



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

矢崎武瑠¹⁾, 眞鍋美祈¹⁾, 瓜生大輔¹⁾, 船津武志¹⁾, 泉原厚史¹⁾,

陳意昕²⁾, 廖奕雅³⁾, 劉康宜³⁾, 葛如鈞³⁾, 檜山敦¹⁾, 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, takeru.yazaki@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: 本稿では屋内の遠隔接続者 (Viewer) と 360 度カメラとマイク・スピーカーモジュールを装着した複数の屋外探索者 (Nodes) によるテレプレゼンスシステム “T-Leap” について報告する. 台湾での検証において, Viewer が現地の情報に詳しい場合, Node を効果的に支援できることが判った. 検証で得られた知見から, 今後 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, 遠隔コミュニケーション, デザインリサーチ

1. はじめに

テレプレゼンス研究領域では, 遠隔会議システムに代表される離れた部屋同士を接続する試み [1] や一対一のモバイルコミュニケーション [2] などが検討されてきた. あるいは人型ロボットを用いてコミュニケーションを円滑にする研究も存在する [3] が, 高価で大型なデバイスは, 日常的に使用可能な遠隔コミュニケーションに利用することは難しい.

従来の先行事例をふまえて, 私たちは買い物や旅行などの様々な場所を複数の人が行き来する状況におけるテレプレゼンス経験に着目する. 具体的には, 一対複数の遠隔コミュニケーションが可能となるシステムと, 市街地で使用できるウェアラブルデバイスの開発を行っている. デバイスを装着した複数の人が屋外を練り歩くさまを遠隔接続者に共有可能な状態で, 相互にコミュニケーションを取りながら街歩きをしたり, ショッピングをしたりする. このような協同作業のなかにテレプレゼンスシステムを介した新たなコミュニケーションの可能性があると考えた.

本研究ではこのような経験を実現するために, 屋内でコンピュータを操作する遠隔接続者 (Viewer) と, 360 度カメラとマイク・スピーカーモジュールを装着した複数の屋外探索者 (Nodes) による遠隔コミュニケーションを可能とする装着型小型テレプレゼンスシステム “T-Leap” を提案する (図 1). また, T-Leap が特定の状況下でどのように作用し, Viewer や Node にどのような体験をもたらすかを確かめるための検証を台北市内で行った. 地理的知識や母国語が異なる参加者を手配し, 市街地での買い物, 公園探索, ガイドなどといった T-Leap が活用されうる場面を想定した. さらに, 検証中に見られた Viewer 及び Nodes の行動や経験を考察し, 解決すべき課題や改善点を議論したうえで, 今後 T-Leap を用いることで実現できる経験やサービスを検討する.



図 1: 屋外の Node (左) と屋内の Viewer (右)

2. 関連研究

2.1 遠隔コミュニケーションを用いた体験の共有

屋内外にいる人同士の遠隔コミュニケーションはトレジャーハント [4] や買い物, 街歩き [5] などの体験共有に広く応用されている. これらの研究では「相互に見えている映像」の共有は両者の一体感を増加させ, 見たいもののリクエストや質問などのインタラクションを誘発するとされる. 本研究は地理的・言語的知識が異なる人たちが参加者となった検証により, 複数の視点を切り替えられる Viewer と屋外を探索する Nodes 同士のインタラクションを報告し, 遠隔コミュニケーション研究領域に新たな知見を提供する.

2.2 360 度カメラを利用した遠隔コミュニケーション

映像を介したコミュニケーションの課題のひとつにカメラワークがある. 現地 (屋外) 探索者はカメラを安定させるように要求され, 遠隔接続者の体験の質は探索者のカメラワークに依存する [6]. この問題に対処するために本研究では 360 度カメラを使用する. 全方位映像の視聴者は見たい方向を任意に切り替えられるため, 探索者にカメラアングルの切り替えを求める必要がない. さらに, 360 度カメラ映像を用いたコミュニケーションは遠隔接続者に高い没入感を与え [7], また, 共同作業を援助する [8]. T-Leap では, 専用のウェブアプリケーションを用いて, 屋内の遠隔

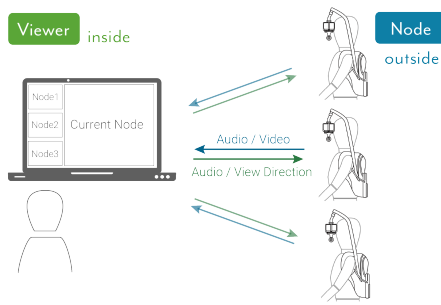


図 2: T-Leap のシステム図

接続者が接続したい屋外探索者の 360 度カメラ映像を選択し、切り替えられる。

2.3 テレプレゼンスとトレイグジスタンス

離れた場所にいる人の存在感を伝えるために、ジェスチャー [9] や表情 [10] を伝えるテレプレゼンスやトレイグジスタンスの研究がある。しかし、トレイグジスタンスロボットの多くは大型で、機動性やウェアラブル性の観点から、屋外での使用には課題が残る。T-Leap は実験参加者が機動性を保って屋外を探索できるように、後述するような小型の装着型システムとして設計した。

3. 検証実験

T-Leap が買い物、公園内探索、ガイドという目的で使用される状況を想定し、屋内から遠隔で接続する人と屋外を移動する人がどのようなふるまいを見せるか観察した。

3.1 システムデザイン

T-Leap システムを図 2 に示す。T-Leap では屋内から操作する Viewer と屋外を探索する Node(s) という 2 種類の参加者が存在し、Viewer と Viewer が選択した 1 つの Node と接続する。接続先は Viewer が任意に切り替えられる。

3.1.1 Viewer

本研究では 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)。本実験では Viewer は屋内で、ラップトップコンピュータのインターネットブラウザから専用のウェブアプリケーションを用いてシステムを操作する (図 3)。Viewer は画面左側の Sub-View にて全ての Nodes の 360 度カメラ映像を見ることが可能であり、接続する Node を画面右上のボタンで自由に切り替え、画面中央の Main-View に表示することができる。見たい方向の切り替え、音声の取得及び会話は接続されている 1 つの Node に対してのみ可能である。

3.1.2 Nodes

本研究では 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 モジュールの構成を図 3 に示す。Node モジュールは 360 度カメラ (Insta360 Air)、マイク・スピーカーモジュール (EMeet M0)、Arduino により制御される LED、そしてタブレット端末 (Surface Go) から成る。Node は LED の点灯場所により Viewer の視線方向を知ることができる。また、Node は Viewer に接続先として選択されているときのみ Viewer

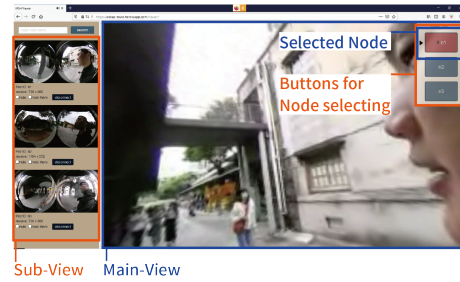


図 3: ブラウザ上でのユーザーインターフェース

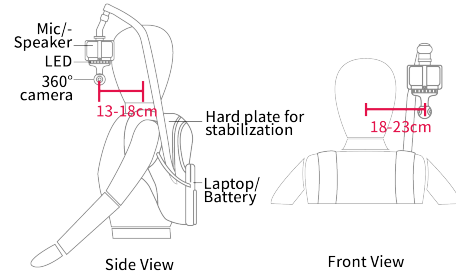


図 4: Node モジュール

と会話することが可能で、Node から Viewer を呼び出すことはできない。

3.2 検証実験詳細

本検証では、Viewer, Nodes, 歩行者などの現地の人々の間に言語的・地理的知識に差異が存在するように参加者を日本人と台湾在住者で構成し、検証中のコミュニケーションは英語で行った。検証設定は実際に起こりうる状況を想定した 3 パターンを用意し、いずれも Viewer は屋内の部屋にあり、Nodes は実験開始時に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ある。いずれの検証も所要時間は 1 時間程度だった。

3.2.1 参加者

参加者 (P1-P7) は 7 名のうち、P1, P2, P3 は中国語を話せない台北市内の知識が乏しい日本人で、P4, P5, P6, P7 は中国語が母国語で現地に詳しい台湾在住者である。

3.2.2 実験 1: 買い物

検証 1 (S1) は Viewer が現地に詳しい Node にお土産を買ってくるように頼むというシナリオである。観光旅行中によく見られる状況におけるふるまいを観察することが目的である。検証 1 には P2 が Viewer として、P5, P6, P7 が Nodes として参加した。Node は中国語を話すことができ、台北にも詳しいため現地の人とコミュニケーションを取ることができる。

3.2.3 検証 2: 共同作業

検証 2 (S2) は Viewer が Node と協力して公園内にある 3 つのオブジェを見つけるというシナリオである。Viewer と Nodes 両方にとって初めて訪れる場所での探索を観察することが目的である。検証 2 には P6 が Viewer として、P1, P3, P7 が Nodes として参加した。V-P6 は中国語を話すことができるが対象の公園には行ったことがない。N-P1,



図 5: Viewer の指示でタクシーを拾う Node

N-P3 は中国語を話せず、現地に詳しくもない。N-P7 は S2 においては英語のみを使用した。公園に行くための交通手段としてタクシーを用いた。

3.2.4 検証 3: ガイド

検証 3 (S3) は現地に詳しい Viewer のガイダンスにより、Node が欲しいものを買うというシナリオである。検証 3 には P5 が Viewer として、P1, P2, P4 が Nodes として参加した。V-P5 は買い物をするエリアについて詳しいが、Nodes は詳しくない。買い物の目的を定めるために Viewer が事前に欲しいものを尋ねたところ、N-P1 は HDMI ケーブル、N-P2 はスマートフォンケース、N-P4 はクリスマスプレゼント用の何かと答えた。

3.2.5 記録方法

検証は Viewer のラップトップコンピュータの画面録画、iPhone XR による検証の様子撮影、Nodes の前方に取り付けた GoPro 8 による録画という 3 つの方法で記録した。

4. 検証結果抜粋

現地情報に詳しい Viewer が Node をサポートする場面と Viewer と Node の協力が必要な場面において特徴的なふるまいがみられた。

4.1 現地情報をもった Viewer のガイダンス

本項では Viewer が地理的知識・言語的知識を持っている場合の Viewer と Nodes のふるまいについて述べる。

4.1.1 Viewer が現地の地理的知識をもっている場合

Viewer が現地の地理情報(店、道など)を知っているとき、Nodes をサポートする場面が何度かみられた。例えば、S2 において、N-P1 はタクシーを拾う場所がわからなかったが、現地に詳しい V-P6 が現地のルールや適切な場所を指示していた(図 5)。また、S3 において、買い物をすべき店がどこにあるか知っている V-P5 が目的地までの行き方や所要時間を計算し、ガイドすることで、その土地に行ったことがなく、現地の言葉も話せない N-P1 が 11 分以内に欲しいものを買うことができた。

4.1.2 Viewer が現地の言語的知識をもっている場合

Viewer が現地の言語を知っているとき、Nodes をサポートする場面も何度かみられた。例えば、S2 において、N-P4 は V-P6 を介して現地の人とコミュニケーションをとり、目的の場所に行くことができた。Viewer も Node もオブジェへの行き方を知らなかったが、Node が公園内の歩行者を見

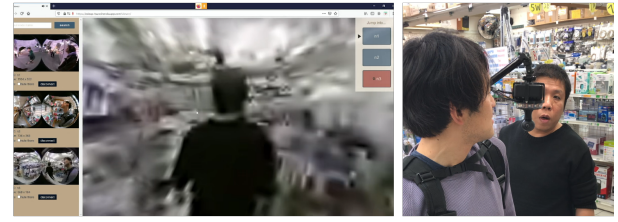


図 6: Viewer による買い物サポート

つけ、Viewer が話しかけるという方法でタスクをこなすことができた。また、S3 において V-P5 が店員に質問することで、N-P1 は欲しい商品を買うことができた(図 6)。

4.2 協力作業

与えられたタスクに明確なゴールがないとき、Viewer と Node が協力する場面が何度かみられた。例えば、S1 において、Viewer と Nodes は買い物をする場所に到着するまで具体的に買うお土産を決めていなかったが、店に到着後、N-P6 はポストカードを買うことを勧め、V-P2 は好みのものを見つけた。また、S3 において、V-P5 は店に何が売っているか知らなかったが、まるで一緒に買い物をしているかのように N-P4 に様々なものを勧めていた。

5. 考察

5.1 課題

5.1.1 ネットワーク環境と映像の解像度

ネットワーク環境次第では Node から送られる 360 度カメラ映像の解像度が大きく低下してしまい(図 6)、Viewer によって状況の理解度に差異が出た。S1 では、現地に詳しくない V-P2 は Node がどこにいるのか 368 × 184 以下の解像度では理解できなかった。一方、S3 では、現地に詳しい V-P5 は低解像度の映像からでもどこにいるのか理解できた。現状の T-Leap を旅行やガイドに使用する場合、Viewer となる人が現地に詳しくない場合、低解像度映像では状況の理解や没入感、臨場感が損なわれてしまう。

5.1.2 Node のコミュニケーションの制限

Viewer は接続している 1 つの Node とのみ会話することができるという T-leap の特徴は Node の行動を制限してしまう。Node は他の Nodes とコミュニケーションをとれないため、Viewer が他の Node に接続しているとき、待機しなければならない。また Node は他の Viewer と Nodes が何を話し、何をしているのか知ることができない。

5.1.3 LED による Viewer の視線方向表示

Viewer の視線方向が表示される LED を取り付けたが、これが有効に作用した場面はみられなかった。しかし、この機能は複数の Viewer が存在する場合効果が表れると考えられる。この場合、Node は声だけで Viewer が誰なのか、何人いるのか判別しにくい、LED 表示により複数の Viewer の存在を認識できると予想される。

5.2 T-Leap が適する状況

Nodes は Viewer に接続されているときだけ会話すればよいのに対し、Viewer はほとんどの時間いずれかの Node

と会話する必要がある。S3のようにViewerがその土地に詳しく、円滑にガイドすることができるならば、Nodeは短時間でタスクを遂行できる。逆にS1のようにViewerが現地に詳しくないならば、ViewerはNodeとの連携に戸惑い、大きな負担となる。これらの特徴と映像の解像度低下時のViewerのふるまいから、現状版のT-Leapでは旅行ガイドや博物館の案内人などのガイドに慣れた人がViewerになると効果的であると考えられる。

5.3 T-Leapにより実現されうるサービス

5.3.1 ウェアラブル 360 度ライブ配信

今後、屋外で使用可能な高速通信ネットワークを利用できる場合、NodeがViewerへのウェアラブルライブ配信システムが実現する。本検証ではViewerは一人としたが、複数人への配信も可能だ。比較的安価なデバイスで構成されるT-Leapは、多くの人に普及し、複数対複数の360度映像ライブ配信サービスを提供する可能性を持つ。S3において、N-P4は目的のものがある場所を一度通り過ぎたが、Nodeの360度映像からそれに気付いたV-P5からの指摘により再び戻ることができた。このように視聴者は360度映像から得られる発見やリアクションを配信者に伝えることができる。また、360度映像の見たい方向を自在に変えられる機能により、視聴者はまるで自分が配信者と一緒にいるかのような感覚を味わうことができる。本検証においてもS1後にV-P2は「お土産を自分で買ったように感じた」と語った。

5.3.2 360 度ガイドの可能性

本検証で得られた「Viewerがガイド先の情報に詳しい場合、T-Leapは効果的である」という知見から、一対複数のコミュニケーションを提供するオンラインガイドサービスへの応用が挙げられる。S3のように、ガイド先の現地情報に詳しいViewerが、現システムよりもNodesを簡単に切り替えられるユーザーインターフェースを用いることにより、より多くのNodesを短時間でガイドできる。一方、ガイドが一度に会話できるNodeは一人だけであるため、各Nodeにあわせた柔軟な対応を行えるほか、Node同士は接続されないため、プライバシーも保護される。また、S3でわかったように「プロのガイド」は低解像度の映像からも状況を把握できるため、現状のネットワーク環境に依存した映像解像度もさほど問題にならない。

6. むすび

本研究では、一対複数の小型テレプレゼンスシステム“T-Leap”を提案し、地理的・言語的知識量に差異がある参加者による市街地での検証を行った。その結果、ViewerとNodeの協力がみられ、特にT-Leapはガイド役のViewerがNodesがいる現地の情報に詳しい場合、効果的に作用することがわかった。また検証で得られた知見から、T-Leapにより実現されうるサービスを構想した。

謝辞 本研究はJST ERATO (JPMJER1701)、JST 国際科学技術協力基盤整備事業、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] Sara A Bly, Steve R Harrison, and Susan Irwin. 1993. Media spaces: bringing people together in a video, audio, and computing environment. *Commun. ACM*, 28–46.
- [2] Sven Kratz, Don Kimber, Weiqing Su, Gwen Gordon, and Don Severns. 2014. Polly: “Being There” through the Parrot and a Guide. In *Proceedings of the 16th International Conference on Human-Computer Interaction with MobileHCI’ 14*. ACM, 625–630.
- [3] Sigurdur O Adalgeirsson and Cynthia Breazeal. 2010. MeBot: A Robotic Platform for Socially Embodied Presence. *Proceedings of the 5th ACM/IEEE International Conference on HRI’ 10*. IEEE Press, 15–22.
- [4] Jason Procyk, Carman Neustaedter, Carolyn Pang, Anthony Tang, and Tejinder K. Judge. 2014. Exploring Video Streaming in Public Settings: Shared Geocaching over Distance Using Mobile Video Chat. *Proceedings of CHI’ 14*. ACM, 2163–2172.
- [5] Sven Kratz, Don Kimber, Weiqing Su, Gwen Gordon, and Don Severns. 2014. Polly: “Being There” through the Parrot and a Guide. *Proceedings of MobileHCI’ 14*. ACM, 625–630.
- [6] Brennan Jones, Anna Witcraft, Scott Bateman, Carman Neustaedter, and Anthony Tang. 2015. Mechanics of Camera Work in Mobile Video Collaboration. *Proceedings of the 33rd Annual ACM Conference on CHI’ 15*. ACM, 957–966.
- [7] Yasamin Heshmat, Brennan Jones, Xiaoxuan Xiong, Carman Neustaedter, Anthony Tang, Bernhard E. Riecke, and Lillian Yang. 2018. Geocaching with a Beam: Shared Outdoor Activities through a Telepresence Robot with 360 Degree Viewing. *Proceedings of CHI’ 18*. ACM, 13.
- [8] S. Kasahara and J. Rekimoto. 2015. JackIn Head: Immersive Visual Telepresence System with Omnidirectional Wearable Camera for Remote Collaboration. *Proceedings of the 21st ACM Symposium on VRST’ 15*. ACM, 217–225.
- [9] Y. Tsumaki, F. Ono, T. Tsukuda. 2012. The 20-DOF miniature humanoid MH-2: A wearable communication system. *2012 IEEE International Conference on Robotics and Automation*, 3930–3935.
- [10] H. Tobita. 2017. Gutsy-Avatar: Computational Assimilation for Advanced Communication and Collaboration. *2017 First IEEE International Conference on Robotic Computing*, 8–13.