



第 21 回日本バーチャルリアリティ学会大会論文集 (2016 年 9 月)

レイグジスタンスの研究 (第 86 報): インターネットを越えた Telexistence を実現するために求められる通信要件について

Study on Telexistence (LXXXVI): Required features for Telexistence system over the Internet

中蔵聡哉^{1),3)}, チャリス・フェルナンド¹⁾, 南澤孝太¹⁾, 加藤朗¹⁾, 砂原秀樹¹⁾, 館暲²⁾

1) 慶應義塾大学 メディアデザイン研究科 (〒 223-8526 横浜市港北区日吉 4-1-1, t.nakakura@kmd.keio.ac.jp)

2) 東京大学高齢社会総合研究機構 (〒 113-8656 東京都文京区本郷 7-3-1)

3) NTT コミュニケーションズ技術開発部 (〒 108-8118 東京都港区芝浦 3-4-1 グランパークタワー 16F)

概要: The purpose of this study is to develop a Telexistence system over the Internet. Communication over the internet will be affected by network obstacles like delay, jitter, packet loss and etc, and we have to establish the connection over the Firewalls and NATs. We have developed a prototype system and evaluate the effect with network simulator. Our goal is the spread of Telexistence applications, therefore our prototype is based on one of the standard protocols named WebRTC in W3C and IETF. This experimental results for Telexistence system will be used for the robust design of any real-time system over the Internet.

キーワード: Telexistence, WebRTC, Virtual Reality

1. はじめに

遠隔地に対してもインタラクションを行いたいというのが人類の長年の夢であり、これまでに様々な取り組みが成されている。例えば、遠方と音声対話を行うために電話が開発され、更に近年ではビデオ通話機能により視覚情報を得ることも可能になった。我々はこれらに加え、遠隔地に対して没入感を伴った身体インタラクションを提供することを目標に Telexistence[1] の研究を進めており、現在までに Telexistence ロボット TELESAR V([2]) の開発を行っている。TELESAR V は操縦者の体の動きに追従して動作し、自身のセンサが捉えた視覚・聴覚・触覚信号を操縦者の五感情報として感じられるように VR(Virtual Reality) 技術を用いて適切に提示することができる。これらの機能を利用することで、操縦者は高い没入感を得ることができ、自身があたかもロボットの位置にいるかのように振る舞うことが可能になる。

ロボットを遠隔地に置いて操作する為には、操縦者とロボットの間で長距離通信が必要であり、市中の商用網において長距離通信を行うための通信機能の開発が必要である。更に、通信品質によっては、映像品質の低下による視覚への影響や、ロボットの操作データの遅延や欠損による操縦性の低下が懸念される。実際に遠距離で Telexistence アプリケーションを利用するためには、これらの影響を緩和し没

入感を損なわないようにするための機構が必要である。しかしながら現時点においては、没入感を持って Telexistence アプリケーションを操作するために、通信品質に対して求められる要件は自明ではない。個々人の認知能力に影響を受けるため個人差があることが想定され、汎用的に利用できる評価基準は現在までに策定されていない。

通信品質に対する要件は、通信経路で発生すると思われる遅延やロスのようなネットワーク特性による受けつつも、Telexistence アプリケーションの利用に十分な品質への影響を押さえることである。まずは、ネットワークを介して動作する Telexistence アプリケーションのプロトタイプ構築を行い、ネットワーク特性による影響を考察することにより、ユーザの主観評価へ影響を与えるメカニズムを把握することが重要である。本稿ではこのメカニズム把握と、これを緩和するためのエンジンの開発を行い、今後の通信要件策定に繋げるものである。

2. プロトタイプで利用する通信プロトコル

最終的には商用公衆網で利用する事を目標としており、プロトタイプでの検証の妥当性を高めるため、プロトタイプ作成に当たってもこの点を念頭に置いて検討を行う。

2.1 前提とする通信要件

2.1.1 市中のネットワークでの活用

2016 年時点において、全世界において最も普及率が高く接続人口の多いネットワークは、IP ネットワークをベースとするインターネットであるため、これの利用を前提とする。

Toshiya NAKAKURA, Charith FERNANDO, Kouta MINAMIZAWA, Akira KATO, Hideki SUNAHARA and Susumu TACHI

総務省による情報通信白書 ([3]) が示す通り、現在のインターネット構成は、外部からの接続を積極的に待ち受ける端末をサーバ、サーバに接続してサービスを享受するクライアントと分類した形で設計がなされており、通信トラフィックもサーバからクライアントへのダウンロードトラフィックが主である。外部からの攻撃を防ぐため、サーバネットワークは管理者しか立ち入れない専用のサーバールームに配置され、クライアントネットワークへ向けたトラフィックは、サーバからの折り返し通信以外の通信は全て Firewall により遮断されるのが一般的である。また IP アドレス枯渇対策の観点から、外部からアクセスするのに利用できるグローバル IP アドレスはサーバ用のネットワークに集中させ、クライアント用のネットワークは NAT の配下に設置し、外部からの通信確立が困難なローカル IP アドレスによって運用されるケースが多い。Telexistence ロボットはその用途から、サーバールームだけではなく日常空間への設置も想定するため、クライアントネットワークに設置された場合においても接続を確立できなければならない。

2.1.2 Telexistence アプリケーションにおける動作要件

Telexistence ロボットは人間の動きに追従して動作し、自身の持つセンサで得られた五感情報をユーザへ提示する。これを動作させるために求められる通信機能は、操作端末からロボットに対して操作信号の伝送と、ロボットから操作端末への五感信号の伝送である。

通信経路が長くなれば長くなるほど、ネットワーク特性の影響を受ける箇所が増加し、通信品質が低下しやすくなるのが懸念される。仮にロボットと操作端末が同じ構内網に接続していれば、構内網に閉じた通信を行えるのが望ましい。更に中継サーバを利用するよりも直接通信の方が一般に低遅延高品質であるため、まずは直接通信を試行し、不可能であればフォールバック系として中継サーバを経由した形で通信できるように設計するのが望ましい。

2.2 WebRTC の活用

我々は、WebRTC([4]) をベースに本取り組みの通信仕様を検討することとした。WebRTC は、W3C 及び IETF によって標準化が進められている通信仕様であり、アプリケーション間に対して通信ソケットを提供するプロトコルである。提供される通信ソケットは、映像と音声の通信を担当する MediaStream と、任意のデータの転送を担当する DataChannel の機能を持つ。WebRTC はサーバからの折り返しを偽装する機能を備えており、NAT のもつグローバル IP アドレスを利用することで、クライアント端末間で直接通信を試行することができる。直接通信が確立できない場合のフォールバック系として、中継サーバを経由して通信を行うことも可能である。

MediaStream に投入した映像や音声は、VP8([5]) 又は H.264([6]) で圧縮された後、SRTP([7]) パケットとしてネットワークに送出される。DataChannel で伝送するデータは SCTP([8]) のパケットに格納されて送信される。これらの通信は DTLS により暗号化が施される。

通信の確立性能については、民間の調査会社 callstats.io のレポート ([9]) で利用統計が報告されている。カメラやマイクを備えていない端末で MediaStream を利用しようとしたケースのようなハードウェアの不備を除けば、通信確立成功/失敗の比率は 85:1 である。確立された通信のうち直接接続できたものの比率は 78% で、残りの 22% がサーバ中継である。この統計には、セキュリティを重視する企業ネットワークのように、想定外の通信を意識的に遮断しているものも含まれており、100% にすることはそもそも不可能であること、その他のメリットとのトレードオフを考慮して、本用途にとっては十分なものとする。

3. プロトタイプのアプリケーション構成

プロトタイプアプリケーションとして、人間の頭部のみを模した Telexistence ロボットを用いた試験実装を作成した。このロボットは、ヨー、ロール、ピッチの三軸で頭を動かすことができ、胴体を固定した際の人間の頭部移動を再現することができるものである。このロボットを用いたプロトタイプモジュール構成を図 2 に示す。

図中左上部、ユーザの装着したヘッドマウントディスプレイ (以下 HMD) のジャイロセンサーを用いてユーザの頭部方向を検出し、DataChannel を通じてロボット駆動用端末へ送信する。このジャイロセンサーはヨー、ロール、ピッチの 3 軸を備えており、ユーザが HMD を装着する際のキャリブレーションにより設定した原点に対し、頭部角度の絶対値になるよう加工を施した上で送信する。ロボット駆動用端末でこの値を受信し、同じ角度までロボットのサーボを動かすことで、ロボットをユーザの頭部に追従させる。

視覚に関しては、ロボットの右目と左目に相当するカメラから、それぞれ横 640 ピクセル縦 480 ピクセルの映像を取得して MediaStream で送信し、ユーザ用の操作端末内で 3D レンダリングを行い、HMD で表示を行う。

WebRTC のエンジンには Google Chrome、カメラには See3Cam CU130、HMD には Oculus Rift DK2、ユーザ操作端末はノート PC (Galleria QSF960HGS)、ロボット駆動用端末には Intel NUC 6i5SYH、ロボットのサーボには Herkulex 0201 を用いた。

4. 標準仕様下における実装と通信品質の及ぼす影響

4.1 模擬環境における評価実験

ネットワーク上で流れるパケットは、通信経路上で様々なネットワーク特性により影響を受ける。まず通信は瞬時には届かず、送信側がデータを送出してから受信側で受け取るまでの間に遅延が発生する。この遅延は必ずしも一定ではなく、通信経路上のネットワーク機器に対する負荷により変動 (ジッタ) が発生する。無線ネットワークの電波状況やネットワーク機器にかかる負荷によってはパケットが損失しロスが発生する可能性がある。通信ケーブル上には電気信号や光信号として送出されるが、それらが物理的干渉を受けることにより内容が変質する場合があります、一部の

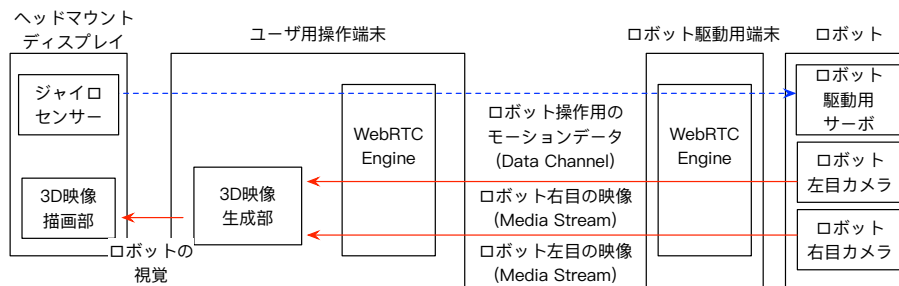


図1: プロトタイプのモジュール構成

データが書き換わってしまうことも想定しなければならない。インターネット上の通信経路は一定ではなく途中で通信経路が変わる可能性があり、それにより後発の packets がより品質のよい経路を通して転送されることで、packets の到着順も前後する可能性がある。途中のネットワーク機器の MTU (Maximum Transmission Unit) 設定を越える大きなサイズの packets は分割されて転送されるので、packets の断片化と再統合を前提にしたプロトコルにする必要がある。回線のバンド幅による全体転送量についても考慮しなければならない。これらが Telexistence アプリケーションに対しどのような影響を与えるのか、評価するための環境を構築した。

インターネット回線を想定したネットワークシミュレータとして動作するソフトウェアルータ (VyOS) を設置し、ローカルセグメントを2つ作成する。操作用端末とロボット駆動用端末を別々のセグメントに接続し、ネットワークシミュレータを経由して通信を行う事で、双方向又は片方向の通信に対してネットワーク特性を設定することが可能である。この評価環境において得られた傾向について、実際に通信に影響を与えたものについて以下に記述する。

4.1.1 遅延・ジッタ

ネットワーク上において遅延が発生すると、操作信号、映像ともに送信から受信までにより多くの時間がかかることになる。IP ネットワークで発生する遅延は、通信負荷やネットワーク機器の負荷、通信経路の変更といった理由により変動するため一定ではなく、ゆらぎ、即ちジッタが発生する。

Telexistence の系全体としては、ユーザが体を動かしてからそれに合わせた信号が送信され、ロボットが駆動し、ロボットの目で観測した映像がユーザへ送られる。通信遅延以外にもセンサによるユーザ動作の取得やロボットの物理駆動による遅延が考えられるが、ネットワーク遅延以外の部分についてはいかなる通信プロトコルにおいても共通のものであるため、本稿の検討からは除外する。

ジッタが発生しない固定遅延においては通信内容は変質せず、ユーザの認識に対して遅れが発生する。開発者を含む数名での評価においては、片道 10ms 以上の遅延が発生するとユーザが遅延を体感し始めた。片道 50ms 以上の通信遅延が発生すると、ほぼ全てのユーザに VR 酔いが発生し、長時間操作できない状態になった。ping コマンドを用いて実網での遅延の計測を行ったが、東京にある我々の研究室

から札幌に設置されたサーバまでの遅延が片道 12ms 程度、東京から米国西海岸に設置されたサーバまでの遅延が片道 75ms 程度であった。海外への通信を想定すると、動作の先読みのようなアプリケーション上の作り込みが必要になると考えられる。

ジッタに関しては、WebRTC の性質から、視聴覚情報とその他の伝送で別々の影響を受ける。視聴覚情報は MediaStream 内で SRTP として転送されるため、packets 自体にタイムスタンプがあり、到着が多少前後しても影響を受けない。DataChannel で送信されるデータについては SCTP で伝送されるためタイムスタンプがなく、アプリケーション側で独自に調整する必要がある。タイムスタンプでの吸収を検討する場合、再生までの間にバッファリングが必要になるが、ジッタが大きくなるにつれてバッファの幅が大きくなり、送信から再生までのタイムラグが大きくなる。実網のジッタについては iPerf コマンドによって計測を行ったが、東京～札幌間で 0.4ms、東京～米国西海岸で 4ms 程度であった。この程度であれば、ジッタがない場合と比較して、有意な操作感覚の悪化は観察されなかったが、ジッタをより大きくすると、自分の動きに対してロボットの応答速度が一定でなくなり、影響が生じることが考えられる。

4.1.2 パケットロス・一部破損

Telexistence アプリケーションにおいては、ロボットとユーザの体の間の同期が求められる。通信 packets が欠損した際に再送を行うのはかえってその同期を乱す元となるため再送は検討せず、SCTP の品質保証オプションは off にして利用する。

まずロボットの動作について述べる。今回の実装においてはロボットの操作信号は絶対座標の連続値として送信したため、操作信号が欠損した際はロボットが沿って動くべき軌跡のうちの特定の点に関する情報が欠損することになり、動作がごちこちなくなる。これを緩和するためには、ロボットを動作させる前に移動平均を計算して動作させるような、アプリケーションの作り込みが必要になると考えられる。

視聴覚情報についてはまず品質の低下が観測された。VP8 や H.264 は圧縮後のデータの一部欠損を想定しており、更に WebRTC には通信品質に応じて自動的にビデオの品質を変更する仕様があるため、1%程度までのロスについては影響は観測されなかった。それ以上のロスが発生した際にはブロックノイズが発生する。

視聴覚情報については、品質の他に映像の同期ズレも観測された。WebRTCにおいて、同一のMediaStreamのトラックとして入力された映像や音声は、SRTPのタイムスタンプで同期される仕様になっているが、それぞれのトラックは別々のパケットで伝送されるため、1つのパケットがロスした際に影響を受けるトラックは1つである。そのためパケットロスによって受ける影響はトラック毎に異なり、残されたパケットから複合された映像の同期ズレの原因となる。

Telexistenceアプリケーションにおいて同期ズレの影響は顕著である。左右の目の同期にズレが発生すると、左右の目で別々の地点を見ることになるため、ユーザの脳内で映像を処理することができなくなる。音声の同期がズレると音声方向の認識が崩れる。更に現在のWebRTCの仕様では触覚情報はタイムスタンプのないDataChannelで送信するため、影響はより顕著である。視覚と触覚の再現の同期がズレると、触覚の認識に影響が出ることが知られており、例えばWuら([10])は、映像に対し力覚のフィードバックが遅れることでバネの硬さの認識が崩れ、より柔らかいバネであるかのように誤認識するという報告を行っている。

4.1.3 パケットの到着順の前後

視聴覚情報については、パケットの到着順が前後してもRTPのシーケンス番号により修復され影響が見られなかった。操作信号については、前述の通りSCTPの品質保証オプションをoffにしたため、到着の前後によりロボットが沿って動作すべき軌跡の点が前後したため、動作がぎこちなくなる挙動が観察された。

5. Telexistence用エンジンの開発

前章での実験結果から、Telexistence用エンジン開発のためにWebRTCに加えるべき改修は2点である。まず、ジッタやパケット到着順の前後にはタイムスタンプやシーケンス番号が有効であり、現状のWebRTCの仕様ではMediaStreamのみがこの恩恵を受けDataChannelは影響を受けることがあげられる。MediaStreamとDataChannelが別々の影響を受けることは、同期の観点から望ましくない。次に、タイムスタンプやシーケンス番号によってネットワークプロトコルレベルで同期が保証されていても、パケットロスは特定のストリームにのみ影響を及ぼすため、アプリケーションレベルでは同期が崩れる場合があることである。以上の点を踏まえ、SRTPの利用をベースに同期が崩れないように工夫を行うのが妥当であると考えられる。

ロスによる同期崩れは別々のパケットに格納されている限り防ぎようがないため、同時刻の入力をチャンクして一つのバイナリペイロードとしてSRTPに格納するエンジンを開発した。

今後はこのエンジンを元にユーザ評価実験を進め、触覚信号も含めた形で、ネットワーク特性が没入感を及ぼす影響を評価し、没入感を担保するために必要なネットワーク要件を定義する予定である。

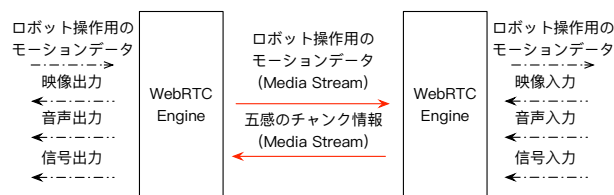


図 2: 開発したエンジンの動作

6. まとめ

本稿は、インターネットを超えてTelexistenceを実現するために必要な通信要件を洗い出すことを目的とし、市中の商用網での動作も検討した上で、テスト用のアプリケーションを作成して評価を行った。インターネット上で想定される遅延やロスといったネットワーク特性がTelexistenceアプリケーションに及ぼす影響を把握するため、ネットワークシミュレータで負荷をかけることで傾向の把握を行い、把握した傾向を元にTelexistenceアプリケーション用のエンジンの開発を行った。このエンジンを用いて詳細なユーザ評価を行い、Telexistenceに求められる通信要件を策定することを目標に継続して検討を進める。

謝辞

WebRTC版のTelexistence評価アプリケーション作成に当たり、一般的なTelexistenceアプリケーションに必要なシステム構成や作り込みについて、技術協力を頂いた慶應義塾大学のムハマド・ヤメン・サライジ氏に感謝します。

参考文献

- [1] Susumu Tachi. *Telexistence*. World Scientific, 2010.
- [2] Charith Lasantha Fernando, Masahiro Furukawa, Tadatashi Kurogi, Sho Kamuro, Katsunari Sato, Kouta Minamizawa, and Susumu Tachi. Design of TELESAR V for transferring bodily consciousness in telexistence. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 5112–5118, 2012.
- [3] 総務省. 平成 27 年版情報通信白書, 2015.
- [4] W3C. WebRTC 1.0: Real-time communication between browsers. <http://dev.w3.org/2011/webrtc/editor/webrtc.html>.
- [5] IETF. Vp8 data format and decoding guide, 2011.
- [6] IETF. Rtp payload format for h.264 video, 2011.
- [7] IETF. The secure real-time transport protocol (srtp), 2004.
- [8] IETF. Stream control transmission protocol, 2007.
- [9] callstats.io. WebRTC - an analytics perspective. <http://www.slideshare.net/callstatsio/webrtc-an-analytics-perspective-by-callstatsio>.
- [10] Bing Wu, Sung Hun Sim, Andinet Enquobahrie, and Ricardo Ortiz. Effects of visual latency on visual-haptic experience of stiffness. In *Quality of Multi-media Experience (QoMEX)*, pp. 1–6, 2015.