期間を通じて、5G 回線における著しい品質低下は確認されておらず、結果としてコミュニケーションを阻害するような問題は皆無であった。

今後の展望として、イベント会場や中継会場での 5G の活用が期待できる。イベントによっては、会場や現場の制約からケーブル配線が困難なケースが想定される。

このような状況において、超高速・低遅延という特長を有する 5G を活用することで、イベント会場における設備敷設の容易性を各段に向上できると期待される。

4. 将来期待される技術

ここまで、現在実施されている 5G と VR 技術の活用事例を紹介した。これらの事例において利用されるコンテンツは、360 度カメラで撮影された実写映像、またはあらかじめ制作された CG コンテンツとなっている。前者に関しては「カメラの撮影位置に VR 視聴者の視点が固定される」、後者に関しては「実写コンテンツの利用が難しい」という課題がある。同課題の解決技術の一つとして、ここでは「自由視点 VR」を紹介する。

自由視点 VR とは、複数の実写カメラ映像を基に被写体の空間を三次元コンピュータグラフィックス (3DCG) モデルとして生成することで、あらゆる位置における 360 度視聴 (≒VR 視聴) を可能とする技術である。また、3DCG モデルという大きなデータ量を扱うことから、超高速伝送を実現する 5G との親和性が高い。その技術詳細及び活用事例を以下に紹介する。

4.1 技術解説

将来的なモバイルデバイス (VR デバイスを含む) での利用を見据えると、その映像入力から自由視点 VR 表示までの計算コスト及びデータ量は小さい方が望ましい。そこで、自由視点 VR 技術の中でも、これらの要件を満足するビルボード型の自由視点 VR 方式 (以下では単に自由視点 VR と呼ぶ) について解説を行う。

同方式の処理フローを図 5 に示す。この方式では、まずカメラごとの映像から選手のテキストチャを抽出し厚みのない 1 枚の 3D の板 (ビルボード) として表現する。そして、図 5 中の「合成処理」のようにユーザが選択した仮想視点に応じて、それに正対するように最近傍カメラの選手のビルボードを回転させ 3D 空間上に配置することで、任意のアングルからの自由視点映像を生成する。

自由視点 VR 生成処理は、主に「(選手) オブジェクト抽出」及び「オブジェクト追跡」により構成される。ここで、自由視点 VR は図 5 に示すように複数台のカメラを使用することが一般的であり、同環境に沿った形で以後の解説を進める。複数カメラの時刻同期については、様々な方法が存在するが、ここではユーザ入力などによって既に時刻同期が行われた映像の入力を前提とする。また、カメラパラメータと呼ばれるカメラの 3D 空間上の位置情報などについても、あらかじめ算出されているものとする。

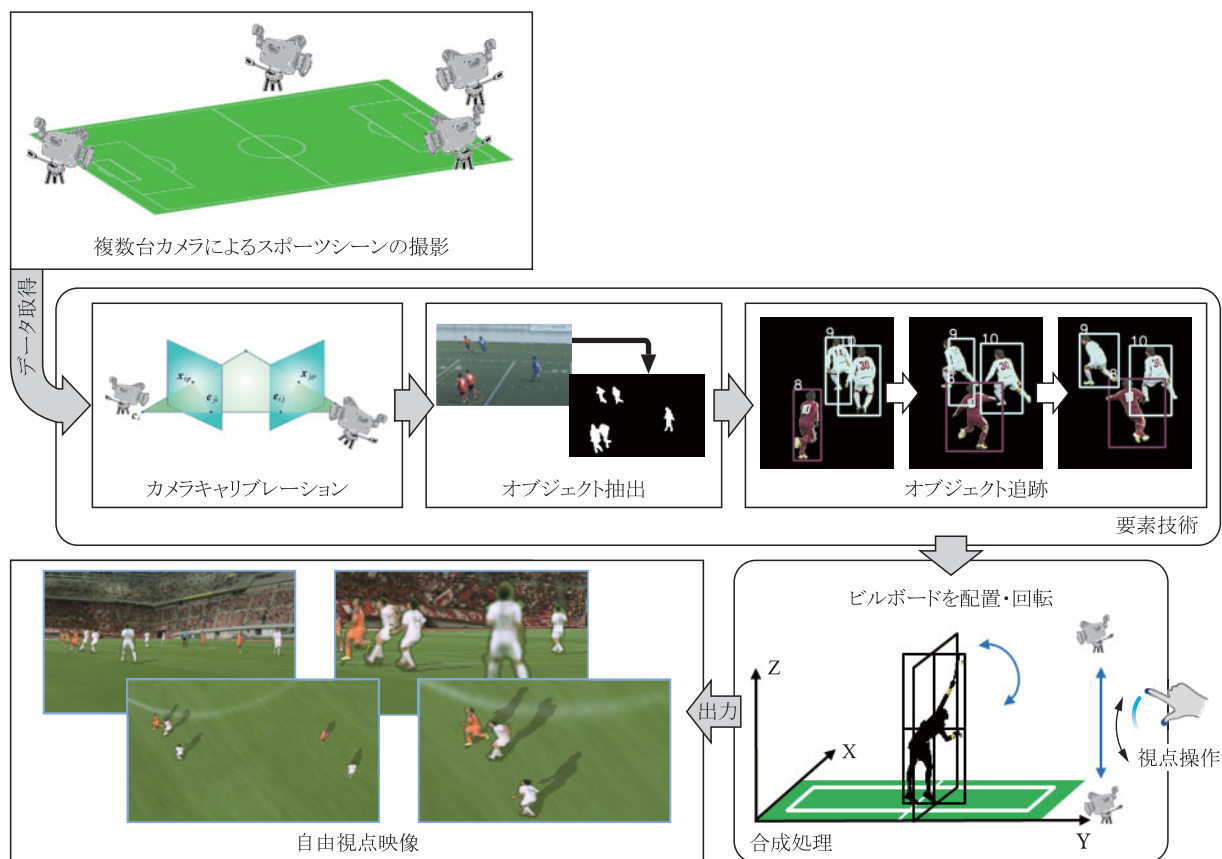


図5 自由視点 VR 映像の制作フロー ユーザーが選択した仮想視点に応じて、それに正対するように最近傍カメラの選手のビルボードを回転させ3D空間上に配置することで、任意の角度からの自由視点映像を生成する。

4.1.1 オブジェクト抽出

自由視点 VR においては、各選手オブジェクトをそれぞれのカメラ映像から抽出した選手領域を用いて表現する。ここで、選手オブジェクトの3D空間上の座標は、前述のカメラパラメータと抽出された選手の映像上の座標によって算出される。このため、オブジェクト抽出の精度は自由視点 VR の視聴映像における選手の動きの精度を決定する重要な要素である。

本抽出処理について、筆者らは選手が試合中におおむね動いていることを前提とした背景差分処理を採用している⁽³⁾。まず、連続したフレームを平均化することで背景画像と呼ばれる選手のいない画像を得る。この背景画像と選手抽出の対象フレームとの差分を取り、あるしきい値以上の差を持つ画素の集合を選手オブジェクトとして抽出する。

4.1.2 オブジェクト追跡

自由視点 VR を再生した際に、選手が配置される座標の時間軸方向の連続性はユーザーの体感品質に影響を与えるもう一つの重要な要素である。この座標の精度を向上させる目的から、フレーム間の情報を用いて選手モデル追跡を行い、同一の選手の動きを滑らかにする処理を

行っている。ここで、追跡とは選手へのID付与及び座標推定を指す。

筆者らは複数映像の情報を統合した3D空間上で追跡を行うことに加えて、文字認識用のAIを利用することで選手映像から背番号を認識し、追跡を誤った場合においても正しく補正するアプローチをとっている⁽⁴⁾。

4.2 自由視点 VR の活用事例

筆者らの自由視点映像で生成されるコンテンツは3DCGに相当する形式となっている。そのため、ビデオゲームのように他のデバイスとの親和性が高いものとなっている。

その一例として、スタジアムに配置した4Kカメラ4台のサッカー映像を用いた自由視点映像をヘッドマウントディスプレイ(HMD)にて視聴するシステムを示す。図6がその利用イメージである。出力された3DCGをHMDにて視聴することで、あたかも自身がサッカーのピッチ上にいるような体験が可能となる。また3D空間全体を構築しているため、視点選択に関する制約はない。同図のシステムではHMD自身が位置測位の機能を備えているため、ユーザーの移動が実スケールで映像内の仮想視点と連動する作りとなっている。

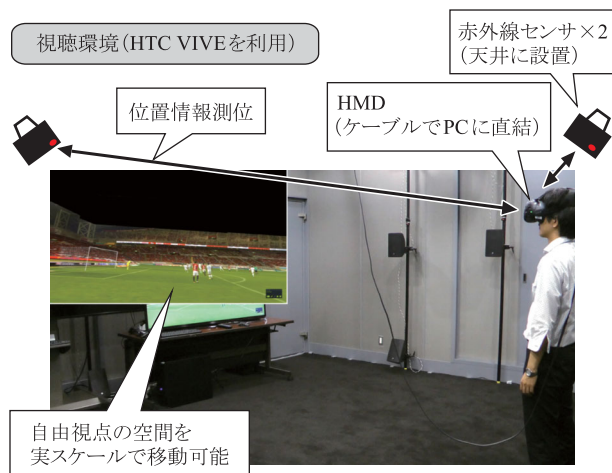


図6 HMDを用いた自由視点VR体験の一例 出力された3DCGをHMDにて視聴することで、あたかも自身がサッカーのピッチ上にいるような体験が可能となる。

自由視点VRの商用化事例としてカラオケVRが挙げられる⁽⁵⁾。カラオケVRでは、8台のフルハイビジョンカメラを用いて撮影したアイドルの映像から自由視点VRのコンテンツ制作を行い、HMDデバイスで視聴できるサービスである。同サービスにより、カラオケ店において実物さながらのアイドルに寄り添う形で歌や踊りを楽しむことができる。当該HMDはPCに直結したHDMIケーブルを介して映像を表示しているが、近い将来に5Gによる超高速伝送が可能となれば、自由視点VR映像をいつでもどこでも視聴できるサービスの展開が可能になると考えられる。

5. む す び

5Gは次世代のモバイル通信技術であり、超高速・低遅延・多接続（大容量）を特長とする。本稿で紹介したトライアル実施例に限らず、5Gネットワークを使用したVR映像伝送デモンストレーションにおいては超高速や低遅延を訴求していることが多い。今後、多接続に関してもスタジアムやイベント会場において、大規模な端末数を想定した実証実験により、5Gのポテンシャルを広げていきたいと考える。またVRについて、実写映像による制作となるとVRカメラの設置場所に限定した3DoF（3 Degrees of Freedom：立ち位置は固定した上で視線方向のみを操作）での映像体験にとどまることから、自由視点VRの実用化を通じた6DoF（6 Degrees of Freedom：3DoFに加えて歩き回る体験が可能）への進化が極めて重要な技術課題である。実写映像を基に

6DoF対応のVRを実現することで、競技者目線での映像による新しいスポーツ観戦、並びに自宅にいながらの海外旅行や宇宙活動など、あらゆる生活シーンでの体験価値創出に強く期待したい。

文 献

- (1) K.K. Sreedhar, A. Aminlou, M.M. Hannuksela, and M. Gabbouj, "Viewport-adaptive encoding and streaming of 360-degree video for virtual reality applications," IEEE, International Symposium on Multimedia, pp. 583-586, Dec. 2016.
- (2) Y. Sanchez, R. Skupin, and T. Schierl, "Compressed domain video processing for tile based panoramic streaming using HEVC," Proc. IEEE International Conference on Image Processing (ICIP), no. COM-P7-2, pp. 2242-2248, Quebec, Canada, Sept. 2015.
- (3) Q. Yao, H. Sankoh, H. Sabirin, and S. Naito, "Accurate silhouette extraction of multiple moving objects for free viewpoint sports video synthesis," Proc. MMSP, pp. 1-6, Oct. 2015.
- (4) 渡邊良亮, 野中敬介, 三功浩嗣, 内藤 整, "畳み込みニューラルネットワークを用いた複数カメラ環境における背番号認識手法," 第100回オーディオビジュアル複合情報処理研究発表会, no. 12, pp. 1-6, March 2018.
- (5) KDDI 株式会社, "カラオケ店舗で気軽にVRが楽しめる！「KDDI カラオケVR」提供開始," <http://news.kddi.com/kddi/corporate/newsrelease/2017/08/08/besshi2624.html> (accessed on 29th May)

(平成30年6月5日受付)



福井 啓允 (正員)

平22京大・工学卒。平24同大学院情報学研究科知能情報学専攻修士課程了。同年、KDDI株式会社に入社。平27から、同社総合研究所にて、主にVR・自由視点などの次世代映像メディアの動画伝送方式の研究・開発に従事。



大塚 裕太

平21静岡同大学院工学研究科博士前期課程了。同年KDDI株式会社入社。携帯電話エリア品質改善、アプリケーションサーバの保守運用を経て、現在は5G実証実験並びに技術開発に従事。



野中 敬介 (正員)

平22東工大大学院理工学研究科集積システム専攻修士課程了。平26同大学院総合理工学研究科物理情報システム専攻博士後期課程了。同年、KDDI株式会社に入社。以来、同社総合研究所にてコンピュータビジョンの研究に従事。博士(工学)。



内藤 整 (正員：シニア会員)

平6早大・理工・通信卒。平8同大学院修士課程了。同年、国際電信電話株式会社に入社。以来、同社研究所にて、主に高精細映像の符号化方式の研究に従事。現在、(株)KDDI総合研究所超臨場感通信グループリーダー。工博。平28年度本会情報・システムソサイエティ活動功労賞受賞。