

特別研究報告

題目

脳波情報から推定した 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

森 大慶

令和 5 年 2 月 7 日

大阪大学 基礎工学部 情報科学科

脳波情報から推定した 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

森 大慶

内容梗概

YouTube、TikTok、Netflix、Amazon Prime Video などに代表される動画配信サービスが広く普及し、それに伴いインターネット上を流れるデータ量が増大している。また、Twitch などのライブ配信サービス、Zoom、Webex などの Web 会議サービスなどの利用も一般的になり、単方向あるいは双方向の、動画によるリアルタイムなコミュニケーションが行われるようになってきている。このように、インターネットに接続するユーザの増加や、アプリケーションの多様化・複雑化によって、インターネットに対する通信品質への要求が高まり、要求される品質（QoS、Quality of Service）を保証するようなシステムの構築が困難になってきている。さらに近年では、ユーザの体感品質（QoE、Quality of Experience）という観点での評価が重要視されるようになってきている。

動画像を対象とするアプリケーションにおける QoE の維持・向上を実現するためには、画質の変化の大きさや頻度、動画の停止する長さや頻度を改善することが重要となるが、これらの中には互いにトレードオフの関係にあるものがあり、かつ、どれがユーザの QoE にとって重要なかは、ユーザによって異なり、さらには同じユーザでも視聴中の気分などで異なることがある。このように、QoE の変化の要因には動画品質や動画コンテンツなどユーザ自身に由来しない外的要因と、動画視聴中のユーザの状態という内的要因が複雑に絡み合っているため、通信のスループットなどの外的要因のみで、ユーザの QoE の維持・向上を図るのには限界がある。

本報告では、さらなるユーザの QoE の維持・向上を目標として、文献 [1] で提案されている手法により脳波情報から推定した QoE により、MPEG-DASH におけるビットレート制御を行うシステムを実装し、その動作検証を行った。文献 [1] では、あらかじめ用意した脳波データに対して 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、リアルタイムで動画視聴中のユーザの脳波を取得し、QoE の推定を行えることが必要になる。そのため本報告では、文献 [1] で作成された QoE 推定プログラムをリアルタイムで推定を行えるように修正し、ビットレート制御に組み込むことでシステムを実装した。

実装したシステムの動作を検証した結果、リアルタイムで QoE の推定を行いつつ、その推定した QoE の値を元にビットレート制御が行えていることを確認した。

主な用語

QoE

脳波情報

動画ストリーミング

MPEG-DASH

適応的ビットレート制御

目次

1	はじめに	6
2	関連技術・研究	8
2.1	MPEG-DASH	8
2.2	脳波を用いた動画視聴中のユーザの QoE 推定	9
2.2.1	特徴量の計算	9
2.2.2	特徴量の選別	10
2.2.3	QoE 推定モデルの形式	10
2.2.4	QoE 推定の精度	10
3	脳波情報から推定した QoE を用いたビットレート制御システムの実装	11
3.1	システム概要	11
3.2	機器構成	12
3.2.1	動画配信用サーバ	12
3.2.2	動画再生クライアント	12
3.2.3	脳波計	12
3.3	ビットレート制御用プログラムの実装	14
3.3.1	脳波情報による QoE 推定の実装	14
3.3.2	ビットレート制御の実装	14
4	実装したシステムの評価	16
4.1	評価手法	16
4.2	評価環境	17
4.3	QoE 推定のためのモデル学習	17
4.3.1	学習用データの作成	17
4.3.2	モデルの学習	18
4.4	評価結果	20
5	おわりに	24
	謝辞	25
	参考文献	26

目次

1	MPEG-DASH における動画ストリーミング	8
2	ビットレート制御の流れ	9
3	システム概観	11
4	Emotiv EPOC X	13
5	リバッファリングの発生と選択ビットレート	20
6	選択ビットレートと推定 QoE	21
7	リバッファリングの発生と選択ビットレート (推定 QoE 固定)	22
8	選択ビットレートと推定 QoE (推定 QoE 固定)	23

表 目 次

1	再生用の計算機	13
2	ping の結果	17
3	iperf3 による帯域測定の結果	17

1 はじめに

近年、YouTube、TikTok などの動画投稿サイト、Netflix、Amazon Prime Video などの動画配信サービスが普及し、インターネット上での動画コンテンツが増加しており、それに伴いインターネットを流れるデータ量が急激に増大している。また、Twitch などのライブ配信サービス、Zoom、Webex に代表される Web 会議サービスなども普及してきており、リアルタイムでの高品質な動画データの通信が要求されるようになってきている。このように、年々インターネットに対する通信品質への要求が高まり、サービスへ要求される品質 (QoS) を保証するようなシステムの構築が困難になってきている。このような背景により、限られた通信資源の中でユーザの体感品質 (QoE) の維持・向上を目指せるようなアプリケーションレベルでの制御手法が重要になってきている。

限られた通信資源の中で、ユーザの QoE の維持・向上を行うため、YouTube などの動画配信サービスでは適応的ビットレート制御とよばれる技術を採用している。これは、常に固定の画質で動画ストリーミングを行うのではなく、1つの動画に対して複数の画質の動画ファイルを用意しておき、さらにそのそれぞれの動画を数秒単位で分割し、直近で計測されたネットワークの帯域などの情報から適宜適切な画質の動画を選択してストリーミングを行う技術である。適切なビットレートの選択は、適応的ビットレート制御 (ABR) と呼ばれる技術によって実現され、制御目的の例としては、動画視聴の継続時間に大きな影響を与える、視聴中の動画の一時停止 (リバッファリングの発生) の削減がある。

動画視聴中のユーザに対し、QoE に影響を与える要素としては、画質の変化の大きさや頻度、動画の停止する長さや頻度などが知られている [2]。文献 [3] では、適応的ビットレート制御において、これらの要素の改善を行うことによって、ユーザの QoE の維持・向上を試みている。しかし、これらの改善要素は一方を改善すると他方は悪化する、すなわちトレードオフの関係にあり、かつどれがユーザにとって重要かはユーザによって異なるため、ユーザの QoE の維持・向上という観点でどの要素を重要視すべきかが容易に求まらないという問題が発生する。ユーザによって重要視する点が異なるというのは、より具体的には、画質が下がっても良いので途切れることなく動画を再生してほしいという人もいれば、動画の再生が途切れても良いが高画質を維持したまま再生してほしいという人もいるということである。このような嗜好は、動画視聴中のユーザの気分などにより同じユーザでも時によって変わることがある。このように動画視聴における QoE は、動画品質や、動画の内容などユーザ自身に由来しない外的要因だけでなく、動画視聴中の気分や嗜好などのユーザ自身に由来する内的要因が複雑に絡み合っている [4] ため、通信のスループットなどの外的要因のみを元にした制御でユーザの QoE の維持・向上を行うことには限界がある。

このような背景から、動画視聴中のユーザの QoE の維持・向上を目的として、我々の研

究グループでは脳波情報から動画視聴中のユーザの QoE を推定する手法 [1] の提案や、瞬きの頻度などの生体情報をもとに推定した QoE を元にビットレート制御を行うシステムの実装 [5] を行い、生体情報を用いたビットレート制御の手法の確立を試みてきた。

我々の研究グループの目標は、様々な生体情報をもとに、リアルタイムで動画視聴中のユーザの QoE を推定し、その推定した QoE を用いてビットレート制御を行うことで、ユーザの QoE の維持・向上を達成することである。本報告では脳波情報を用いて推定した動画視聴者の QoE をもとに、ビットレート制御の実現を行うシステムを実装し、その評価を行う。そのために、文献 [1] で提案された手法を拡張し、リアルタイムで脳波情報から QoE を推定するプログラムの実装を行う。また、適応的ビットレート制御のプロトコルの 1 つである MPEG-DASH を対象に、生体情報によるビットレート選択を行なっている文献 [5] の機能を拡張し、脳波から推定した QoE に基づきビットレート選択を行うシステムを実装する。

2 関連技術・研究

2.1 MPEG-DASH

MPEG-DASH (MPEG Dynamic Adaptive Streaming over HTTP) は、HTTP プロトコルを用いた適応的ストリーミング (HTTP Adaptive Streaming、HAS) 技術のうち、もっとも標準的なものの一つであり、国際標準化機関 ISO/IEC によって規格化されている。通信品質やクライアントの再生バッファに残っているセグメントの量などのいくつかの指標にもとづき動的にビットレート制御を行う。

適応的レート制御の仕組み MPEG-DASH では、MPD (Media Presentation Description) と、セグメントと呼ばれる配信する動画を細かく分割したファイルの、2種類のファイルの形式が規格化されている。セグメントは一本の動画を短い秒数で分割した動画ファイルの1つを指す。MPD は動画のエンコード時における符号化方式やビットレート、セグメントの分割単位や取得方法など、動画のストリーミング再生に必要なメタデータを保持する XML ファイルである。ストリーミングサービスを提供するためには、提供する動画をあらかじめ複数の画質 (解像度、ビットレート) でエンコードし、それぞれの画質の動画を同じ時刻の位置で分割し、それらをすべてサーバに保存しておく。これらの画質の情報や、分割位置などの情報は MPD ファイルに記載しておく。

MPEG-DASH の仕組みを表す図を図 1、2 に示す。再生クライアントでは、まず再生する動画の MPD ファイルをサーバから取得し、セグメントの情報を得る。この情報を元に、ネットワークの通信品質などを考慮して再生クライアントはサーバに対して各分割位置におけるセグメントの画質を指定し要求する。サーバからクライアントへの動画ファイルの送信も含め、これらの通信には HTTP プロトコルが使用される。

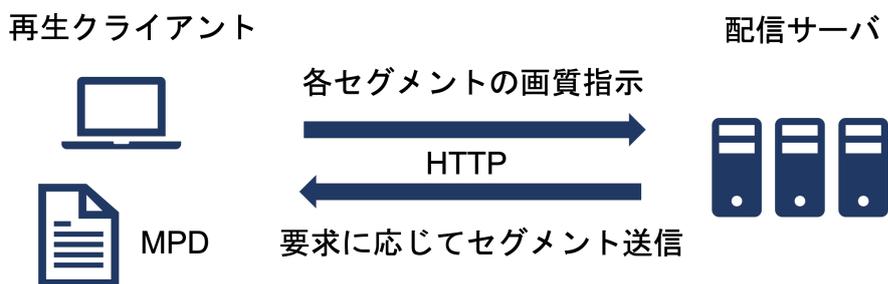


図 1: MPEG-DASH における動画ストリーミング

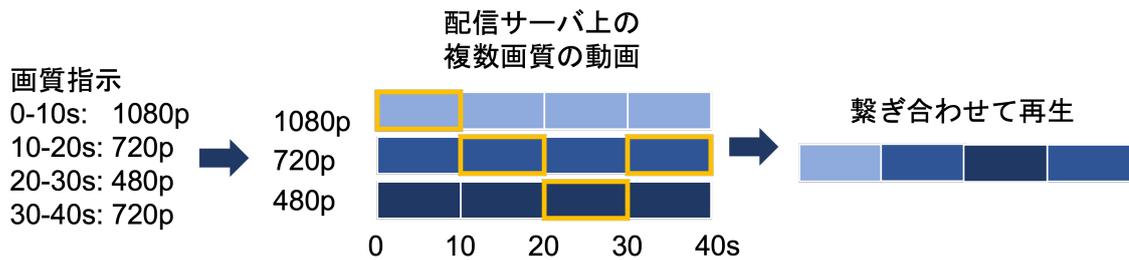


図 2: ビットレート制御の流れ

2.2 脳波を用いた動画視聴中のユーザの QoE 推定

本報告で利用している、脳波情報に基づき QoE を推定する手法は、文献 [1,7] において提案されたものである。この節ではこの文献で提案されている、推定モデルの入力に使用する特徴量の種類、特徴量のうち適切な組み合わせを決定する手法、推定モデルについて説明を行う。

2.2.1 特徴量の計算

入力特徴量には、Emotiv 社の Emotiv EPOC X [8] と呼ばれる 14 チャンネルのセンサを持った脳波計から得た 14 種類 (AF3、F7、F3、FC5、T7、P7、O1、O2、P8、T8、FC6、F4、F8、AF4) の脳波データ (EEG データ) を用いる。EEG データはデバイス内部で 2048sps (サンプリング/秒) でサンプリングされ、出力時に設定に応じて 128 または 256sps にダウンサンプリングされる。測定時の電圧解像度は 14bit、測定帯域幅は 0.16Hz-43Hz である。また、EPOC はモーションセンサ (加速度計、ジャイロセンサ、磁力計) を備えており、アーチファクトの除去の補助のために使用可能である。サンプリングしたデータは Bluetooth 接続された計算機に送信される。EPOC から送信された EEG データおよびモーションデータを記録するソフトウェアとして同社 EmotivPro を使用する。

特徴量の計算の前に前処理としてまずこれらの波形に 4 つの帯域のバターワースバンドパスフィルタを適用する。脳波の周波数成分は運動や認知による脳の活動で変化することが知られている。一般に用いられる周波数成分は 30Hz 以下の成分であり、本報告において、特徴的な周波数成分毎に特徴量の抽出を行う (α , β , θ 波)。各帯域毎の信号の抽出にバターワースバンドパスフィルタを使用する。バターワースバンドパスフィルタは特定帯域を通過させ、それ以外の帯域を遮断または減衰させる特性を持つフィルタである。前処理を行った 14 種類の脳波データに対し、バンドパワー、パワースペクトル密度、離散ウェーブレット変換の値などの特徴量を計算する。特徴量には、最大値、最小値、中央値、分散を用いてお

り、合計 546 種類の特徴量になる。

2.2.2 特徴量の選別

文献 [1, 7] では前述した 546 種類の特徴量をすべて入力とするのではなく、いくつかの特徴量に絞ってモデルの入力としている。これは、推定にかかる計算時間を減らすこと、過学習を防ぐことを目的としている。特徴量の選別は、選択した特徴量で学習させたモデルの交差検証のスコアを目的関数とした、遺伝的アルゴリズムによって行われる。

2.2.3 QoE 推定モデルの形式

QoE 推定のモデルは、546 種類の特徴量の一部を前述の手法によって選別したものを入力とし、出力を QoE が高い・普通・低い の 3 分類にするサポートベクターマシンである。文献 [1] では、学習用のデータにおいては QoE は 0~100 の値で記録されているが、この値のうち上位 25% を QoE が高い、下位 25% は低いとし、それ以外は普通としてラベル付けをおこなっている。

2.2.4 QoE 推定の精度

文献 [1] では、複数人の被験者に対して QoE 推定のモデルを作成し、検証用データに対して QoE が低いときを判別するときの再現率を計測している。文献 [1] によると、再現率は最大 74% で平均は 49.3%、最小 19% であったとし、個人差が大きいことが示唆されている。

3 脳波情報から推定した QoE を用いたビットレート制御システムの実装

3.1 システム概要

実装するシステムの外観を図 3 に示す。2.1 節で説明したように、システムは動画配信用のサーバと、動画セグメントを受け取って再生するクライアント側に分類される。システムは以下の順序で動作する。

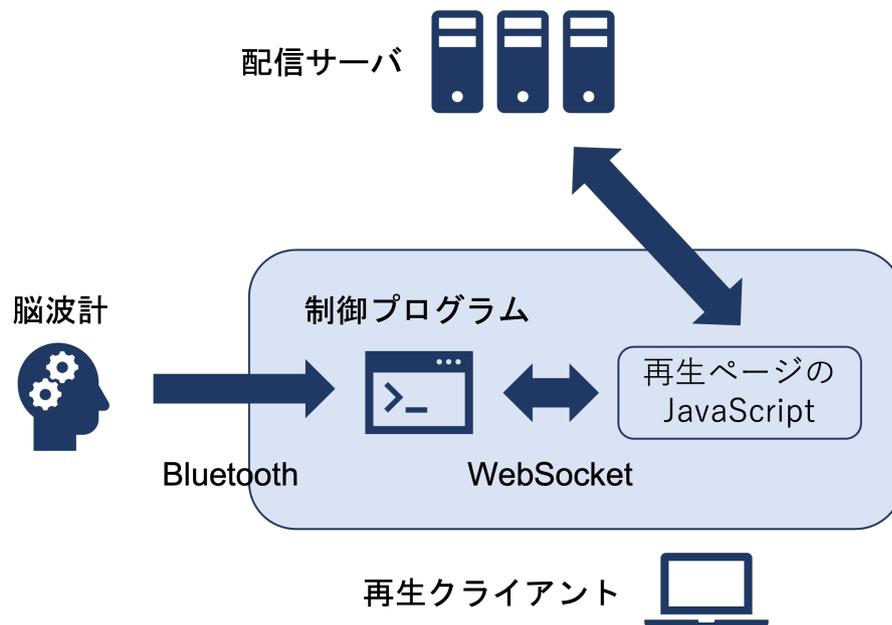


図 3: システム概観

1. 配信サーバへのアクセス クライアントは動画配信用のサーバに http によりアクセスを行う。再生用のページの HTML と JavaScript がダウンロードされる。ここでダウンロードされた JavaScript が MPEG-DASH のクライアント用ライブラリを利用して配信サーバとの通信を行い、ビットレートの制御や通信のスループットなどの情報の取得を行う。

2. DASH クライアントと制御用プログラムの接続 ダウンロードした JavaScript ファイルの初回読み込み時に、ローカルであらかじめ起動している QoE 推定結果に基づくビットレート制御を行うプログラムへの接続を行う。制御用プログラムと DASH クライアントは WebSocket によって通信を行う。

3. 動画再生開始 再生ページ上の再生開始用のボタンをクリックすると、動画の再生が開始される。また、DASH クライアント側から制御用のプログラムに対し現在の通信スループットやリバッファリングの発生などの情報の送信を開始する。

4. ビットレート制御 ビットレート制御用のプログラムは、脳波計から受け取る脳波データと、DASH クライアントから受け取るスループットやリバッファリングの発生などの情報をもとにビットレートを決定し、DASH クライアントに対し WebSocket を通してビットレートの変更を指示する。

なお、図 3 の DASH クライアントプログラム、DASH サーバの実装方法であるが、これは文献 [5] で提案されたものを利用する。文献 [5] では制御用プログラムと DASH クライアントとの通信は WebSocket を用いて行われており、そのプロトコル上で、いくつかの通信フォーマットが定められているとしている。今回の実装では、"nextBitrate,1" という文字列を WebSocket を通して制御用サーバから DASH クライアントに送信することで要求ビットレートを変更することを可能としている。制御用プログラムの通信インターフェースを既存の DASH クライアント側に合わせることで実装を行い、その他のアルゴリズムは文献 [5] のものと同様とした。

3.2 機器構成

3.2.1 動画配信サーバ

配信のサーバは Ubuntu を利用し、配信のサーバプログラムには Apache を利用する。配信する動画はあらかじめ ffmpeg で複数画質にエンコードし、サーバ上に保存している。クライアントからのリクエストに応じて、適宜指定されたビットレートのセグメントをクライアントに送信する。

3.2.2 動画再生クライアント

再生用の計算機には、ノート型計算機を用いた。詳細な性能は表 1 に示す。動画再生クライアントと脳波計との接続には Bluetooth を用いた。

3.2.3 脳波計

脳波の計測には、2.2 節と同様に、Emotiv 社の Emotiv EPOC X [8] と、そのソフトウェアの EmotivPro を使用した。使用した脳波計を図 4 に示す。EPOC X は AF3、F7、F3、

表 1: 再生用の計算機

OS	macOS Monterey バージョン 12.6.3
プロセッサ	2.4 GHz 8 コア Intel Core i9
メモリ	64 GB 2667 MHz DDR4
グラフィックス	AMD Radeon Pro 5600M 8 GB

FC5、T7、P7、O1、O2、P8、T8、FC6、F4、F8、AF4 の 14 個のセンサーを持っており、センサーの精度は 14 ビットで、サンプリングレートは 128Hz である。またビットレート制御部ではリアルタイムに脳波データを取得する必要がある。本報告で使用した脳波計である EPOC X では、Lab Streaming Layer [9] と呼ばれる時系列データのリアルタイム送受信システムを使用することで、脳波のストリーミングデータを受け取ることができる。Lab Streaming Layer には各種プログラミング言語用のパッケージが用意されており、本報告の制御用のプログラムでは Python を利用しているため、Python 用の pylsl パッケージを使用した。



図 4: Emotiv EPOC X

3.3 ビットレート制御用プログラムの実装

3.3.1 脳波情報による QoE 推定の実装

脳波情報による QoE 推定手法では、関連研究の節で説明したように取得した脳波データに数種類のバンドパスフィルタを掛け、その波形に対して 546 種類の特徴量を計算する部分と、計算した特徴量から QoE の高い・普通・低いを分類するサポートベクターマシンのモデルを学習する部分に分割することができる。前者は文献 [1] では MATLAB を使用したものであり、ビットレート制御プログラムとの接続を考慮し、本報告では Python を用いた実装を採用した。特徴量から QoE の分類を行うサポートベクターマシンの学習を行うプログラムの実装では、学習したモデルの保存が可能となるように `pickle` モジュールでモデルを保存することとした。リアルタイムでの QoE の推定を行う際には、あらかじめ保存しておいた `pickle` 形式のモデルファイルを読み込み、逐次 `pylsl` のモジュールを通して脳波データを受け取り、特徴量計算処理を呼び出して特徴量を計算し、計算した特徴量を読み込んでおいたモデルに渡して QoE の分類を行うといった手順となる。

3.3.2 ビットレート制御の実装

ビットレート制御部分のプログラムは、文献 [5] のビットレート制御アルゴリズムを参考に実装を行った。文献 [5] では、瞬きの頻度を QoE の推定値とみなして制御アルゴリズムを実装しているため、この部分を脳波により推定した QoE に置き換えることで制御方法の実装を行った。ビットレート制御は、WebSocket を通して `nextBitrate,{bitrate_id}` (`bitrate_id` は 0~4 のいずれか) という文字列を DASH クライアントに送信することで行われる。ここで送信する `bitrate_id` はあらかじめ設定してあるビットレートの値と対応しており、`bitrate_id` が 0、1、2、3、4 のときはそれぞれ、0.5 Mbps、1 Mbps、1.5 Mbps、3 Mbps、5 Mbps (bps: bit per second) に対応するように実装している。以下では、`bitrate_id` を 1 上げることを、ビットレートを一段階上げると表すことにする。

本報告では、リバッファリングによる QoE の低下を検出した際に、可能な限り高い画質を提供しつつリバッファリングを抑制するようなビットレートの選択を行う。一方で、QoE の低下によってビットレートを低下させる機能のみでは、一度下がったビットレートが増加することがない。一般的に、画質の向上は動画視聴者の QoE の向上と正の相関があると言われており、本報告においても、ビットレートを増加させる制御を加えることとする。脳波情報から推定した QoE によるビットレート制御のアルゴリズムは以下の通りである。

- DASH クライアントから送られてくる通信のスループットを最新 5 回分記録しその平均値を算出する。平均値が連続して 20 回、一段階高いビットレートよりも大きくなっ

た場合、ビットレートを一段階上げる

- DASH クライアントからリバッファリングの発生の通知を受け取ると、ユーザの QoE の推定を行う。このとき、QoE が低いと推定された場合は、直近 5 回のスループットの平均をネットワークの推定スループットとし、そのスループットを超えない最大のビットレートを以降のビットレートとして設定する

上記のアルゴリズムの 2 つ目において、リバッファリングが発生したときにおいても、QoE が高いと判断されたときにビットレートをそのまま維持するには、QoE が悪化する原因の 1 つである画質の変化をなるべく避けるという目的がある。

4 実装したシステムの評価

4.1 評価手法

本報告において実装したビットレート制御システムの動作検証では、実装したシステムにおいて動画を視聴しながら、各時刻における通信スループット、動画のビットレート、推定 QoE とリバッファリングの発生時刻を記録し、その記録した各種データにより、3章で述べたビットレート制御が正しく動作していること確認することで行う。本実験は2名の被験者を対象に行なった（本報告の報告者および、所属研究室教員の1名）。

本報告で行った実験環境では、通常通信のスループットを計測すると約 900 Mbps となった。この環境下で動画視聴を行うとリバッファリングが全く発生しない。そこで、動画配信サーバにおいて、ネットワークの帯域と遅延を制御する `tc` コマンドを実行することで、意図的に帯域を指定した値に設定することとした。時刻の経過に応じて帯域を変更するために、ソースコード 1 に示すシェルスクリプトを作成し、一定時間ごとに帯域を変化させた。以下のシェルスクリプトでは、はじめ 5 Mbps に設定し、60 秒後に 1 Mbps に、その 120 秒後に 10 Mbps に設定する動作となっている。また、以下のスクリプトでは、評価時点における帯域の情報を記録するために、`date` コマンドを用いて適宜帯域を変更した時刻を出力している。

ソースコード 1: `tc` コマンドによって帯域のスケジューリングを行うシェルスクリプト

```
1 tc qdisc add dev ens160 root handle 1:0 tbf rate 5mbit burst 2.5kb limit
   25kb # 5Mbps
2 echo '5Mbps'
3 date +%s
4 sleep 60
5 tc qdisc add dev ens160 root handle 1:0 tbf rate 1mbit burst 0.5kb limit
   5kb # 1Mbps
6 echo '1Mbps'
7 date +%s
8 sleep 120
9 tc qdisc add dev ens160 root handle 1:0 tbf rate 10mbit burst 5kb limit
   50kb # 10Mbps
10 echo '10Mbps'
11 date +%s
12 sleep 180
13 tc qdisc del dev ens160 root
```

本実験において、`tc` コマンドでは 1 Mbps~10 Mbps の間で帯域を変化させるが、これは本評価実験で用いる動画の最大のビットレートが 5 Mbps であるためである。なお、使用した動画は Big Buck Bunny [10] であり、これは CC-BY-3.0 ライセンスで公開されてお

り、動画の全体の長さは4分ほどである。動画配信用サーバーでは、5 Mbps の動画以外にも 0.5 Mbps、1 Mbps、1.5 Mbps、3 Mbps のビットレートの動画も用意されており、クライアント側で指定したビットレートの動画が配信される。

4.2 評価環境

動画再生用のクライアントと配信サーバは研究室のローカルネットワークを介して相互に接続した。いずれもネットワークの接続には有線ケーブルを用いた。評価実験時の、再生クライアントから配信サーバへの遅延時間、配信サーバから再生クライアントへのデータ送信におけるネットワークの帯域は、それぞれ ping コマンドと iperf3 コマンドによって事前に計測した。計測結果はそれぞれ表 2、表 3 であった。いずれも、tc コマンドを使用しない場合には、最高画質の動画の送受信が問題なく行える環境となっている。

表 2: ping の結果

	min	avg	max	stddev
ping	0.432ms	0.489ms	0.535ms	0.035ms

表 3: iperf3 による帯域測定の結果

Transfer	1.10 GBytes
Bitrate	941 Mbits/sec

4.3 QoE 推定のためのモデル学習

実装したシステムを動作させるにあたって、視聴するユーザ毎に QoE を推定するためのモデルをあらかじめ用意しておく必要がある。ここでは、本報告の実験で使用した QoE 推定モデルの学習方法について説明する。

4.3.1 学習用データの作成

QoE 推定モデルを学習するために必要なデータは、動画視聴中のユーザの脳波データと、動画視聴中に随時入力した QoE の値である。このような学習用データを作成をする際、文献 [1] では動画視聴中の人が QoE を回答するために、スマートフォン上で動作する QoE 入力用のアプリケーションを用いている。本報告における実験と類似した入力用のアプリケー

ションを開発する。本報告で QoE 推定を行うモデルは QoE の 2 値、あるいは 3 値分類を行うものである。そこで、学習用データにおいても、3 値の計測を行うものとした。

作成したアプリケーションは QoE の入力を、視聴用クライアントのキーボードの J を押したとき QoE を下げ、K を押したときに上げるという方法を取り、その値として 0、50、100 の 3 値をとるものとした。動画視聴中に視聴者が随時入力した QoE の値を 0.5 秒の間隔で記録することで、文献 [1] で用いられているように、学習用のラベル情報を取得する。

学習時に視聴した動画には、LFOVIA [11] の提供する動画データセットを使用した。このデータセットにはゲームの映像、アニメーション、街中の風景映像や自然映像など様々な種類の動画が含まれており、途中で一時停止が発生したり、画質が変化したりする動画が複数本収録されている。本実験ではこのうち 18 本の動画を学習用データの取得のために使用した。

4.3.2 モデルの学習

作成した学習データをこれまでに説明した QoE 推定のためのプログラムに渡して実行することで、モデルの学習が完了する。学習において発生した問題について、以下のように検討を行った上で最終的なモデルを作成した。

入力した QoE の値についてのアノテーション 文献 [1] では、入力した 0~100 の範囲の QoE の値の、上位 25% を QoE が高い、下位 25% を QoE が低い、その他を普通という分類にしているが、本報告では QoE の値は 0、50、100 のいずれかとなるため、文献 [1] で計測した脳波が違うクラスに分類されてしまう可能性がある。この点については QoE の計測値が各値となった個数が均一となるのか、不均一となるのかについても考慮する必要があり、実験的に、0 を QoE が低い、1 が QoE が普通、2 が QoE が高いとラベル付けする方法をそのまま採用することとした。

推定 QoE の値が常に一定になる 前述のように QoE の分類の仕方を変更し、学習を行うと、QoE が高い、または低いどちらか一方の値を常に出力するモデルが学習結果となってしまいう結果が得られた。この原因は、学習データのうち QoE が低いデータと高いデータの割合がおおよそ 3:7 と一方に大幅に偏ってしまったためであり、その偏った側の値を常に出力するモデルが作成されたためである。この際の精度は見かけ上 70% となるものの、QoE 制御に使用するとリバッファリングの回避が一切行われないうこととなる。モデルの不均一性の問題を解決するために、学習用のデータにおいて、QoE の分類が低い、あるいは高いラベルがつけられたデータの数が同数になるように、データ数が多い側のデータを減らすこととした。

学習パラメータの調整 モデルの学習を行ったところ、学習に使用したデータでは精度が 70% 程度となるものの、検証用のデータでは精度が 40% 程度まで低下した。これは過学習が原因であると考えられるため、文献 [1] では 546 種類の特徴量から 10 種類の特徴量を選択するようになっていたが、本報告の実験では 2 種類にまで削減することで検証用データでの推定結果の低下を抑制した。ただし、脳波を用いた QoE の推定には個人差が多く現れることがこれまでに確認できており、この調整は被験者ごとに行う必要がある可能性も考えられる。

4.4 評価結果

事前に作成した学習モデルを用いて、本報告で実装したビットレート制御を有効にした状態で、動画視聴を行う実験を実施した。動画を視聴する被験者には視聴中の動画に対してQoEの回答を行わせた。同時に視聴中におけるネットワークの帯域の変化、選択されたビットレート、リバッファリングの発生などアプリケーションに発生したイベントを記録した。

この実験により取得したデータのうち、各時刻における推定スループット、選択ビットレートと、リバッファリングの発生時刻を示したグラフを図5に示す。図の青線はビットレート、緑色はスループット、図のグレーの範囲はリバッファリングが発生し、動画が一時停止していた時間帯を示している。図5のリバッファリングの発生時刻における選択ビットレートに着目すると、リバッファリング発生時にビットレートが低下し、スループットを超えないビットレートに切り替えができていていることが分かる。

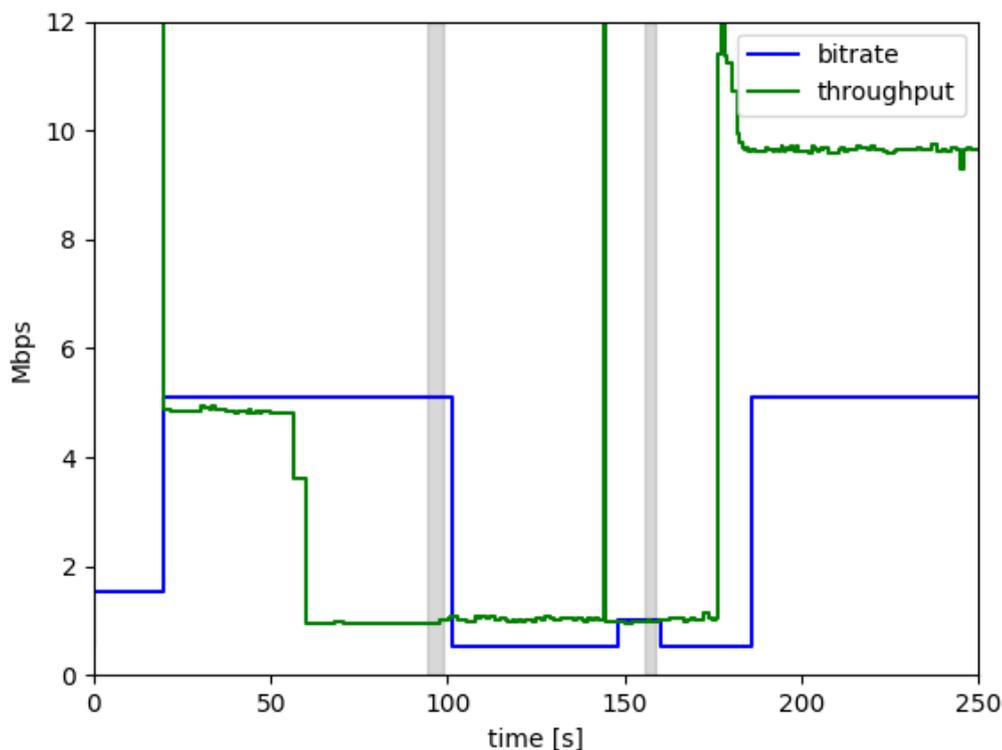


図5: リバッファリングの発生と選択ビットレート

推定 QoE に応じてビットレート制御が行われていることを確認するため、同実験の、各時刻における推定 QoE、選択ビットレートに合わせて、リバッファリングの発生時刻をプロットしたグラフを、図 6 に示す。図の青線は先程のグラフと同じくビットレート、橙色はスループット、グレーの範囲も先程と同じくリバッファリングが発生し、動画が一時停止していた時間帯を示している。

図のリバッファリングの発生時刻の部分（グレーの範囲の左端）に注目すると、2つのグレーの範囲のいずれも推定された QoE が 0 となっており、QoE が低いと判定されているため、その直後に低いビットレートの動画が送られていることが分かる。ただし、この実験では、リバッファリングが発生した際に QoE が高くなっているケースが観測できなかったため、リバッファリングが発生してもビットレートの変更を行わない場合の動作確認については十分ではない。

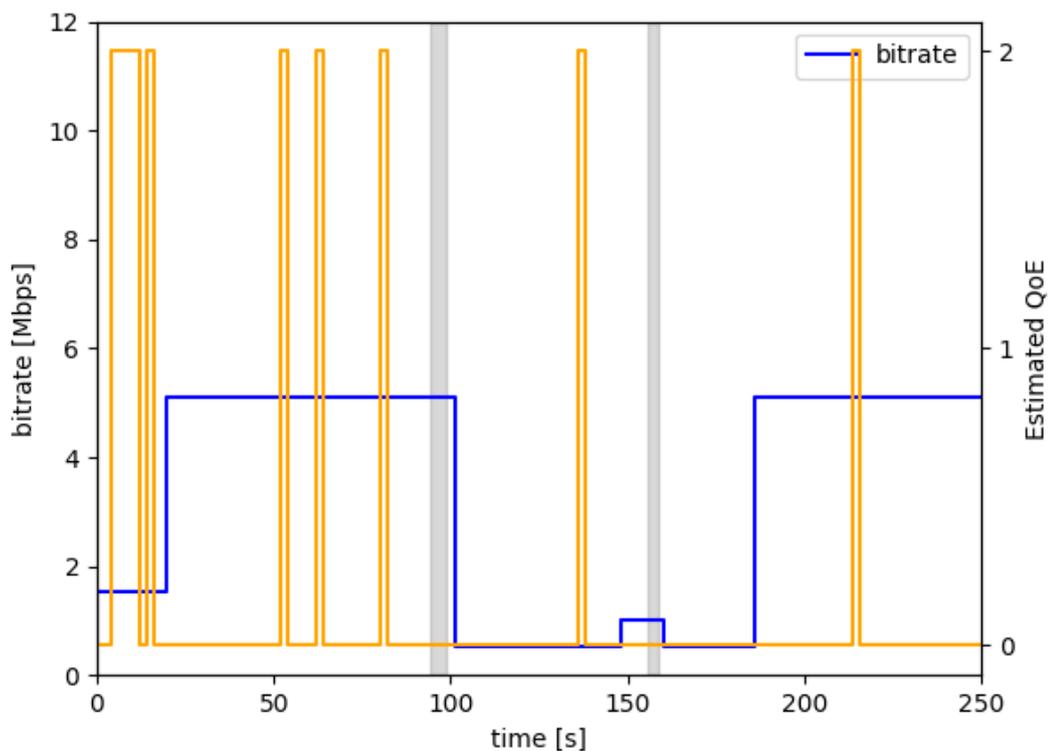


図 6: 選択ビットレートと推定 QoE

そこで、制御プログラムにおいて、QoE の推定プログラムが常に 2 を返す（常に QoE が高いと推定する）ように修正し、再度実装した動画配信システムにて、各時刻の選択ビットレート、スループット、推定 QoE と、リバッファリングの発生時刻を記録しながら動画を再生する実験を行った。その結果を図 7、8 に示す。グラフの線の色分けは 1 つ目の実験と同じである。

図 7 を見るとリバッファリングが多数発生しているが、ビットレートは変化せず、3 Mbps を保っていることが分かる。その後、通信スループットが 10 Mbps まで回復し、一定時間経過した後は画質の最も高い 5 Mbps までビットレートが引き上げられていることも確認できる。

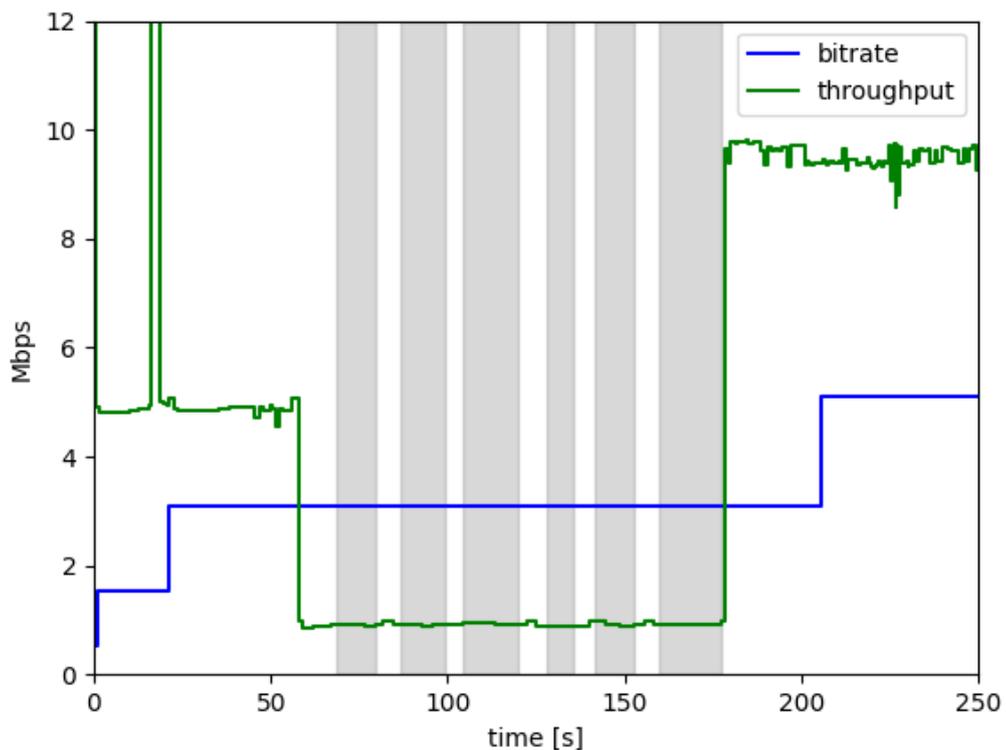


図 7: リバッファリングの発生と選択ビットレート（推定 QoE 固定）

図 8 は選択ビットレートと推定 QoE、リバッファリングの発生時刻の関係を示したグラフであるが、QoE が常に一定で 2 となっており、推定 QoE が高いと判定されるときはビットレート制御が行われないことが確認できる。

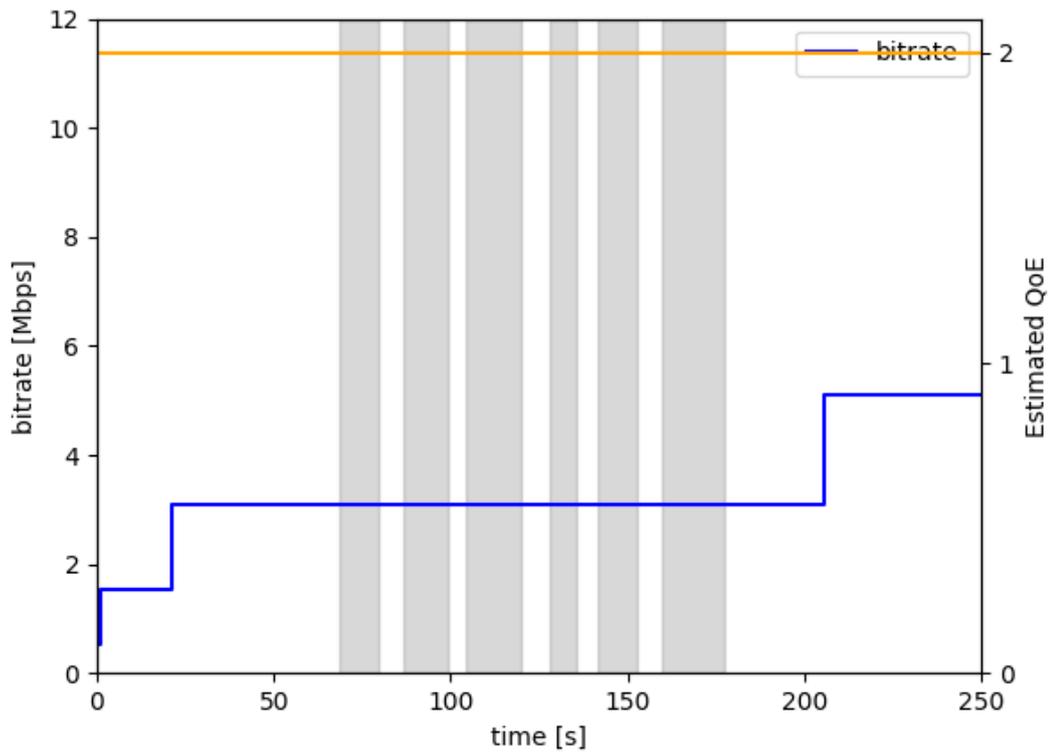


図 8: 選択ビットレートと推定 QoE (推定 QoE 固定)

5 おわりに

本報告では、動画視聴中のユーザの QoE の維持・向上を目的とし、脳波情報から推定した QoE を用いたビットレート制御システムを実装した。我々の研究グループではこれまでに脳波情報から QoE を推定する手法についての研究を行っており [1, 7]、本報告ではこの脳波情報により推定した QoE を用いた、動画配信システムにおけるビットレート制御アルゴリズムを実装した。実装したビットレート制御システムの動作を検証した結果、リアルタイムで脳波情報から QoE を推定し、その値を用いてビットレート制御を行えていることが確認できた。また今回実装したアルゴリズムによって、ユーザの QoE が高いときはビットレートの変更を行わないことにより、不要なビットレートの変更を減らせるため、QoE の維持・向上を期待できる。今後は、文献 [1] の手法を改善することによる脳波情報を用いた QoE の推定の精度向上に加え、文献 [5] にあるような脳波以外の生体情報を複合的に利用することによるさらなる QoE の推定精度の向上を図る。さらに、QoE とその性質に基づく適切なビットレート制御方法の検討、実装に取り組む。構築したシステムの評価のためにより多くの被験者実験を行い、システムの改善を行う。

謝辞

本報告を終えるにあたり、大阪大学大学院情報科学研究科の村田正幸教授には、ご多忙の中多大なるご指導賜りましたことを心より深く感謝申し上げます。ならびに、大阪大学大学院情報科学研究科の小南大智助教には平素より手厚いご指導いただきましたこと、ここに深く感謝いたします。また、貴重なご指導を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、深謝の意を表します。最後に、日々の学生生活をご支援いただいた研究室の皆様に対し感謝の意を表し、ここに謝辞とさせていただきます。

参考文献

- [1] 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.
- [2] 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.
- [3] Y. Yuan, S. Lin, and G. Zhou, “QoE control for dynamic adaptive video streaming over HTTP at access point,” in *Proceedings of IEEE International Conference on Industrial Internet (ICII)*, Nov. 2019, pp. 268–277.
- [4] S. Jiarun, Y. Fuzheng, Z. Yicong, W. Shuai, and W. Hong Ren, “QoE evaluation of multimedia services based on audiovisual quality and user interest,” *IEEE Transactions on Multimedia*, vol. 18, no. 3, Mar. 2016.
- [5] 児玉 大暉, “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,” *大阪大学基礎工学部情報科学科特別研究報告*, Feb. 2022.
- [6] 西村 敏, “MPEG-DASH,” *映像情報メディア学会誌*, vol. 71, no. 1, pp. 78–81, 2017.
- [7] K. Kitao, D. Kominami, and M. Murata, “GA-based feature selection for QoE estimation using EEG during video viewing,” in *Proceedings of International Conference on Emerging Technologies for Communications*, Dec. 2020.
- [8] Emotiv©. Emotiv epoc x — 14 channel mobile eeg headset. <https://www.emotiv.com/epoc-x/>. Accessed: 2023-01-19.
- [9] sccn. sccn/labstreaminglayer. <https://github.com/sccn/labstreaminglayer>. Accessed: 2022-11-22.
- [10] “Big buck bunny.” [Online]. Available: <https://peach.blender.org/>
- [11] LFOVIA group. Lab for video and image analysis – lfovia. <https://people.iith.ac.in/lfovia/index.html>. Accessed: 2023-01-25.