

5G時代に相応しい 8KVR映像ライブ配信・ 視聴システムの開発

移動機開発部 まとば なおと
的場 直人

ドコモはVR技術を用いた全天周8K映像のライブ配信・視聴システムを世界で初めて開発した。本システムでは、複数のカメラの映像をつなぎ目のない全天周8K映像に合成するスティッチング装置と、全天周8K映像を圧縮し配信サーバにアップロードするエンコード装置をFPGAで実現することにより、リアルタイム動作を可能にした。またパノラマ超エンジン技術と高解像度液晶ディスプレイを用いたHMDにより全天周8K映像の視聴を可能とした。

1. まえがき

近年、仮想現実（VR：Virtual Reality）^{*1}技術の進展により、臨場感、没入感の高い全天周映像^{*2}の視聴環境が整いつつある。全天周映像を撮影可能なカメラについてはフルHD^{*3}から4K^{*4}と高解像度化が進み、さらには業務用では全天周8K^{*5}映像の撮影が可能となるカメラも登場している。

また全天周映像を視聴するHMD（Head Mounted Display）^{*6}についても高解像度化が進んでいる。今後5Gの登場により、高速通信が可能となるため、4Kや8Kといった情報量の多いデータであっても、ストリーミング^{*7}、さらにはライブ配信が可能となってくる。そのため、このようなVR技術と5Gの

組合せにより、スポーツやライブのようなエンターテインメントを、遠隔地においてもまるでその場にいるような高臨場感で体験するといったサービスが登場することが期待されている。

これまで全天周映像のライブ配信はほとんど行われていなかった。それは、通常の映像に比べて表示範囲が広がる全天周映像では、解像度不足が生じてしまい、ユーザが満足できる品質での配信ができないことが原因と考えられる。

全天周映像では1つのカメラで撮影できないため、複数カメラからの映像信号をリアルタイムにスティッチング^{*8}し、エンコード^{*9}、アップロードすることが必要になる。しかしながらこのリアルタイムスティッチングやエンコードは非常に処理負荷が

©2019 NTT DOCOMO, INC.
本誌掲載記事の無断転載を禁じます。

- *1 仮想現実（VR）：まるであたかも仮想世界にいるかのように錯覚を与える技術。近年はHMD（*6参照）を用いて主に視覚に働きかけ、このような錯覚を実現する手法が主流である。
- *2 全天周映像：前後左右頭上から足元まで全視野を覆うような映像。
- *3 フルHD：縦横が約1,000×2,000画素で構成される画面。
- *4 4K：縦横が約2,000×4,000画素で構成される画面。
- *5 8K：縦横が約4,000×8,000画素で構成される画面。

高く、業務用の製品を用いたとしても全天周4Kが限界である。また、全天周8K映像をサーバから配信する場合、それを視聴するための8K映像のデコード*10は専用装置が必要であり、非常に高価となるため、ユーザが利用可能な範囲での機器構成で実現することができない。さらにユーザが用いるHMDについても、全天周8Kの映像を再生するためにはディスプレイの解像度が不足していた。

そこで本開発では、今まで行われていなかった全天周8K映像のライブ配信サービスの実用化に向け、全天周8K映像の撮影、リアルタイムでの配信、視聴を可能とする装置の試作を行った。

具体的には、①複数の4Kカメラからの映像信号をリアルタイムで合成し全天周8K映像を生成するステッチング装置、②全天周8K映像をリアルタイムに圧縮しサーバにアップロードするエンコード装置、③全天周8K映像をユーザが視聴可能な形式で配信するための技術であるパノラマ超エンジン®(PCE: Panorama Cho Engine)*11技術を採用したリアルタイムPCEエンコード装置、④PCEエンコードされたデータを再生するPCEプレイヤー、⑤片目分の解像度2Kの液晶ディスプレイを採用したHMDを開発した。さらにこの装置の可用性を検証するために、屋外にカメラを設置し、5G装置と接続してライブ配信を行う、全天周8K映像ライブ配信実験を行った。

本稿では、開発した全天周8K映像のライブ配信・視聴システムの構成とライブ配信実証実験について解説する。

2. 全天周8K映像ライブ配信・視聴システムの構成

2.1 システム目標と構成

全天周映像の配信ではエクイレクタングラー形

式*12という中間フォーマットを用いることが一般的となっている。これは全天周映像の圧縮や配信をする際に、既存のソフトウェア、ハードウェアを活用して構成するために、通常の映像伝送で使われている映像のフォーマットに合わせる必要があること、またHMDで視聴する際のプレイヤーでの処理が容易となることが理由である。このエクイレクタングラー形式に変換した際の解像度で全天周映像の解像度を示すことが一般的となっている。

本開発は、2020年頃の全天周映像ライブ配信の商用化をめざし、この配信で用いる全天周映像の解像度を向上させることを目的としている。全天周映像の表示に用いるHMDディスプレイの解像度は、現在は片目で1Kから1.5Kのものが多く用いられている。したがって、2020年頃には片目2Kのディスプレイ解像度が一般的に用いられると想定される。

HMDディスプレイの解像度と映像の解像度を比較した場合、水平方向の視野角を90度程度と仮定すると、片目2K程度のディスプレイ解像度で表示する全天周映像はエクイレクタングラー形式で8Kの解像度が必要ということになる。したがって本開発では全天周8K映像のライブ配信を技術目標として設定することとした。

全天周映像ライブ配信では、このエクイレクタングラー形式に変換した映像を通常のライブ配信と同様に伝送し、視聴側でHMD向けに変換し再生を行う。そのため、通常のライブ配信視聴よりも多くの映像処理プロセスがあり、映像の解像度が大きくなると処理負荷も増大することになるが、ライブ配信ではこれらをリアルタイムで行うことが必要となる。ここでいうリアルタイムとは、映像の1フレーム*13ごとの処理を視聴装置で表示する際のフレームレート(1秒当りに表示するフレーム数)の逆数以下で行うということを意味している。つまりフレームレートが30fpsであれば1/30秒以内にそれぞれの装

*6 HMD: 頭部に装着し、視界を覆い内部のディスプレイに仮想世界を映し出すもの。視野角が広いと没入感、臨場感が上がる。

*7 ストリーミング: NW上で音声や映像データを送受信するときの通信方法の一種。データを受信しながら、同時に再生を行う。

*8 ステッチング: 複数のカメラで撮影された映像からつなぎ目のない全天周映像を作り出すこと。出力される全天周映像はエ

クイレクタングラー形式(*12参照)で表されることが多い。

*9 エンコード: 本稿では大容量の映像データを伝送のため、容量圧縮を行うこと。

*10 デコード: エンコード装置により圧縮されたデータを復元して映像データに変換すること。

*11 パノラマ超エンジン®(PCE): NTTメディアインテリジェンス研究所が開発した配信技術を利用している。パノラマ超エンジンはNTTテクノクロス株式会社の登録商標。

置で映像1フレーム分の処理を終える必要がある。

本開発で試作したシステムの構成図を図1に示す。まず複数のカメラで撮影された映像をつなぎ合わせ、エクイレクタングラー形式に変換、さらに伝送可能なデータ容量に圧縮するためのエンコードを行い、サーバにアップロードを行う。エンコードされたデータをサーバから配信し、視聴側でデコード、さらにパノラマ形式に変換しHMDのディスプレイで再生を行う。

2.2 全天周8K映像の撮影・ステッチング・エンコード

まず全天周映像の撮影方法について解説する。カメラ1台では180度から200度以上撮影が可能な魚眼

レンズを用いたとしても全天周をカバーすることができないため、複数のカメラを用いて全天周をカバーする映像を撮影する。[1] また、あまり大きな業務用のカメラはカメラ間の距離が大きくなり、ステッチングすることができなくなるため、使うことができない。そのため現状、一般に使われている小型の4Kカメラを組み合わせるようになる。

本開発では全天周8K映像を撮影するため、通常の撮影では縦方向の解像度が足りない。そこでカメラを90度傾け縦にした複数の4Kカメラを円周上の水平外向きに配置した。またレンズは視野角180度撮影可能な円周魚眼レンズを採用した。撮影対象は水平付近に存在することが多く、天頂や底部方向にカメラを設置する必要性は低いため、水平方向のみ

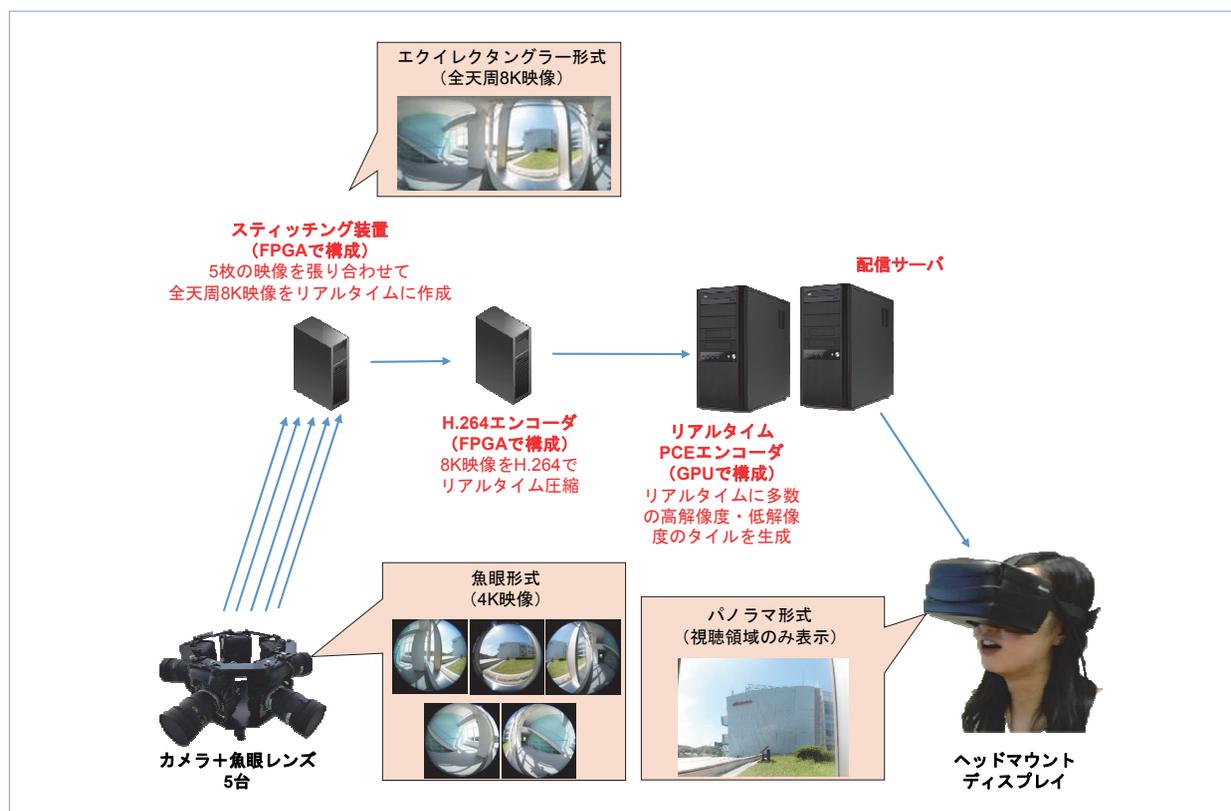


図1 装置全体構成図

- *12 エクイレクタングラー形式：球面の全天周映像を平面に投影する一形式。通常縦横比が1：2であり、緯度経度が垂直に交差するように投影されることを特徴としている。
- *13 フレーム：映像の元になる静止画像の1コマ。

に配置し、魚眼レンズで天頂までカバーする。また、魚眼レンズの円周付近の映像は歪が大きいので、カメラの台数を多くしてなるべく魚眼レンズの中心近くの映像を用いることが望ましいが、台数が増えるとスティッチング装置に入力する映像信号の数が増えるため、スティッチング装置での処理負荷が増大する。そのため今回はカメラ台数を5台とした。

次にスティッチング装置について説明する。リアルタイムスティッチングについては、商用装置を用いると出力映像は4Kが限界であり、その上の解像度をリアルタイムに変換し映像信号を出力可能な装置は現状では存在しない。また4K映像を複数スティッチングする方法 [2] も提案されているが、魚眼レンズ映像からのスティッチングは考慮されていないため、本開発へそのまま適用することは困難である。今回スティッチングに用いたアルゴリズム

を図2に示す [3]。まずカメラが縦置きになっているため、撮影された映像を90度回転し、魚眼形式からエクイレクタングラー形式に変換する。次にカメラを固定する治具の誤差があるため、映像をレンズ光軸の上下左右方向、回転方向に動かしてカメラ間の映像のずれが目立たなくなるように補正を行い5つの映像を合成する。さらにカメラ間の境界が目立たないようにカメラ間の映像の重複部分についてブレンディング処理を行う。

このスティッチング処理をリアルタイムで行う必要があるが、検討の結果、高速演算処理が可能なFPGA (Field Programmable Gate Array)*14を用いることで30fpsを実現できることが分かったため、これを用いることとした。

8Kエンコード装置については、[4] などの実用例があるが、デコード装置も含めたコスト面を考慮

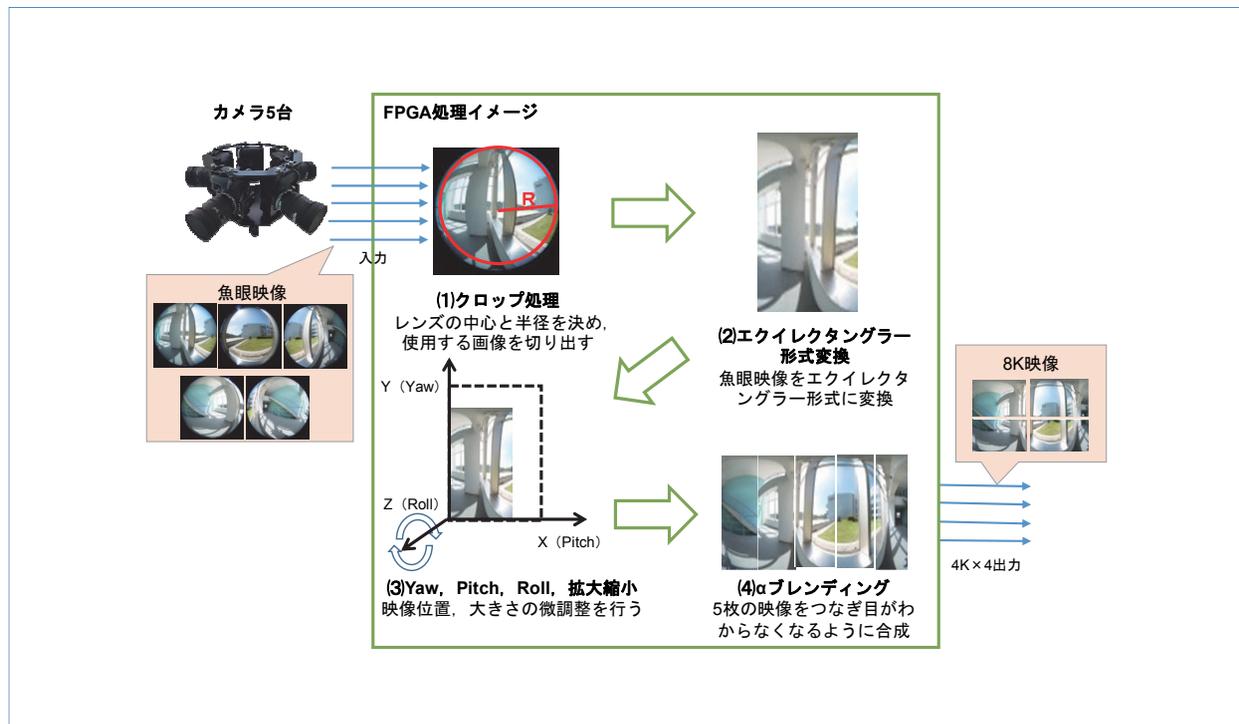


図2 スティッチングのアルゴリズム

*14 FPGA：製造後に構成を変更可能な集積回路。

した結果、本開発ではH.264*15のIPコア*16を用いたFPGAで実現することとした。

カメラ、スティッチングFPGA、エンコードFPGAの間の映像信号の伝送については、SDI (Serial Digital Interface)*17を採用している。全天周映像の場合、撮影範囲が全方向のため、カメラのそばにスティッチングやエンコード装置を設置することができない。また、屋外の実験での機材の配置を現場の状況に合わせて柔軟に行うことが必要となる。SDIを採用することにより、機器間を同軸ケーブルで接続することができ、さらに最大100mまで間隔をとることができる。

試作したスティッチング装置とエンコード装置の外観を図3に示す。

2.3 全天周8K映像の配信・デコード

次に全天周8K映像の配信・デコード方法について述べる。

サーバから配信される8K30fpsのライブ映像を視聴する場合、デコードにかかる負荷が大きいため、現状、デコード性能が高いハイエンドのPCを用い

てもリアルタイムでの再生が困難である。また画質を大きく落とさず伝送するためにはエンコードでのビットレートをあまり低くすることができず、80～100Mbps程度の伝送能力を必要とする。そのため2020年頃の実用化をめざすために、必要な伝送能力要件と視聴装置での処理負荷の両方を下げる必要がある。

本開発ではこの課題を解決するために、PCE技術を採用した。このPCEの原理を図4に示す。全天周映像をHMDで視聴する場合、水平視野角が90度とすると、それ以外の270度方向の映像は見えていない。また、全天周映像のうち、ユーザがどちらの方向から見ているのかをジャイロなどのセンサを用いて計測し、デコード後の映像の視聴方向のみディスプレイで表示している。そこで、PCEではこの計測した方向データを配信サーバへ送信し、見えている方向の映像のみ全天周8K映像と同じ解像度（高解像度部分タイル）で伝送する。またユーザが急に頭を動かした場合を考慮し、全天周の映像も解像度を下げて同時に送信する（低解像度全天周タイル）。これにより、静止したユーザの視線方向の映像は常



図3 スティッチング装置 (左), エンコーダ装置 (右)

- *15 H.264：ITUによって勧告された映像データの圧縮符号化方式標準の1つ。放送やインターネット配信などに幅広く用いられている。
- *16 IPコア：FPGAなどを開発する際に用いられる機能単位でまとめられた部分的な回路情報。
- *17 SDI：主に業務用映像機器で広く使われている、非圧縮の映像信号を同軸ケーブルで伝送可能なビデオ信号伝送の規格の1つ。

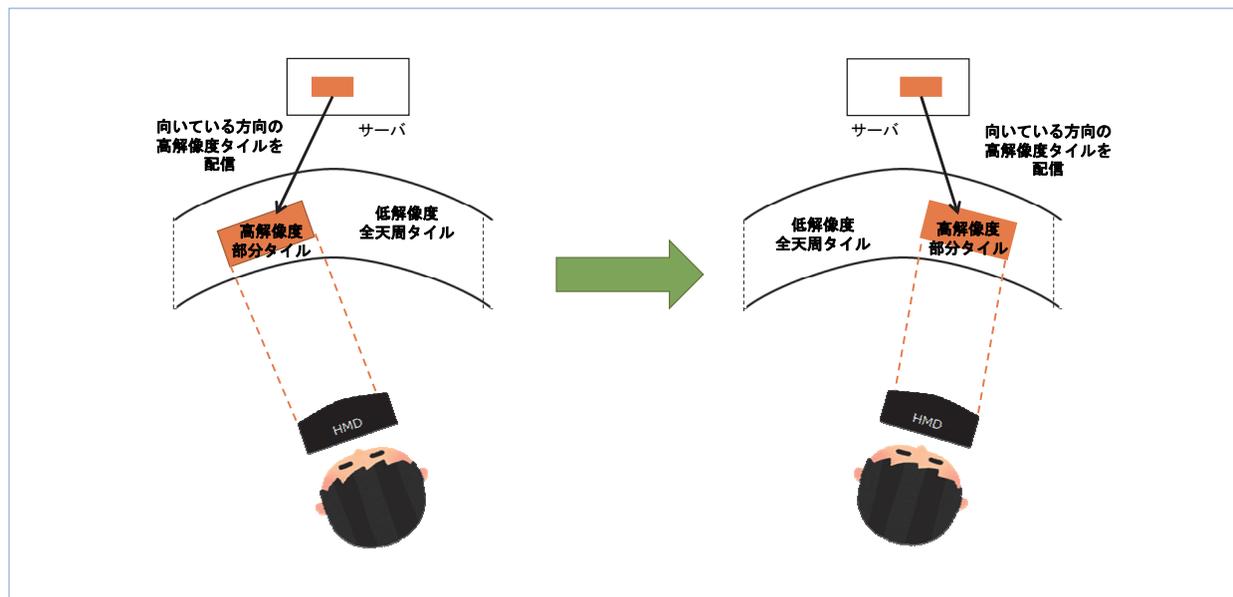


図4 PCEの原理

に8K解像度となり、急に頭を動かした場合でも、低解像度の映像が担保される。

本開発では視野角を考慮し、高解像度部分タイル、低解像度全天周タイルの解像度を両方とも2Kとしている。これにより2K解像度の映像2つ分のデコード負荷で済むため、8K映像の場合よりも大幅に処理負荷を下げることができる。これにより、最新のスマートフォンと同等以上の処理能力のあるHMDであれば再生が可能である。

またこれらのタイルをサーバ側でリアルタイムに生成する必要があるが、このサーバ側の処理負荷を考慮した結果、本開発では高解像度部分タイルをさまざまな方向に16枚、低解像度全天周タイルを1枚、合計17枚のタイルをリアルタイムで生成している。

このPCE技術により8K映像と同じ品質の映像の視聴が可能となる。

2.4 HMDによる視聴

全天周映像の視聴に用いるHMDは映像を表示す

るディスプレイと視野角を広げピントが合うように調節されたレンズと、頭部に装着し適切な位置にディスプレイとレンズを保持するための筐体で構成される。前述のように市中では全天周8K映像の視聴に適したディスプレイを採用したものがいないため、新たに試作を行った。

このHMD（図5）では全天周8K映像の解像度に合わせ、片目分で2Kの解像度を持つ液晶ディスプレイを採用した。画素密度は1,008ppi (pixel per inch)^{*18}となっている。

また、HMDではディスプレイだけでなく、レンズ性能も視聴時に体感する映像の品質に影響を与えるが、既存のレンズの組合せでは適したものがなかったため、この液晶ディスプレイに合わせて新たなレンズの設計・開発を行った。

HMDに用いるレンズは、没入感、臨場感を高めるため、視野角を広くすること、頭に装着するため、小さく軽量化すること、また周辺の収差を小さくすることなどが要求条件として列挙できる。

*18 ppi：1インチ当りの画素数。

本開発では、軽量化のため、高屈折プラスチックレンズを採用し、レンズ周辺部分の収差改善のために、非球面レンズを設計した。そして、このレンズを複数枚組み合わせることにより、周辺の収差はさ



図5 試作HMD

らに改善されるが、HMD全体の重量と収差の改善幅のバランスを考慮し、レンズの枚数を3枚とした。

3. 5G装置を用いた8K全天周映像のライブ配信実証実験

前述のように、5G時代にはこのような全天周8K映像の伝送に必要な高速通信が可能となり、高品質な全天周映像のライブ配信の実現が期待されている。

この技術実証のため、新潟市の実証実験事業の一環として、2018年9月16日に開催されたにいがた総おどりにおいて5Gを用いて全天周8K映像のライブ配信実証実験を行った。この実証実験の様子を図6に示す。

実験では、ほぼ終日に渡ってライブ配信を行ったが、終始安定した配信を行うことができ、来訪された約400名のお客様に全天周8K映像のHMDによる視聴体験をしていただくことができた。

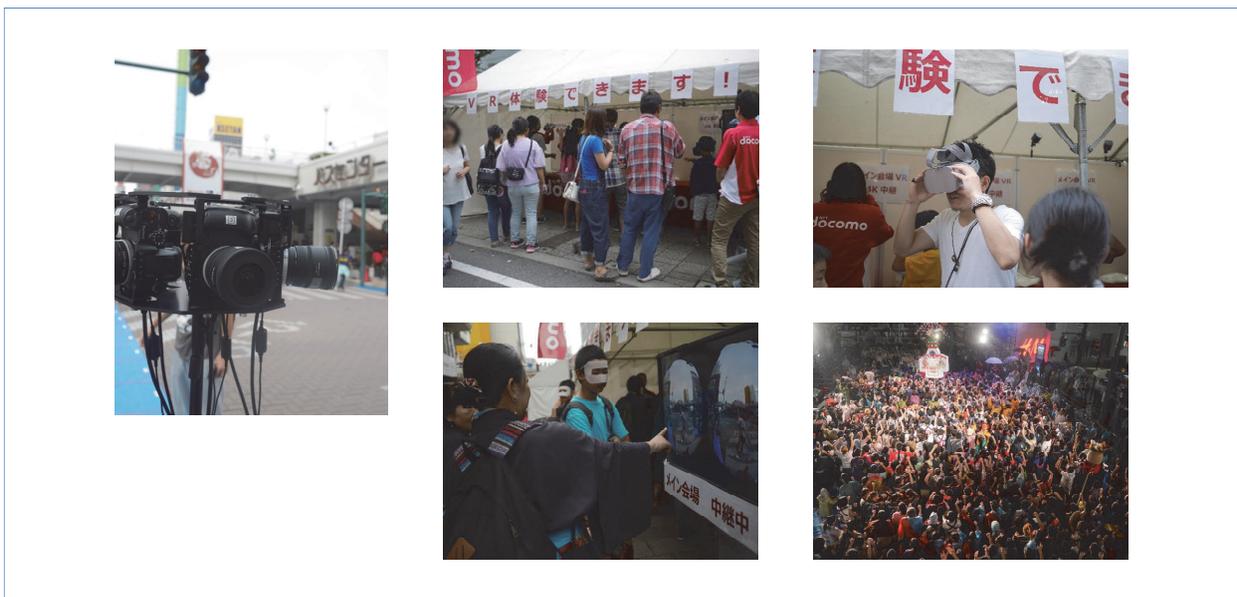


図6 にいがた総おどりでの実証実験の様子

4. あとがき

5G時代に相応しい全天周映像配信の高品質化に取り組み、世界初となる全天周8K映像ライブ配信・視聴システムの試作開発を行った。またこの試作した装置を用いて屋外での技術実証実験を実施した。これらにより全天周8K映像ライブ配信の実現性・有用性について多くの知見が得られ、同時に商用化に向けたさまざまな課題が明らかになった。今後はこれらの課題解決に取り組み、まず、ユーザーの体感を向上させるためのさらなる高品質化に向けた開発を進めていく。さらに、システム全体の低コスト化についても並行して検討を進める予定である。

文献

- [1] 庄原 誠, 佐藤 裕之, 山本 勝也: “特集A 新映像音響体験 2章 全天周撮影,” 映像情報メディア学会誌, Vol.69, No.7, pp.652-657, Sep. 2015.
- [2] 佐藤 孝子, 難波 功次, 小野 正人, 菊池 由実, 山口 徹也, 小野 朗: “競技空間全体の高臨場ライブ中継に向けたサラウンド映像合成・同期伝送技術,” NTT技術ジャーナル, Vol.29, No.10, Oct. 2017.
- [3] Kolor: “Autopano - Tutorials - Quick Start Guide.”
http://www.kolor.com/wiki-en/action/view/Autopano_-_Tutorials_-_Quick_Start_Guide
- [4] 中島 靖之, 西田 享邦, 池田 充郎, 中村 健, 大西 隆之, 佐野 卓, 岩崎 裕江, 清水 淳: “HEVCリアルタイム符号化LSIによる8Kエンコーダの開発,” 映像学技報, Vol.40, No.35, BCT2016-76, pp.13-16, Oct. 2016.