



複数の 360 度カメラ間を移動できる テレプレゼンスシステムの開発

泉原厚史¹⁾, 檜山敦²⁾, 稲見昌彦²⁾

1) 東京大学 工学系研究科 (〒 153-8904 東京都目黒区駒場 4-6-1 izumihara@star.rcast.u-tokyo.ac.jp)

2) 東京大学 先端科学技術研究センター

概要: ロボットを用いないテレプレゼンスサービスでは移動ができないため、インフォーマルコミュニケーションに制約がある。我々は実運用可能な機材構成でこの問題を解決すべく、一部屋に複数設置した全天周カメラの映像を伝送し、3次元空間に配置した上でその間を移動しながら同じ領域が見えている二つの映像間をクロスフェードすることで、カメラの設置位置間を自分の身体が素早く移動したような体験の出来るシステムを構築した。

キーワード: Telepresence, VR, Cameras

1. はじめに

通常のビデオチャットをはじめとして、高臨場感通信やテレグジスタンスなど多くのテレプレゼンス関連研究が行われてきた。しかし今日の社会を見てみると、企業等の会議や親密な家族友人間でビデオチャットが行われていることを除けば、これらの技術が日常的に利用されているとはいえない。

本研究で提案するシステムでは、全天球カメラ映像を伝送し、遠隔地の利用者はヘッドマウントディスプレイまたは通常のディスプレイをもちいて現地の映像をリアルタイムに見ながら会話することが出来る。さらに図 1 に示すように、全天球カメラのノードを室内に複数配置することで、利用者は視点間を素早く簡単に移動し、室内の広い範囲を見たり多くの人に話しかけることが出来る。

このシステムは人々に日常的に利用してもらうことを主眼に置き、普及性（導入しやすさと入手性および価格）と可用性（運用保守の簡便性・同時利用性）を重視しながらも、固定視点間の移動という形である程度の移動機能を備えている。これにより、いつでも即座に利用出来、気になる物や人があったら近くまで行って見たり話しかけたりする、

Atsushi IZUMIHARA, Atsushi HIYAMA,
Masahiko INAMI

という偶発的なコミュニケーションを可能としている。

2. 関連研究

2.1 高臨場感研究

高い臨場感を追求した研究として Fuchs らは The Office of the Future という研究シリーズにおいて、離れたオフィス空間同士でリアルタイムに空間を三次元的に再構成し伝送し合うことで、接続した空間で一緒に作業するようなビジョンを示している。 [3]

高臨場感にまつわる研究は環境の再現性の高さについて議論されることが多いが、大掛かりな設備・計算リソースが必要になること等が課題である。

2.2 テレグジスタンスとテレプレゼンスロボット

館らは長年にわたり、テレグジスタンスと呼ばれる領域を打ち立て研究を続けている。人間に近似した身体・感覚を持つロボットを介することで、操縦者があたかも現地に行ったような高い臨場感を得て作業やコミュニケーションすることを目指すものである。 [4]

大がかりになりがちなテレグジスタンスの考え方から機能の一部を取り出すことでサービスとして実際に利用され始めているのが、テレプレゼンスロボットである。首振りや移動に特化した物が一般的であるが、移動出来ること

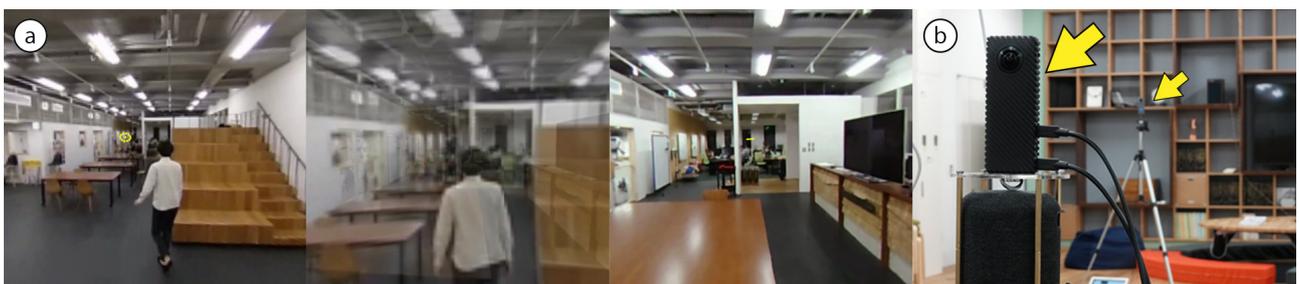


図 1: Multi Spherical View Hopping a. 目標点を定めて移動する様子 b. ノード設置の例



図 2: テレプレゼンスの時空間的な適用可能性の広がり

により室内の誰かに話しかけに行ったり、話さずとも室内をぶらつき雰囲気を感じるなどが出来る。[5]

これらは実際に商品として販売されているが、活用していると下記のような課題に気がつく。

- 充電がされていないなどの理由で使いたい時にすぐ使えないことが多い。
- 環境が整頓されていないと思ったようにスムーズには動けない。用法次第では故障する。

2.3 全天球カメラによるテレプレゼンスと視点の複数化

機材の低廉化と普及、ネットワークインフラの拡充に伴い、全天球カメラとHMDなどを用いた自由に見回すことの出来るライブ配信や仮想旅行体験のサービスが存在している。[2] ライブ配信については複数カメラで行われた物も存在しているが、カメラ間の移動や、身体的な自在感について議論した物はない。Google Street View では、大量の視点で撮影された静止画をつなぎ合わせることで、連続的に移動したような体験を実現している。これはあくまで過去に記録された静止画像に限定される。[6]

2.4 テレプレゼンスのUXにおける自在感

渡邊は利用状態への移行しやすさの尺度のことを「アプローチャビリティ」と呼び、「使おうとし易い」ことの重要性をうたった。[1] テレプレゼンスにおけるアプローチャビリティとは「行きたいときにすぐ行ける・大したことなくなくても気軽に行ける」というふうな障壁の無さであり、これはテレプレゼンスを前提とした暮らしの効果を最大化する上で非常に重要である。

一方で我々はテレプレゼンスを「身体を遠隔地に移動出来る能力の獲得」と言う側面からもとらえている。道具を一時的に利用するのではなく、自らの身体能力として遠隔地と行き来するという実感を持つことが出来れば、それは人の身体観の変容を意味する。使いこなす対象であるという印象を無くすためにもやはり、準備などのコストなく即時に能力の行使が可能であることが重要であると考えられる。

このような考えに基づき我々は、時間・空間的な適用可能性の広がりやテレプレゼンス可能な身体の自在感を構成すると考える。図2にその模式図を示す。ここではそれら要素を、身体から空間側に向けて順に示している。以下に図の詳細について述べる。

物理的干渉可能性

身体が遠隔地に移動した際に手などで物理的な干渉が可能であると、具体的な作業という観点で大きく自在感が増す。トレイグジスタンスの研究では、ロボットの腕や手と操縦者のそれを同体のように動かす技術や、指先の触覚を伝送再現する技術などが研究されている。[4]

方向選択性

見たい方向を見ることが出来ること。人は注意を向ける際にまず首や体を使ってその方向を見ようとすることから重要であると考えられる。テレプレゼンスロボットでは筐体の振り向きによって実現している。昨今は360度カメラとVR HMDの低廉化普及にともない、それらを用いたシステムも提供されている。

移動可能性

遠隔地において、存在位置を移動出来ること。話題に応じた場所に移動したり、移動しながらのコミュニケーションを可能にする。移動によって偶発的なコミュニケーションが発生することも期待できる。テレプレゼンスロボットでは車輪を用いる場合が多い。

普及性

様々な場所にすぐに行けるということを実現する上で、装置が数多く導入されることが必須であり、そのためには導入のしやすさと価格面などの入手性は極めて重要である。これは3の移動可能性と併せて空間的な自在感を作る。

可用性

先述の通りテレプレゼンスを身体的能力として身につける上で、「いつでも使える」ことは極めて重要である。そのため設計の指針としてシステムの安定性ならびに保守の簡便性が挙げられる。また、ロボット型のトレイグジスタンスのほとんどは一人ずつしか利用出来ないものであり、誰かが使っている間は他の人が使えなくなってしまう。同時利用可能性も可用性を上げる要素である。

2.5 MSVH の設計指針

今日のテレプレゼンス研究をこの5要素に照らし合わせたのが表1である。本論文で我々が提案するMulti Spherical View Hopping (MSVH) は以下の指針で設計される。

- 物理的干渉可能性と移動操作性、つまり可動部を排除することにより普及性を上げつつ、機械破損などのリ

表 1: テレプレゼンス技術の自在性の比較

	物理干渉	方向	移動	普及	可用
ビデオチャット	×	×	×	○	○
テレプレゼンス ロボット	△	○	○	△	×
3次元再構成	×	○	△	△	△
MSVH (目標)	×	○	△	○	○

スクをさげ可用性を高くすることを選択する。

- 全地球カメラを用いることで方向操作性を持ちつつ普及性をあげ、数で中距離の移動機能をカバーする。
- 全地球カメラを用いることで多人数同時利用が可能なシステムとし、可用性を高める。
- 利用者は HMD を用いることで没入感の高い体験を得ることが出来るが、同時に通常のディスプレイとマウスでも操作できるようにし利用障壁を下げる。
- 用途次第では常時接続にすることでアプローチチャビリティを高めることが出来る。

このように MSVH は、テレプレゼンスロボットの方向選択・移動の良さを可能な限り取り込みながら、普及性・可用性にフォーカスしてデザインされている。次の章ではその具体的なデザインについて述べる。

3. Multi Spherical View Hopping (MSVH)

3.1 MSVH の体験デザイン

全地球カメラ・マイクスピーカーならびに制御 PC から構成されるノードを室内に複数配置し、インターネット経由でクライアント PC に伝送する。クライアント側では複数の全地球映像間を空間的に移動しながら、ノード視点において自由に辺りを見回し、近くの人とコミュニケーションを取ることで出来るシステムである。利用者の体験する映像を図 1-a に示す。

ノードの配置の例を図 1-b に示す。このように、室内にノードが互いに見える位置に配置することをメインの利用ケースと想定している。これにより利用者は、ノードが配置された箇所に限定されるが、気軽に素早く室内を移動することができる。ノードの映像音声情報は個別にインターネットで伝送されるので、たとえば複数箇所にオフィスのある会社において、オフィス A の中を移動した後、連続的にオフィス B に移動する、といった体験も可能であるし、複数人で同時に利用することも可能である。

図 3 に示すように、利用者は HMD での没入的な体験、あるいはディスプレイとマウスを用いて気軽な体験が可能である。HMD を利用する事で現地に行ったような体験の質が向上するが、HMD の装着は気軽な利用には向いていないので、ディスプレイで利用出来ることは重要であると考え

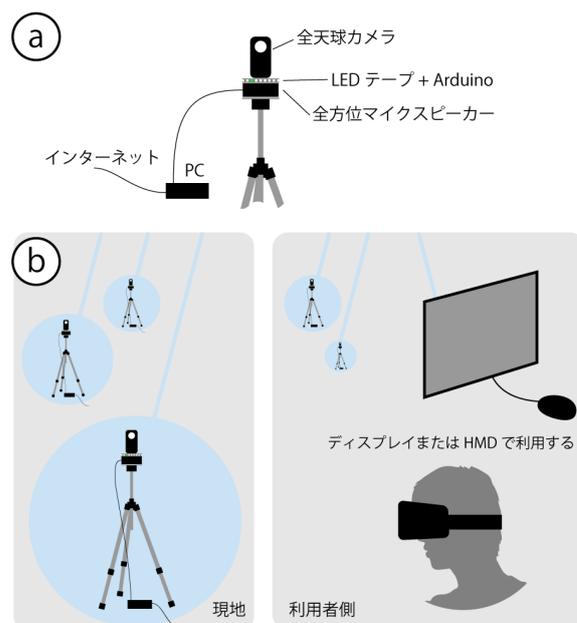


図 3: a. 各ノードの構成 b. システム全容

ている。後述する予備実験で構築したシステムにおいては、ノードを常時アクティブにし、クライアント側も専用マシンとして常設しておくことで、ふと気になったときに使う、という高いアプローチチャビリティを実現している。

今回提案する MSVH のノードにはディスプレイは無い。代わりにノードの外周を巻くように LED のサインを取り付け、現地にいる人にいま利用者がいることと、その人が向いている方向を提示する。これにより一方的に監視されているような感覚を和らげ、コミュニケーションを円滑にする狙いがある。

3.2 MSVH の実装

改めて図 3 にノードおよびシステム全体の構成を示す。

MSVH の各ノードは民生品で構成されており、安価で入手性も高く、設置も容易である。また、可動部を廃したことで高い安定動作性が期待できる。

各ノードとクライアントは WebRTC で通信しており、クライアントは WebVR で実装してある。これにより、クライアントはプラットフォームに限定されず利用可能であり、スマートフォンで出先から利用することもできる。ノードは複数に増やすことができ、それは場所をまたがったり、図 3-b のように今利用者のいる空間であっても良い。

移動中の様子を図 1-a に示す。現在のノードから見た移動可能なノード位置が示されており、そこを視野中心（画面中心）に入れた状態でトリガーを引く（現在はマウスクリック）ことで移動を行う。

装置をインストールする際には、室内におけるノードの相対位置をあらかじめ計測し、システムに定数として与える。ノード間を移動したような画面遷移は図 4 のような方法で実現している。ユーザがノード A にいるときには、ノード A の全地球映像に囲まれているが、ノード B に移り始め

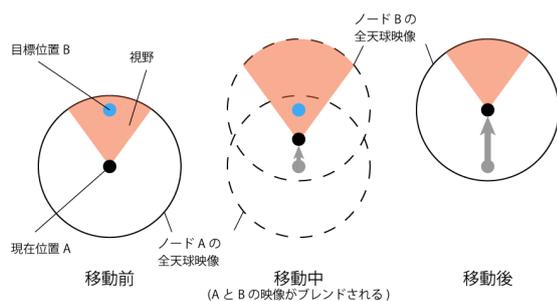


図 4: 全天球映像間のトランジション

ると、WebVR によって作られた 3 次元空間中を移動しながら、次第にノード B の映像とクロスフェードし、最終的にノード B の全天球映像の中心にたどり着く。行き先ノードのカメラ位置が視野中心にある状態で遷移するため、遷移前と遷移後において視線の中心領域が一致する。これによりほぼ連続的なトランジションが実現している。

4. 予備実験

予備実験として、本システムを著者らの所属する研究室に設置運用した。我々の研究室は二箇所のキャンパスに分かれて所在しており両方の居室が活発に利用されている。そのため研究室居室間をつなぐテレプレゼンス環境が強く求められている。

4.1 遠距離条件

初めの条件では、キャンパス A にクライアントを設置し、キャンパス B 居室の三つのノードと映像音声をやりとりした。10 名ほどに利用してもらったところ、移動操作は一切迷い無く気軽に利用してもらえた。学生 S さんがキャンパス A のクライアントの前を通りかかった時に、画面の向こうに話したい友人がいることに気がついたのでディスプレイで利用し始め、キャンパス B のノード 1 に入り、ノード 3 までささっと移動して、友人と短時間会話をすぐ抜けるという計 1 分程度の従来のテレプレゼンスシステムとは異なるコミュニケーションも発生した。本システムがこのような日常的なライトなコミュニケーションを強くサポートする可能性を示唆している。

4.2 近距離条件

2 つ目の条件では、キャンパス A にクライアントを設置し、同じくキャンパス A 居室に三つのノードを設置し映像音声をやりとりした。使用感については初めの条件と同じであるが、利用者が利用中の自分自身を見ることも可能であったことから、利用者の何名かから、いわゆる遠隔地のテレプレゼンスの利便性より、瞬間移動感や幽体離脱しているような感覚が得られて面白いというコメントを得られた。

4.3 議論

4.3.1 利用者の存在感

テレプレゼンスされる側からは、監視感や利用者が突然現れる驚きについてコメントをもらった。LED での方向提示はコミュニケーションが始まった後には充分効果がある

が、そもそも気が付かないケースがあった。また、利用者が入ってきた直後には、誰が利用しているのかもぱっとわからない、という課題も寄せられた。このことから、何かしらの方法で、利用開始を周囲の人に気づかせる方法や、利用者が誰であるかをよりわかりやすく提示する方法が必要であると考えられる。

4.3.2 移動可能性の向上

今回提案した移動法の気軽さ、気持ちよさに関してはおおむねポジティブにとらえられたが、何名かからはもっと移動出来る範囲を拡げたいとコメントされた。3. 章で述べたとおり割り切った部分ではあるが、一つの実装法として、MSVH のシステムにテレプレゼンスロボットや 3 次元再構成のシステムを組み合わせることが考えられる。

5. おわりに

本研究では、普及性と可用性に着目して、テレプレゼンスにおける自在感を議論した上で、複数の全天球カメラ視点間を移動することのできるテレプレゼンス環境を提案、開発した。予備実験として著者らの研究室に配備し利用してもらい、その有効性を確認した。

本論文はシステム構築と簡単な予備実験までをレポートしたが、気軽に始められるライトなテレプレゼンス環境が長期間にわたって安定的に利用出来ることで、コミュニケーションの仕方にどのような影響を与えるか、プライバシー感覚や存在感に変容を起こすか、移動出来ることを前提とした身体観の変化はあり得るのか、という点に興味をもって研究を続けていきたいと考えている。

謝辞

本研究の一部は、JST 国際科学技術協力基盤整備事業およびコベルコ建機株式会社の支援を受けている。

参考文献

- [1] 渡邊恵太 (2016). インタラクションコストとアプローチチャビリティ, EKRICS. <http://ekrics.jp/2016/05/2041/>
- [2] BIGLOOK360 (2011). Use multiple camera views & 360 degree video to enhance live event broadcasts. <http://biglook360.com/2011/11/360-degree-video-3/>
- [3] Raskar, Ramesh, et al. "The office of the future: A unified approach to image-based modeling and spatially immersive displays." Proceedings of the 25th annual conference on Computer graphics and interactive techniques. ACM, 1998.
- [4] Tachi, Susumu, et al. "Telexistence—from 1980 to 2012." Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on. IEEE, 2012.
- [5] Desai, Munjal, et al. "Essential features of telepresence robots." Technologies for Practical Robot Applications (TePRA), 2011 IEEE Conference on. IEEE, 2011.
- [6] Anguelov, Dragomir, et al. "Google street view: Capturing the world at street level." Computer 43.