

特別研究報告

題目

生体情報から推定した QoE を用いた
MPEG-DASH におけるビットレート制御手法の実装と評価

指導教員

村田 正幸 教授

報告者

児玉 大暉

令和 4 年 2 月 8 日

大阪大学 基礎工学部 情報科学科

生体情報から推定した QoE を用いた MPEG-DASH におけるビットレート制御手法の
実装と評価

児玉 大暉

内容梗概

近年、動画配信サービスや遠隔 Web 会議システムの普及、インターネットに接続するエンドホスト数の増加、性能の向上、アプリケーションの多様化・高度化に伴い、ユーザやサービス提供者の通信品質 (QoS) に対する要求はより高くなっているが、ネットワークを流れるトラフィック量は年々増大し、ユーザに対して一定の QoS を保証するようなシステムの運用は困難となっている。近年では従来の QoS という観点のみならず、ユーザ自身の体感品質 (QoE) が重要視され、限られた通信資源の中で、ユーザーの QoE を向上させるようなアプリケーションレベルでの制御が、ユーザ、サービス提供者の両者にとって非常に重要となっている。一般に、QoE はサービスに対してユーザの感じる主観的指標であるため定量化が困難である。動画視聴中のユーザの QoE に関して、既存研究ではネットワークの QoS や動画品質を用いた推定を行っているが、これらのユーザ自身に依存しない外的要因だけでなく、ユーザ自身の気分や嗜好といった内的要因も QoE には複雑に絡み合っている。動画ストリーミング視聴中のユーザの QoE を向上させるためには、ユーザの QoE を向上・低下させる要因を明らかにする必要があるが、これもユーザごとに異なると考えられる。

限られた通信資源の中で、個々のユーザにとっての QoE を最大化するという観点から、YouTube や Netflix などの動画ストリーミングサービスプロバイダーの多くは、適応的なビットレート制御を行う技術を採用している。ビットレートの制御には、ネットワークの QoS や動画品質を用いており、個々のユーザにとっての QoE を向上するものとはなっていない。我々の研究グループでは、これまでに生体情報の一つである脳波情報を用いた QoE の推定手法を構築してきた。映像視聴に対する QoE を脳波によって推定することで、個々のユーザの QoE を直接的に把握することが可能となる。本報告では、脳波情報を用いて推定した QoE を用いて、MPEG-DASH のビットレート制御を行うシステムを実装する。先行研究ではオフラインでの QoE 推定を行っているが、動画ストリーミングの制御に利用するためには、オンラインかつリアルタイムで動画を視聴しているユーザの QoE を推定することが必須となる。生体情報から推定した QoE を用いたビットレート制御を実装し、実機

上での動作を通して提案する制御方式が実現可能であることを示す。本研究では、生体情報として瞬き情報を用いることで、ユーザが動画視聴に集中する度合いを推定して利用した。動画視聴に対する集中度は、体感品質である QoE を把握するための一つの要素になると考えて用いている。実装したシステムにより映像の視聴実験を行った結果、動画に対する集中度合いが低下しない限り高いビットレートが維持され、集中度が低下した際にはビットレートが変更される動作がリアルタイムに実行されることを確認した。このような動作により、たとえ動画にリバッファリングによる一時的な停止が生じたとしても、ユーザがそれによって動画への集中度を削がれない限りは、高いビットレートの映像が提供される。

主な用語

HTTP adaptive streaming

適応ビットレート制御

体感品質

まばたき

目次

1	はじめに	6
2	関連研究	8
2.1	生体情報を用いた QoE 推定手法	8
2.2	適応的ストリーミング技術	9
3	生体情報から推定した QoE を用いたビットレート制御手法の実装	10
3.1	実装するシステムの概要	10
3.2	動画視聴プレイヤーの実装	10
3.3	瞬き情報を用いた QoE 推定手法の実装	11
3.4	レート制御機能の実装	16
4	実装したビットレート制御手法の動作検証および評価	18
4.1	評価手法	18
4.2	評価環境	19
4.3	実装したプレイヤーの動作検証	19
5	おわりに	24
	謝辞	25
	参考文献	26

目 次

1	実装するシステムの概要	10
2	プログラム間の情報の送受信	11
3	顔のランドマークの推定	13
4	瞬き検出プログラムの動作画面	14
5	瞬きが検出されるまで	15
6	リバッファリングが起きた際の動作	16
7	視聴する動画 (big buck bunny)	18
8	瞬きから推定された QoE を用いなかった場合の実験結果	20
9	瞬き情報から推定した QoE を用いた場合の実験結果 (推定スループット、動 画のビットレート)	21
10	瞬き情報から推定した QoE を用いた場合の実験結果 (動画のビットレート、 直近 5 回の瞬きに要した時間)	22
11	瞬き情報から推定した QoE を用いた場合の実験結果	23

表 目 次

1	瞬き検出精度の実験結果	15
---	-----------------------	----

1 はじめに

近年、YouTube、Netflix、Hulu などの動画配信サービスの普及により、インターネットを流れるデータ量が増大している [1]。さらに、新型コロナウイルスの影響によって Zoom、Cisco WebEx などの遠隔 Web 会議システムが急速に普及し、映像トラフィックの増加は著しい。また、インターネットに接続するエンドホスト数の増加、性能の向上、アプリケーションの多様化・高度化に伴い、ユーザやサービス提供者の通信品質 (QoS) に対する要求はより高くなっている。しかしながら、ネットワークを流れるトラフィック量は年々増大しており、ユーザに対して一定の QoS を保証するようなシステムの運用は困難となっている。その中で近年では従来の QoS という観点のみならず、ユーザ自身の体感品質 (QoE) が重要視されている。限られた通信資源の中で、ユーザの QoE を向上させるようなアプリケーションレベルでの制御が、ユーザ、サービス提供者の両者にとって非常に重要となっている。

動画ストリーミング視聴中のユーザの QoE 向上に関して、これまでに多くの研究が行われてきた [2,3]。QoE はサービスに対してユーザの感じる主観的指標であり、同じサービスに対しても、ユーザー一人ひとりにとって QoE は異なる。QoE の定量化に関して、既存研究ではネットワークの QoS を用いて推定する手法や、アプリケーションレベルでの動画品質を用いて推定する手法がある。しかしながら、同じ動画ストリーミングサービスを使用していたとしても、ユーザが利用している端末やサービス、端末の通信状況、温湿度や照度などユーザの置かれている状況が多種多様である。一般に、ユーザの QoE は、動画品質や通信品質、視聴環境、コンテンツそのものなど、ユーザ自身に依存しない外的要因だけでなく、ユーザ自身の気分や嗜好といった内的要因も複雑に絡み合っている [4]。

動画ストリーミング視聴中のユーザの QoE を向上させるためには、ユーザの QoE を向上・低下させる要因を明らかにする必要があるが、これもユーザごとに異なると考えられる。動画視聴中のユーザであれば、画質の変化の大きさや頻度、動画の停止する長さや頻度が QoE に影響を与えることが知られている [5]。ただし、これらの要因が個々のユーザの QoE にとってどの程度影響を与えるのかについてはモデル化が非常に困難である。限られた通信資源の中で、個々のユーザにとっての QoE を最大化するという観点から、YouTube や Netflix などの動画ストリーミングサービスプロバイダーの多くは、QoS に応じて適応的なビットレート制御を行う技術である HTTP Adaptive Streaming (HAS) を採用しており、特に HAS の標準規格として Dynamic Adaptive Streaming over HTTP (MPEG-DASH) が広く普及している。MPEG-DASH では、動画ファイルを複数の異なる品質にエンコードし、それぞれを固定長のセグメントに分割する。セグメントごとに動画ファイルの品質を変更することができるため、クライアント端末は動画再生中においても動画品質を動的に切り替えることができる。クライアント端末による動画品質選択は、クライアントに実装されて

いる適応ビットレートアルゴリズム（ABR アルゴリズム）に基づいて行われる。一般的な ABR アルゴリズムでは、クライアント端末と動画配信サーバー間の QoS の推定や動画再生中のプレイヤーから取得可能な情報の収集を行い、これらの情報に基づいて次にダウンロードするセグメントを決定する。QoE の向上を目的とした ABR アルゴリズムも近年では提案されている [6]。

前述の通り、従来の手法のように動画品質のみを用いて QoE を推定することは困難である。また、推定された QoE を動画ストリーミングの制御に利用するためには、リアルタイムで動画を視聴しているユーザの QoE を取得する必要がある。これらの技術的課題を解決するために、計測機器の進歩によりリアルタイムで取得できるユーザの生体情報を用いて QoE を推定する方法を用いることとする。我々の研究グループでは、これまでに脳波情報を用いた QoE の推定手法を構築してきた。文献 [7] では、14 チャンネルの脳波情報から得られる特徴量をもとに、QoE の 2 値分類を行う手法、文献 [8] では、脳波における事象関連電位の一つである P300 に着目し、動画視聴ユーザが動画品質の変化を認知しているのかどうかの判定に用いた。P300 は、特に注目を促している刺激に対しては大きな振幅を生じることが知られている。動画品質が悪化してからおよそ 300 ms 後に P300 がしきい値を超えるか否かを判定し、その際に QoE の低下が生じている場合に、動画品質の変化が QoE の低下を招いたと判別する。文献 [8] での脳波を用いた QoE の推定手法では、平均 49.3%、最大 75.3%の再現率で QoE の低下を推定可能であり、被験者間の個人差が大きいことが示されていた。

本報告では、動画視聴中のユーザの QoE を生体情報からリアルタイムに推定し、推定した QoE を用いたビットレート制御手法を提案する。ビットレート制御に関しては MPEG-DASH をベースとしたシステムを実装し、その有効性を検証する。生体情報と QoE の関係性を調査した文献 [9] を参考に、動画視聴中に容易に取得が可能であり、動画視聴に関連すると考えられる瞬き情報を用いる。瞬きは集中度合いに関連が深いことが知られており、一般に集中するほど頻度が下がる。動画視聴に集中していることは、QoE が高いことに由来すると想定し、瞬き頻度をビットレート制御に用いた。

2 関連研究

本章では関連研究として、生体情報を用いて QoE 推定を行っている研究（脳波を用いた手法、視線や瞬きを用いた手法）、および MPEG-DASH を中心としたストリーミング技術について記述する。

2.1 生体情報を用いた QoE 推定手法

QoE 推定に利用できる可能性のある生体情報の手法について調査した文献 [9] では、生体情報を中枢神経系、自律神経系、目の測定の 3 種類に分類している。中枢神経系に関しては、脳波の測定が挙げられており、そのうち、驚くような新しい刺激に反応する事象関連電位や持続的で周期的な刺激に反応する定常状態視覚誘発電位に注目することで QoE 推定に利用可能であることが述べられている。自律神経系に関しては、心臓の電気的な活動の様子をグラフの形に記録する心電図、筋肉で発生する微弱な電場の変化の様子をグラフの形に記録する筋電図、人間の皮膚抵抗を測定することで汗腺の活動の変化を測定するガルバニック皮膚反応が挙げられている。目の測定に関しては、視線情報により視覚的注意を記録するゲイズ・トラッキング、認知的負荷・推論・視覚的注意・言語処理に関与するといわれている瞳孔測定、集中度合いに関与しているといわれている瞬きが挙げられている。自律神経系の手法を適用するためには特殊な装置が必要なものが多く、我々の研究には適さないと考えており、測定が比較的容易に行える脳波情報、視覚情報に注目している。

瞬きに関しては、文献 [10] で頻度と集中度合いについて関連性を評価する実験が行われている。実験では、Web カメラを用いた瞬き検出システム（検出精度は 92%であると述べられている）を用いて、以下の 3 つのタスクを同じ時間行った際の瞬きの回数を測定することで考察を行っている。

タスク 1 1~3 桁の足し算または引き算の計算式を表示し、その答えの最上位の桁の数字の偶奇を判定する

タスク 2 1 桁の足し算の計算式を表示し、その答えの偶奇を判定する

タスク 3 白い画面を見る

仮説では、タスク 1 が最も集中力が必要であり、タスク 3 が最も集中力の必要がないタスクである。被験者は 4 人で、そのうち 3 人はタスク 1 で最も瞬きの回数が少なくなり、タスク 3 で最も回数が多くなった。また、残りの 1 人はタスク 1 で最も瞬きが多くなったが、これはタスク 1 で問題が切り替わる瞬間に瞬きの回数が増加したためであった。実験の結果から、集中力が高いほど瞬きの回数が減少することが示されている。

脳波に関しては、文献 [7,8] で脳波情報を用いた QoE の推定手法が提案されている。文献 [7] では、14 チャンネルの脳波情報から得られる特徴量をもとに、QoE の 2 値分類を行う手法が提案されている。文献 [8] では、脳波における事象関連電位の一つである P300 に着目し、動画視聴中のユーザが動画品質の変化を認知しているのかどうかの判定に用いている。P300 は、特に注目を促している刺激に対しては大きな振幅を生じることが知られている。動画品質が悪化してからおよそ 300 ms 後に P300 がしきい値を超えるか否かを判定し、その際に QoE の低下が生じている場合に、動画品質の変化が QoE の低下を招いたと判別する。文献 [8] での脳波を用いた QoE の推定手法では、平均 49.3%、最大 75.3% の再現率で QoE の低下を推定可能であると述べられており、また、被験者間の個人差が大きいことが示されていた。

2.2 適応的ストリーミング技術

QoE 向上を目指した ABR アルゴリズムに関する研究は盛んに行われている。その一例として、文献 [6] が挙げられる。文献 [6] では、既存の ABR アルゴリズムの問題点を挙げ、それを解決する ABR アルゴリズムの作成方法を提示している。既存の ABR アルゴリズムはネットワークのスループットや空き帯域に対する再生バッファの占有率をもとに固定化された制御ルールに基づいてビットレートを決定するが、ネットワークの帯域の変動に対応しなければならないこと、QoE を最大にするためにビデオ品質の最大化やリバッファリングの最小化など複数の条件を考慮する必要があること、これらの条件は相反的であり一方を達成しようとするとは他方が連鎖的に達成されないこと、が問題になると述べられており、最適な性能の達成が困難となる。そのために、文献 [6] では ABR アルゴリズムを生成するシステムである Pensieve を提案している。Pensieve は、事前にプログラムされたモデルや環境に関する仮定に基づかず、クライアントの動画視聴プレーヤーから収集されたネットワークのスループットや、再生バッファ、再生時間などの観測結果に基づいて学習を行い、制御方針をニューラルネットワークで表現してビットレートの選択を行う。Pensieve を既存の最新 ABR アルゴリズムと比較実験を行った結果、動画視聴ユーザの平均 QoE を 12~25% 向上させることが示されていた。

動画視聴中のユーザの QoE は非常に主観的かつ複雑な要因に依存していると考えられ、スループットや、再生バッファ、再生時間に基づき推定を行う既存の多くの手法では、真の体感品質を把握するには不十分であると、我々は考えている。生体情報を用いることで真の体感品質を把握できることを期待しており、そのようにして把握した QoE を用いたビットレート制御が、動画視聴中のユーザの QoE 向上に有効になると考えている。

3 生体情報から推定した QoE を用いたビットレート制御手法の実装

3.1 実装するシステムの概要

実装するシステムは、動画配信サーバと動画再生用のクライアントからなる（図 1）。動画画像の提供と再生に関しては、動画配信サーバで HTTP サーバが動作し、MPEG-DASH フレームワークで動画配信サービスを提供する。また、動画再生用のクライアントでは、MPEG-DASH プレーヤーである Dash.js が動作し、サーバへの動画画像の要求および、動画画像の再生を行う。

生体情報の収集及び QoE の推定はクライアント側に実装する。Dash.js と websocket を使用して接続することで、QoE 推定プログラムは動画再生時に生じたイベントを取得する。さらに、QoE 推定プログラムが、次にサーバに要求するセグメントのビットレートの決定も行う。動画視聴中のユーザの QoE を常時推定し、画質低下や動画停止によってユーザの QoE の低下が生じないように、ビットレート選択アルゴリズムがレート選択を実施する。

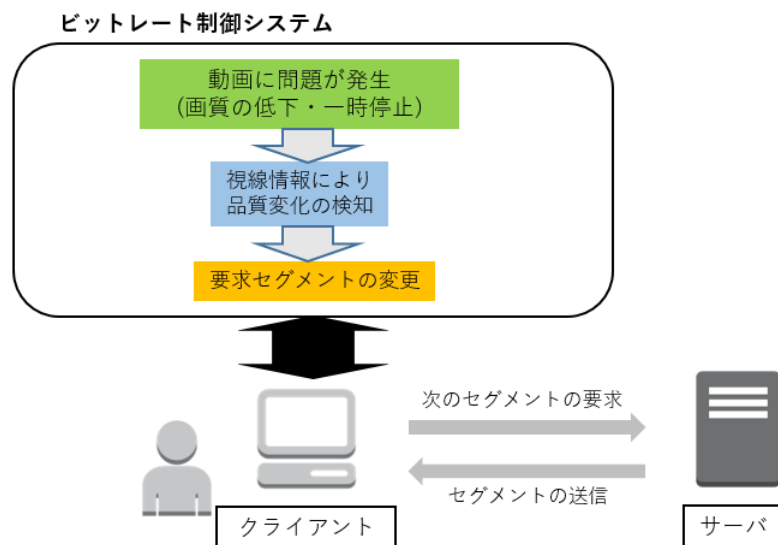


図 1: 実装するシステムの概要

3.2 動画視聴プレーヤーの実装

動画視聴プレーヤーには、JavaScript で記述された、オープンソースの MPEG-DASH ビデオプレーヤーである Dash.js [11] を用いる。

クライアント側で動作する QoE 推定プログラムでは QoE 推定に加えてビットレートの決定を行う。ここではこの二つの機能を分けて説明する。ビットレートを決定するために必要な情報として、Dash.js から直近にダウンロードしたセグメントのサイズとそのダウンロードに要した時間から算出した推定スループット、プレーヤーの再生バッファに残っている映像の長さ（秒）、リバッファリングの発生時刻を取得する。情報を取得するために、Dash.js が WebSocket 接続の受け付けを開始し、ビットレートの決定を行う機能側から接続を行う。Dash.js およびビットレートの決定を行う機能では、WebSocket を介したメッセージ受信が生じるとイベントリスナが受信イベントをハンドルし、定められた動作を行う。また、ビットレートの決定を行う機能は QoE 推定機能の推定結果を受け取る。以上のプログラム機能間の情報の送受信を図 2 に示す。

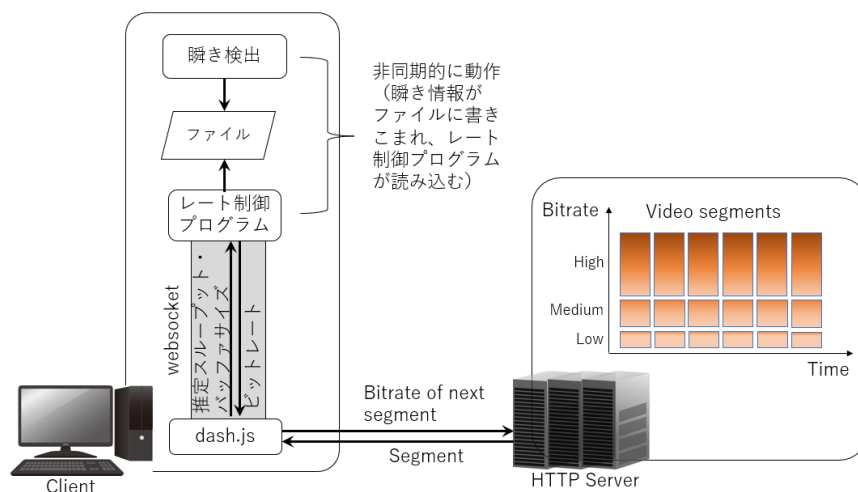


図 2: プログラム間の情報の送受信

Dash.js ではセグメントのダウンロードを終えるごとに、推定スループット、プレーヤーの再生バッファに残っている映像の長さ（秒）を送信する。また、リバッファリングが発生した際に、その発生時刻を送信する。ビットレートの決定を行うプログラム機能では、これらの受け取った情報に加え、推定した QoE を用いてビットレートの決定を行う。一つ前に要求したセグメントのビットレートから変更がある場合にのみ、Dash.js にビットレートの変更を伝える。実際のサーバーへのセグメント要求は Dash.js が行う。

3.3 瞬き情報を用いた QoE 推定手法の実装

本報告では、QoE を推定するための生体情報として視覚情報を用いる。文献 [9] では、QoE 推定に利用可能な視覚情報として視線情報 (ゲイズ・トラッキング)、瞳孔測定、瞬きを挙げ

ている。

今回の実験で想定するシステムでは、QoE の実際の値を推定することに着目しているわけではなく、QoE が向上・低下したかどうかを識別することが重要である。視線トラッキング情報を用いる場合に関しては、視線がディスプレイの外を長時間見ている場合や視線の移動が激しかった場合に集中度が低下したと考えることができ、ディスプレイ内を集中して見ている場合は集中度が高く維持されていると考えることができる。瞬き情報を用いる場合に関しては、目に異物が入るなどの要因を考慮しなければ、文献 [10] で述べられているように、瞬きの頻度が多い場合は集中度が低いと考えることができ、瞬きの頻度が少ない場合は集中度が高いと考えることができる。このように基準がわかりやすく、比較的測定が容易な瞬き情報を本研究では使用する。

動画視聴に集中している際には、動画視聴に満足しているあるいは少なくとも不満がない状態であると考えられ、QoE が高いと想定できる。瞬きから推定される集中度が、動画視聴中の QoE に関連が深いという想定から、瞬きの頻度に基づきビットレートを制御する。瞬きの頻度を計測する方法として、Web カメラから瞬きを検出して回数を測ることとする。

3.3.1 Web カメラを用いた瞬き検出手法

瞬き検出手法およびプログラムは多く公開されている。本研究での要件としては、動画視聴中のユーザーの瞬きを検出するために Web カメラを用いて瞬きを検出することが可能であり、かつリアルタイムな検出が可能であることがある。github 上で公開されており、瞬きを検出して回数を測定することができる Python で実装されたプログラム [12] を用いて瞬きの頻度を測定することとした。ただし、瞬き検出精度の向上のためのチューニングを行い、レート制御部へ情報を渡す機能をこのプログラムに追加した。

まず、プログラム [12] について説明する。プログラム [12] は、part.1~part.4 の 4 つの部分からなっているが、今回使用したのは、視線の向きと瞬きを検出するプログラムである part.3 である。part.3 は、Web カメラから取得した顔画像から目を抽出する部分、抽出した目の情報から視線の向きを推測する部分、抽出した目の情報から瞬きを検出して回数を測定する部分に分かれている。

Web カメラから取得した顔画像から目を抽出する部分では、顔や動体を検出する Python のライブラリである Mediapipe の中で、動画から 468 個の顔のランドマークを推測することのできる Face Mesh を使用している。図 3 は、Web カメラから録画した動画から顔のランドマークを推測し、表示している。この顔画像から推測したランドマークから目に関する部分のみ抽出し、以降の視線の向きの推測や瞬きの検出で使用する。



図 3: 顔のランドマークの推定

抽出した目の情報から視線の向きを推測する部分では、抽出した目の中の黒目のピクセル数に注目している。黒のピクセル数が右に寄っていた場合、右を向いていると判断し、黒のピクセル数が左に寄っていた場合、左を向いていると判断している。しかし、この視線の向きの推測は精度が低く、公開されているものは QoE 推定に使用することができなかった。

抽出した目の情報から瞬きを検出して回数を測定する部分は、抽出した目の縦と横の長さを測り、その比率に注目して瞬きを判定している。瞬きをしているかどうかの判定は目の縦幅と横幅の比率が閾値を上回っているかで判別する。瞬きをしているとき、目は閉じているため、目の横幅は変わらないが、目の縦幅は小さくなる。今回は、片目ずつの比率を (比率) = $\frac{\text{目の横幅}}{\text{目の縦幅}}$ で計算し、両目の平均を全体の比率とすることで計算していたため、瞬きをしている瞬間は目の縦幅のみ小さくなり、比率は大きくなる。github に挙げられているソースプログラムでは、比率の閾値が 5.5 に設定されており、これを 4 フレーム以上続けて超えた後に閾値を下回ったときに瞬きが起こったとカウントする。



図 4: 瞬き検出プログラムの動作画面

3.3.2 瞬き検出手法のチューニング

[12] のプログラムで変更した部分について説明する。

まず、瞬き検出の精度を向上させるために、瞬きを判定するための指標である比率の閾値と瞬きと判別するための条件を変更した (図 5)。比率の閾値に関しては、元のソースプログラムでは 5.5 であったが、この閾値が大きいため瞬きを検出しないことが多かったため、閾値を 4.0 に変更した。また、元のソースプログラムでは、瞬きと認識するためには閾値を 4 フレームを超える必要があったが、4 フレームが長くて瞬きをしているのにもかかわらず検出できないことがあったため、2 フレームに変更した。しかし、これにより 1 回の瞬きの間に目の比率が閾値を前後することによって、2 回以上カウントする現象が発生した。そのため、目の比率が 3.5~4.0 の間は閾値を下回ったとしても瞬きが起こったとカウントしないように変更することで、目の比率が瞬きとカウントするフレーム数だけ閾値を超えていたとしても、目がある程度開かなければ瞬きとカウントされないようにした。

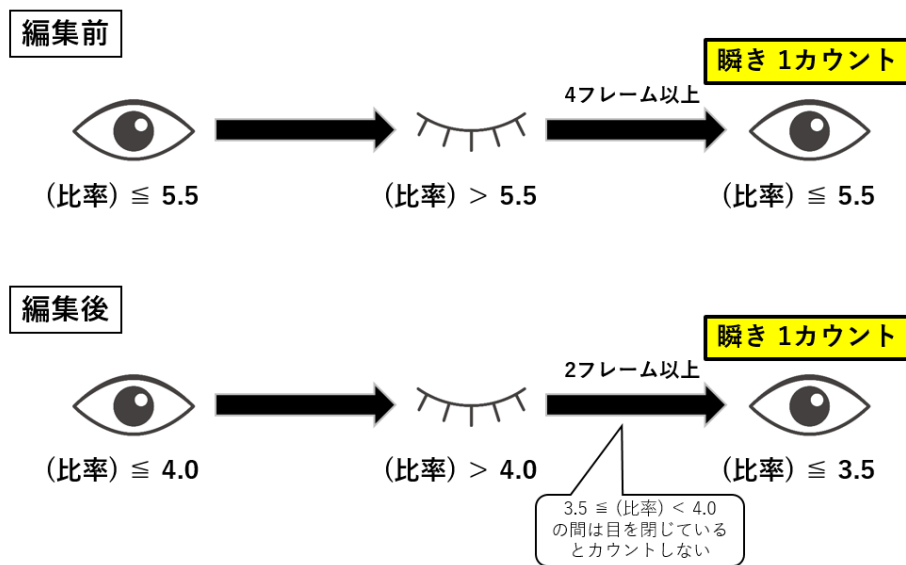


図 5: 瞬きが検出されるまで

変更の前後の瞬き検出精度に関して、両プログラムで 100 回瞬きをしたときに検出した回数を測定し、検出率を求めた。瞬きは頻度が高い場合と低い場合でそれぞれ 100 回検証し、顔の検出精度を高めるために、顔周りの装着物（マスク・眼鏡など）は外した状態で実験を行った。実験を行っている様子を図 4 に示す。結果は表 1 に示した。

表 1: 瞬き検出精度の実験結果

瞬き回数	瞬きの頻度	
	高い	低い
編集前	7	13
編集後	84	100

プログラムの変更によって大きく精度が向上した。これは、元のプログラムの作成者との実験環境の差が考えられる。プログラム変更後の実験環境では、まぶたが完全に閉じていても目の縦幅は 0 にはならなかったため、元のプログラムの作成者の実験環境との照度の差によって目の検出に影響を与えている可能性がある。また、編集後において瞬きの頻度に差が生じたのは、カメラのフレームレートの低下によって瞬きが認識されていなかった可能性があると考えられる。これは、コンピュータのスペックによる問題なので、使用するコンピュータ次第で精度が向上する可能性がある。

次に、生体情報から推定した QoE に基づきビットレート制御を行うために、プログラ

ム [12] に、瞬きの頻度出力する機能を追加した。レート制御部へ瞬きの頻度の情報を渡すために、今回は、QoE 推定プログラムの出力結果をテキストファイルに毎フレーム書き込み、レート制御部ではそのテキストファイルを読み込んでレート制御を行う。この瞬き検出による QoE 推定プログラムでは、過去 5 回の瞬きにかかった合計時間が一定の閾値を超えた場合を集中度が低いと判定する。そのため、テキストファイルには過去 5 回の瞬きにかかった時間を出力し、レート制御部でレート制御を行っている。

3.4 レート制御機能の実装

MPEG-DASH により、QoE 情報をもとにリアルタイムにレート制御を行う機能の実装をした。QoE を低下させる要因として、画質低下やリバッファリングが挙げられるが、本研究では、QoE を低下させる要因としてリバッファリングに着目する。実際にリバッファリングが発生した際に行われる制御についてまとめたものが、図 6 となる。

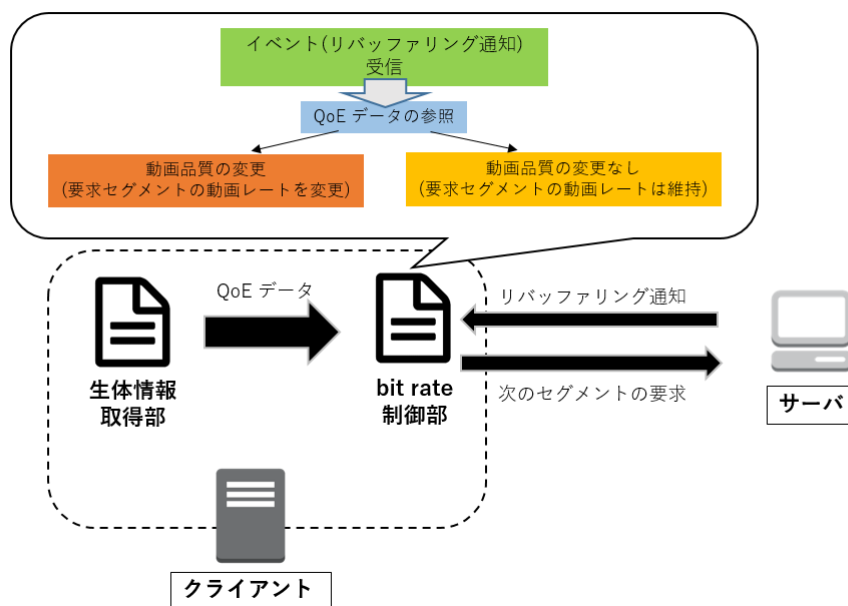


図 6: リバッファリングが起きた際の動作

リバッファリングが起これると、最初にサーバからクライアントへリバッファリング発生のお知らせが送信される。その通知は次のセグメントを要求するビットレート制御部が受信し、リバッファリングが起きたことをクライアント側で知ることができる。このイベントが起きると、ビットレート制御部は生体情報を取得しているプログラムが更新している QoE データが入力されたテキストファイルを参照し、QoE が低下していないかどうかを確認する。低

下していなかった場合は、前回と要求セグメントの動画レートは変更しないが、低下していた場合は、要求するセグメントの動画レートを変更する。

本研究で実装した制御システムにおいて、リバッファリングが発生した際にどのレートにするかについては、そのリバッファリングでユーザの QoE が閾値を下回ったか否かで決定する。ユーザの QoE が閾値を下回らなかった場合は、ユーザはこのリバッファリングを許容できるとし、画質の変更は行わない。しかし、ユーザの QoE が閾値を下回った場合は、画質の変更が行われる。画質は QoE 低下時の動画視聴プレイヤーのスループットの推定値を元にこれを上回らない最大の画質の動画を Dash.js に要求する。今回はサーバー側に、動画の品質を異なる 5 段階で用意している。最も画質の高いものから 5 Mbps、3 Mbps、1.5 Mbps、1 Mbps、0.5 Mbps (bps: bit per second) である。本研究で実装した制御システムでは、要求している動画のビットレートに対して推定スループットに余裕がある場合に、動画の品質を上げるようにしている。現在要求している動画のビットレートの 1 つ上の動画ビットレートが推定スループットより 20 回データを受け取っている間に連続して上回っていた際、要求する動画のビットレートを 1 段階上げる。

4 実装したビットレート制御手法の動作検証および評価

4.1 評価手法

3章では、QoEの推定結果を用いてビットレートを制御する動画視聴プレイヤーの実装方法について説明した。この章では、実装した動画視聴プレイヤーが適切に動作しているかを実験により検証し、ユーザのQoEに基づいて動画のビットレートが変更できているかを確認する。また、QoEの推定結果から決定される動画のビットレートと、QoEに無関係にリバッファリングが発生した際に必ず動画のビットレートを落とす制御でのビットレートと比較する。

本研究の実験の被験者は特別研究報告の著者自身で行った。実験では、被験者がQoEを測定するためのデバイスを装着した状態で、本研究で実装した動画視聴プレイヤーで動画を視聴する。視聴する動画は図7に示した、著作権フリーのBig Buck Bunny [13]である¹。

この動画を視聴している間、以下のデータを記録する；この動画を視聴している間の動画レートを変更されるときに使用されるユーザのQoE指標、ネットワークの推定スループット、選択されたビットレート、リバッファリングが発生した時刻。QoEの推定結果から変更された動画のビットレートが適切なものとなっているか、リアルタイムにビットレートの変更がなされているかを確認する。



図7: 視聴する動画 (big buck bunny)

また、ネットワークの利用可能帯域について、今回使用する動画は、前述のとおり最大でも動画のビットレートは5 Mbpsであるため、その値に比較的近い値に設定する。この設定

¹Big Buck Bunny Copyright (C) 2008 Blender Foundation — peach.blender.org Some Rights Reserved. Creative Commons Attribution 3.0 license. <https://peach.blender.org/>

は linux の tc コマンドを用いることで行う。サーバーとクライアント間は 1Gbps のイーサネット上で互いに有線接続しておき、初期状態ではサーバの外向きの通信速度を 10 Mbps に制限する。動画を視聴している途中に利用可能帯域を 5 Mbps に落とし、リバッファリングを発生させる。さらにその後に利用可能帯域を 10 Mbps に戻し、その際の動画のビットレートの変化も確認する。

4.2 評価環境

HTTP サーバーとして、Ubuntu16.04 上に Apache を準備した。動画像のエンコーディングには ffmpeg を使用した。DASH クライアントはラップトップ型計算機を用いており、ディスプレイ上部に付属の Web カメラを利用した。この実験では前述のとおりユーザの QoE を Web カメラを用いた瞬きの検出で行うため、特別なデバイスの装着はない。

4.3 実装したプレーヤーの動作検証

はじめに、瞬きから推定された QoE を用いなかった場合の結果について示す。この場合、ネットワークの利用可能帯域が、要求する動画のビットレートを下回った場合にリバッファリングが発生する。そして、リバッファリングが発生した際に動画のビットレートが推定スループットを超えない程度の品質まで低下する。さらに、動画のビットレートに対して推定スループットが一定時間高かった場合には、動画のビットレートが推定スループットを超えない品質まで上昇する。この視聴プレーヤーが動作していた際のネットワークの推定スループット、動画のビットレートを図 8 に示す。

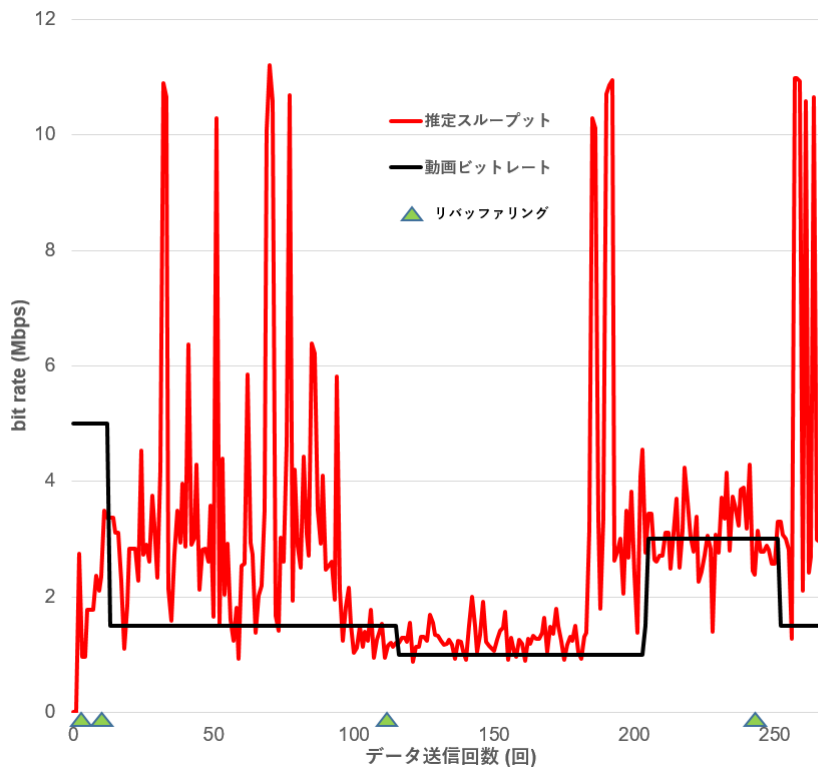


図 8: 瞬きから推定された QoE を用いなかった場合の実験結果

このグラフは、赤色の実線がネットワークの推定スループットを、黒色の実線が要求した動画のビットレートを、横軸の緑色の逆三角の位置でリバッファリングの発生を表している。このグラフから、推定スループットに対して動画のビットレートが高かったときにリバッファリングが発生していることが確認できる。また、リバッファリングが発生した際に動画のビットレートが低下しており、推定スループットが動画のビットレートに対して十分高い状態が継続したときには動画のビットレートが向上している。

次に、瞬き情報から推定された QoE を用いた場合の結果について示す。ネットワークの推定スループットが要求する動画のビットレートを下回った場合にリバッファリングが発生するのは同様である。リバッファリングが発生した際には、ユーザの直近 5 回の瞬きに要した時間が、集中度の低下検出とみなすためのしきい値（今回は 200 f; フレーム）を下回っていた場合のみ、動画のビットレートを推定スループットを超えない品質まで低下させる。さらに、動画のビットレートに対して推定スループットが一定時間高い値を継続した場合に動画のビットレートが推定スループットを超えない品質まで向上する。

この視聴プレーヤーが動作する際のネットワークの推定スループット、動画のビットレートを図 9 に示す。

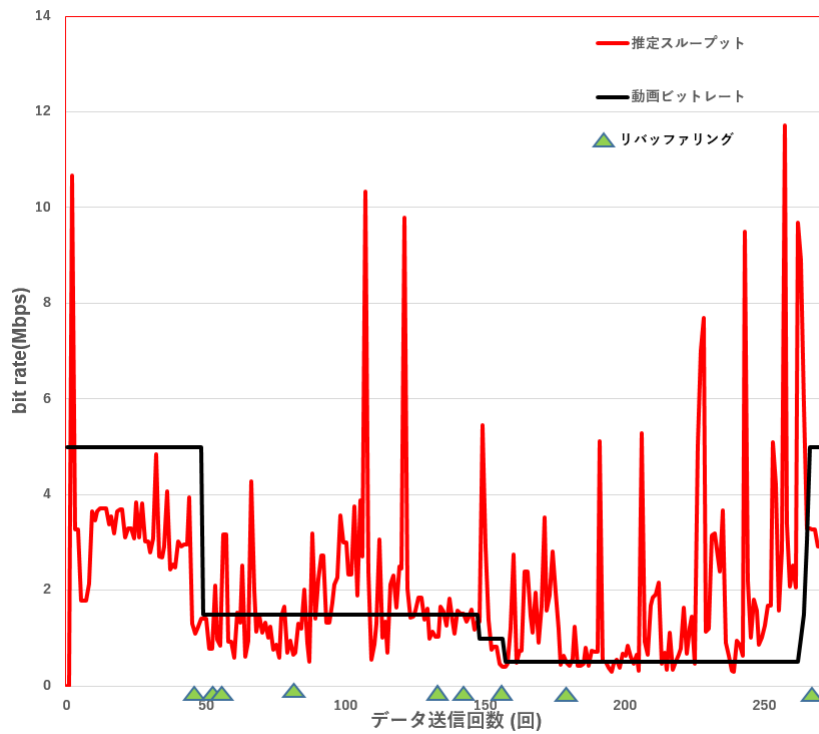


図 9: 瞬き情報から推定した QoE を用いた場合の実験結果 (推定スループット、動画のビットレート)

図 9 と図 8 を比較すると、いずれも推定スループットが動画ビットレートを下回った場合にバッファリングが発生している。しかしながら、リバッファリングが発生した際に必ずしも動画のビットレートを低下させているわけではなく、ユーザの QoE が低下していないと判断されたときは動画のビットレートが低下しない。これを示しているのが、選択された動画のビットレートと直近 5 回の瞬きに要した時間を示した図 10 である。

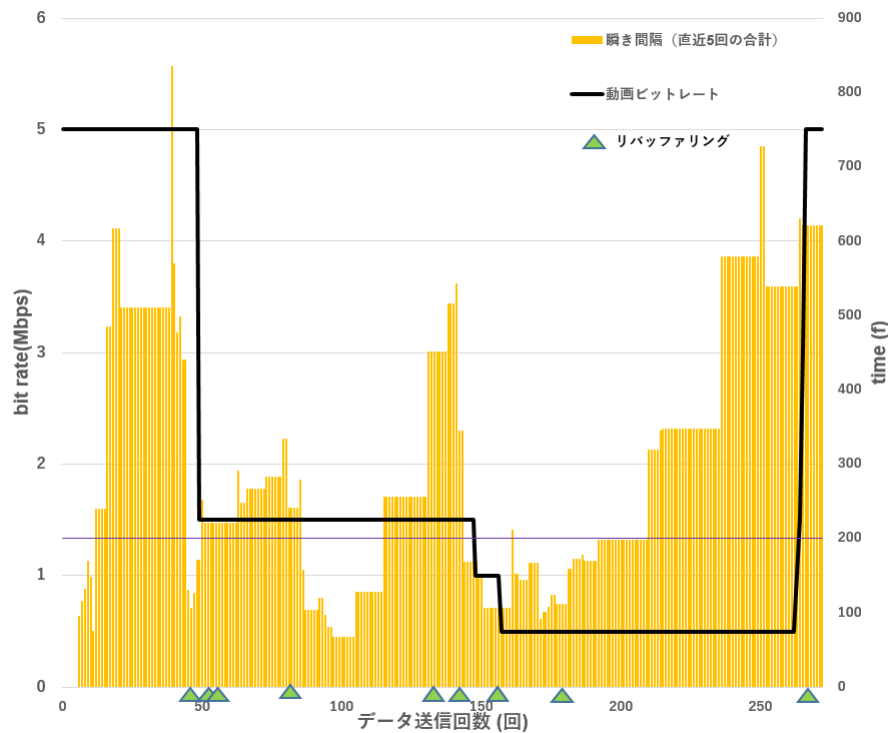


図 10: 瞬き情報から推定した QoE を用いた場合の実験結果 (動画のビットレート、直近 5 回の瞬きに要した時間)

図 10 は、黄色の棒グラフが直近 5 回の瞬きに要した時間を、黒色の実線が要求した動画のビットレートを、横軸の緑色の逆三角の位置でリバッファリングの発生を表している。紫色の横線が集中度の低下を判定する閾値 (200 f) である。推定スループットの減少によってリバッファリングが生じていた場合においても、瞬きに要した時間が 200 f を超えているときには動画のビットレートは変化していない。一方で、リバッファリングが発生し、かつ QoE の低下が検出されたときのみ動画のビットレートを低下していることがわかる。以上より、瞬き情報から推定された QoE をもとにリアルタイムに動画のビットレートを制御することができていることが確認できた。

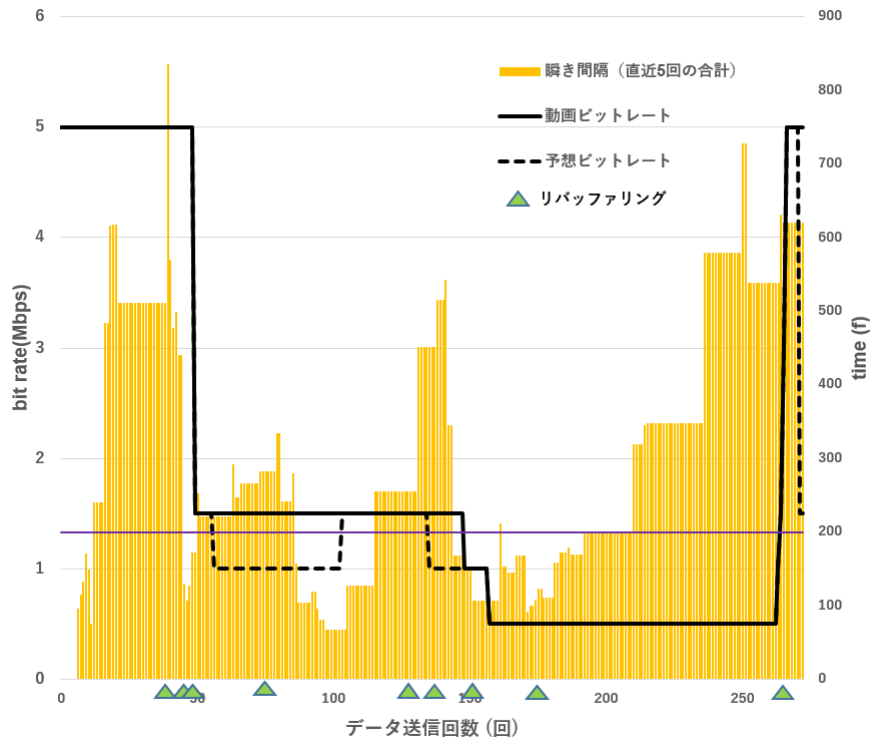


図 11: 瞬き情報から推定した QoE を用いた場合の実験結果

図 10 においてリバッファリングが発生した際に、より低いビットレートが存在するのであれば必ず下げることにしたときのビットレートを黒の破線で追加したものを図 11 に示す。提案したビットレート制御手法では動画に対する QoE が低下しない限り高いビットレートが維持される。そのため、平均的なビットレートは高い値となる。実装したビットレート制御を用いることでリバッファリングによる QoE の低下は、前述の必ずビットレートを下げる方法と比較してもそれほど低下しないことが期待できる。一方で平均的なビットレートの高さに起因してユーザの平均的な QoE が高くなることが期待される。また、推定 QoE を用いた場合はビットレートの変更回数が少なくなる傾向にあり、このことも一般に QoE の低下を抑制することが知られている。以上より、実装したシステムはユーザの QoE を維持・向上に繋がる可能性があると考えている。

5 おわりに

体感品質を表す QoE の向上が動画配信サービスにとって非常に重要なものとなっている。QoE は主観的な値であり、様々な要因に依存するため定量化が困難である。本報告では、生体情報から推定した QoE の利用に着目し、生体情報から推定した QoE を用いたビットレート制御方式を実装した。具体的には、生体情報として web カメラにより瞬きを観測し、その回数を測定することで集中度を推定し、ビットレートの制御に用いた。実機上での動作検証を通して、生体情報を用いた QoE を用いたリアルタイムなビットレートの制御が実現可能であることを示した。本報告では動画の視聴への集中度合いが QoE に深く関連していると想定して、生体情報として瞬き情報を用いた。集中度の推定には、瞬きの他にも視線トラッキング情報の利用が可能である。また、QoE を直接把握する手段としての脳波の利用に関しても我々の研究グループでは取り組んでいる。今後の課題として、他の生体情報を用いて推定した QoE の利用があげられる。瞬き情報も含めて、個人差に対応可能な QoE の推定方法の確立が重要であり、その上で被験者を募集し、体感品質を評価する実験を行うことで手法の有効性を示す。

謝辞

本研究をするにあたり、忙しい中丁寧にご教授くださいました、大阪大学大学院情報科学研究科の村田正幸教授に心より深く感謝を申し上げます。ならびに、研究を進めるにあたり、研究内容や方針、論文の執筆方針について手厚く指導していただいた大阪大学大学院情報科学研究科の小南大智助教、平素よりご指導いただきました大阪大学大学院情報科学研究科の荒川伸一准教授、大下裕一准教授に心より感謝申し上げます。最後に日々の研究生活でお世話になった村田研究室の皆様感謝の意を表し、謝辞とさせていただきます。

参考文献

- [1] Cisco, “Cisco annual internet report (2018–2023),” Tech. Rep., Mar. 2019 (Date Accessed : 2021-12-20). [Online]. Available: <https://www.cisco.com/c/en/us/solutions/executive-perspectives/annual-internet-report/index.html>
- [2] C. Wang, A. Rizk, and M. Zink, “SQUAD: a spectrum-based quality adaptation for dynamic adaptive streaming over HTTP,” *Proceedings of the 7th International Conference on Multimedia Systems, MMSys 2016*, pp. 1–12, 2016. [Online]. Available: <http://doi.acm.org/10.1145/2910017.2910593>
- [3] X. K. Zou, J. Erman, V. Gopalakrishnan, E. Halepovic, R. Jana, X. Jin, J. Rexford, and R. K. Sinha, “Can accurate predictions improve video streaming in cellular networks?” in *Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications - HotMobile '15*, Feb. 2015, pp. 57–62.
- [4] J. Song, F. Yang, Y. Zhou, S. Wan, and H. R. Wu, “QoE evaluation of multimedia services based on audiovisual quality and user interest,” *IEEE Transactions on Multimedia*, vol. 18, no. 3, pp. 444–457, 2016.
- [5] H. Nam, K. Kim, and H. Schulzrinne, “QoE matters more than QoS: Why people stop watching cat videos,” in *Proceedings of IEEE International Conference on Computer Communications*, Apr. 2016, pp. 1–9.
- [6] H. Mao, R. Netravali, and M. Alizadeh, “Neural adaptive video streaming with pensieve,” *Proceedings of the Conference of the ACM Special Interest Group on Data Communication (SIGCOMM)*, pp. 197–210, Aug. 2017.
- [7] K. Kitao, D. Kominami, and M. Murata, “GA-based feature selection for QoE estimation using EEG during video viewing,” in *Proceedings of the 2020 International Conference on Emerging Technologies for Communications (ICETC 2020)*, Dec. 2020, pp. 1–4.
- [8] K. Kitao, “Real-time QoE estimation method using EEG for video delivery services,” Master’s thesis, Graduate School of Information Science and Technology, Jan. 2021.

- [9] U. Engelke, D. P. Darcy, G. H. Mulliken, S. Bosse, M. G. Martini, S. Arndt, J.-N. Antons, K. Y. Chan, N. Ramzan, and K. Brunnström, “Psychophysiology-based QoE assessment: A survey,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 11, no. 1, pp. 6–21, 2016.
- [10] 兜森仁志, 安彦智史, 長谷川大, 佐久田博司, “Web カメラを用いた瞬き検出による集中度評価,” 第 77 回全国大会講演論文集, vol. 2015, no. 1, pp. 931–932, 2015.
- [11] “DASH Industry Forum.” [Online]. Available: <https://github.com/Dash-Industry-Forum/dash.js>
- [12] “Eyes-Position-Estimator-Mediapipe.” [Online]. Available: <https://github.com/Asadullah-Dal17/Eyes-Position-Estimator-Mediapipe>
- [13] “Big buck bunny.” [Online]. Available: <https://peach.blender.org/>