

# 遠隔地間車いす操作ソフトウェアの開発と WebRTCによる処理遅延計測

馬場 哲晃<sup>1,a)</sup> 栗原 渉<sup>2</sup> 韓 旭<sup>1</sup> 串山 久美子<sup>1</sup>

**概要:**我々は2021年よりWHEE(ウィー) Projectとして、車いすユーザのための屋外活動支援システムを開発してきた。医療や介護分野等で遠隔地からの支援手法が広まる中、特別なソフト、デバイスを必要としないブラウザベースな開発環境は社会実装における大きなコストダウンを実現する可能性がある。そこで我々はWebRTC技術を用いて、電動車いすWHILLを遠隔操作するためのビデオ、オーディオ、およびその他のデータ通信を統合したマルチモーダル操作システムを開発し、オープンソースとして公開した。本稿では遠隔地からの車いす操作を実現するためのプラットフォームとして開発したTELEWHEELに関する簡単な報告を行い、基盤技術となるWebRTCにおける処理遅延に関する計測結果を示す。計測の結果から、ネットワーク品質が高ければHD画質程度でリアルタイム操作に十分な遅延を処理遅延が実現できることが示された。

## 1. はじめに

障害支援において、ビデオ通信は遠隔地介護の観点から頻繁に利用されている。遠隔からの医療や臨床応用の包括的な評価フレームワークや遠隔医療による影響、情報通信技術へのサポート等[1]に関する報告がされる中、2018年には日本国内において厚生労働省が「オンライン診察指針」を発表し、その後の2023年6月に「オンライン診察その他の遠隔医療の推進に向けた基本方針」が策定されている<sup>\*1</sup>。遠隔医療やオンライン診察等のいわゆるTelemedicineだけでなく、Telecareと呼ばれる支援やケアを遠隔地から実施する手法も広く知られている。高齢者や認知症患者の見守り、遠隔リハビリテーションや視覚・聴覚障害者支援等、遠隔からの障害当事者支援には多用な需要が考えられる。例えば、電話リレーサービス<sup>\*2</sup>やアイコサポート<sup>\*3</sup>はサービスとしてそれぞれ聴覚障害者と視覚障害者の遠隔支援手段として知られている。このような支援アプリケーションの開発には次に述べるような環境構築が必須であり、一定の開発コストが避けられない。

遠隔地間における映像音声伝送サービスには多くのものが存在している。ZOOMやGoogle Meet等は広く知られた映像音声伝送手段であり、アカウントさえ作成すればユーザは手軽にビデオチャットサービスを利用できる。この他RTP(Real-time Transport Protocol), RTMP(Real-Time Messaging Protocol), SRT(Secure Reliable Transport)等の映像音声伝送プロトコルを利用することで、類似機能を開発することも可能である。一方でこのような映像音声伝送サービスを開発する場合、ユーザや開発者側は専用サーバやそれに準ずる環境構築が必要となる。特にインターネット間通信においては様々な知識や開発経験が必要となり、容易にプロトタイピングすることは困難である。

ここまでの問題に対し、サーバを利用しないP2P方式<sup>\*4</sup>による映像音声伝送方式もいくつか存在している。NDI(Network Device Interface)<sup>\*5</sup>は映像制作や作品、広告分野で頻繁に利用されており、低遅延や映像の高圧縮、さらには無圧縮、可逆圧縮等、品質を重視する現場で頻繁に利用されている。この他、WebRTC(Web Real-Time Communication)はブラウザ上で利用可能なP2P型リアルタイム通信手法として広く知られている。WebRTCではPCやスマートフォン端末上のブラウザで動作するため、Google Meet, Slack, discord等様々な商用アプリケーション

<sup>1</sup> Tokyo Metropolitan University, 6-6, Asahigaoka, Hino, Tokyo, 191-0065, Japan

<sup>2</sup> Tokyo University of Technology, 1404-1, Katakura-machi, Hachioji, Tokyo, 192-0982, Japan

<sup>a)</sup> baba@tmu.ac.jp

<sup>\*1</sup> <https://www.pwc.com/jp/ja/knowledge/column/hospital-management/vol17-1.html>

<sup>\*2</sup> <https://www.nftsr.or.jp/>

<sup>\*3</sup> <https://eyecosupport.prime-as.co.jp/>

<sup>\*4</sup> 実際にはシグナリングサーバ等、中継用のサーバが必要になることもあるが、映像伝送そのものはクライアント端末同士で直接実現している点からP2Pとして説明する。

<sup>\*5</sup> 米国 Vizrt (NewTek) 社によって開発された低遅延高品質の映像音声伝送プロトコルの名称

ンにおいて活用されている。インターネット間での通信が前提となっているため、NDI で提供するような無圧縮映像転送や可逆圧縮等は利用できないが、低遅延且つ様々なプラットフォームに容易に映像音声伝送システムを提供できるという点において、開発者において大きなメリットとなる。Nelson ら [9] はこのような背景において、WebRTC を用いることで従来の http ベースな映像配信システムを超低遅延の映像配信サービスへの移行が可能であることを述べている。Jansen ら [10] はダミーネットワーク上における WebRTC ビデオ通話の詳細な性能報告をしている。圧迫されたネットワーク帯域内での品質低下や、極端に制限された帯域幅 (250kbps ダウンリンク / 100kbps アップリンク) では RTT (Round Trip Time) が 2 倍以上増加することを報告している。

遠隔地間における支援技術開発は設計が多様化しており、プロトタイピングのようなシステム要件が定まっていないデザイン作業において、システム設計・構築コストが一般的なインタラクティブアプリケーションと比較して高い。これに対して本稿では WebRTC を利用したアプリケーション開発を通じ、支援技術開発者支援の観点から開発知見をまとめる。本稿では遠隔地間におけるローカル 5G 回線を活用した遠隔地間車いす操作ソフトウェア TELEWHEEL の簡易な報告と、WebRTC 映像配信において生じる処理遅延に関する計測結果を報告する。処理遅延とは、ネットワーク環境とは関係のない動画のエンコード、デコード、画面等の処理によって生じる遅延として、本稿では定義する。

## 2. 関連研究

車いすの遠隔操作に関し、Gunderson らは VR ヘッドセットを利用した探査ローバーの操作技術応用として、車いす遠隔操作を障害支援に応用した事例を報告している [2]。また同様の機能に、YOLO による物体検出、SLAM による自己位置推定を実装した自動車いす操作システム [4] を追加した Telewheelchair や VR ヘッドセットの代わりに 360 半球体ディスプレイを用いた NUNACHair [3] 事例が報告されている。いずれの研究においてもシステムの実装には特別なデバイスや、高計算コストに耐えうる PC を用意する必要がある。これに対して本研究ではより手軽に遠隔地間車いす操作を実現するために必要なシステムをファブリケータブルな形で支援技術開発者に提供する点が先行研究と異なる。また遠隔操作による映像遅延に関する詳細な考察が行われていない。

遠隔地操作による車いすは、当事者の QoL 向上だけでなく、上肢操作に制約がある車いすユーザにとっても応用価値は高い。このようなユーザに対しては車いすの操作 UI の多様化、安全制御の追加、自動運転化等のアプローチが存在している。Mohammed ら [5] は車いす搭乗者の頭部動

作で制御可能な電動車椅子を開発し、連続的な速度や方向制御に関する被験者実験を行っている。この他、眼球運動を電動車椅子操作に応用した事例 [6] や自動運転車いすに関する事例も多い [7][8]。本研究では第一段階として遠隔地間の機能実装を実現するが、安全装置や遠隔地間の協調車いす操作等においてこれら多様なインタフェースが必要となると考えている。

## 3. TELEWHEEL

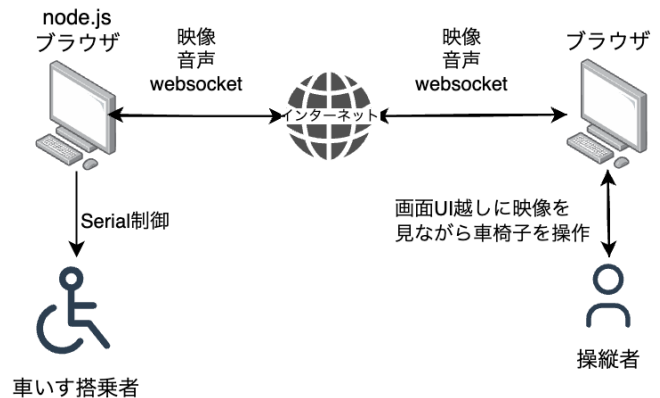


図 1 TELEWHEEL システムの接続概略図。車いす搭乗者、操縦者ともに PC を利用して通信を行う。node.js は WebRTC のシグナリングサーバとして、また操縦者から送信されてくる websocket データに基づき車いす WHILL をシリアル制御する。

本研究では制御用車いすに WHILL 社製 WHILL Model CR (以後 whill) を利用した。WHILL 社の Github\*<sup>6</sup>において制御用仕様をまとめた whill control system protocol specification が公開されているため、このプロトコルに従い Serial 制御にて whill を操作することができる。whill には BLE も搭載されており、WHILL 社が提供するアプリケーションでは BLE 無線制御も可能であるが、BLE 仕様は非公開とされている為、今回の実装では利用しない。WebRTC を利用する点から、本システム TELEWHEEL は javascript で記述し、シグナリングサーバ及び、Serial 制御を受け付けるコマンドとして、websocket を利用する。図 1 に TELEWHEEL システムの接続概要を示す。

本システムを利用した実験は文献 [11] を、体験イベントの様子は動画配信サイトから閲覧できる\*<sup>7</sup>。

## 4. WebRTC の処理遅延計測

本システムでは端末間の通信技術として WebRTC を利用しているが、開発コストやユーザビリティの観点からブラウザで動作するシステムは長所が明確である一方、システムのリアルタイム性については他の開発環境に劣る可能

\*<sup>6</sup> <https://github.com/WHILL>

\*<sup>7</sup> <https://www.youtube.com/watch?v=3pyb3P2dnLk>

## Latency Logger



図 2 計測アプリケーションの画面。送信ビデオより Timestamp の進んでいるのは、計測時に送信ビデオ及び受信ビデオを一時停止しているためである。

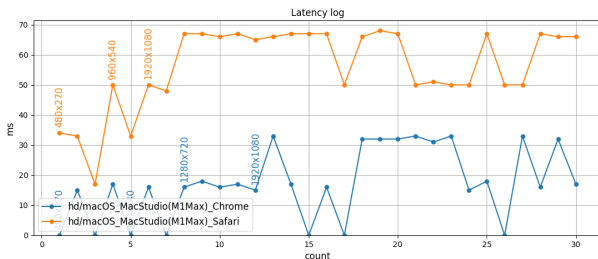


図 3 1920x1080@60fps を目標設定値として、徐々に目的とする品質に近づく様子がわかる。またブラウザによって遅延時間にやや違いが見られる。横軸のカウントは 1 カウント 3 秒の単位となるため、90 秒間のデータとなる。

性がある。TELEWHEEL システムにおいては主に、1. カメラデバイスの入力遅延、2.WebRTC における画像転送に生じる遅延、3. 利用するインターネット環境における遅延の 3 つに大別できる。1,3 に関しては環境依存が大きいが、2 については画像の品質によってどの程度の遅延が生じるかをあらかじめ設計要件に組み込むことができると考え、WebRTC における動画送信、及び受信処理における処理遅延を計測することとした。

図 2 に開発した計測アプリケーションのスクリーンショットを示す。ネットワーク環境の遅延時間を排除するため、計測アプリケーションは単一ブラウザ内で、配信用キャンパスの作成、エンコード、送信、受信、デコード、表示までの手続きを行う。

WebRTC で転送する映像にはタイムスタンプを記載し、受信した映像情報から OCR 処理をすることで処理遅延を計測する。図 2 中央の送信ビデオの数値が受信ビデオの数値よりもタイムスタンプが進んでいる様子がわかる。これは送信側からの映像が受信時にはすでに遅延が生じていることを示す。計測 Start ボタンを押下することで結果がブ

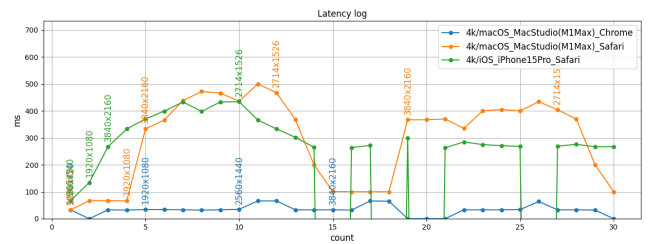


図 4 4K(3840x2160),30fps での遅延計測結果

ラウザ内のテキストエリアに保存される。なお、ORC 処理に一定の時間を有するため、毎フレームごとに ORC 処理は行わず、今回の計測では 3 秒毎に ORC 処理をし、処理遅延を計測している。この他ウェブカメラ動画を送信する機能も付与されているが、今回の計測には利用していない。

ITU-T による G.114 勧告 [12] では通信ネットワークにおける一方の伝送遅延に関するガイドラインを提供しているが、Highly Interactive なタスクである、speech, video conferencing interactive data applications では 150[ms] 以下が理想的な遅延であり、150-400ms が容認可能な遅延範囲として示されている。

本アプリケーションを利用し、macOS (Apple M1 Max) 環境においてそれぞれ Chrome, Safari ブラウザを利用して計測した結果を図 3 に示す。1920x1080, 60fps を送信側の設定値にしており、通信開始時の受信動画や送信動画よりも小さな解像度となっているが<sup>\*8</sup>、これは WebRTC の仕様によるもので、配信を継続することで、ネットワーク環境に応じて徐々に品質を目標品質に近づける動作のためである。一般的なウェブカメラでは 30fps がほとんどであるが、処理遅延をより細かく観察するために今回は 60fps としている。ORC 計測の処理時間の都合上、横軸のカウントは 3 秒おき増加しており、3 では、Firefox においても計測を実施したが、OCR ライブラリ、tesseract.js<sup>\*9</sup>が正しい結果を返さないことから、今回は計測対象ブラウザから除外している。

macOS, Windows, Linux, iOS, Android での計測結果をまとめたグラフを図 5 に示す。ほとんどの計測端末上は 100ms 以下に処理遅延が抑えられているが、Android スマートフォンのみ、大きく処理遅延が生じている。これは CPU スペックによるものであるため、利用端末の演算能力次第では、WebRTC の処理遅延のみで G.114 勧告による容認可能な遅延を超えてしまう可能性があることがわかる。TELEWHEEL システムでは PC を利用しているが、将来的にスマートフォンやタブレット端末を利用する場合、それだけで処理遅延が大きくなる点を十分注意する必要がある。

<sup>\*8</sup> fps 情報は含まれていないが、fps も回線状況に応じて変化している

<sup>\*9</sup> <https://github.com/naptha/tesseract.js>

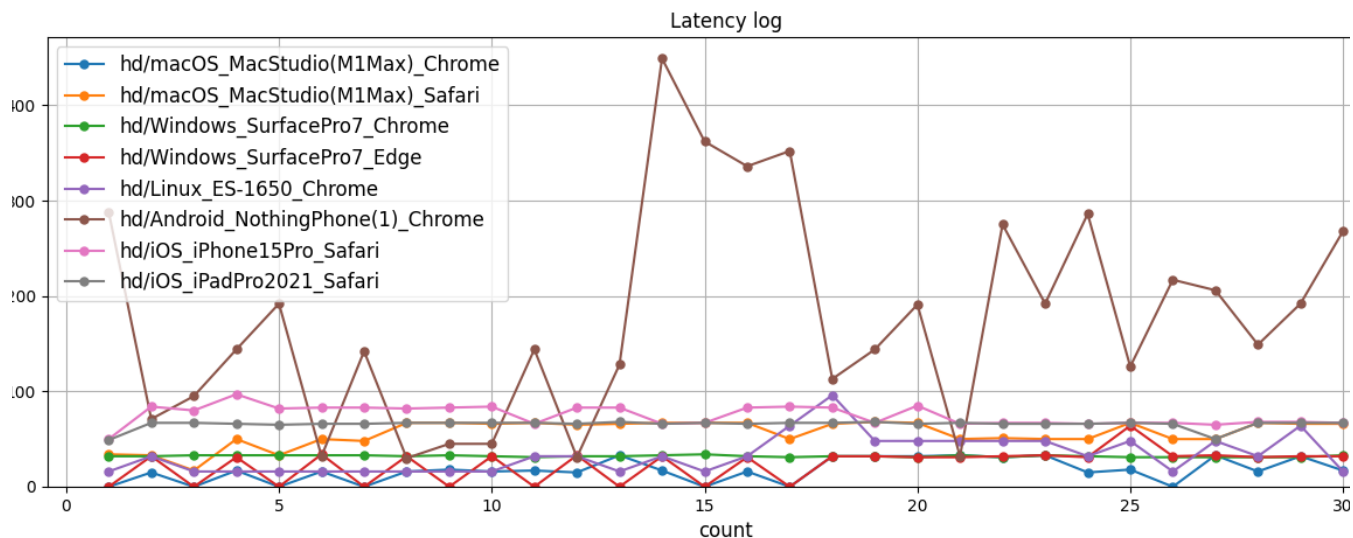


図 5 様々な端末上での処理遅延計測結果。Android スマートフォンのみ、大きく遅延が生じている様子がわかるが、これは他の端末と比較して計測に使用した Android 端末は CPU 処理能力がやや劣るため。

遠隔地からのビデオ支援においては解像度及びに fps の品質が高ければ高いほど好ましい。WebRTC では 4K 動画転送もサポートされているため、3840x2160, 30fps 環境下での処理遅延計測結果を 4 に示す。利用するブラウザや端末によっては 4K 解像度が維持できずに 2K になる他、処理落ちによる計測データ欠損が生じているが、仕様機器や環境によっては適切な処理遅延でビデオ通信を実現できる可能性が示された。

**謝辞** 本研究は「東京都立大学ローカル 5G 環境活用研究支援社会実装型研究（研究代表者：串山久美子）」の助成を受けて実施された。

## 参考文献

- [1] Institute of Medicine (US) Committee on Evaluating Clinical Applications of Telemedicine. Telemedicine: A Guide to Assessing Telecommunications in Health Care. Field MJ, editor. Washington (DC): National Academies Press (US); 1996. PMID: 20845554.
- [2] Gunderson, RW & Smith, Stephen & Abbott, Ben. (1996). Applications of virtual reality technology to wheelchair remote steering systems. Proceedings of the 1st European Conference on Disability, Virtual Reality and Associated Technologies (ECDVRAT'96), Maidenhead, UK. 8-10.
- [3] Luna Takagi, Shio Miyafuji, Jefferson Pardomuan, and Hideki Koike. 2022. LUNACHair: Remote Wheelchair System that Links Up a Remote Caregiver and Wheelchair Surroundings. In Adjunct Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology (UIST '22 Adjunct). Association for Computing Machinery, New York, NY, USA, Article 40, 1-3. <https://doi.org/10.1145/3526114.3558729>
- [4] Satoshi Hashizume, Ippei Suzuki, Kazuki Takazawa, Ryuichiro Sasaki, Yoshikuni Hashimoto, and Yoichi Ochiai. 2017. Telewheelchair: a demonstration of the intelligent electric wheelchair system towards human-machine. In SIGGRAPH Asia 2017 Emerging Technologies (SA '17). Association for Computing Machinery, New York, NY, USA, Article 15, 1. <https://doi.org/10.1145/3132818.3132834>
- [5] Mohammed Kutbi, Xiaoxue Du, Yizhe Chang, Bo Sun, Nikolaos Agadakis, Haoxiang Li, Gang Hua, and Philippos Mordohai. 2020. Usability Studies of an Egocentric Vision-Based Robotic Wheelchair. J. Hum.-Robot Interact. 10, 1, Article 4 (March 2021), 23 pages. <https://doi.org/10.1145/3399434>
- [6] Erik Wästlund, Kay Sponseller, and Ola Pettersson. 2010. What you see is where you go: testing a gaze-driven power wheelchair for individuals with severe multiple disabilities. In Proceedings of the 2010 Symposium on Eye-Tracking Research; Applications (ETRA '10). Association for Computing Machinery, New York, NY, USA, 133-136. <https://doi.org/10.1145/1743666.1743699>
- [7] Harkishan Grewal, Aaron Matthews, Richard Tea, and Kiran George. 2017. LIDAR-based autonomous wheelchair. In 2017 IEEE Sensors Applications Symposium (SAS). IEEE Press, 1-6. <https://doi.org/10.1109/SAS.2017.7894082>
- [8] Baltazar, A.R., Petry, M.R., Silva, M.F. et al. Autonomous wheelchair for patient's transportation on healthcare institutions. SN Appl. Sci. 3, 354 (2021). <https://doi.org/10.1007/s42452-021-04304-1>
- [9] Nelson Francisco, Olie Baumann, Julien Le Tanou, and Richard Fliam. 2024. Ultra-low Latency Video Delivery Over WebRTC Data Channels. In Proceedings of the 3rd Mile-High Video Conference (MHV '24). Association for Computing Machinery, New York, NY, USA, 88-89. <https://doi.org/10.1145/3638036.3640247>
- [10] Bart Jansen, Timothy Goodwin, Varun Gupta, Fernando Kuipers, and Gil Zussman. 2018. Performance Evaluation of WebRTC-based Video Conferencing. SIGMETRICS Perform. Eval. Rev. 45, 3 (December 2017), 56-68. <https://doi.org/10.1145/3199524.3199534>
- [11] 栗原 渉、韓 旭、阪口 紗季、馬場 哲晃、難波 治、串山 久美子、遠隔地からの車椅子移動操作を伴うコミュニケーションシステム、第 28 回バーチャルリアリティ学会、2B1-11, 日本バーチャル・リアリティ学会、2023
- [12] One-way transmission time. ITU-T, G.114 (May2003).