

# エッジコンピューティングによるネットワーク型 複合現実アプリケーションのサービス体感品質の向上性に関する評価

高木 詩織<sup>†</sup> 金田 純一<sup>†</sup> 荒川 伸一<sup>†</sup> 村田 正幸<sup>†</sup>

<sup>†</sup> 大阪大学 大学院情報科学研究科 〒565-0871 大阪府吹田市山田丘 1-5

E-mail: †{s-takagi,j-kaneda,arakawa,murata}@ist.osaka-u.ac.jp

**あらまし** 近年、仮想現実 (VR: Virtual Reality) 技術や複合現実 (MR: Mixed Reality) 技術が発展し、これらの技術を用いて臨場感のある体験を提供するネットワークサービスが展開されつつある。仮想現実技術や複合現実技術を用いたネットワークサービスでは、ユーザが一方的に音声や映像などの体験を享受するサービスと比べて、ネットワークの遅延やパケットロスの増大に対するユーザ体感品質の変化が大きくなる可能性がある。そこで、本稿では、複合現実技術を用いたアプリケーションを作成し、エッジコンピューティングの導入によるユーザの体感品質の向上性を検証する。検証の結果、遅延が 720 [ms] から 920 [ms] の間でユーザの体感品質が急激に悪化するため、クラウドコンピューティング環境で約 1 秒の遅延が発生する場合、エッジコンピューティングの導入によりユーザの体感品質が向上することが明らかとなった。

**キーワード** 複合現実 (MR)、エッジコンピューティング、サービス品質 (QoS)、ユーザ体感品質 (QoE)、ネットワークロボット

## On the Improvement of Service Quality by Edge Computing in Network-oriented Mixed Reality Application

Shiori TAKAGI<sup>†</sup>, Junichi KANEDA<sup>†</sup>, Shin'ichi ARAKAWA<sup>†</sup>, and Masayuki MURATA<sup>†</sup>

<sup>†</sup> Graduate School of Information Science and Technology, Osaka University

Yamadaoka 1-5, Suita, Osaka, 565-0871 Japan

E-mail: †{s-takagi,j-kaneda,arakawa,murata}@ist.osaka-u.ac.jp

**Abstract** In recent years, VR (Virtual reality) and MR (Mixed Reality) technology are developed, and network services that provide a realistic experience using these technologies are being developed. In the network service using VR or MR, QoE (Quality of Experience) change with network delay and packet loss may increase compared with the service in which the user enjoy voice and video unilaterally. In this paper, we create an application using MR and evaluate QoE improvement by introducing edge computing. The results of our experiments reveal that QoE suddenly gets worse when network delay is between 720 [ms] and 920 [ms] and QoE is expected to be improved by introducing edge computing when network delay is about 1 [second] in cloud computing environment.

**Key words** Mixed Reality (MR), Edge Computing, Quality of Service (QoS), Quality of Experience (QoE), Network Robot

### 1. はじめに

スマートフォンやタブレット端末が普及するとともに、インターネットはますます人々に身近な存在となっている。インターネットを通して提供されるサービスは人々の要求に応じて多様化しており、数多くの新しいアプリケーションやサービスが登場している。

近年、ネットワークの研究分野では、エッジコンピューティング (MEC: Multi-access Edge Computing) の研究開発が進められている。エッジコンピューティングでは、通信網のエッジである基地局などにコンピューティングリソースやストレージを配備することで、エンド端末が必要とする処理をよりエンド端末に近い場所で行う。文献 [1] では、エッジコンピューティングのユースケースとして、AR (Augmented Reality) コンテ

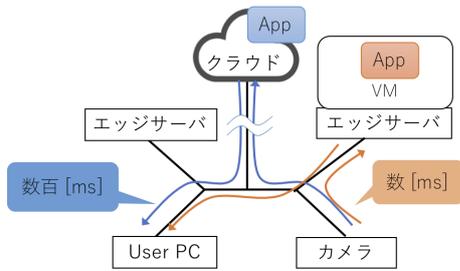


図 1: エッジコンピューティングによる情報処理の例

の配信アプリケーションを挙げている。AR コンテンツをエッジサーバにキャッシュしておくことで、RTT を最小限に抑え、スループットを最大化することができる。これにより、遅延を抑え、アプリケーションの応答性が改善されることが期待されている [2]。例えば、図 1 のようにカメラやセンサーなどを搭載した端末において取得された実世界の情報を、別拠点のデータセンターへ転送することなく、局所的に統合・処理を行うことが期待される。しかし、エッジコンピューティング技術は研究開発の段階であり、利用アプリケーションを含めて検討段階にある。

最近では、仮想現実 (VR: Virtual Reality) 技術や複合現実 (MR: Mixed Reality) 技術が発展し、これらの技術を用いて臨場感のある体験を提供するネットワークサービスが展開されつつある。例えば遠隔地にいる VR 用ヘッドセットを装着した親族に、結婚式の様子を届けるサービス [3] や、スポーツ競技の会場から離れた会場に選手が現れたように感じる「超高臨場ブリックビューイング」[4] などが考えられている。

このような新たなアプリケーションの登場に際し、情報ネットワークでは、運ぶ情報の特性に応じて、例えばリアルタイム性が要求される情報伝達には RTP (Real-time Transport Protocol) などのプロトコルを導入したり、サービス品質 (QoS: Quality of Service) を高めるプロトコル制御等の導入も行われている [5]。しかし、QoS はネットワーク提供者が測定可能な性能品質であり、QoS のみではユーザのサービス体感品質は測定できない。これは、ユーザの体感品質が通信の品質だけでなくアプリケーション自体の使い勝手やユーザの心理的要因や状況要因によっても決定されるためである。そのため、ユーザの体感品質 (QoE: Quality of Experience) [6] に関する研究がなされている。例えば、文献 [7] では、Web ブラウジングにおいて、ネットワーク品質がユーザの行動に与える影響を明らかにしており、ユーザの行動が変化するネットワーク品質の条件は、平均オピニオン評点 (MOS: Mean Opinion Score) が低下する点と一致することを述べている。

仮想現実技術や複合現実技術を用いたネットワークサービスの登場に際し、情報ネットワークにおいても新たなプロトコルの導入・開発が求められる可能性はある。しかし、新たなプロトコルを導入するにあたっては、ユーザの体感品質が何によって決定づけられるのかを明らかにする必要がある。仮想現実技術や複合現実技術を用いたネットワークサービスでは、ユーザ

は Web ブラウザ越しに情報閲覧を行うのではなく、五感情報の臨場体験の享受を可能とすることが求められる。また、ユーザによるジェスチャーなどの直感的な運動で近接端末や遠隔端末を操作することも考えられる。従って、ユーザが一方的に音声や映像などの体験を享受するサービスと比べて、ネットワークの遅延やパケットロスの増大に対するユーザ体感品質の変化が大きくなる可能性がある。

そこで、本稿では、複合現実技術を用いたアプリケーションとして、局所的に取得される実世界の情報を局所的に統合・処理を行いつつ情報伝達を行い、ユーザに対して新たな臨場体験を提供するアプリケーションを作成した。さらに、エッジコンピューティング環境を導入し、クラウドコンピューティング環境での遅延を模擬的に発生させることで、エッジコンピューティングによるユーザ体感品質の向上性を実機を用いて検証した。

本稿の構成は以下の通りである。まず、2 章で想定するサービスについて述べる。次に、3 章で実験方法と評価方法、および実験結果について述べる。最後に、4 章でまとめと今後の課題について述べる。

## 2. ネットワークを利用した複合現実型サービス

本章では、本稿で想定するネットワーク型複合現実アプリケーションの内容と、実際に実装したアプリケーションの仕様について述べる。

### 2.1 想定するサービス

想定するサービスは、複合現実技術を用いたショッピングモール体験サービスである。現実世界のショッピングモールにロボットを配置し、ユーザが遠隔地からそのロボットを操作することで実際の店舗を見て回って買い物をしている体験を提供する。

ロボットは実際の店舗の様子を撮影し、例えば以下の環境情報を映像に付加して配信する。

- 商品の在庫
- 売れ筋の商品
- 新発売の商品
- 店舗別特典の内容

環境情報はロボット側のエッジサーバに蓄積されおり、適宜更新される。また、エッジサーバではロボットの移動に合わせて随時環境情報を映像に挿入してユーザに送信する。このとき、映像や音声だけではなく、匂いや商品の触感のような五感の情報もユーザに送信することで、臨場感のある体験を提供することも考えられる。

一方、ユーザは MR ヘッドセットを装着しており、その MR ヘッドセットには、例えば以下のユーザ側の環境情報が表示される。

- 冷蔵庫にあるもの・ないもの
- 冷蔵庫にあるものの賞味期限
- 電池などの普段使っている日用品の大きさや種類

ユーザ側の環境情報はユーザの自宅にあるエッジサーバに蓄積されており、MR ヘッドセットがエッジサーバに映像を送信

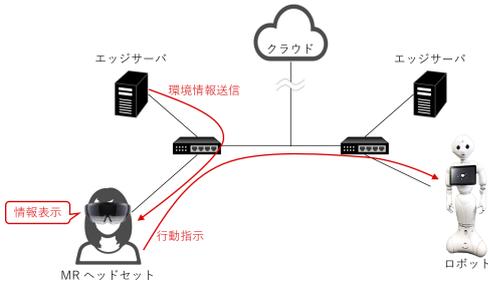


図 2: ユーザ側のアプリケーションでの処理

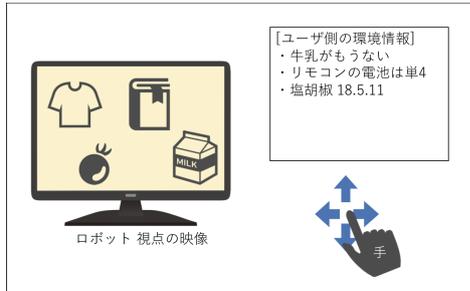


図 3: ユーザの視界イメージ

することで自動的に環境情報が更新される。

ユーザは、遠隔地のロボットから送られてくる映像と MR ヘッドセットのレンズ上に表示される情報を見ながらジェスチャーや音声によりロボットを操作する。

## 2.2 実装したアプリケーションプログラム

実装したアプリケーションプログラムは、ロボットのカメラで撮影している映像とその映像上に表示された情報、MR ヘッドセット上に表示された情報を見ながらジェスチャーでロボットを移動させるプログラムである。以下、ユーザ側、ロボット側のそれぞれのプログラムの概要を述べる。

### 2.2.1 ユーザ側

実装したユーザ側のアプリケーションにおける処理を図 2 に示す。また、ユーザがアプリケーションを使用中の MR ヘッドセット越しの視界のイメージを図 3 に示す。

環境情報を表示するオブジェクトは、ホワイトボードのようなオブジェクトで、アプリケーションの起動時に MR ヘッドセット上に表示される。MR ヘッドセットは定期的に HTTP で取得し、その内容をホワイトボード上に表示する。

また、MR ヘッドセットで指の動きを検知し、スワイプ操作によってロボットに歩行移動をさせる。一本の指を上から下に下げ、指を下げたまま上下左右のいずれかの方向に動かしたのち、指を上へ上げると、ジェスチャーが確定する。指の動いた距離が閾値を越えると表 1 に示す方向にロボットが歩行する。ロボットが一度に動く距離は、指の動いた距離の 6 倍である。閾値を超えないスワイプ操作およびタップ操作では、ロボットが時計回りに 90 度回転する。

本稿では、ロボットとしてソフトバンクロボティクスの Pepper を用いる。本プログラムでは、MR ヘッドセット上で動作するアプリケーションから Pepper に接続し、Pepper の移動を

表 1: 指ジェスチャーとロボットの方向の対応

指	Pepper
上	前
下	後
左	左
右	右

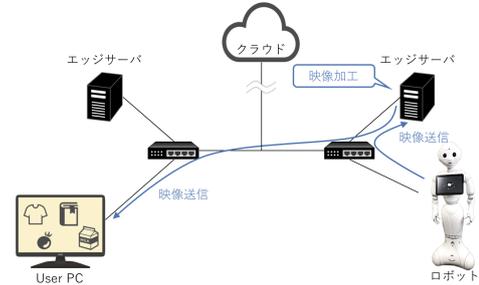


図 4: ロボット側の処理

提供する API である ALmotion API を用いて Pepper の移動を実現している。

また、MR ヘッドセットとして Microsoft HoloLens を用いる。HoloLens とは、ヘッドマウントディスプレイ型の自己完結型ウェアラブルコンピュータである。Hologram と呼ばれる仮想オブジェクトを現実世界と重ね合わせて表現することができる。また、Hologram は視線やジェスチャー、音声によって操作することが可能である。

HoloLens アプリケーションの開発は Unity を用いて C# で行う。C# と Pepper の連携は標準ではサポートされていないため、Baku.LibqiDotNet ライブラリ [8] を用いる。プログラム内でロボットの IP アドレスを指定し、TCP で接続する。HoloLens には手の位置やタップ操作を検出する機能が搭載されており、タップ開始時と終了時それぞれの手の位置から手の移動量距離を計算し、Pepper の移動距離を決定する。

### 2.2.2 ロボット側

本稿でのロボット側の処理内容を図 4 に示す。ロボット側では、撮影した映像を加工して配信する。

撮影した映像の取得・加工・配信には、映像の記録や加工などを行うことができるフリーソフトウェアである ffmpeg [9] を用いる。ffmpeg は Pepper の OS である NAOqi に標準搭載されているが、標準搭載されている ffmpeg ではネットワーク経由で映像をストリーミングすることができないため、ネットワーク経由でのストリーミングが可能となる ffmpeg0.9.0 を再コンパイルし、置き換えて使用する。ffmpeg での映像の送信には UDP を使用する。

映像のストリーミング配信には、ffmpeg に付属するストリーミングサーバである ffserver を用いる。

ffserver で配信される映像の再生には、ffmpeg 付属の再生ソフトウェアである ffplay を用いる。ストリーミング配信される映像をバッファリングする機能を持つ。しかし、本稿では、映像のバッファリングによる遅延を抑えるため、バッファサイズ

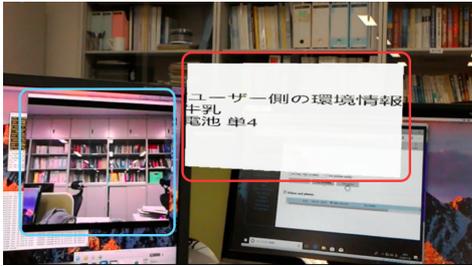


図 5: アプリケーションの実行例: 情報更新前



図 6: アプリケーションの実行例: 情報更新後

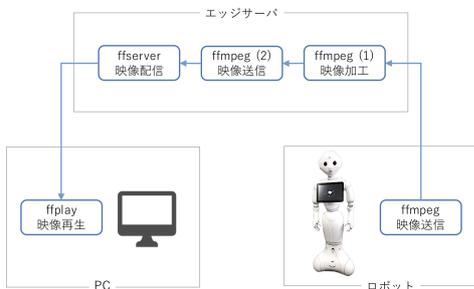


図 7: ロボット側の映像処理の流れ

を最小に設定して使用する。

### 2.3 アプリケーションの実行例

実際にアプリケーションを使用している様子を図 5 に示す。青色で囲んだ領域がロボット視点の映像、赤色で囲んだ領域が MR ヘッドセットのレンズ上に表示されるユーザー側の環境情報である。図 6 は、ユーザー側、ロボット側の環境情報が挿入されたときのユーザーの視界画像である。

### 2.4 エッジコンピューティングの導入

本稿で想定するサービスでは、ロボットの移動や MR ヘッドセットを装着したユーザーの環境情報の更新に合わせて映像加工処理を行う必要があるため、高い応答性が必要となる。そのため、エッジコンピューティングを導入する。

ユーザーとロボットは地理的に離れた場所に存在するため、ユーザー側・ロボット側のそれぞれにエッジサーバを設置する。ユーザー側のエッジサーバにはユーザー側の環境情報を蓄積する。また、ロボット側のエッジサーバには、ロボット側の環境情報を蓄積し、映像加工や映像配信を行う ffmpeg と ffserver を配置する (図 7)。エッジサーバの ffmpeg (1) では、ロボットから受け取った映像にテキストを挿入する。ffmpeg (2) では映像を ffserver に送信する。

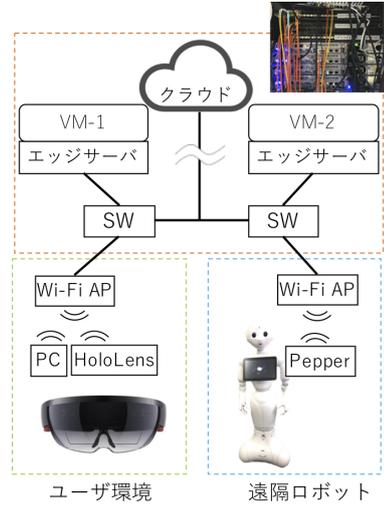


図 8: 実験用ネットワーク構成

表 2: ユーザに与えるタスクと参照する環境情報

タスク	参照する情報	操作内容
1	なし	2 [m] 程度先の目印に移動
2	ユーザー側	環境情報で与えられた目標地点に移動
3	ロボット側	環境情報で与えられた目標地点に移動
4	ユーザー側・ロボット側	環境情報で与えられた目標地点に移動

## 3. ユーザ体感品質の評価

本章では、ユーザー体感品質を評価するための実験環境、ユーザー体感品質の評価方法、および実験結果について述べる。

### 3.1 実験環境

実験用のネットワーク構成を図 8 に示す。ロボット、MR ヘッドセット、ユーザー PC はそれぞれ無線でアクセスポイントに接続する。本実験ではクラウドは用意せず、クラウドコンピューティング環境で発生すると考えられる遅延をエッジサーバで模擬的に発生させる。ネットワーク環境の構築には、仮想化環境構築ソフトウェアである OpenStack [10] を用いている。

なお、ロボットが映像を撮影してエッジサーバに送信するまでの処理に、420 [ms] 程度の遅延が発生することが明らかとなっている [11]。この遅延と通信にかかる遅延に加え、クラウドコンピューティング環境で発生する遅延を VM で模擬的に発生させる。遅延の発生には、netem [12] を用いる。

実験では、被験者に表 2 のタスクを与え、遅延の大きさが変化した場合のタスク完了時間の変化を測定した。タスク 1 から 4 の順に、それぞれ 1 回目に遅延なしの場合、2 回目に遅延ありの場合でロボットを操作している。まず、予備実験として、VM での遅延が 0 [ms]、100 [ms]、200 [ms]、300 [ms]、400 [ms]、500 [ms] の場合のタスク完了時間を測定した。次に、本実験として、8 人の被験者を 4 人ずつのグループに分け、一方のグループでは 300 [ms] のネットワーク遅延、もう一方のグループでは 500 [ms] のネットワーク遅延を発生させた場合のタスク完了時間の変化を測定した。

表 3: 評価項目

E1	ロボットが撮影する映像の品質
E2	ロボット操作の快適性
E3	没入感
E4	MR ヘッドセットに表示された情報の見やすさ
E5	ロボットが撮影する映像に表示される情報の見やすさ

表 4: 評価カテゴリ

カテゴリ	評点
非常に良くなった (Much Better)	3
良くなった (Better)	2
わずかに良くなった (Slightly Better)	1
同じ (About the Same)	0
わずかに悪くなった (Slightly Worse)	-1
悪くなった (Worse)	-2
非常に悪くなった (Much Worse)	-3

### 3.2 評価方法

ユーザの体感品質を主観的に評価する方法として、平均オピニオン評点 (MOS: Mean Opinion Score) [13] がある。ただし、文献 [13] で述べられている評価方法はビデオやオーディオ向けの評価方法であり、複合現実型サービスにおける体感品質の評価方法は確立されていない。そこで、本稿では、以下の方法で評価を行う。

- 客観評価として、各タスクの完了に要した時間を計測する。
- ユーザによるシステムの主観評価として、すべてのタスクの実行が終了した後、表 3 の項目に関して 1 回目 (遅延なし) の実行と比較して 2 回目 (遅延あり) の実行がどのように変わったと感じるかを表 4 のカテゴリで評価し、平均オピニオン評点で表す。

### 3.3 実験結果

#### 3.3.1 客観評価

予備実験における各ネットワーク遅延時間とタスク完了時間の関係を図 9 に、本実験における各ネットワーク遅延時間とタスク完了時間の関係を図 10 に示す。それぞれ横軸はネットワーク遅延時間、縦軸はタスク完了時間を表している。図 9 に示すように、予備実験におけるタスク完了時間は、ネットワーク遅延の大きさに対してほぼ線形に増大している。また、図 10 より、本実験でも予備実験と同様の結果が得られていることがわかる。タスク完了時間について、遅延増大に対する非線形性が観察されなかったことは、映像の遅れに伴う操作ミスがほとんどなかったことを示唆している。これはタスクで与えているロボットへの操作が比較的単純であり、被験者が映像の遅れを予測して操作を行っている可能性がある。複雑なロボット操作や高精度のロボット操作における遅延の影響については今後精査していく必要がある。

#### 3.3.2 主観評価

ユーザの主観評価の平均値を図 11 に示す。横軸は評価項目の種類、縦軸は平均オピニオン評点を表している。

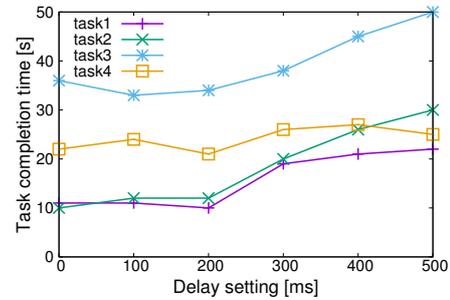


図 9: 予備実験の結果

図 11 の実験結果から、以下のことが言える。

- E1 に関して、遅延が 300 [ms] の場合と 500 [ms] の場合の両方でやや悪化している。本実験において被験者が映像の劣化を感じるのは、ジェスチャーが確定してからロボットが移動しているように見えるまでの時間が増大するという点のみであるため、遅延の増大に対して線形に悪化するのは妥当である。

- E2 に関して、遅延が 300 [ms] の場合では上昇している。これは、同じ内容のタスクを 2 回繰り返したために、操作への慣れやタスク内容を覚えたことによる効果が、遅延による影響を上回ったためと考えられる。一方、500 [ms] の場合は、同じ内容のタスクを繰り返していても、慣れによる効果を遅延による影響が上回ったため、評点が低下したと考えられる。

- E3 に関して、遅延が 300 [ms] の場合は変化がなかったのに対し、遅延が 500 [ms] の場合では著しく悪化している。このことから、遅延が 300 [ms] から 500 [ms] の間でユーザの体感品質が急激に低下すると考えられる。本稿で用いたロボット内での処理遅延は約 420 [ms] であることから、ユーザ・ロボット間の遅延が 720 [ms] から 920 [ms] の間でユーザの体感品質が急激に低下すると考えられる。したがって、クラウドコンピューティング環境で遅延が約 1 秒となるサービス環境では、エッジコンピューティングを導入することでユーザの体感品質が向上すると見込まれる。

- E4 および E5 では、遅延の大きさが変化しても評価は変化しなかった。本実験では、ユーザ側で表示される情報とロボット側で表示される情報は操作中に変化しないため、ネットワーク遅延の影響を受けない。

客観評価では遅延の増大に対してタスク完了時間が線形に増大しているが、主観評価では約 1 秒の映像遅れを境界に体感品質が急激に低下していた。つまり、ネットワーク事業者もしくはサービス提供者が、操作ログにもとづく客観的指標によって体感品質を推定することは容易ではない。特に、エッジコンピューティング環境におけるサービス機能配置を考える上では、人間の体感メカニズムの深い理解が望まれる。

## 4. おわりに

仮想現実技術や複合現実技術を用いたネットワークサービスでは、ユーザは Web ブラウザ越しに情報閲覧を行うのではなく、五感情報の臨場体験の享受を行ったり、ジェスチャーなどの直感的な運動で近接端末や遠隔端末を操作することが考えられ

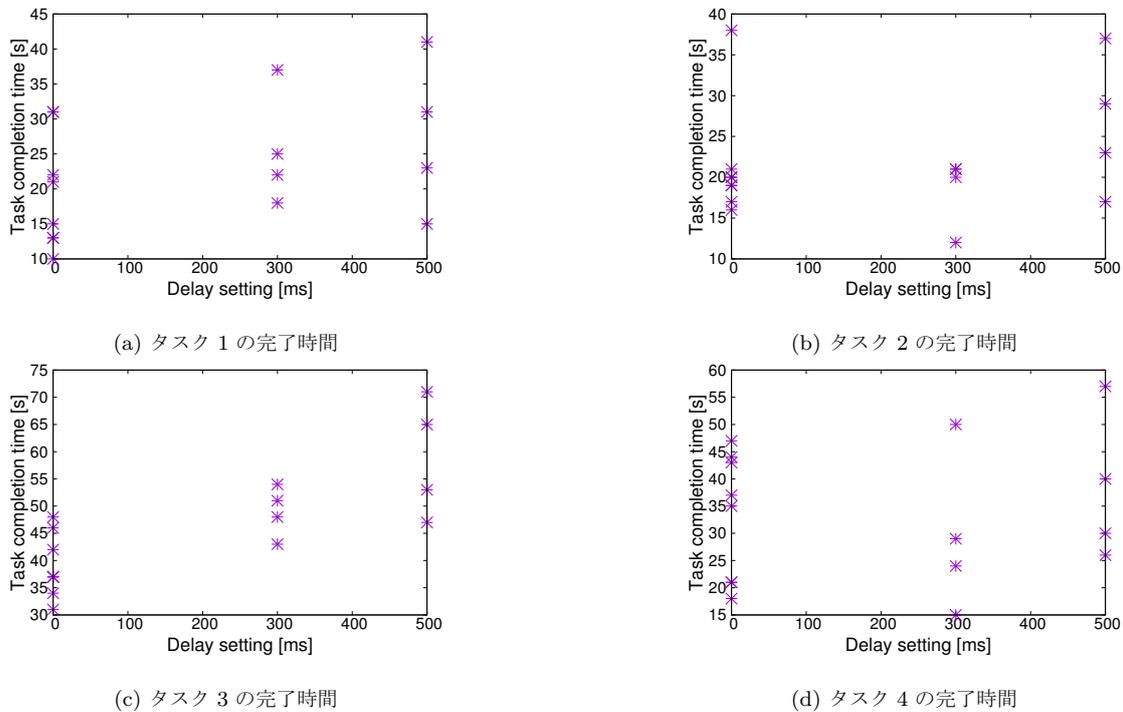


図 10: 各タスクの完了時間

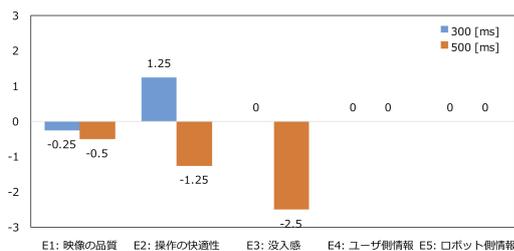


図 11: ユーザの主観評価

る。しかし、そのような状況でのユーザの体感品質が悪化する要因や体感品質の変化量はまだ検証されていない。また、アプリケーションの応答性を高めるためにエッジコンピューティングの標準化が進められている。以上を踏まえ、本稿では複合現実技術を用いたアプリケーションを作成し、エッジコンピューティング環境でクラウドコンピューティング環境で発生する遅延を模擬的に発生させ、ネットワーク遅延の大きさとユーザの体感品質の関係を検証した。検証の結果、タスク完了時間はネットワーク遅延の増大に対して線形に増大するが、遅延が 720 [ms] から 920 [ms] の間でユーザの体感品質が急激に悪化することがわかった。したがって、ロボット操作アプリケーションにおいて、クラウド利用時に約 1 秒の遅延が発生するサービス環境では、エッジコンピューティングの導入によりユーザの体感品質が向上すると見込まれる。

今後の課題としては、ユーザ側・ロボット側で表示される情報が動的に変化する場合の体感品質の評価、アプリケーションのユーザインタフェースの改良、実験で発生させるネットワーク遅延を細分化しユーザの体感品質が低下する閾値を検証することや、多数の被験者での検証が挙げられる。

**謝辞** 本研究の一部は NICT 委託研究「未来を創る新たなネットワーク基盤技術に関する研究開発」によるものである。ここに記して謝意を表す。

## 文 献

- [1] ETSI, “Mobile-edge Computing Introductory Technical White Paper,” Sept. 2014.
- [2] 田中 裕之, 高橋 紀之, 川村 龍太郎, “IoT 時代を拓くエッジコンピューティングの研究開発,” *NTT 技術ジャーナル*, pp. 59–63, Aug. 2015.
- [3] “HUG PROJECT.” <https://hugproject.net/>.
- [4] 窪菌 竜二, 阿久津 明人, 松浦 宣彦, 南 憲一, 小野 朗, “Beyond 2020:高臨場 UX サービスの創出,” *NTT 技術ジャーナル*, pp. 6–9, Oct. 2017.
- [5] I. Busse, B. Deffner, and H. Schulzrinne, “Dynamic QoS control of multimedia applications based on RTP,” *Computer Communications*, vol. 19, pp. 49–58, Jan. 1996.
- [6] International Telecommunication Union, “Recommendation ITU-T P.10/G.100, Vocabulary for performance, quality of service and quality of experience.”
- [7] Hideyuki Koto, Norihiro Fukumoto, Sumaru Niida, Hidetoshi Yokota, Shin’ichi Arakawa, and Masayuki Murata, “Users’ reaction to network quality during web browsing on smartphones,” Sept. 2014.
- [8] “Bakulibqidotnet.” <https://github.com/malaybaku/BakuLibQiDotNet>.
- [9] “FFmpeg.” <https://www.ffmpeg.org/>.
- [10] “Openstack.” <https://www.openstack.org/>.
- [11] 金田 純一, 荒川 伸一, 村田 正幸, “エッジコンピューティング環境におけるサービス機能の配置がユーザの通信品質に与える効果の評価,” *電子情報通信学会技術研究報告 (IN2017-33)*, vol. 117, pp. 61–66, Sept. 2017.
- [12] “netem.” <https://wiki.linuxfoundation.org/networking/netem>.
- [13] International Telecommunication Union, “Recommendation ITU-T P.800, Methods for subjective determination of transmission quality.”