

クラウドゲーミングにおける映像伝送の低トラヒック化に向けた初期的検討

岡出 紳太郎[†] 石岡 卓将[‡] 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

1 はじめに

ネットワーク技術の高度化とともに、クラウドゲーミングサービスへの注目が高まっている。クラウドゲーミングとは各プレイヤー端末における各種ゲーム処理を必要とせず、各プレイヤー端末に対してサーバ上で実行しているゲーム映像を伝送することで、小型端末や携帯端末におけるゲーム体験を可能とするサービスである。5G ネットワークの広がりとともに、クラウドゲーミングの市場規模は4年間で約11倍の成長が見込まれている[1]。本稿では、クラウドゲーミングにおいて課題となるゲーム映像伝送のトラヒック削減に向けた検討を行う。

2 クラウドゲーミングシステムとその課題

現在のゲーミングシステムはローカルゲーミングとクラウドゲーミングに大別される。ローカルゲーミングシステムとはプレイヤーが保持する端末(例えば、PCや家庭用ゲーム機)上にゲームエンジンをダウンロード・インストールしてゲームを実行するシステムを指す。クラウドゲーミングシステムではプレイヤー端末の代替としてゲーム情報の更新およびゲーム画面の生成を遠隔のクラウドゲーミングサーバが担当するため、限られた処理能力を持つ小型端末や携帯端末に対して臨場感あふれるゲームを提供できる。より具体的には、各プレイヤーが入力した操作信号を、ゲームエンジンをインストールした遠隔のクラウドゲーミングサーバに対して伝送する。サーバは受信した操作信号に基づいてゲームエンジンを通して各プレイヤーの位置や状態、周辺状況を計算する。計算結果はキューに保持するとともに、キューから取得した計算結果に応じてゲーム画面を更新する。更新したゲーム画面をサーバに接続する各プレイヤー端末に対して伝送することで各プレイヤーは自身の操作に応じたゲーム画面を取得できる。

クラウドゲーミングにおける主な1課題として多大なトラヒックが挙げられる。例えば、ゲーム映像の解像度が4K 240fpsである場合、無圧縮伝送に要するトラヒックは約48 Gbpsとなる。通常はゲーム映像を符号化するために映像符号化技術 H.264/Advanced Video Coding (AVC) などを用いる。H.264/AVCは Group of Picture (GoP) ごとにゲーム映像を符号化する。GoPとは複数のビデオフレームの集合体のことであり、通常は8枚のビデオフレームから構成される。GoPに含まれるビデオフレームは大きくIフレーム、Pフレーム、Bフレームとして符号化される。Iフレームはビデオフレームを静止画として符号化したもの、Pフレームは符号化済みのフレームと符号化対象となるフレームとの差分情報を符号化したもの、Bフレームは複数の符号化済みのフレームと符号化対象となるフレームとの差分情報を符号化したものである。H.264/AVCはビデオフレーム間の差分情報を元にして映像情報を符号化することでゲーム映像の伝送に要するトラヒックを低減する。現在では後継の映像符号化規格として H.265/High Efficiency Video Coding (HEVC) や H.266/Versatile Video Coding (VVC) も設計されているが、エンコード遅延が著しく増大するため導入されていない。また、未だ多くのトラヒックが必要となるため、限られたネットワーク帯域下でゲーム映像全体を符号化する場合、受信するゲーム映像の品質低下を招く。ゲーム映像品質の低下は各プレイヤーの体験品質低下を招く可能性がある。

また、複数のプレイヤーがクラウドゲーミングサーバに接続して共通のゲームをプレイする場合、接続するユーザー数の増加とともにトラヒックが増加する場合がある。クラウドゲーミングサーバは各プレイヤーの操作信号に対応するゲーム映像を生成すると、それぞれのプレイヤーに対してゲーム映像をユニキャスト伝送する。このとき、ゲーム映像内にプレイヤー間で共通の映像が含まれている場合、冗長な伝送によってトラヒックの増大を招く。

3 提案手法

前述のとおり、クラウドゲーミングを対象とした映像伝送においては1) ゲーム映像の高解像度化に起因するトラヒック増大、2) 接続ユーザー数の増加に起因するトラヒック増大の2課題を解決する必要がある。本稿ではこれらの課題に対処するため、1) ゲーム画面内の注目領域に応じた適応符号化手法、2) ゲーム画面内の同期領域・非同期領域に応じた映像伝送手法を組み合わせた新たなトラヒック削減手法を提案する。

3.1 想定環境

図1に提案手法の全体像を示す。本稿で対象とするゲームとしてプレイヤー2名による対戦格闘ゲームを想定する。対戦格闘ゲームの画面は各プレイヤーが操作するキャラクタ、背景、ゲーム情報などで構成される。プレイヤー間でキャラクタの位置情報やゲーム情報は同期される一方、背景は非同期となる。

各プレイヤーはネットワークを通じて共通のクラウドゲーミングサーバに接続する。プレイヤー*i*は時刻*t*におけるゲーム画面 $f_t^{(i)}$ にしたがって、自身の操作情報 $a_t^{(i)}$ および注目点 $p_t^{(i)}$ をクラウドゲーミングサーバに対して送信する。ゲーム画面の解像度は $W \times H$ 画素とする。クラウドゲーミングサーバは受信した各プレイヤーの操作情報 $a_t^{(i)}$ に基づいてキャラクタ情報やゲーム情報を更新するとともに、時刻 $t+1$ におけるゲーム画面 $f_{t+1}^{(i)}$ をプレイヤー*i*に対して生成する。その後、クラウドゲーミングサーバはゲーム画面 $f_{t+1}^{(i)}$ を映像符号化技術 H.264/AVC を用いて符号化するとともに各プレイヤーに対して伝送する。このとき、クラウドゲーミングと各プレイヤーとの間の帯域は R [bps] とした。各プレイヤーは受信した映像情報を復号して自身の画面上に表示する。

3.2 注目点に応じた適応符号化

各プレイヤーは時刻*t*におけるゲーム画面 $f_t^{(i)}$ に応じて、ゲーム画面のうち、一部の領域を注目しながらキャラクタを操作する。例えば、既存研究[2]では、各プレイヤーはゲーム画面のうち、中心部分を注目しやすいことが報告されている。提案手法では、視線情報やキャラクタの位置情報などを元にしてゲーム画面 $f_t^{(i)}$ のうち、各プレイヤーの注目点 $p_t^{(i)} = (x_t^{(i)}, y_t^{(i)})$ を取得する。取得した各プレイヤーの注目点はサーバに対してそれぞれ通知する。その後、サーバは注目点 $p_t^{(i)}$ を中心とする $c_w \times c_h$ 画素の矩形領域 f_{ROI} をゲーム画面内の注目領域、それ以外の領域 f_{NROI} をゲーム画面内の非注目領域として取得する。

提案手法では、注目領域 f_{ROI} および非注目領域 f_{NROI} に対して異なる映像符号化パラメータを設定する。より具体的には、あるプレイヤーに対して映像全体に割当可能なビットレートが R であるとき、注目領域に対するビットレートを R_{ROI} 、非共通領域に対するビットレートを R_{NROI} と定めて映像符号化技術 H.264/AVC を用いてそれぞれ符号化する。このとき、 $R_{ROI} + R_{NROI} \leq R$ を満たす必要がある。

3.3 同期領域・非同期領域に応じた映像伝送

提案手法では、ゲーム画面を注目領域・非注目領域に分割すると同時に、それぞれの領域に含まれる同期領域・非同期領域を抽出する。図2に、Street Fighter V Champion Edition (CAPCOM Co., Ltd., Japan)

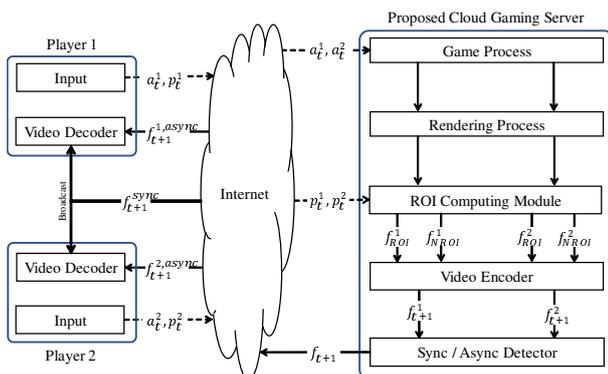


図1: 提案手法の全体像



図2: 提案手法が想定する注目領域・非注目領域および同期領域・非同期領域の一例。Street Fighter V Champion Edition (CAPCOM Co., Ltd., Japan)

を対象とした同期領域・非同期領域の例を示す。例えば、非注目領域に含まれる同期領域としてはゲーム画面上部にある体力ゲージなどが挙げられる。また、注目領域に含まれる同期領域としては各プレイヤーが操作するキャラクターなどが挙げられる。

提案手法では、同期領域に含まれるゲーム画面はプレイヤーが共通して必要とするものとしてマルチキャスト伝送する。一方で、非同期領域に含まれるゲーム画面はプレイヤーそれぞれが必要とするものとしてユニキャスト伝送する。同期領域が広がるにつれて複数プレイヤーへのゲーム映像の伝送に要するトラフィックを削減できる。

4 性能評価

実際のゲーム映像を用いた実験評価を通して、提案手法による映像品質への効果、トラフィック削減効果を議論した。

4.1 評価環境

実験環境: ゲーム映像を取得するために本実験ではプレイヤー端末として DELL G3 2 台を用意した。プレイヤー端末間で 1vs1 のオンライン対戦を実施して各端末上でゲーム映像をキャプチャした。取得したゲーム映像はデスクトップ PC である GALLERIA ZA9C を用いて映像処理、映像符号化を実施するとともに映像品質およびトラフィックを取得した。デスクトップ PC の OS は Ubuntu 20.04.3 LTS, CPU は Intel(R) Core(TM) i9-10850K CPU@3.60GHz, GPU は NVIDIA GeForce RTX 3080 を利用した。

映像符号化: デスクトップ PC における映像処理や映像符号化にはフリーソフトウェア ffmpeg を用いた。より具体的には、ゲーム映像全体もしくは注目領域・非注目領域を H.264/AVC で符号化する場合には `ffmpeg -i filename -r 30 -b:v ${bitrate}` コマンドを使用した。

ゲーム映像: 評価対象となるゲーム映像は対戦格闘ゲームであるストリートファイター V から取得した。各プレイヤーはリュウをキャラクターとして選択するものとした。ゲーム映像の解像度は 1920 × 1080 [px], フレームレートは 30[fps] とした。

比較手法: 比較手法として各プレイヤーに対してゲーム映像全体を符号化する手法および提案手法を用いた。より具体的には、ゲーム映像全体を符号化する手法では ffmpeg を用いて各プレイヤーに対するゲーム映像のビットレートを R [bps] に設定する。一方で、提案手法では ffmpeg を用いて注目領域に対するビットレート R_{ROI} , 非共通領域に対するビットレート R_{NROI} が $R_{ROI} + R_{NROI} = R$ となるように設定した。

評価指標: 評価指標として、注目領域内におけるゲーム映像の Structural Similarity (SSIM) を導入した。SSIM は元映像と復号後の映像との類似度を示す指標で [0, 1] までの値を取りうる。SSIM が 1 であるとき、元映像と復号後の映像が完全に一致することを表している。

4.2 注目領域の割合に対する復元品質

評価内容: 本節では、ゲーム映像内に含まれる注目領域の割合がもたらす影響について評価する。ここで、ゲーム映像全体を符号化する手法においては、クラウドゲーミングサーバと各プレイヤーとの間の帯域が 2.0 [Mbps] および 10.0 [Mbps] である場合をそれぞれ評価した。提案手法においては、クラウドゲーミングサーバと各プレイヤーとの間の帯域が 2.0 [Mbps] であるとき、注目領域に対してビットレートを 1.0 [Mbps], 1.2 [Mbps], 1.4 [Mbps], 1.6 [Mbps], 1.8 [Mbps] を割り当てた場合の映像品質をそれぞれ評価した。同様に、クラウドゲーミングサーバと各プレイヤーとの間の帯域が 10.0 [Mbps] であるときは、注目領域に対してビットレートを 5.0 [Mbps], 6.0 [Mbps], 7.0 [Mbps], 8.0 [Mbps], 9.0 [Mbps] を割り当てた場合の映像品質をそれぞれ評価した。また、注目領域の横幅は 1920 [px] に固定するとともに、縦幅は 100 [px] から 800 [px] ずつ変化させてその影響を評価した。

評価結果: 図 3 にクラウドゲーミングサーバと各プレイヤーとの間の帯域が 2.0 [Mbps], ゲーム映像内の注目領域の割合を変化させたときにおける、ゲーム映像全体を符号化する手法の映像品質と提案手法の映像品質を評価した。また、図 4 にクラウドゲーミングサーバと各プレイヤーとの間の帯域が 10.0 [Mbps], ゲーム映像内の注目領域の割合を変化させた

ときにおける、ゲーム映像全体を符号化する手法の映像品質と提案手法の映像品質を評価した。

評価結果から以下の 2 つのことがわかる。

- 注目領域の割合が小さくなるにつれて提案手法は同程度のトラフィックでゲーム映像全体を符号化する手法と比較して高い映像品質を達成できること。
- 帯域が広がったとき、注目領域の割合が 50%を超えた場合においても、提案手法はより高い映像品質を達成できること。

例えば、注目領域の割合がゲーム映像全体のうち約 65%, 帯域が 2 [Mbps] であるとき、提案手法では注目領域に対して約 1.6 [Mbps] 以上を割り当てた場合、高い映像品質が達成できる。

4.3 同期領域のマルチキャスト伝送に起因するトラフィック削減効果

評価内容: 本節では注目領域・非注目領域に含まれる同期領域の割合がトラフィック削減効果にもたらす影響を評価する。ここで、ゲーム映像内の注目領域の割合は 65%, すなわち、1920 × 700[px] に固定した、また、注目領域に対して割り当てたビットレートを 1.6[Mbps], 非注目領域に割り当てたビットレートを 0.4 [Mbps] とした。比較手法としてゲーム映像全体を符号化した手法、非注目領域における同期部分だけを考慮してマルチキャスト伝送した手法、注目領域・非注目領域における同期部分を考慮してマルチキャスト伝送した手法を用意した。ここで、注目領域に含まれる同期領域の割合はゲーム映像内のキャラクタ領域などを元に 45.3%, 非注目領域に含まれる同期領域の割合はゲーム映像内の体力ゲージなどを元に約 48.4% と定めた。

評価結果: 図 5 に、各手法においてクラウドゲーミングサーバが映像伝送に要するトラフィックを示す。評価結果から、ゲーム映像全体を符号化する手法と比較して、提案手法は同期領域のマルチキャスト伝送を実現することでクラウドゲーミングサーバが映像伝送に必要なトラフィックを約 15%削減できることが分かった。また、非注目領域におけるマルチキャスト伝送はトラフィック削減効果が小さいものの、より多くのビットレートが割り当てられている注目領域に対するマルチキャスト伝送は効果が高いことがわかった。

5 おわりに

本研究ではクラウドゲーミングにおけるトラフィック削減を達成するための新たな伝送手法を提案した。提案手法では、各プレイヤーの注目点およびゲーム内の同期領域・非同期領域を考慮してゲーム映像を符号化、伝送することで複数プレイヤーへの映像伝送に要するトラフィック削減を達成した。

謝辞

本研究は JSPS 科研費 (JP19H01101, 20K19783), NTT アクセスサービスシステム研究所の支援の下で行った。

参考文献

- [1] Newzoo, “The Global Cloud Gaming Market Is on Track to Generate Revenues of \$3.2 Billion by 2023 - Newzoo”, 2020. [online] Available at: “https://newzoo.com/insights/articles/cloud-gaming-business-market-revenues-and-ecosystem/” [Accessed 17 May 2021].
- [2] Gazi Illahi, Thomas van Gemert, Matti Siekkinen, Enrico Masala, Antti Oulasvirta, Antti Ylä-Jääski, “Cloud Gaming with Foveated Video Encoding,” ACM Transactions on Multimedia Computing, Communications, and Applications, vol. 16, pp. 1-24, 2020.

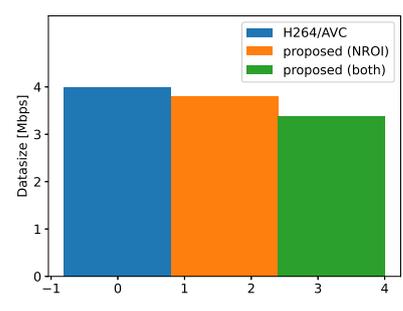
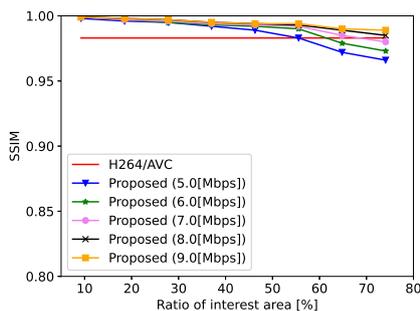
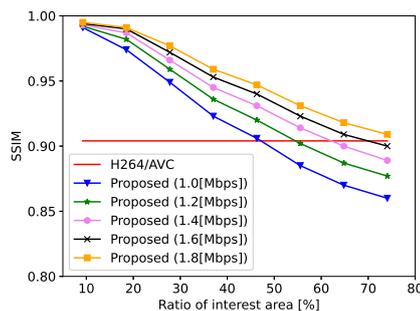


図 3: クラウドゲーミングサーバと各プレイヤーとの間の帯域が 2 [Mbps] である場合の映像品質

図 4: クラウドゲーミングサーバと各プレイヤーとの間の帯域が 10 [Mbps] である場合の映像品質

図 5: 同期領域のマルチキャスト伝送によるトラフィック削減効果