



ウェアラブル・テレプレゼンスシステム “T-Leap” を用いた 東京・渋谷での検証結果と今後の課題

平田智也¹⁾, 瓜生大輔¹⁾, 矢崎武瑠¹⁾, 陳意昕²⁾,

郭聖文³⁾, 劉康宜³⁾, 廖奕雅³⁾, 葛如鈞³⁾, 陳建旭²⁾, 稲見昌彦¹⁾

1) 東京大学 先端科学技術研究センター (〒 153-8904 東京都目黒区駒場 4-6-1, hirata@star.rcast.u-tokyo.ac.jp)

2) 国立成功大学 (台湾台南市大学路 1 号)

3) 国立台北科技大学 (台湾台北市大安区忠孝東路 3 段 1 号)

概要: 本稿では、装着型の 360 度カメラとマイク・スピーカーモジュールを用いたテレプレゼンスシステム “T-Leap” の改良版について報告する。小型・軽量化によるユーザビリティの向上と、通信・画像処理能力の強化を図った T-Leap を、5G 通信エリアが広がる東京・渋谷を舞台に試用した。検証結果の報告と分析を通し、市街地で T-Leap が提供できる体験やサービス、実用に向けた課題や改善点を議論する。

キーワード: テレプレゼンス, ウェアラブル

1. はじめに

著者らは、専用のウェアラブルデバイスを装着した複数の屋外探索者 (Node) と、屋内でコンピュータを操作する遠隔接続者 (Viewer) から成る装着型小型テレプレゼンスシステム “T-Leap” を開発し、屋外での実証実験を通じた研究開発を行っている [1]。私たちは以前、屋外の実用的な場面として買い物や市街地の探索を想定し、台北市内での検証を行った [2]。本検証を通して、Viewer が Node 周辺の地理的情報や現地の言語に詳しい場合には、閲覧する Node の映像が低画質であっても効果的なナビゲーションが成立することが分かった。しかし当時の T-Leap は、通信環境によっては Viewer が視認できる映像が著しく低画質となる点や、Node モジュールの大きさや重さなどにおいてデザイン的な改善の余地があった。

本稿では、渋谷を舞台として、小型・軽量化と通信・画像処理能力の改良によって Node 側のユーザビリティと Viewer 側の臨場感の向上を図った改良版 T-Leap を用いた検証について報告する。検証地として渋谷を選択した理由は主に以下の二点である。一点目は、通信・画像処理能力の改良を施した T-Leap が、一部の地域に高速・低遅延な 5G 通信が整備された環境において、高画質の映像配信とそれによる Viewer 側の高い臨場感を実現するかを調査するためである。二点目は、様々な観光スポットや商業施設、人が密集する環境で、T-Leap が Node や Viewer にどのような体験をもたらすかを調査するためである。さらに本論文では、今回の検証で見られた Viewer と Node の行動や経験を考察し、改良すべき点を議論した上で今後 T-Leap を用いることで実現できる体験やサービスを検討する。

2. T-Leap のコンセプト及びシステムデザイン

2.1 T-Leap のコンセプト

T-Leap システムの概要を図 1 に示す。T-Leap システムでは、屋内からコンピュータを操作する Viewer(一人) と屋外を探索する Node(複数人) という 2 種類の参加者が存在する。そして、Viewer と Node は、WebRTC を用いて繋がれており、映像・音声の通信を行うことが出来る。

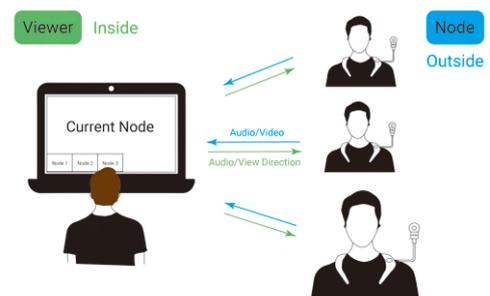


図 1: T-Leap のシステム図

2.1.1 Viewer

本稿では、Viewer となる参加者は V-P[参加者番号] と表記する (参加者 1 が Viewer のときは V-P1)。Viewer が操作するブラウザ上には、画面の左下の Sub-View に接続可能な Node のリアルタイム映像が表示される (図 2)。また、Viewer は Sub-View 上の Node 映像をクリックすることで Node を選択でき、選択された Node の映像が画面中央の Main-View に表示される。そして、Main-View に表示される 360 度映像は、カーソルを用いて回転させることで、任意の方向を見ることが出来る。



図 2: Viewer が用いるユーザーインターフェース

2.1.2 Node

本稿では、Node となる参加者は N-P[参加者番号] と表記する（参加者 1 が Node のときは N-P1）。Node は、360 度カメラ、マイク、スピーカー、LED、コンピュータから構成される Node モジュールを装着して市街地を歩いた。

2.2 今回の検証におけるシステムデザイン

今回の検証において、Viewer と Node の通信は、ブラウザを介さずに WebRTC で映像や音声配信できるネイティブクライアント（株式会社時雨堂の Momo）を用いた。Momo では、ハードウェアエンコーダの利用による高画質な映像の配信ができるため、これを利用して画像処理能力の向上を図った。

また Node モジュールは、360 度カメラ（Insta360 Air）、マイク、スピーカー、Arduino により制御される LED、小型の組込コンピュータ（Jetson Xavier NX）、オーディオインターフェース、スピーカーアンプ、モバイルルーター、モバイルバッテリーから構成される（図 3）。このように、カメラ・スピーカーが埋め込まれたモジュールを首にかけ、その制御機器をショルダーバッグに収納することで小型のモジュールを実現した。

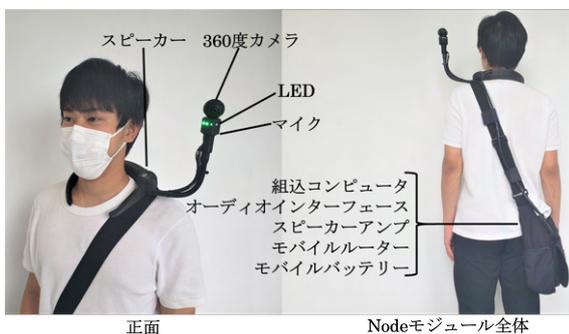


図 3: Node モジュール

3. 検証実験

3.1 検証実験詳細

本検証は、検証地である渋谷の中で、109 周辺と宮下公園の二つのエリアで合計三つの検証（S1～S3）を行った。

いずれの検証も、Viewer は屋内におり、Node(3 人) は実験開始時に渋谷駅周辺のホテルの部屋から屋外に向かい、指定のタスクを行った後に再び部屋に戻ってくるという流

れで行った。この検証の際、システムが停止した時に備えてそれぞれの Node、Viewer には復旧の為のサポーターが一人ずつ付いていた。各タスクの所要時間はいずれも 1 時間程度だった。

3.1.1 参加者

参加者（P1-P10）は 10 名のうち、P1・P2 は日本語を話せない渋谷周辺の地理的知識が乏しい台湾人で、P3 は英語が流暢で渋谷に詳しいコスタリカ人、P4 は英語が流暢で渋谷周辺の知識が乏しいメキシコ人である。そして P6・P7・P9 は渋谷に詳しい日本人で、P5・P8・P10 は渋谷周辺の知識が乏しい日本人である。

3.1.2 検証 1：現地に詳しい Node による Viewer のガイド

検証 1（S1）は、109 周辺に詳しい Node が Viewer の要望を聞き、現地に点在する商業施設でお土産を購入したり観光スポットをガイドするというシナリオである（図 4）。S1 では 109 周辺の地理的知識に乏しい台湾人である P1 が Viewer として、109 周辺に詳しい P3・P6・P7 が Node として参加し、英語を用いたコミュニケーションが行われた。

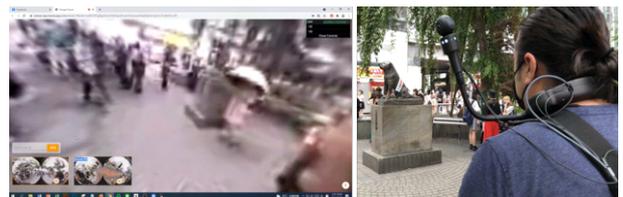


図 4: 現地に詳しい Node による観光ガイド

3.1.3 検証 2：現地に詳しい Viewer による Node のガイド

検証 2（S2）は現地に詳しい Viewer が Node の要望を聞き、その要望に応えるために 109 周辺にて各 Node をガイドするというシナリオである（図 5）。S2 では 109 周辺に詳しい P9 が Viewer として、P5・P8・P10 が Node として参加した。Viewer、Node はいずれも日本人であり、日本語を用いたコミュニケーションが行われた。

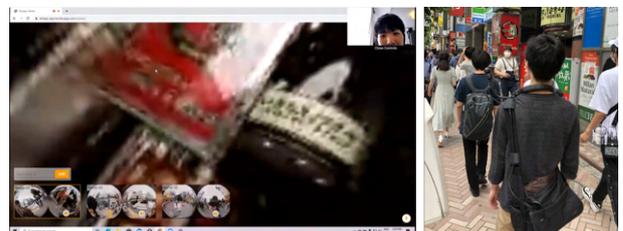


図 5: Viewer によるラーメン店へのガイド

3.1.4 検証 3：共同作業

検証 3（S3）は、Viewer と Node のどちらも現地に詳しくない宮下公園において、事前に指定したオブジェの写真を手掛かりに共同でオブジェを見つけるというシナリオである（図 6）。S3 では P2 が Viewer として、P4・P9・P10 が Node として参加し、Viewer と Node の間のコミュニケーションでは英語が用いられた。

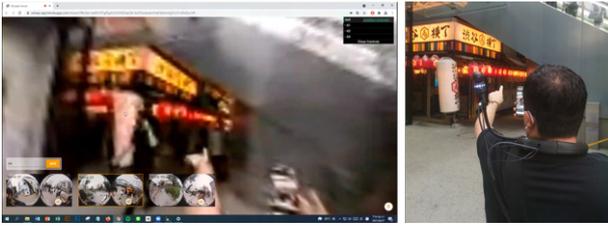


図 6: オブジェの探索

3.2 記録方法

検証は、以下の三つの方法で記録した。一つ目は、Viewer のラップトップコンピュータの画面録画で、検証が続く限り録画が行われた。また、各検証の録画時間はそれぞれ一時間程度であった。二つ目は、三脚で固定した iPhone XR を用いた Viewer の録画 (S2 のみ) であり、これも同じく検証が続く限り録画が行われた。三つ目は、Node に付き添うサポーターによる iPhone XR を用いた Node の撮影であり、Viewer と Node が話し合いながらガイドや買い物に取り組む場面を適宜撮影した。

4. 検証結果抜粋

ここでは、S1～S3 で見られた検証結果から抜粋して、いくつかの特徴的なふるまいを紹介する。

4.1 現地に詳しい Node による観光スポットのガイド

現地に詳しい Node が、人混みの多い環境で観光スポットをガイドする場面で、個人的なツアーガイドをしてもらうような特別な経験を Viewer に提供できることが分かった。例えば、N-P3 が忠犬ハチ公の銅像から 5m 程の位置でその紹介を行った時、V-P1 はハチ公像そのものを認識することはできなかったが、その観光スポットがどのような場所かは理解できた (図 4)。

V-P1 は検証後のインタビューで、ハチ公像を認識することができなかった理由として「映像の遅延や雑音が発生する中で、英語で話す Node との会話に集中するのが精一杯で、映像を見渡す余裕が無かった。」と語っている。一方で V-P1 は、N-P3 によるガイドの経験について、「現地のツアーガイドが自分を観光に連れて行ってきているようだった。特別に興味深い経験だったので機会があればまた参加したい。」と語った。

4.2 Node が騒音の大きい環境にいる場面

渋谷のようにあらゆる場所で大きな騒音に見舞われる環境において、Node が自身で音量調整できないことで Viewer とのやり取りが遮断される場面が見受けられた。一方で、付き添うサポーターが Node の音量調整をすることで問題なく音声コミュニケーションが行われる場面があった。

例えば S1 では、N-P3 がハチ公前からスクランブル交差点へ向かったところ、広告の大きな音声によって V-P1 の音声聞き取れなくなる場面があった。一方、N-P3 に付いていたサポーターが音声の調整を行ったところ、スクランブル交差点と比較してより大きな騒音のある地域でも V-P1 の音声を完全に認識しながら要望に応えるようにガイドを

遂行できた。

4.3 360 度映像を活用した状況把握

T-Leap の特徴のひとつに Node から送られる全方位映像を Viewer が任意の角度を選択して確認できる機能がある。いくつかの場面において、Viewer が本機能を効果的に活用した。例えば、S2 において V-P9 が N-P5 をガイドする際、N-P5 の周りを 360 度見渡すことで、現在位置と進むべき方向を決定していた。また、目的の店舗に辿り着いた N-P5 が商品の選択を決めかねている時に、V-P9 が商品棚を見渡しおすすめの商品を紹介する場面があった。

いずれの場面においても Viewer 側の画質は不十分であり、即時に状況を把握することは難しかった。しかし V-P9 は、画面をスクロールし 360 度を見渡すことにより、その場所についての自身の記憶と照らし合わせながら、Node 周辺の状況を把握していた。

一方で、仮に十分な画質が担保されていたとしても、特定の画角のみから Node の位置を把握することが困難な場合がある。例えば、S2 において、3 人の Node がいずれもスクランブル交差点周辺に位置していた際、多くの歩道が多方向に位置する交差点の中で、特定の画角のみから各 Node の位置を把握することは困難であった。そこで、V-P9 は本機能を活かすことで即座に各人の位置を把握し、適切な指示を送った。全方位映像を用いたテレプレゼンス環境が活きる状況のひとつである。

4.4 Viewer によるガイドすべき Node の選択

S2 では、Viewer が複数の Node をガイドをする中で、ガイドを必要とする Node を集中的にガイドする場面があった。例えば、N-P8 をラーメン店へ誘導した V-P9 は、その後、時折食事中の N-P8 にアクセスして様子を見守りつつ、他の二つの Node (N-P5、N-P10) に労力を割いた。また、N-P5、N-P10 の誘導においても、まず N-P10 に所望の店舗までの道のりを伝えた後、N-P10 がそこまで移動している間に、N-P5 の様子を見守る場面があった。これらの場面は、屋外の実用的な用途において、T-Leap を用いた一対複数のガイドの可能性を示唆している。

5. 考察と今後の展望

5.1 課題と改善のための設計案

5.1.1 映像の画質

今回の検証では、通信・画像処理能力を強化した T-Leap を用いたにも関わらず、画質に関して前回の検証から大きな変化は見られなかった。ここでは、この原因について考察する。

まず一点目として、Viewer 側の通信環境が不十分であったことが挙げられる。S1 と S3 の Viewer を担当した台湾側の通信速度はダウンロードが 50Mbps 前後であり、S2 の Viewer については 100Mbps 程度であった。

次に、二点目として、機器からの熱の発生への対策不足が挙げられる。今回の Node モジュールでは、ショルダーバッグの中に Jetson Xavier NX やモバイルルーターなど

の発熱する機材を収納したことで、バッグの内部に熱が留まった。これによって、一部のルーターでは発熱により性能が低下するような挙動が確認された。これについては、ペルチェ素子や保冷剤を用いた冷却や、バッグ内の換気性能を上げることで熱対策を行うことができると考えられる。

最後に三点目として、渋谷における人混みが挙げられる。検証を通して、検証開始時のホテルの部屋では映像が安定していた一方で、スクランブル交差点や渋谷ハチ公前など、人混みの多い場所では通信が途切れたり映像の画質が低下する場面が散見された。このことから、電波の発する通信機器や人が極端に密集することで、T-Leap の通信に何らかの悪影響を及ぼすことが考えられる。

5.1.2 音量調整に関するユーザビリティの向上

本検証では、4.2 で述べたように、Node が騒音の大きい地域に差し掛かった場合、音量を調整する機能が必要であることが示唆された。その一方で、Node に付いていたサポーターが音声の調整を行ったところ、より大きな騒音のある地域においても V-P1 の音声を完全に認識しながら要望に応えるようにガイドを遂行することが分かった。従って、例えばショルダーバッグ内部の配線を最適化した後、音量を調整する機器をショルダーバッグの外部に出すように T-Leap をデザインすれば、Node が気軽に音量を調整できるようになり、騒音が多い繁華街においても Node と Viewer が問題なく音声コミュニケーションを行えるだろう。

5.1.3 Node から Viewer に配信される映像の確認

現在の T-Leap では、Node は自身が Viewer に配信する映像を確認することができない。これによって、特に Node と Viewer のお互いが現地の情報に乏しい場合に問題が生じた。S3 において、事前に指定されたオブジェを見つけた N-P4 が V-P2 に報告している際、配信される映像が低画質であったため、V-P2 はオブジェを認識できなかった。一方、N-P4 は、自分の配信する映像の画質を知る術がなかったため、目の前にあるオブジェをなぜ V-P2 が認識できないのか理解できなかった (図 6)。この Node と Viewer のふるまいから、Node が自身から Viewer へ配信される映像を確認できる機能の必要性が示唆された。

現在の Viewer と Node 間の通信は P2P である為、Viewer として一人だけが WebRTC にアクセスできる。従って、参加者の間で zoom を立ち上げ Viewer が画面共有を行えば、現在の仕様のまま Viewer 側から見た Node の配信の様子が伺える。

5.1.4 Node の視線方向の提示

本検証で用いた T-Leap は、Node 側の視線方向は Viewer 側へ提示していない。従って、Node がある方向へ注意を向けようとしても Viewer はどちらの方向を向けばよいか困惑する場面があった。後のインタビューでは、S2 で Viewer を担当した P9 が「Node がこっちにいても、どちらの方向か分かりにくかった。」と話しており、Node の視線方向を Viewer へ提示する機能の必要性が明らかになった。

例えば Viewer の視点を、ボタン一つで Node に対して

正面方向に戻すような機能を加えれば、Node がどちらの方向を見ながら話しているかを確認することが容易になるだろう。

5.2 T-Leap により実現されるサービス

5.2.1 現地に詳しい Node による繁華街での遠隔ガイド

台北での検証では見られなかった現象として、S1 において、現地に詳しい Node が、人混みにおいてガイドを試みた際、雑音による音声対話に困難が生じる一方で、Viewer が現地に詳しくなくとも観光スポットについて理解が深められた。5.1.2 で述べたとおり、音量調整に関する機能の改良により、渋谷のような騒音過多な繁華街においても Node と Viewer が問題なく音声コミュニケーションを行える可能性が示唆された。さらに 5.1.3 で言及したとおり、Viewer に配信される映像を Node が確認できれば、Node が Viewer をガイドするためのコミュニケーションをより円滑化できると考えられる。今後、これらの機能の付加により、Viewer による Node のガイドのみならず、Node による Viewer のガイドの環境を整えたい。

6. むすび

本稿では小型・軽量化と、通信・画像処理能力の改良を施した T-Leap を用い、一部の地域に 5G 環境が整備され、多くの商業施設が密集する渋谷を舞台として検証について報告した。残念ながら、当初想定していた画質の改善は実現できなかったほか、音量調整に関するユーザビリティなどの改善点を確認した。一方で、現地に詳しい Node による観光スポットのガイド、Viewer による 360 度映像の活用や状況に応じたガイドすべき Node の選択など、T-Leap がもたらす特有の経験を確認した。今後、本検証で得られた改善項目を反映し類似する条件下における反復的検証を進めると共に、T-Leap の特徴が生かされる他の用途についても検討を重ねていく。

謝辞 本研究は JST ERATO (JPMJER1701) および TIS 株式会社からの助成を受けたものである。

参考文献

- [1] 矢崎武瑠, 眞鍋美祈, 瓜生大輔, 船津武志, 泉原厚史, 陳意昕, 廖奕雅, 劉康宜, 葛如鈞, 檜山敦, 稲見昌彦: 「ウェアラブル・テレプレゼンスシステム “T-Leap” を用いた市街地でのナビゲーション検証結果と応用可能性の検討」, 第 25 回日本バーチャルリアリティ学会大会論文集 (2020)
- [2] M. Manabe, D. Uriu, T. Funatsu, A. Izumihara, T. Yazaki, I. Chen, Y. Liao, K. Liu, J. Ko, Z. Kashino, A. Hiyama, M. Inami: “Exploring in the city with your personal guide: Design and user study of T-Leap, a telepresence system”, in 19th International Conference on Mobile and Ubiquitous Multimedia, pp.96-106 (2020)