

令和 4 年 5 月 31 日現在

機関番号：32689

研究種目：基盤研究(C) (一般)

研究期間：2019～2021

課題番号：19K11932

研究課題名(和文) AVコンテンツ視聴時におけるユーザ主観評価推定に関する多角的研究

研究課題名(英文) Study on Estimation of User Subjective Evaluation while Watching AV Contents from Various Angles

研究代表者

亀山 渉 (Wataru, KAMEYAMA)

早稲田大学・理工学術院・教授

研究者番号：90318858

交付決定額(研究期間全体)：(直接経費) 3,400,000円

研究成果の概要(和文)：AVコンテンツ視聴時におけるユーザの映像品質に対する主観評価を効率よく、かつ、精度良く生体情報等から推定する手法に関する研究を行った。(1)固定されたモニタ環境で視聴中の、被験者の生体情報、動画像の興味度と満足度、及び再生ビットレートから、被験者がスライダを使って評価した主観評価値を推定する手法と、(2)モバイル端末(スマートフォン)環境で視聴中の、被験者の生体情報、興味度、視聴コンテキスト、動画解像度から、視聴後に答えてもらう主観動画像品質を推定する手法を検討した。実験の結果、両者ともに、機械学習を適用することで高い精度で主観動画像品質を推定できることが分かった。

研究成果の学術的意義や社会的意義

従来、ユーザの主観映像品質評価取得には、国際標準化されたアンケートによるMOS評点(Mean Opinion Score)が広く用いられている。しかしながら、人の特性による回答内容の正当性担保、多大な評価時間、リアルタイム回答が不可等の問題が存在する。また、ユーザ毎の違い及びユーザコンテキストは考慮されず、映像内容を考慮した主観評価は行えない。本研究により、ユーザの生体情報、並びに、使用するデバイス及びコンテンツから得られる情報から主観映像品質が高い精度で推定できる可能性が明らかになったことは、提案方式が、これらの問題点を解決する有効な新しい主観映像品質評価手法となりえることを示唆している。

研究成果の概要(英文)：We have conducted a study on methods for efficiently and accurately estimating users' subjective evaluations of video quality during AV contents watching from users' bio-signals. We study two methods: (1) a method for estimating the subjective evaluation values rated by subjects using a slider, from their bio-signals, interest in and satisfaction of contents, and video bitrates, in a fixed monitor environment, and (2) a method for estimating subjective video qualities answered by subjects after watching contents on a mobile terminal (smartphone), from bio-signals, interest in contents, viewing contexts, and video resolutions. Experimental results show that, for both methods, estimating subjective video quality with high accuracy can be achieved by using machine learning.

研究分野：情報通信工学

キーワード：AVコンテンツ 主観動画像品質評価推定 ユーザ満足度 ユーザ興味度 視聴コンテキスト センサ情報 生体情報 機械学習

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

1. 研究開始当初の背景

各種の情報通信サービス等の提供においては、サービス等に対するユーザの主観評価を得て、サービス及びネットワーク等の管理や制御に利用する必要がある。ユーザ主観評価を得る方法としては、従来、国際標準化されたアンケートによる MOS 評点 (Mean Opinion Score) [1] が広く用いられている。しかしながら、アンケート等による主観評価には、設問内容の妥当性担保と人の特性による回答内容の正当性担保の問題があり、取得に多大な時間を要するだけでなく、一般にリアルタイムに回答を得ることはできない[2]。

一方、ネットワーク制御の分野では、QoS (Quality of Service) に代わるものとして、実際にユーザがサービスを体感する品質である QoE (Quality of Experience) を利用する研究が盛んに行われている。QoS が等しい場合であってもユーザ毎、並びに、ユーザがおかれた状況等 (以下、ユーザコンテキスト) によって QoE は異なる。例えば、文献[3]では、QoE に影響を与えるものとして、「アプリケーションあるいはサービスのタイプや特性、使用コンテキスト、ユーザの期待や満足感、ユーザの文化的背景、社会的及び経済的観点、心理的側面、ユーザの情動」等が挙げられている。

また、映像品質の主観評価においては、長年の研究成果を基に、数学モデルによる HDTV 映像の客観評価方式[4, 5]が国際標準化され、検証実験によって PSNR (Peak Signal to Noise Ratio) よりも MOS 評点に近い結果が得られることが分かっている。しかしながら、この方法ではユーザ毎の違い及びユーザコンテキストは考慮されておらず、映像内容を考慮したユーザの主観評価は推定できない。

以上より、各種の情報通信サービスに対するユーザ主観評価の取得においては、アンケートによる MOS 評点ではなく、ユーザ毎の違いとユーザコンテキストを考慮に入れつつ、リアルタイムに精度よく推定できる手法が求められていた。

2. 研究の目的

そこで、本研究では、ネットワークを介した AV コンテンツ視聴時のユーザ主観評価を、ユーザに装着させる各種の生体情報測定器及びセンサ、並びに、ユーザが使用するデバイスから得られる情報から、精度よく推定する手法について明らかにすることを目的とした。従来のユーザコンテキスト推定に関する研究では、ユーザに装着させるセンサ (加速度センサ、方位センサ、位置センサ等) やユーザが使用するデバイスから得られる情報 (各種のサービス利用に関する情報等) のみを使用していたが、本研究ではこれらに加え、心電計等によるユーザの生体情報を取得し、先の情報と併せてユーザコンテキスト推定も行った。これにより、AV コンテンツ視聴の際の、ユーザ毎の異なった情動反応と個別のユーザコンテキストを考慮したユーザ主観評価推定を目指した。

また、従来のユーザ主観評価推定の研究では、対象となるサービス品質のユーザ主観評価推定のみを静的に行っていた。本来、ユーザ主観評価は、品質だけではなく、文献[3]に定義されているように多角的な側面を持ち、時々刻々と変化するものである。よって、本研究では、先に述べた生体情報を併用することで、サービス品質の主観評価のみならず、ユーザコンテキストによって変化するユーザ満足度やユーザ興味度も併せて推定に利用することで、多角的な観点からユーザ主観評価をリアルタイムに推定することを目的とした。

3. 研究の方法

環境毎に使用できる生体情報測定器、センサ機器、デバイスから得られる情報が異なるため、以下の二つの研究課題を設けて研究を遂行した。

(1) 研究課題 1 固定環境での主観評価推定: 固定されたモニタで異なったビットレートの AV コンテンツを視聴する際の生体情報 (顔特徴点、脳波、心電等) を取得し、併せてアンケートによる主観評価 (品質、満足度、AV コンテンツに対する興味度等) も取得した。機械学習によって、それらの関係を解析した。

(2) 研究課題 2 モバイル環境での主観評価推定: スマートフォンで異なったビットレートの AV コンテンツを視聴する際の、各種センサ情報 (頭部の加速度・方位・位置等)、スマートフォンから得られる各種の情報 (加速度、方位、位置、アプリの使用状況等)、並びに、生体情報 (歩行状態、心電等) を取得し、研究課題 1 と同様のアンケートによる主観評価も取得した。機械学習によって、それらの関係を解析した。

なお、本研究は、実験協力者 (被験者) として学生や一般人等の協力を仰ぐこととなり、これら協力者の個人関連情報を伴うデータ (脳波・瞳孔径等の生理指標、端末から得られる個人情報、アンケート回答等) を取り扱う。そのため、人権、個人情報保護及び安全対策について、以下の対策を講じて行った。まず、早稲田大学の人権保護規定・倫理規定、並びに、研究代表者と研究分担者が所属する学会における同様の規定の遵守と共に、研究開始に先立ち、早稲田大学「人を

対象とする研究に関する倫理委員会」から研究計画及び内容の承認を得た。更に、早稲田大学「人を対象とする研究に関する倫理規定」に基づき、実験協力者に対して、研究目的、内容、並びに、実験の安全性と考えられる危険性等について、書面及び口頭で十分に説明を行った上で、実験協力者から実験参加の同意を得て実験を行った。

4. 研究成果

(1) 研究課題 1 固定環境での主観評価推定[6]

本研究課題では、様々に実験条件を変化させ、推定精度の高い方式を試行錯誤的に探る研究アプローチを取った。そのため、中間的な研究結果は省略し、最終的な実験条件での研究成果を以下に述べる。

実験には、男性 8 名及び女性 9 名の被験者（平均年齢 21.7 歳，標準偏差 1.27）が参加した。試料動画は 16 本の有音で異なるジャンルの 1920×1080[pixel]の動画を 10 分程度にトリミングしたものをを用いた。トリミング後の動画は 10 分割して、それぞれを再生ビットレートが 300[kbps]、500[kbps]、1000[kbps]、2500[kbps]、5000[kbps]のいずれかとなるよう H.264 で再エンコードした。各試料動画の 10 分割中で、これらの 5 段階の再生ビットレートを少なくとも一つ含むように再生ビットレートを割り振った。また、16 本の動画の総再生時間のうち、5 段階の再生ビットレートの合計再生時間、並びに、300[kbps]から 1[Mbps]へといった再生ビットレートの変化パターンの頻度はそれぞれ等しく設定した。

取得する生体情報は、顔画像、脳波（AF3、F7、F3、FC5、T7、P7、O1、O2、P8、T8、FC6、F4、F8、AF4 の 14 チャンネル）心電であり、視聴後に 5 段階の興味度及び満足度をアンケートによって取得した。また、視聴中には文献[1]に準拠した自作のスライダを操作してもらい、主観動画品質を被験者から取得した。取得データは前処理を経て、RRI（R-wave R-wave Interval: 心電の R 波間隔）1 次元、脳波 140 次元、顔特徴点 48 次元を 2[Hz]のデータとして用意した。これに、興味度及び満足度の 2 次元、並びに、視聴動画の再生ビットレート 1 次元を加え、計 192 次元変数を説明変数とした。一方、スライダによる評価値を 1-5 の整数値に量子化したものと、0-5 の連続値に変換したものを用意し、XGB（eXtreme Gradient Boosting）によって、前者を目的変数としたクラス分類、後者を目的変数とした回帰分析を行った。結果の一部を表 1 に示す。なお、表中の「同一動画内解析」は、各被験者について、各動画のデータからランダムに選択した 80% を訓練データ、残りの 20% をテストデータとしたものを指す。また、「未学習動画解析」は、各被験者の 1 本の動画のデータをテストデータ、それ以外の 15 本の動画からテストデータの 4 倍の量となるようランダムにデータを選択して訓練データとしたものを指す。

表 1 XGB による主観動画品質の推定結果

条件	値
同一動画内解析の回帰分析(平均決定係数)	0.894
同一動画内解析のクラス分類(平均精度)	0.890
未学習動画解析のクラス分類(平均精度)	0.635

実験の結果、同一動画解析では、生体情報のみでも高い精度の推定が可能であったが、表 1 に示すように興味度、満足度、視聴動画の再生ビットレートを説明変数に加えることでより高い精度での推定が可能であった。特に、再生ビットレートを説明変数に加えることで従来よりも高い精度での推定が可能となった。一方、未学習動画解析では、再生ビットレートのみで推定した場合と比較して、生体情報、興味度、満足度、再生ビットレートを説明変数に加えて推定することで推定精度が向上した。以上から、可変ビットレート動画に対して生体情報を用いた QoE 推定の可能性が示唆された。今後の課題としては、異なったジャンルの動画に対する生体情報の有効性の検討、未学習動画解析の精度向上が挙げられる。

(2) 研究課題 2 モバイル環境での主観評価推定[7]

本研究課題でも研究課題 1 と同様の研究アプローチを取ったため、最終的な実験条件での研究成果を以下に述べる。

実験には、男性 5 名及び女性 2 名の被験者（平均年齢 22.9 歳，標準偏差 0.639）が参加した。被験者は約 1 週間の実験期間内に、生体信号を取得する実験器具を装着した状態で、約 5 分の動画をスマートフォンで計 54 本視聴した。動画は 9 本を 1 リストとしてまとめた上でオフライン保存し、解像度は 144p、360p、1080p の 3 種類をそれぞれ 18 本用意し、動画のフレームレートは 30fps 以下とした。なお、1 リストには各解像度の動画が 3 本ずつ含まれ、動画 - 解像度の組み合わせと動画リスト - 視聴コンテキストの組み合わせを 3 種類ずつ用意し、偏りが発生しないよう条件を定めた。被験者は実験期間内に 1 リストの動画を視聴する実験を各自で 6 度行う。動画を視聴する際のコンテキストは、自宅、移動中（公共交通機関利用時）、外出先の 3 種類とし、被験者はそれぞれのコンテキスト下において 2 リストずつ動画を視聴した。動画を視聴する際、動画の巻き戻し、スキップ及び一時停止は緊急時を除き禁止した。被験者には、各動画視聴中に、スライダを用いて自身の興味度についてリアルタイムで回答するように求めた。各動画視聴後には、5 段階の主観動画品質、再視聴欲、他人への推薦度合

い、興味度を、また、前述した3種の視聴コンテキストをアンケートによって取得した。

取得する生体情報は、眼球の動き、頭部加速度、頭部角加速度、瞬目、歩行状態、心電(RRI)である。取得データには前処理を施した後、10[Hz]のデータに整形した。これらのデータに、アンケートで取得した再視聴欲、他人への推薦度合い、興味度、コンテキスト、スライドによる興味度、並びに、動画画像解像度を加え、計32次元の説明変数とした。一方、目的変数はアンケートで取得した主観動画品質である。RF(Random Forest)及びXGBによる結果の一部を表2と表3に示す。なお、表3の「未学習データに対する解析結果」は、テストデータのある動画画像1本に関するデータとし、残りの動画画像に関するデータを全て学習用データとして解析を行ったものを指す。ここで、解析は全ての動画画像が1度ずつテストデータとなるように行い、精度は推定が成功した動画画像の割合とした。

表2 RFによる層化分割交差検証結果(推定精度)

被験者	コンテキストあり	コンテキストなし
1	0.948	0.925
2	0.893	0.883
3	0.958	0.950
4	0.919	0.927
5	0.910	0.889
6	0.905	0.870
7	0.921	0.914

表3 XGBによる未学習データに対する解析結果(推定精度)

被験者	コンテキストあり	コンテキストなし
1	0.407	0.444
2	0.620	0.560
3	0.460	0.426
4	0.590	0.592
5	0.673	0.558
6	0.679	0.434
7	0.537	0.519

主観動画品質推定に関して、層化分割交差検証ではRFで9割以上の精度、未学習データに対する解析では層化分割交差検証に比べ精度は低かったもののXGBによってチャンスレベルの2倍以上の精度が得られた。これらの結果より、実環境下での実験においても生体情報やアンケート情報を用いることで主観動画品質が推定可能であることが示された。また、過半数の被験者において視聴コンテキスト情報を入力に加えることで推定精度が向上した。よって、視聴コンテキストはユーザの主観評価に影響を有すると考えられる。今後の課題としては、未学習データ推定の精度向上、解析手法の最適化、コンテキストの細分化が挙げられる。

本検討ではコンテキストを被験者に入力してもらったが、被験者が利用するスマートフォンのセンサ情報(加速度,角速度,気圧,ネットワーク使用量,緯度経度,高度)並びに、被験者に装着してもらった心電計から取得するRRIから、コンテキストの自動認識を検討した[8,9]。時系列データを考慮したCNN(Convolutional Neural Network)によるコンテキスト推定実験の結果、98%以上の高い推定結果を得ることができ、また、生体信号であるRRIの有効性も確認した。よって、本検討にこの成果を適用することで、コンテキスト情報の入力を省略できることが期待される。

文献

- [1] ITU-R Rec. BT.500-13, "Methodology for the subjective assessment of the quality of television pictures", 2012年
- [2] 亀山 渉, "生体信号情報の情報通信工学への適用—生体信号によるユーザ主観評価の客観的推定—", ITU ジャーナル, Vol.48, No.10, pp.26-pp.29, 2018年
- [3] ITU-T Rec. P.10/G.100, "Vocabulary for performance, quality of service and quality of experience", 2017年
- [4] ITU-T Rec. J.341, "Objective perceptual multimedia video quality measurement of HDTV for digital cable television in the presence of a full reference", 2016年
- [5] ITU-R Rec. BT.1907, "Objective perceptual video quality measurement techniques for broadcasting applications using HDTV in the presence of a full reference signal", 2012年
- [6] 大池 健太郎, 菅沼 睦, 亀山 渉, "可変ビットレート動画視聴時の生体情報を用いた QoE 推定," 信学技報, Vol.121, No.421, CQ2021-109, pp.49-54, 2022年3月(電子情報通信学会コミュニケーションクオリティ研究会 学生優秀発表賞受賞)

- [7] 三品 翔太郎, 菅沼 睦, 亀山 渉, “視聴コンテキストを考慮した実環境下におけるスマートフォン使用者の生体情報による主観動画像品質推定,” 信学技報, Vol.121, No.421, CQ2021-110, pp.55-60, 2022 年 3 月(電子情報通信学会コミュニケーションクオリティ研究会 学生優秀発表賞受賞)
- [8] Hiromi SHIMIZU, Mutsumi SUGANUMA, Wataru KAMEYAMA, “Context Analysis and Estimation of Mobile Users Considering the Time Series of Data,” 2020 IEEE 9th Global Conference on Consumer Electronics (GCCE), Oct. 13-16, 2020, Kobe, Japan, pp.496-498, 2020 年 10 月
- [9] 清水 寛生, 菅沼 睦, 亀山 渉, “時系列を考慮したモバイルユーザのコンテキスト分析と推定手法の比較検討,” 信学技報, Vol.120, No.392, CQ2020-107, pp.1-5, 2021 年 3 月

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 Shotaro MISHINA, Mutsumi SUGANUMA, Wataru KAMEYAMA	4. 巻 Vol.9, No.12
2. 論文標題 Estimating subjective video quality while using smartphone	5. 発行年 2020年
3. 雑誌名 IEICE Communications Express	6. 最初と最後の頁 pp.632-635
掲載論文のDOI（デジタルオブジェクト識別子） 10.1587/comex.2020COL0035	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計10件（うち招待講演 1件 / うち国際学会 2件）

1. 発表者名 大池 健太郎, 菅沼 睦, 亀山 涉
2. 発表標題 可変ビットレート動画視聴時の生体情報を用いたQoE推定
3. 学会等名 信学技報, Vol.121, No.421, CQ2021-109, pp.49-54
4. 発表年 2022年

1. 発表者名 三品 翔太郎, 菅沼 睦, 亀山 涉
2. 発表標題 視聴コンテキストを考慮した実環境下におけるスマートフォン使用者の生体情報による主観動画品質推定
3. 学会等名 信学技報, Vol.121, No.421, CQ2021-110, pp.55-60
4. 発表年 2022年

1. 発表者名 松本 真奈, 大池 健太郎, 菅沼 睦, 亀山 涉
2. 発表標題 Autoencoderと生体情報を用いた可変ビットレート動画に対するQoE推定に関する検討
3. 学会等名 2022年電子情報通信学会総合大会, B-11-17
4. 発表年 2022年

1. 発表者名 Hiromi SHIMIZU, Mutsumi SUGANUMA, Wataru KAMEYAMA
2. 発表標題 Context Analysis and Estimation of Mobile Users Considering the Time Series of Data
3. 学会等名 2020 IEEE 9th Global Conference on Consumer Electronics (GCCE), Oct. 13-16, 2020, Kobe, Japan, pp.496-498 (査読有) (国際学会)
4. 発表年 2020年

1. 発表者名 清水 寛生, 菅沼 睦, 亀山 涉
2. 発表標題 時系列を考慮したモバイルユーザのコンテキスト分析と推定手法の比較検討
3. 学会等名 信学技報, Vol.120, No.392, CQ2020-107, pp.1-5
4. 発表年 2021年

1. 発表者名 亀山 涉
2. 発表標題 生体情報を利用した映像視聴者のQoE推定
3. 学会等名 2019年電子情報通信学会ソサイエティ大会, BI-8-2 (招待講演)
4. 発表年 2019年

1. 発表者名 Hiromi SHIMIZU, Mutsumi SUGANUMA, Wataru KAMEYAMA
2. 発表標題 Context Analysis and Estimation of Mobile Users by Using Bio-signals and Sensor Data
3. 学会等名 2019 IEEE 8th Global Conference on Consumer Electronics (GCCE 2019), Oct.15-Oct.18, Osaka, Japan, pp.267-270 (査読有) (国際学会)
4. 発表年 2019年

1. 発表者名 坂本 貴幸, 菅沼 睦, 亀山 渉
2. 発表標題 実際のながら視聴状況想定環境下における生体情報を用いた主観映像品質評価推定についての検討
3. 学会等名 信学技報, Vol.119, No.455, CQ2019-135, pp.5-10
4. 発表年 2020年

1. 発表者名 大池 健太郎, 菅沼 睦, 亀山 渉
2. 発表標題 可変ビットレート動画像に対する生体指標を用いたQoE推定
3. 学会等名 2020年電子情報通信学会総合大会, B-11-28
4. 発表年 2020年

1. 発表者名 三品 翔太郎, 菅沼 睦, 亀山 渉
2. 発表標題 モバイル端末使用時における主観動画像品質の推定に関する検討
3. 学会等名 2020年電子情報通信学会総合大会, B-11-29
4. 発表年 2020年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

電子情報通信学会コミュニケーションクオリティ研究会で2022年3月に発表した「可変ビットレート動画視聴時の生体情報を用いたQoE推定」及び「視聴コンテキストを考慮した実環境下におけるスマートフォン使用者の生体情報による主観動画像品質推定」は、それぞれの筆頭著者が、2022年3月の「学生優秀発表賞」を受賞した。

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分 担 者	菅沼 睦 (SUGAMUMA Mutsumi) (50399507)	早稲田大学・理工学術院 (国際情報通信研究科・セン ター)・客員主任研究員 (32689)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関