

ヘッドマウントディスプレイ情報を利用した 360 度映像伝送に関する初期的検討

進藤 雅史[†] 木藤 嵩人[‡] 藤橋 卓也^{‡†} 猿渡 俊介[‡] 渡辺 尚[‡]
[†] 大阪大学工学部 [‡] 大阪大学大学院情報科学研究科 ^{‡†} 愛媛大学大学院理工学研究科

1 はじめに

近年、ユーザに没入感を与える技術として VR (Virtual Reality) に対する関心が高まっている。各ユーザは 360 度カメラで撮影されたパノラマ映像を HMD を介して視聴することによって VR 体験が可能である。一方で、各ユーザに対して 360 度カメラで撮影したパノラマ映像全体を伝送した場合、ビデオトラフィックの増大による受信映像品質の低下を招く。近年では、ビデオトラフィックの増大を抑制するため、パノラマ映像のうち、ユーザが視聴すると予測される映像領域 (ビューポート) に注目した 360 度映像配信技術に関する研究が進められている。本稿では、同コンテンツに対する複数ユーザの過去の視聴データから新たなユーザが視聴しうるビューポートを予測してパノラマ映像を 2 種類の品質に符号化することで、ビューポートの受信映像品質を高く維持しつつ、360 度映像配信に要するビデオトラフィックを削減する手法を提案する。

2 360 度映像伝送

2.1 配信モデル

図 1 に配信モデルを示す。サーバには、360 度カメラを用いて撮影されたパノラマ映像が保存されている。ユーザはサーバからインターネットを介してパノラマ映像を受信しながら HMD を介して受信映像を再生する。このとき、各ユーザの HMD はサーバに対してビューポート情報を逐次送信する。ビューポート情報とは、パノラマ映像の中で各ユーザの HMD がユーザに対して表示している映像領域を表した情報である。ビューポート情報は、通常、複素数を拡張した数体系の四元数 (quaternion) で表される。また、サーバは全てのユーザから受信したビューポート情報を保持する。

2.2 課題と既存手法

360 度カメラを用いて撮影されたパノラマ映像は、その解像度の高さが要因となって、高いビデオトラフィックを招く。例えば、Oculus 社の Rift では、HMD を通して再生されるビューポートの解像度が 2160×1200 画素となる。このとき、HMD では、パノラマ映像全体の 4 分の 1 の映像領域が表示されている。すなわち、パノラマ映像全体を送信する場合、ビューポートの解像度と比較して約 4 倍の解像度の映像情報を送信する必要がある。

映像品質を高く維持しつつ、360 度映像を伝送する方法として、文献 [1] の手法が提案されている。文献 [1] では、ユーザのビューポート情報から元にして、パノラマ映像の中で高品質化させる領域 (QER:Quality Emphasized Region) を決定する。その後、QER に含まれる映像領域を高品質で、それ以外の映像領域を低品質でエンコードしたパノラマ映

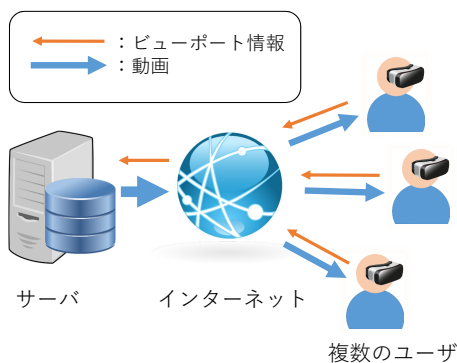


図 1: 配信モデル

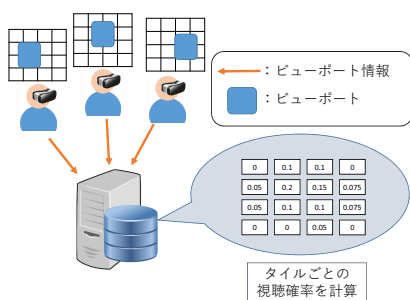


図 2: サーバにおける視聴確率の求め方

Algorithm 1

```

1:  $R_{\Sigma} \leftarrow R_s$ 
2: for  $i = 1$  to  $m$  do
3:   for  $j = 1$  to  $n$  do
4:      $u_{i,j} \leftarrow \sum_{k=1}^s T_{k,i,j}$ 
5:      $p_{i,j} \leftarrow u_{i,j}/R_{\Sigma}$ 
6:   end for
7: end for
    
```

Algorithm 2

```

1: for  $l = 1$  to  $F_{\text{length}}$  do
2:   for  $i = 1$  to  $n$  do
3:     for  $j = 1$  to  $m$  do
4:       if  $p_{i,j} < T$  then
5:         Encode in low quality
6:       else
7:         Encode in High quality
8:       end if
9:     end for
10:  end for
11: end for
    
```

像を伝送する。しかし本手法では、ユーザの多様な視聴要求に応えるために、QER が異なる複数のエンコード後のパノラマ映像をサーバに用意する必要があるため、サーバのストレージを圧迫してしまう。また、360 度映像配信のトラフィックを削減する手法として、パノラマ映像をタイル化して伝送する手法が提案されている。文献 [2] では、パノラマ映像を複数のタイルに分割し、ユーザが HMD を通して視聴すると予測されるタイルのみを伝送する。本手法はパノラマ映像の一部のみを伝送することでビデオトラフィックを削減できるが、もし、サーバ側の予測と異なるタイルをユーザが視聴しようとしたとき再生が停止する。

3 提案手法

提案手法では、再生停止を抑制しつつ、映像品質の高維持、トラフィックの削減を達成するために、サーバがパノラマ映像を複数のタイルに分割する。その後、収集した複数のユーザのビューポート情報から全てのタイルに対するユーザの視聴確率を求める。このとき、視聴される確率が高いタイルを高品質で、視聴する確率が低いタイルを低品質でエンコードして、ユーザに対して全てのタイルを伝送する。図 3 に、提案手法におけるサーバ側・ユーザ側の処理を示す。サーバでは $H \times W$ 画素のパノラマ映像を $m \times n$ 個のタイルに分割する。それぞれのタイルの解像度は $H/m \times W/n$ 画素となる。

3.1 ユーザ視聴履歴に基づく視聴確率の算出

パノラマ映像をタイルに分割した後、サーバでは、Algorithm 1 を用いて s 人のユーザのビューポート情報を元に、1 フレームごとに各タイルの視聴確率を取得する。ここで、 $T_{k,i,j}$ は k 人目のビューポートに含まれる画素のうち、 i 行 j 列目のタイルに含まれる画素の数、 $u_{i,j}$ は全ユーザにまたがった $T_{k,i,j}$ の合計、 R はビューポートの画素数、 R_{Σ} は全ユーザのビューポートに含まれる画素の総和、 $p_{i,j}$ はユーザが i 行 j 列のタイルを視聴する確率である。

3.2 視聴確率に基づく映像符号化

各タイルの視聴確率を求めた後、サーバは、Algorithm 2 を用いて、Algorithm 1 から取得した各タイルの視聴確率に基づいて 1 フレームごとに各タイルの品質を変えてエンコードする。ここで、 F_{length} は動画のフレーム数、 T はあるタイルの視聴確率 $p_{i,j}$ がそれ以上であれば高い品質にエンコードし、それ以下であれば低い品質にエンコードする閾値の定数を表す。タイルの品質は量子化パラメータを変化させることによって設定する。

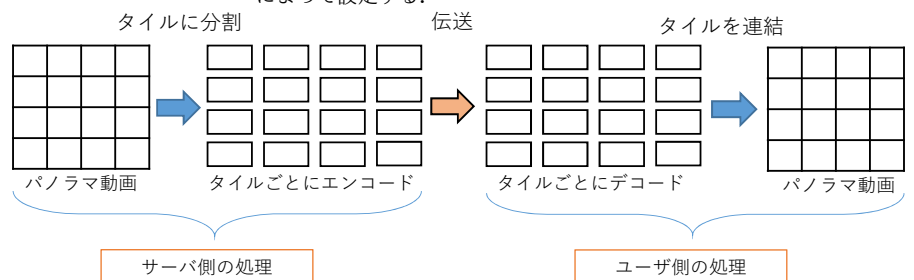


図 3: サーバとユーザの処理



図 4: 評価に使用した 360 度映像

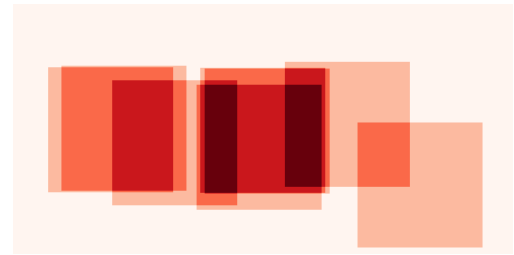


図 5: 8 人の被験者から得られた視聴確率のヒートマップ

サーバは、**Algorithm 2** で生成した高品質/低品質タイルをユーザに対して伝送する。フレームを受信したユーザは、タイルごとにデコードを行った後、タイルを組み合わせてフレームを再構築する。各ユーザは HMD を通して再構築したフレームを順に再生することで VR 体験が可能となる。ユーザが HMD でビューポートを変化させると、**Algorithm 2** に基づいて決定された品質に基づくビューポート映像を視聴することができる。

4 評価

性能評価では、実際の 360 度パノラマ映像と HMD を用いて、ビューポート内の映像品質とユーザへの映像伝送に要するビデオトラフィックを評価した。評価に利用した 360 度パノラマ映像は、リコー社の ThetaS を用いて吹奏楽の練習風景を撮影したものである。パノラマ動画のフレーム数は 2788 フレーム、フレームレートは 29.97、解像度は 1920 × 960 画素であった。また、動画再生、ビューポート情報収集のために、HMD として Oculus 社の Rift を使用した。

タイルごとの視聴確率を求めるためのトレーニングデータとして、8 人の被験者が Rift を通して 360 度映像を視聴したときのビューポート情報を利用した。また、得られた視聴確率から提案手法における映像品質とビデオトラフィックを評価するために、新たに 1 人の被験者のビューポート情報を用いた。8 人のビューポート情報から、**Algorithm 1** を用いて視聴確率を求める際、Oculus Rift のビューポートの角度は、対角線上に約 110 度であるから、簡単化のために、ビューポートは縦横 90 度の正方形と仮定した。具体的には、パノラマ映像全体のうち、高さ方向の 2 分の 1、幅方向の 4 分の 1 の映像領域が HMD を通して視聴できるとしたため、ビューポートの解像度は 480 × 480 画素とした。

映像品質の評価は各フレームにおけるビューポートのピーク信号対雑音比 (PSNR: Peak signal-to-noise ratio) を用いた。ビューポート内の PSNR は以下の式を用いて求められる。ここで、 X はビューポート内の画素番号の集合、 (i, j) は上から i 行、左から j 列の画素を表す。また、 $I(i, j)$ は、元のフレームの画素の輝度、 $I'(i, j)$ は、復号後のフレームの画素の輝度を表す。

$$\epsilon_{\text{MSE}} = \frac{1}{480^2} \sum_{(i,j) \in X} \{I(i,j) - I'(i,j)\}^2 \quad (1)$$

$$\text{PSNR} = 10 \log_{10} \frac{255^2}{\epsilon_{\text{MSE}}} \quad (2)$$

本実験では、パノラマ映像全体を 4 × 4 個のタイルに分割した。また、視聴確率の閾値 (T) は 0.15 とした。比較対象として、全てのタイルを低品質にエンコードする手法と、全てのタイルを高品質にエンコードする手法を用いた。ここで、映像品質が高いタイルの量子化パラメータは 24、映像品質が低いタイルの量子化パラメータは 36 とした。また、本実験ではすべてのフレームを I フレーム (キーフレーム) としてエンコードした。

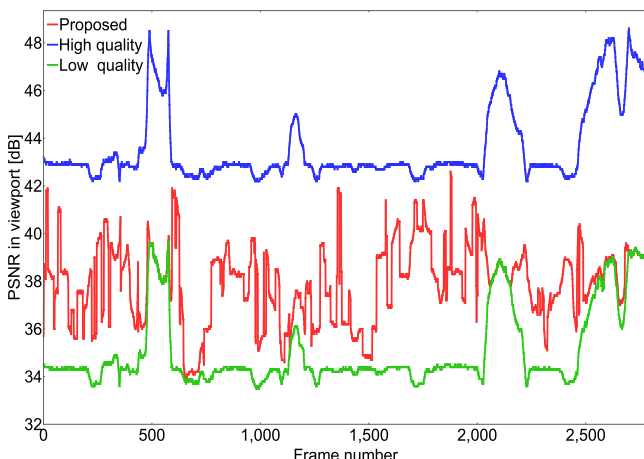


図 6: 各フレームにおけるビューポート内の PSNR の比較

図 6 に、比較手法および提案手法におけるビデオフレームごとのビューポート内の PSNR を示す。評価結果から、全てのタイルを低品質にエンコードする Low quality と比較して、提案手法は多くのフレームにおいて高品質を達成できることが分かった。提案手法では、過去のユーザのビューポート情報に基づいて視聴される可能性が高いタイルを高品質にエンコードすることで受信映像を高く維持している。より具体的には、ビデオフレーム全体を通して、提案手法におけるビューポート内の平均 PSNR が 37.93 dB である一方で、全てのタイルを低品質にエンコードした手法の Low quality におけるビューポート内の平均 PSNR は 35.05 dB であった。また、全てのタイルを高品質にエンコードした手法の High quality におけるビューポート内の平均 PSNR は 43.58 dB であった。

図 7 には、比較手法および提案手法におけるビデオフレームごとのデータサイズを示す。評価結果から、全てのタイルを高品質にエンコードする場合と比べて、提案手法はデータサイズを大幅に削減できることを明らかにした。具体的には、ビデオフレーム全体を通して、提案手法における合計データサイズが 284MB であったが、全てのタイルを高品質にエンコードする方式の合計データサイズは 450MB であった。このとき、提案手法は約 38% のデータサイズを削減している。また、全てのタイルを低品質にエンコードした場合における合計データサイズは 230MB であった。

提案手法は、High quality と比較して約 38% のデータサイズを削減し、Low quality と比較して PSNR が高くなっているため、映像品質を高く維持しつつ、ビデオトラフィックを削減するといえる。

5 おわりに

本稿では、複数のユーザのビューポート情報から求められる全てのタイルに対するユーザの視聴確率に応じてタイルの映像品質を変化させることで、ビューポートの受信映像品質を高く維持しつつ、360 度映像配信に要するビデオトラフィックを削減する手法を提案した。提案手法は多くのフレームにおいて高品質を達成し、比較手法と比べて約 38% のビデオトラフィックを削減した。

謝辞

本研究は JSPS 科研費 JP16H01718, JP15K12018, JP17KT0042, JP17K12672 の助成を受けたものです。

参考文献

- [1] Corbillon, X., Simon, G., Devlic, A. and Chakareski, J.: Viewport-adaptive navigable 360-degree video delivery, *Communications (ICC), 2017 IEEE International Conference on*, IEEE, pp. 1-7 (2017).
- [2] Qian, F., Ji, L., Han, B. and Gopalakrishnan, V.: Optimizing 360 video delivery over cellular networks, *Proceedings of the 5th Workshop on All Things Cellular: Operations, Applications and Challenges*, ACM, pp. 1-6 (2016).

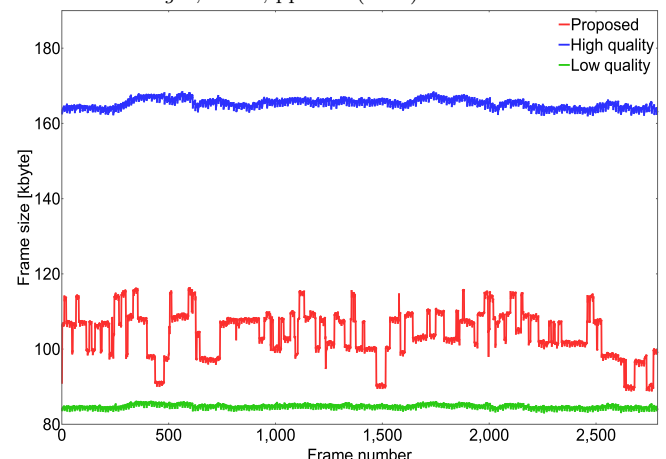


図 7: 各フレームのデータサイズの比較