

視野領域を考慮した 360 度映像配信の 適応レート制御手法の性能評価

篠原裕矢 金井謙治 甲藤二郎

早稲田大学基幹理工学部 〒169-8555 東京都新宿区大久保 3-4-1 55N-6F-09

E-mail: {y_shinohara, kanai, katto}@katto.comm.waseda.ac.jp

概要: 近年, AR/VR コンテンツの普及に伴い, 高品質な 360 度映像配信の需要が高まっている. 360 度映像は, コンテンツが高解像度だけでなく, 視野角も広がるため, より大容量になりやすい. そのため, 低通信量で効率的な配信手法が求められている. そこで, 本稿では, 視野領域を考慮した適応レート制御手法を検討する. この適応レート制御手法では, ユーザーの視野領域を注目領域 (ROI 領域) と定義し, ネットワーク品質のみならず, その注目領域も踏まえたものである. これらを MPEG-DASH やヘッドマウントディスプレイと組み合わせ, 視野領域を考慮した適応レート制御手法の性能評価を実施する.

キーワード: 360 度映像配信, MPEG-DASH, 適応レート制御, ROI 符号化

Performance Evaluation of Adaptive Rate Control Method for 360-degree Video Distribution Considering Viewing Area

YUYA SHINOHARA KENJI KANAI
JIRO KATTO

Fundamental Science and Engineering, Waseda University, 3-4-1 Okubo, Shinjyuku-ku, Tokyo, 169-8555 Japan

E-mail: {y_shinohara, kanai, katto}@katto.comm.waseda.ac.jp

Abstract: Recently, AR / VR contents are spreading, and the demand for high-quality 360-degree video delivery is increasing. The 360-degree video requires a larger network capacity because the content has higher resolution to support the wide viewing angle. Therefore, an efficient delivery method, such as reducing network traffic while keeping high visual quality, is mandatory. In this paper, we introduce an adaptive rate control method by considering the viewing area. In this method, the viewing area is defined as a region of interest (ROI), and the video encoding/transcoding rate is controlled by taking care of not only the network quality but also ROI. We evaluate the performance of the method by applying MPEG-DASH.

Keywords: 360-degree video delivery, MPEG-DASH, Adaptive rate control, Region of Interest coding

1. はじめに

近年, Augmented Reality (AR) /Virtual Reality (VR) コンテンツの普及が急速に進み, それに伴って高品質な 360 度映像配信の需要も高まっている. 2020 年までにその市場規模は 2017 年のおよそ 6 倍に達すると予想されている[14]. また, 2016 年は VR 元年と呼ばれ, Oculus Rift[11]や HTC Vive[8], PlayStation VR[12]などの比較的安価なヘッドマウントディスプレイ(HMD)も一般層向けに普及し始めている. さらに 2020 年の東京オリンピックに向けて, 360 度映像によるリアルタイム配信も検討されており, 今後は 360 度映像の放送も普及していくことが予想される.

一方で 360 度コンテンツは, 4K/8K といった高解像度化の流れに加え, 広い視野角に対応することも求められているため, コンテンツが大容量化してしまうことが懸念され

ている. そのため, 360 度映像の配信において, より効率的な適応レート制御手法が求められている.

そこで[2]で提案されている手法を参考にし, 筆者らは [10]にて, 360 度映像配信の効率化のため視野領域とそれ以外を分割して配信する分割配信の Quality of Service (QoS) 評価を実施してきた. この手法では, 映像中のユーザーの視野領域を注目領域 (Region Of Interest (ROI)領域) と定義し, その注目領域を高ビットレートで, 映像全体を低ビットレートで圧縮し, 共に分割配信することで, 360 度映像配信の効率化を図っている. さらに MPEG-Dynamic Adaptive Streaming over HTTP (MPEG-DASH)[1]と組み合わせ, ネットワーク品質評価および注目領域や映像全体の客観画質評価を実施してきた. 本稿では, [3]で提案されている手法を参考に, 分割配信手法の拡張版として, タイルベース符号化制御手法を適用する. タイルベース符号化制御

手法では、画面全体をいくつかのタイルに分割し、ユーザーの視野領域に応じて各タイルの品質を変える手法として知られている。適応レート制御では、実際にHMDを用いてユーザーに360度映像を視聴してもらい、ユーザーの視野領域をトラッキングし、その視野領域を少しでも含むタイルは高ビットレートに、それ以外のタイルを低ビットレートとしている。このようにすることで、視野領域の画質を落とすことなく全体としての総配信データ量の削減を達成し効率化を図っている。性能評価実験では、表示領域の画質評価および分割タイル数を変化させた際の総配信データ量を比較評価する。

2. 関連技術および関連研究

2.1 Region of Interest 符号化

映像において、ユーザーの注視している領域は注目領域 (ROI 領域) と呼ばれており、ROI 領域、非 ROI 領域に対して適応的に符号化率を変化させる ROI 符号化が提案されている。ROI 符号化では、一般的に ROI 領域は高ビットレートを、非 ROI 領域は低ビットレートを選択することで、ROI 領域の画質を向上させ、全体の符号化率を維持、あるいは削減する制御が行われている。

近年、この ROI 符号化の考え方が 360 度映像配信に適用されている。図 1 のように、ユーザーが 360 度映像を視聴する際は、視聴領域と視野外領域がある。この視野領域を ROI 領域、視野外領域を非 ROI 領域ととらえて、視野領域を高品質に、視野外領域を低品質に適応制御する DASH 配信手法が提案されている。詳細は次の節にて紹介する。

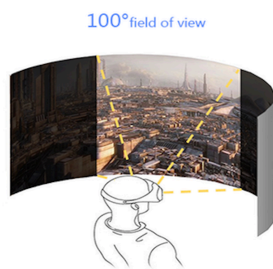


図 1. 360 度映像視聴の様子[13].

2.2 360 度映像の MPEG-DASH 配信

MPEG-DASH は、近年の映像配信技術において主流となっており、ネットワークの通信状況に応じて映像コンテンツの配信ビットレートを動的に変更する「適応レート制御」採用している。MPEG-DASH では、1つのコンテンツを数秒のセグメントごとに分割し、各セグメントに対して、解像度、ビットレートを複数用意することで、コンテンツの階層化を実現している。各階層はレプリゼンテーションと定義している。これらは、Media Presentation Description (MPD) という XML 形式のファイルで管理しており、この

ファイルをセッション開始時に取得、参照することで、通信帯域に応じて最適なレプリゼンテーションを選択することを可能としている。

360 度映像の DASH 配信拡張として、前述した ROI 符号化を適用している手法が提案されている[2,3]。[2]では、ユーザーの視野領域も考慮して、DASH コンテンツの階層化を行っている。そのため、ネットワーク品質だけでなく、ユーザーの視野領域にも応じて、レプリゼンテーションを適応制御することで、配信品質の向上を実現している。[3]では、この視野領域の定義として、画面全体を細かなタイルに分割し、視野領域に応じて、各タイルの符号化率を適応変化させ、より柔軟な品質制御を実現している。本稿では、[2,3]を参考にし、360 度映像の DASH コンテンツ化を行い、360 映像の DASH 映像配信評価を行う。

2.3 360 度コンテンツ投影手法

360 度映像は、一意のポイントからあらゆる方向にキャプチャされたものであるため、本質的には球形の映像となっている。しかし実際にサーバーで保管する際は、通常の動画と同じように矩形の映像として保存されている。その映像を HMD といったクライアントで表示する際、幾何学変換を行っている。このような、球面から平面へ（あるいは逆）の幾何学変換の一つとして、図 2 のような正矩円筒図法 (Equirectangular) [4]が知られている。

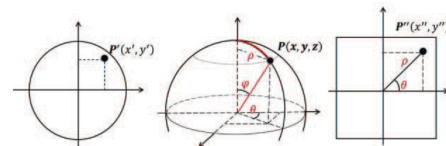


図 2. 球面から平面への変換の流れ[4].

図 2 に示されている球面上の黒点 $P(x, y, z)$ は、右の図の平面上の黒点 $P''(x'', y'')$ にマッピングされる。その際の変換式は(1)式、(2)式のように計算される。

$$x'' = 0.5W + \frac{H}{f} \tan^{-1} \frac{\sqrt{(\sin \varphi \sin \theta)^2 + \cos^2 \varphi^2}}{\sin \varphi \cos \theta} \cos \theta \quad (1)$$

$$y'' = 0.5H + \frac{H}{f} \tan^{-1} \frac{\sqrt{(\sin \varphi \sin \theta)^2 + \cos^2 \varphi^2}}{\sin \varphi \cos \theta} \sin \theta \quad (2)$$

ここで、 W は平面座標にしたときの幅、 H は高さである。また、 f は焦点距離、 φ と θ は図 2 の球面で示されている角度である。この例では、正矩円筒図を紹介したが、他にもピラミッド型やキューブマップ型といった様々な方式の幾何学投影の方法が知られている。

3. 視野領域を考慮した DASH 配信

360 度映像配信では、映像の高解像度化に伴い、より効率的な適応レート制御が必要とされる。一方で、実際にユ

ユーザーが見ている視野領域は、映像全体と比較するとごく一部であると言える。そこで、視野領域を ROI 領域と定義し、ROI 符号化技術を活用し 360 度映像配信を行う。本稿では、分割 DASH 配信とその拡張版であるタイルベース符号化 DASH 配信の 2 つの配信手法を評価する。

3.1 分割 DASH 配信 [10]

筆者らは[10]にて、360 度映像に対して、想定した視野領域を高ビットレートで、画面全体を低ビットレートで圧縮し、MPEG-DASH 配信を行うことで、効率的な適応レート制御を実現している (図 3)。視野領域は高品質で、画面全体を低品質で配信することを目的としている。単純化のため、視野領域はすでに与えられているものとし、また、冗長性は含まれてしまうが、視野領域と、映像全体を別々に圧縮し、同時配信することで分割配信を実現している。比較手法として、映像全体を高ビットレートで圧縮した従来の MPEG-DASH 配信と比較評価してきた。

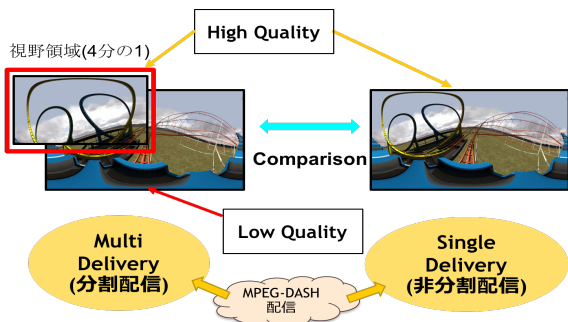


図 3. 分割配信の概要図。

3.2 タイルベース符号化 DASH 配信

[10]のように、360 度映像に対して、視野領域には高品質な映像を、視野外領域には低品質な映像をそれぞれ配信すれば、ユーザーの見ている部分の品質を落とさずに、全体の総配信データ量を削減することができる。ただし、正確にユーザーの視野領域を算出し、動画像符号化処理を加えるのは難しい。そこで本稿では新たに、映像全体をいくつかの矩形のタイルに分割する手法を適用する(図 4)。これをタイルベース符号化 DASH 配信と呼ぶことにする。視野領域が少しでも含まれているタイルは高品質に、視野領域が全く含まれていないタイルは低品質にすることで、視野領域の形状が複雑な場合でも対応することができる。図 4 において、赤い楕円領域は視野領域を表しており、オレンジのタイルは高品質領域、白色のタイルは低品質領域を表している。本稿では、タイル分割数として、分割なし、2 分割、4 分割、8 分割の 4 通りで評価実験を行う。各タイルは高ビットレートで圧縮したものと、低ビットレートで圧縮したものの 2 種類を用意し、視野領域が含まれているかどうかで各タイルは品質レベルの切り替えを行う。

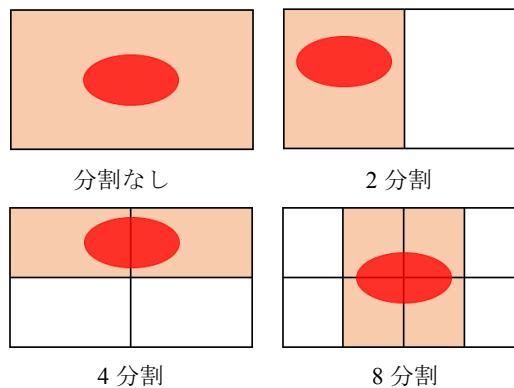


図 4. タイルベース符号化の概要図。

4. 評価実験

4.1 分割 DASH 配信実験

4.1.1 実験概要

本実験では、DASH 配信サーバーとして、[6]で提供されている dash.js を、動画の再生クライアントとして Google Chrome ブラウザを用い、5GHz 帯の IEEE 802.11n を利用して配信サーバーと接続する。また通信環境として、高速な通信帯域(100Mbps)の場合と混雑時の低速な通信帯域(16Mbps)の場合の 2 つのシナリオを想定する。本稿で用いる映像の情報を表 1 に示す。なお、表 1 に示されているように、非圧縮の 360 度映像を用意できず、16Mbps で圧縮済みの動画を利用する。本稿では映像全体の圧縮として、H.264/AVC を利用し、0.5~2 Mbps で再圧縮 (トランスコード) する。また視野領域の動画として、画面全体の左上 4 分の 1 を視野領域と定義し、FFmpeg [7]により映像から切り出し、H.264/AVC を利用して、2~4 Mbps でトランスコードする。なお、これらの選択されたビットレートの映像を DASH レプリゼンテーションと定義する。これらの高品質な注目領域の映像と低品質な全領域(圧縮レート 2~4Mbps)の映像を同時に分割 DASH 配信 (Multi Delivery) する。比較手法として、視野領域と同程度の PSNR となるような圧縮レート (8~12 Mbps) を用いて映像全体をトランスコードし、DASH 配信する場合とする。この手法を非分割 DASH 配信(Single Delivery)と呼ぶことにする。この 2 つの配信手法に対して、配信時の選択されたレプリゼンテーションの PSNR、再生バッファ推移、スタートアップ遅延時間を評価する。なお本稿では、再生開始してから最小の再生バッファがたまり、実際に映像が流れ始めるまでの時間をスタートアップ遅延時間と定義する。

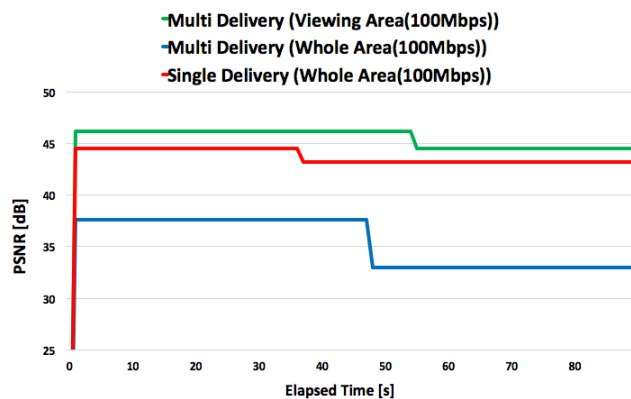
表 1. 使用した 360 度映像の情報。

コンテンツ名	Roller Coaster [5]
解像度	3840×2160
フレームレート [fps]	30

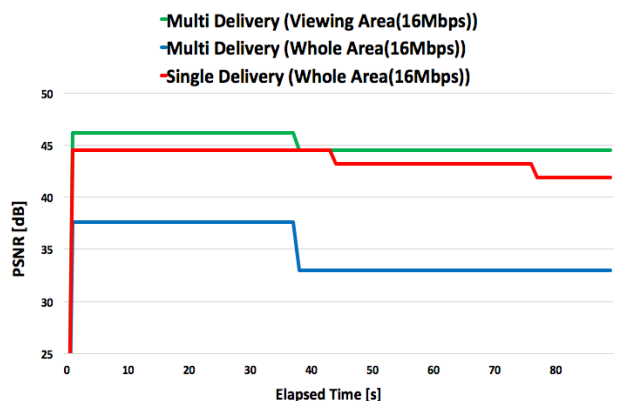
時間 [s]	90
元ビットレート [Mbps]	16
コーデック	H.264/AVC

4.1.2 実験結果

両通信シナリオにおいて、選択されたレプリゼーションに対する PSNR の時間推移を図 5 に、再生バッファの時間推移を図 6 に、スタートアップ遅延時間結果を図 7 にそれぞれ示す。

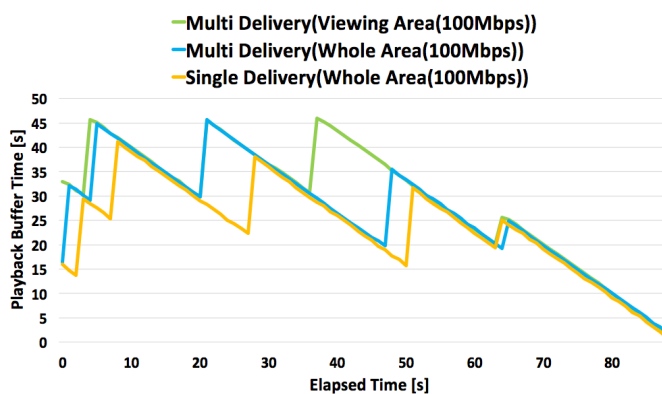


(a)通信帯域 100Mbps の場合

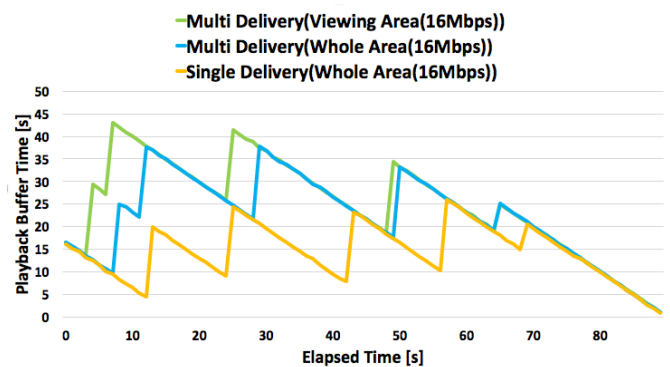


(b)通信帯域 16Mbps の場合

図 5. 選択レプリゼンテーションに対する PSNR 結果.

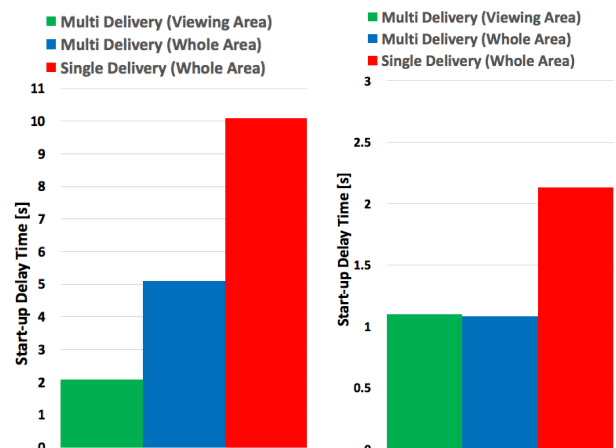


(a)通信帯域 100Mbps の場合



(b)通信帯域 16Mbps の場合

図 6. 再生バッファの時間推移結果.



(a) 通信帯域 100Mbps

(b) 通信帯域 16Mbps

図 7. スタートアップ遅延時間結果.

図 5 より分割 DASH 配信(Multi Delivery)と非分割 DASH 配信(Single Delivery)の結果を比較すると、視野領域の PSNR は 45dB 近くあり、高画質を維持できていることが分かる。通信帯域別に見ると 16Mbps のときの方が比較的低い PSNR になっている。また分割 DASH 配信は、非分割 DASH 配信に比べると合計の配信ビットレートが小さくなることから、特に通信帯域 16Mbps の場合、分割 DASH 配信によって余裕を持って再生バッファを溜められるという結果となった(図 6)。同様に合計の配信ビットレートが小さいという理由から、図 7 においても、分割配信により、スタートアップ遅延時間が平均して約 50%縮まるといった結果となった。

4.2 タイルベース符号化 DASH 配信実験

4.2.1 実験概要

本実験では、先の評価実験とは異なり、視野領域のトラッキングに HMD を用いる。表 1 の Roller Coaster の映像を HTC Vive [8] により、60 秒間視聴する。その際に opentrack [9] を用いて、ユーザーの頭部のトラッキングを行い、それを 2.3 節で紹介した幾何学変換で正距円筒図法形式に変換

する(図 8). opentrack では Yaw 角, Pitch 角, Roll 角, x 座標, y 座標, z 座標の情報を取得できる. 360 度映像を視聴している際の頭部の横振りは Yaw 角, 縦振りは Pitch 角で表される. 今回は単純化のために, Yaw 角と Pitch 角の 2 つを用いて視野領域を算出している.

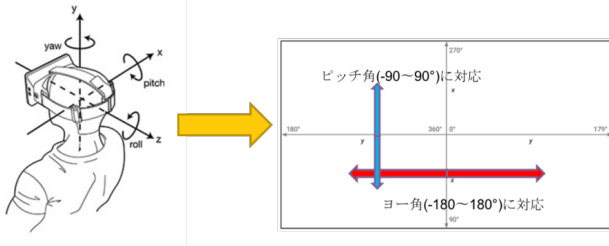


図 8. トラッキング結果の幾何学変換.

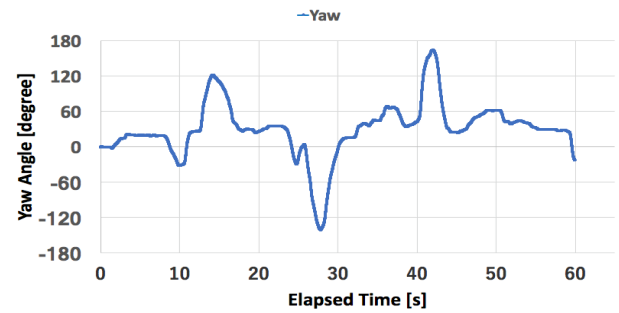
この視野領域を用いて, タイルベース符号化を適用し DASH コンテンツ化する. 分割なし, 2 分割, 4 分割, 8 分割の場合の高品質タイル, 低品質タイルの圧縮レートを表 2 に示す. なお, 利用したコンテンツは表 1 のもので, 公平性のために, すべてのタイルが高品質の場合, またはすべてのタイルが低品質の場合には, 合計のビットレートがほぼ同じになるような圧縮レートに設定している. タイルが視野領域を少しでも含んでいれば高品質に, 全く含まなければ低品質に切り替えることとする. その際の表示領域 PSNR(Viewport PSNR)と総配信データ量を評価する.

表 2. 圧縮レート一覧.

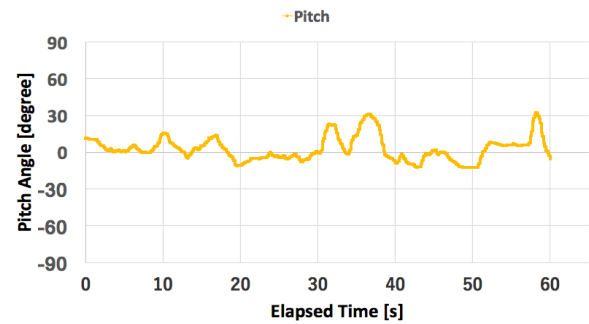
分割数	高品質タイル	低品質タイル
分割なし	16 [Mbps]	8 [Mbps]
2 分割	8 [Mbps]	4 [Mbps]
4 分割	4 [Mbps]	2 [Mbps]
8 分割	2 [Mbps]	1 [Mbps]

4.2.2 実験結果

HTC Vive によるトラッキング結果を図 9 に示す. なお, Yaw 角の正方向は左に向くことを表しており, Pitch 角の正方向は上に向くことを表している. また, 表示領域 PSNR の時間推移を図 10 に, 分割数ごとの総配信データ量を図 11 にそれぞれ示す. なお, 表示領域 PSNR は, 各フレームにおける表示領域を切り出して, フレームごとに算出した.



(a) Yaw 角の時間推移結果.



(b) Pitch 角の時間推移結果.

図 9. トラッキングの時間推移結果.



図 10. 表示領域 PSNR.

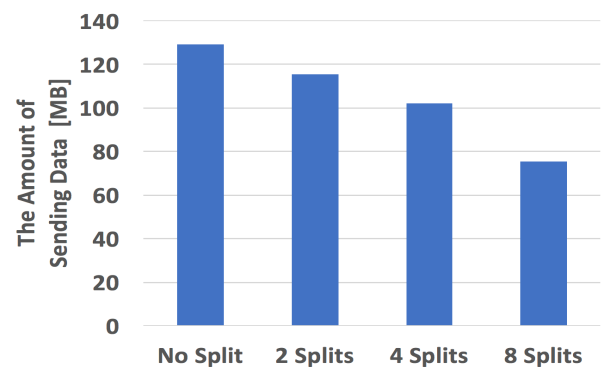


図 11. タイル分割数別の総配信データ量.

図 9 を見ると, Pitch 角に比べて Yaw 角のほうがよく動いていることが分かる. これは, 視聴ユーザーは首を上下

に振ることよりも、左右に振ることのほうが多かったためである。ただし、この結果はあくまで一例であり、ユーザーごとにこの傾向は変わると推測される。図 10 より、いずれの分割数においても、ほとんど変化はなく、表示領域の PSNR はどの時間においても 40dB 以上となっており、高画質を維持できていることが分かる。また図 11 を見ると、分割数を増やすほど総配信データ量を削減できていることが分かる。さらに分割数を増やせば、さらなるデータ量の削減が期待される。ただし、分割数を増やすと、画像符号化の前処理のコストがかかるという面もある。以上より、タイルベース符号化を適用することで、表示領域の画質を維持しつつデータ量削減を達成した。

5. まとめと今後の予定

本稿では、360 度映像の効率的な配信を目的に、視野領域を利用した適応レート制御手法として、分割 DASH 配信とタイルベース符号化 DASH 配信を紹介した。各 DASH 配信評価において、分割 DASH 配信では、視野領域の画質を維持しながらも、再生バッファはより溜まり、スタートアップ遅延時間はより短縮できることを示した。さらに、タイルベース符号化 DASH 配信では、表示領域の画質を維持しながらも、総配信データ量を削減できる結果を示した。

今後は、360 度映像配信の視野移動に対してリアルタイムで画像符号化処理をすることや、360 度映像の視聴データのサンプル数を増やすことが課題である。また、より多彩な評価指標を利用することで 360 度映像 DASH 配信の性能を検証していく予定である。

謝辞 本研究成果は JSPS 科研費 15H01684, 15H02688, 17K12681 の支援を受けている。

参考文献

- [1] I. Sodagar, "The MPEG-DASH Standard for Multimedia Streaming Over the Internet," IEEE Multimedia, vol.18, Issue 4, pp. 62-67, Apr. 2011.
- [2] X. Corbillon, et al. "Viewport-adaptive navigable 360-degree video delivery," IEEE ICC, May 2017.
- [3] Duc V. Nguyen, et al. "A New Adaptation Approach for Viewport-adaptive 360-degree Video Streaming" 2017 IEEE ISM, Dec 2017.
- [4] Tuan Ho, Madhukar Budagavi, "DUAL-FISHEYE LENS STITCHING FOR 360-DEGREE IMAGING", IEEE ICASSP, March 2017.
- [5] Roller Coaster 360 Video 4K – YouTube [online]: <https://www.youtube.com/watch?v=1oOlg6MimeU>
- [6] DASH Industry Forum [online]: <http://dashif.org>
- [7] FFmpeg [online]: <https://ffmpeg.org/ffmpeg.html>
- [8] VIVE™ 日本 | 想像を超えたバーチャルリアリティの体験 [online]: <https://www.vive.com/jp/>
- [9] opentrack/opentrack: Head tracking software for MS Windows, Linux, and Apple OSX [online]: <https://github.com/opentrack/opentrack>
- [10] 篠原, 金井, 甲藤, "360 度映像配信における視野領域を考慮

- した適応レート制御の性能評価," 映像情報メディア学会, 2-2017 年冬季大会, 2017 年 12 月.
- [11] Oculus Rift [online]: <https://www.oculus.com/rift/>
 - [12] PlayStation VR [online]: <http://www.jp.playstation.com/psvr/>
 - [13] Shanghai-based VR Headset "eyeForce" Starts Kickstarter Campaign [online]: <https://medium.com/vr-today-magazine/shanghai-based-vr-headset-eye-force-starts-kickstarter-campaign-2bdabe81aa26>
 - [14] Augmented/Virtual Reality to hit \$150 billion disrupting mobile by 2020 | NEWS >> Digi-Capital [online]: <http://www.digi-capital.com/news/2015/04/augmented-virtual-reality-to-hit-150-billion-disrupting-mobile-by-2020/>