

ADVNET2024

国立研究開発法人情報通信研究機構委託研究（採択番号：02901）

【革新・シーズ創出型プログラム】

人間拡張・空間創成型遠隔作業支援基盤の研究開発

2024年10月4日

代表研究者：国立大学法人東京大学

研究分担者：TOPPAN株式会社

発表者： TOPPAN株式会社 荻野 孝士

◆「遠隔共有三次元空間」の実現

- ① 遠隔地の空間を実時間で計測・記録・伝送・再構築・共有する
- ② 遠隔・オンライン社会の基盤技術
テレワーキング、遠隔共同作業、遠隔教育、技能伝達、遠隔観光、エンターテインメント
- ③ 空間の壁・時間の壁を超える： 4Dアーカイビング
計測記録された実空間を任意のタイミングで再生・閲覧する → 教育・技能伝達・技能継承

※ 単なるVR(仮想)空間ではなく、遠隔地の実世界状況を共有することを目的とする。

◆技術課題と研究テーマ

- ① 空間計測・再構築技術：三次元空間再現精度と計測インフラのトレードオフ・伝送データ量の削減
→ 1. 拡張可能型空間共有
- ② インタラクション：遠隔三次元空間とのインタラクション技術
→ 2. 遠隔作業空間とのインタラクション技術

[URL] https://www.nict.go.jp/collabo/commission/B5Gsokushin/B5G_02901.html

(シーズ029)

人間拡張・空間創成型遠隔作業支援基盤の研究開発

研究概要：時空間の障壁を越えて遠隔地で作業する人間やアバターロボットと空間を共有し、遠隔地の全貌を把握しながら遠隔共同作業を支援する環境を実現する。予め三次元計測された遠隔地の静的な空間情報と、センサーにより動的に取得される空間情報、人間行動情報の融合を実時間でシームレスかつスケーラブルに実現し、低遅延ネットワークと深層学習による身体行動予測を融合したゼロレイテンシー空間共有技術と融合する。遠隔作業者の一人称視点と、三次元空間での自由・俯瞰視点とを自由に行き来することのできる空間作業支援ユーザインタフェースを構築する。



静的に計測・蓄積された
遠隔空間ジオメトリ



各種センサで動的に計測される
遠隔三次元空間情報



高精度に再構築された遠隔空間情報



**三次元空間サイズ 100mx100m以上
再構築処理時間 500ms 以下**

遠隔空間再構築例

【契約期間】 令和3年度～令和4年度（ステージゲート評価予定）【契約総額】 約1.0億円

【受託者】 国立大学法人東京大学（代表研究者）、凸版印刷株式会社

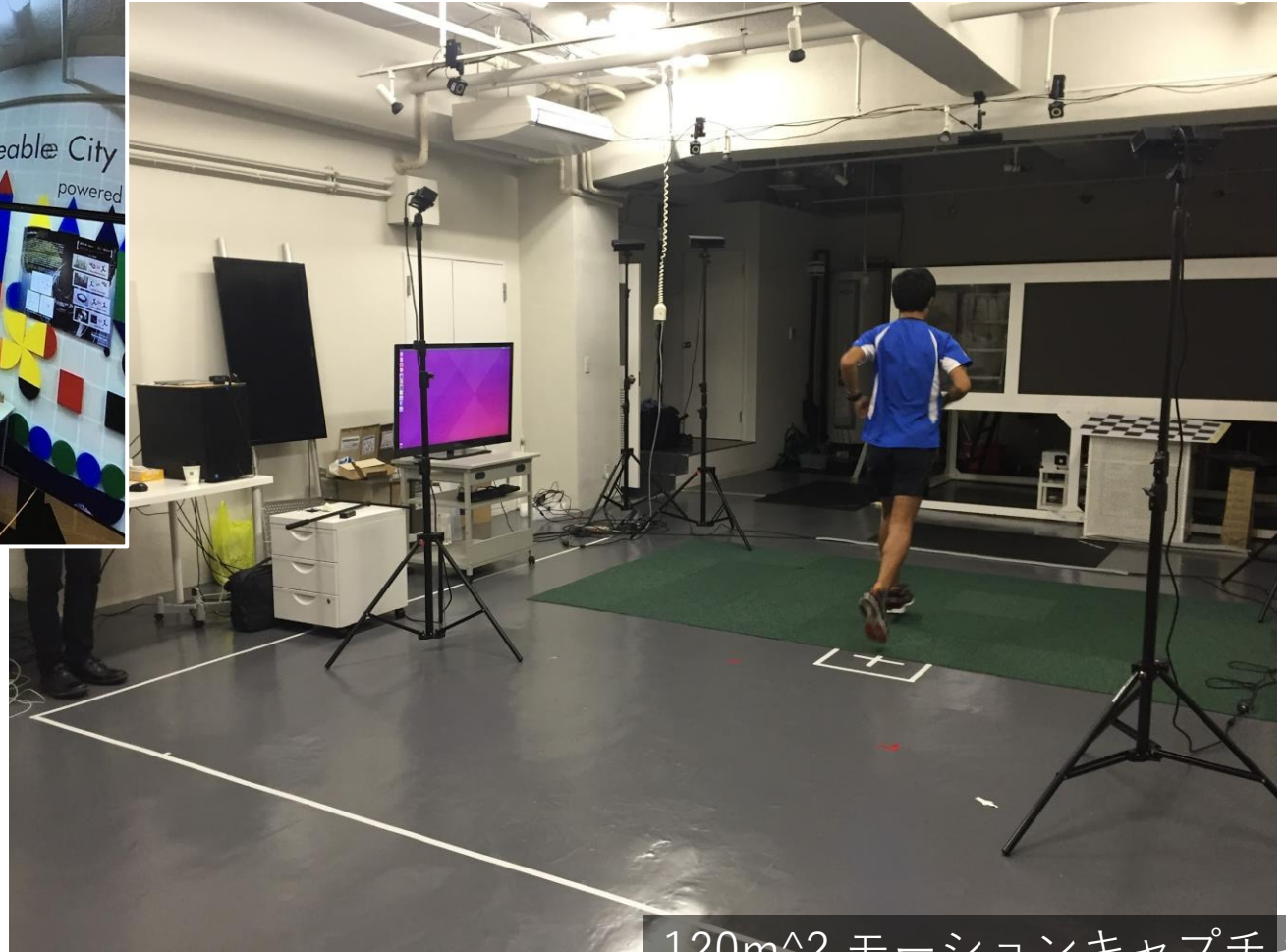


*TMIP（Tokyo Marunouchi Innovation Platform）

暦本研究室 人間拡張実験スタジオ（文京区本郷）



- ・ TIMIPとの丸の内街区での
Local5G 実証実験(当初予定)
⇒ Beynd5Gを想定しJGNによる
広域伝送性能評価に変更
- ・ 東京大学人間拡張社会連携講座との連携



120m² モーションキャプチャー

複数の情報源（静的情報と動的情報）の融合により遠隔三次元空間再構築の拡張性と精度を確保

屋内空間

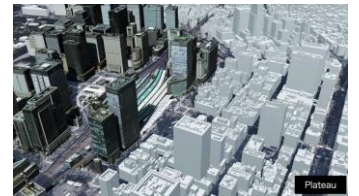


計測・蓄積された
遠隔空間ジオメトリ

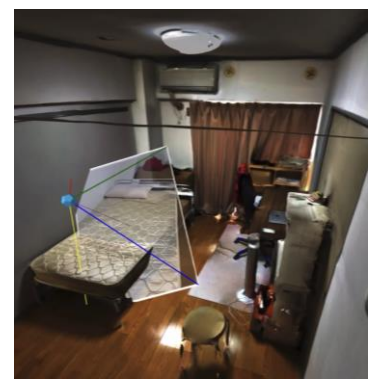
Photogrammetry



PLATEAU等の
3次元データベース



遠隔作業者周辺の状況



一人称映像と
遠隔空間の融合

ウェアラブルカメラ



複数センサー設置箇所 (10m^2)



Depth camera / LiDAR
で動的に計測される三次元情報

Point-cloud計測
身体ポーズ認識

=

高精度に再構築された空間情報



ネットワーク伝送
(Beyond5G)



遠隔地における三次元再構成



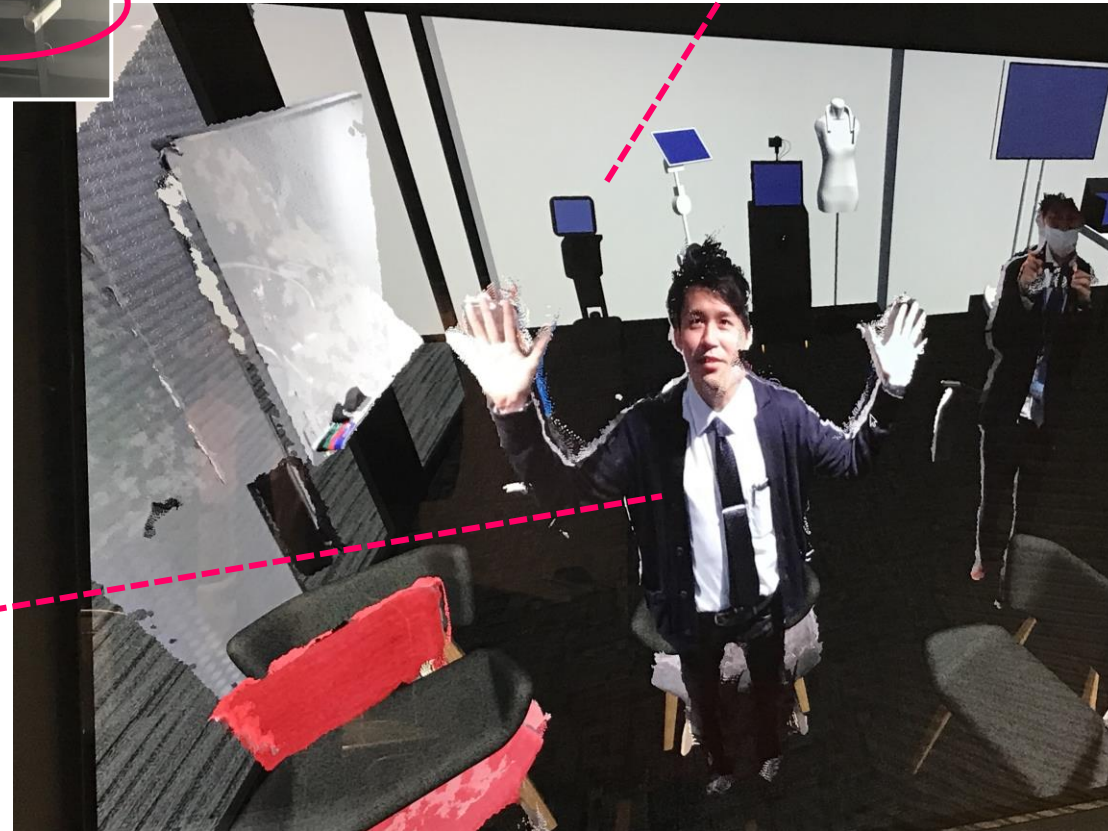
複数の情報源（静的情報と動的情報）の融合による三次元空間再構築



複数デプスセンサー

静的にジオメトリセンシングされた空間情報
(背景部分)

実時間合成・伝送された三次元情報
(人物部分)



複数の情報源（静的情報と動的情報）の融合により遠隔三次元空間再構築の拡張性と精度を確保

屋内空間

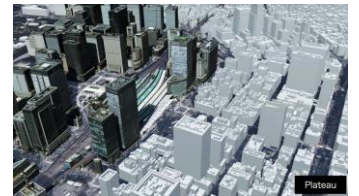


計測・蓄積された
遠隔空間ジオメトリ

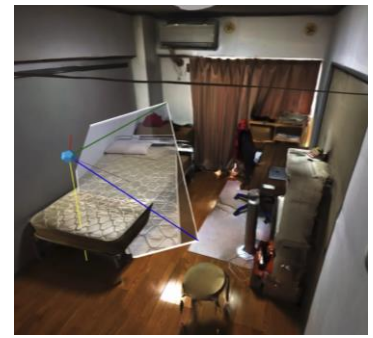
Photogrammetry



PLATEAU等の
3次元データベース



遠隔作業者の移動に伴い
動的に拡張



一人称映像と
遠隔空間の融合

ウェアブルカメラ



複数センサー設置箇所 (10m²)



Depth camera / LiDAR
で動的に計測される三次元情報

Point-cloud計測

=

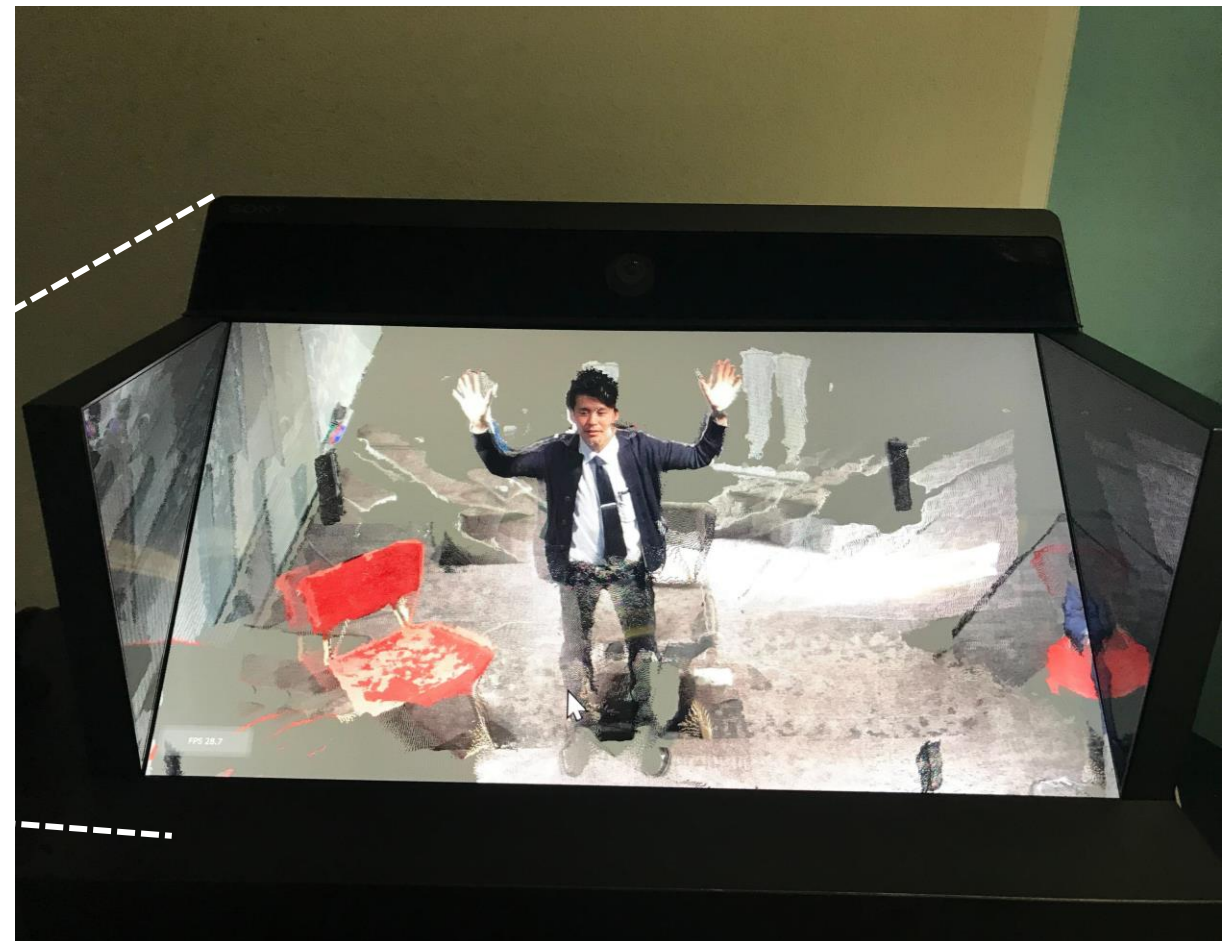


高精度に再構築された空間情報

一人称・任意視点を経由した
遠隔作業者視点間の移動

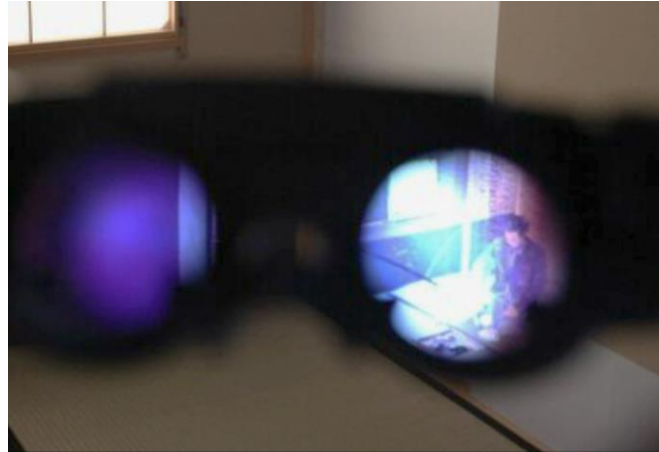


複数人による作業支援実現のため、ディスプレイによる構成を採用



TOPPAN(丸の内)、東京大学(本郷)に設置

簡易利用も想定し、シースルー型のHMDによる構成も検討
技能教示者と同じ身体空間を共有する場合に適合



Optical See-Through (MagicLeap), Video See-Through (Meta Quest3)で稼働

本研究開発の実証実験を2つの視点で実施

1) 技術視点

目的) Beyond 5 Gを見据えたネットワークインフラ下における、本技術による伝送評価

方法) N I C T 超高速研究開発ネットワークテストベッド (J G N)

N I C T 小金井－東京大学の2拠点を相互接続し、伝送評価

評価項目は、受信側のFPS、遅延時間、画質を予定

伝送路の比較としてLAN (1000BASE-T) 環境でも実験

2) 利用者視点

目的) Beyond5G 時代に求められる技能伝承の在り方について、利用者視点での有効性評価

方法) 業界で著名な伝承者の方のご協力をいただき、ピアノの演奏指南をテーマとした

利用者視点での三次元アーカイブの評価実験

- ・ 伝承者： フォルテピアニスト 小川加恵様
- ・ 実証場所： TOPPAN HALL (東京都文京区)
- ・ 参加者： ピアノ経験者及びホールスタッフ
(総勢10名)



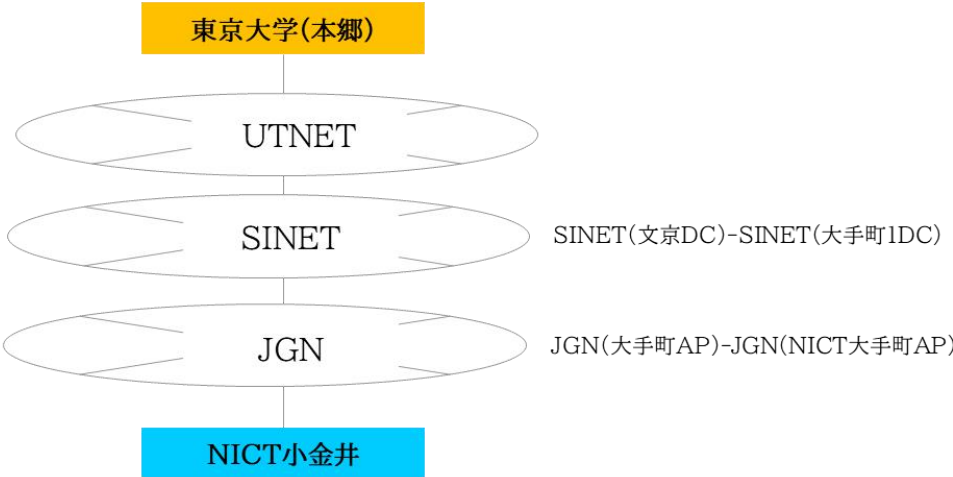
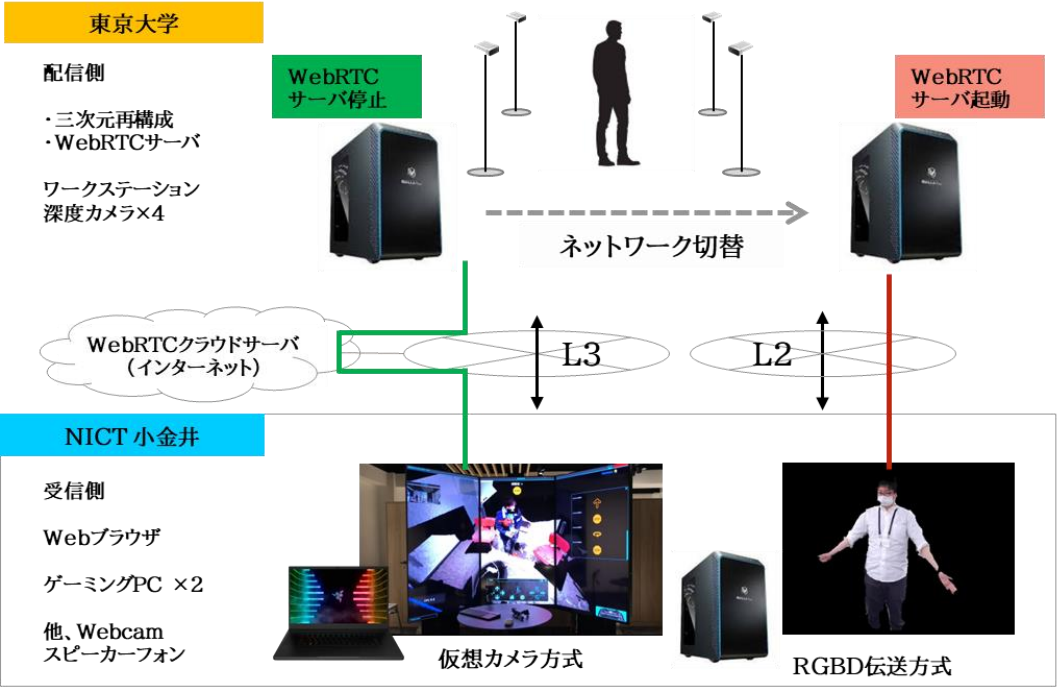
©Shigeto Imura



<https://www.toppanhall.com/>

広域2拠点間（東京大学：本郷 ⇄ NICT：小金井）での3次元センシング・伝送・再構成の性能評価

- ① 環境：JGN L2(イーサネット), L3(インターネット)
- ② 接続拠点：東京大学 ⇄ NICT小金井
- ③ 伝送テスト内容： 3次元再構成データの伝送テスト（配信：東大 受信：小金井）
2種類の伝送方式を予定(仮想カメラ方式、RGBD伝送方式 ※1)



※1 RGBD伝送方式： 特許出願中（2022年度）
映像データ（RGB）と深度データ（D）を分けて伝送する方式
RGBは汎用コーデック（VP9）、Dはテキストデータで伝送
各データにタイムスタンプを付加し受信側で結合
メリット： 受信側でも3Dデータを保持
デメリット： データ結合のオーバーヘッド／現状インターネット通信不可

3/6 初回計測、3/21 追加計測を実施

JGN L3 | 仮想カメラ方式

※ フレームレート、遅延時間は計測時間の平均値を記載
※ 受信画質は目視評価

	フレームレート (fps)		遅延 (ms)	受信画質	考察
	送信	受信			
LAN (1000BASE-T)	30	29	140	良 (配信側に比べて)	送信時に 3D⇒2D変換 動画コーデックにより、データ削減。配信側に比べ先鋭さに欠ける
JGN L3	30	29	140	良 (配信側に比べて)	LANと同様

JGN L2 | RGBD伝送方式

※ フレームレート、遅延時間は計測時間の平均値を記載
※ 受信画質は目視評価

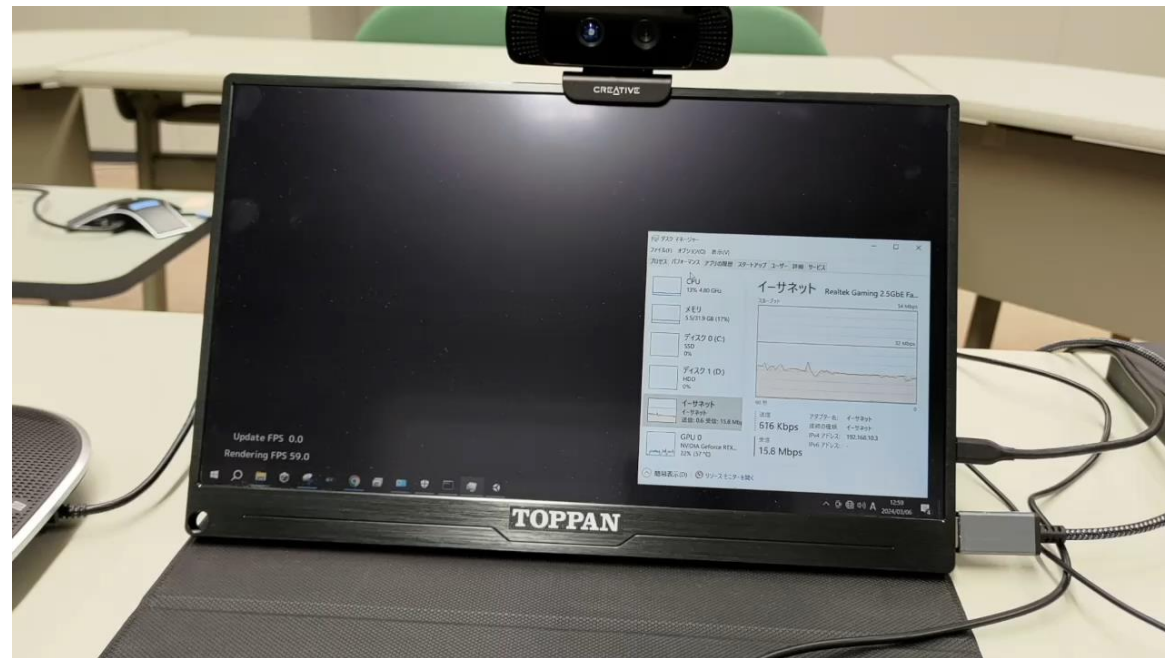
	フレームレート (fps)		遅延 (ms)	受信画質	考察
	送信	受信			
LAN (1000BASE-T)	30	20	500	良好 (送信側とほぼ同様)	RGBデータとDデータの一致確認・合成処理によるオーバーヘッド。 (CPU性能向上によりfps向上を見込める)
JGN L2	30	映像表示されず			送信：約120Mbps 受信：約15Mbps 通信経路でパケットドロップが発生したと思われる RGBデータとDデータが一致しないため映像が表示されなかった。

本郷で学生2名にコントを実演していただき、小金井で三次元鑑賞



JGN L3 | 仮想カメラ方式

※ LAN環境同様にスムーズに再生される



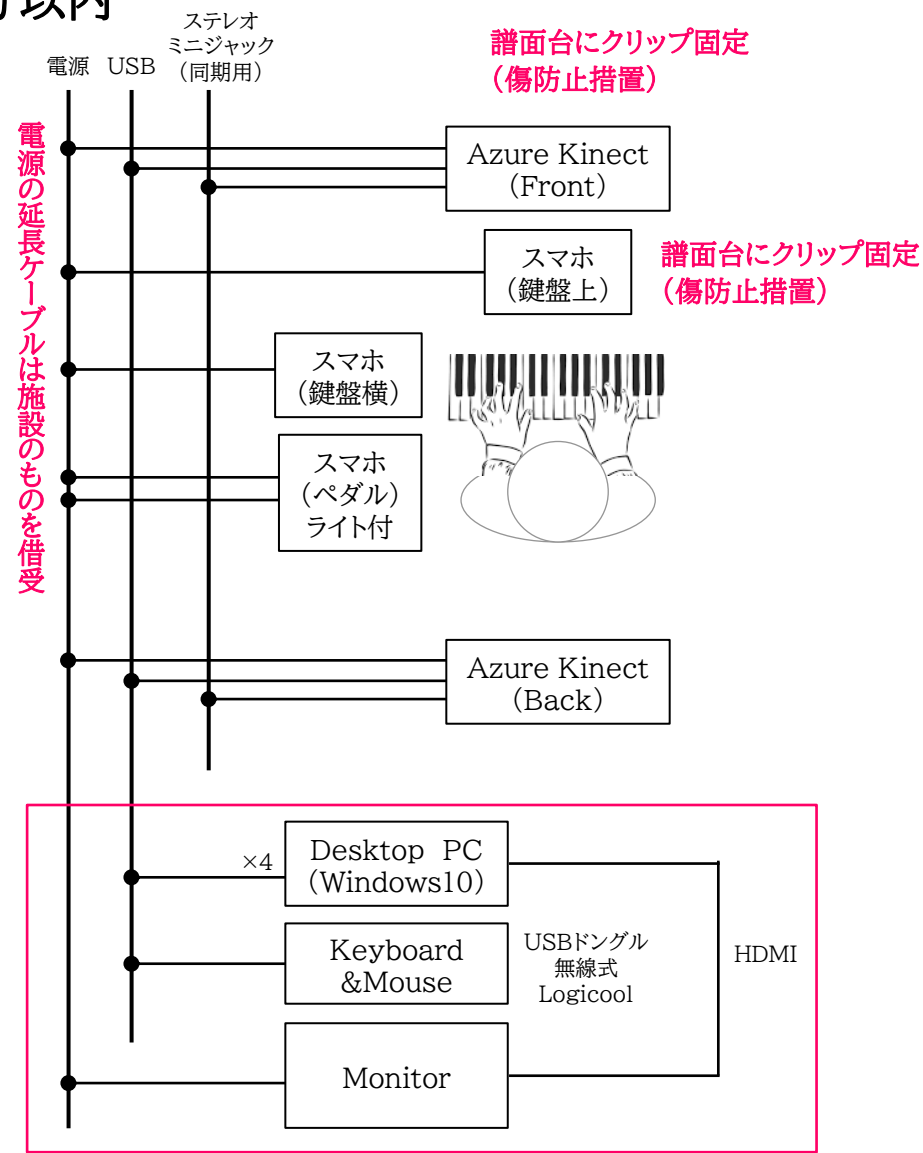
JGN L2 | RGBD伝送方式

※ データ受信はされているが、同じタイムスタンプのRGB・Dのペアが見つからないため、映像が復元できない

- ① アーカイブのみ収録。伝送なし
- ② 3曲の収録(指・手の動き: ゆっくり/ふつう/はげしい)を予定。 各曲3分以内



機器構成



点群サイズ変更



Update FPS 29.0
Rendering FPS 59.7

演奏姿勢3次元再構築例: 身体全体を点群(Point Cloud)、表情は映像から再構築、手指は骨格計測から再構築

収録アーカイブの評価

※ アーカイブに音声を追加し、アーカイブを見ながら10名程のユーザーにインタビューを実施。質問に対する回答を集約
※ 改善が必要な項目は「可」判定。ポジティブな項目は「良」判定とした。

No.	質問内容	回答内容	評価
1	見た目の印象	・近くによると画面が荒くなる ・見たい部分（演者の指先）がよく見えないのでその部分だけでも改善できないか	可
2	演奏指南の有効性	・二次元とは違い自由視点ができるのが良い ・演者と自分を重ねてトレーニングができる。子供の頃にピアノを習っていたがその時、こういったものがあれば練習の励みになったかもしれない ・表面から判別できない要素を加えられないか 例) 打鍵の強さ、上半身の筋肉や骨格の動き ・二次元とは違い自由視点ができるのが良い	良
3	表示媒体（平面ディスプレイ）	・小さな画面よりも大型の画面の方が見やすい	可
4	表示媒体（裸眼立体視ディスプレイ）	・ミニチュアみたいで面白い ・平面ディスプレイよりも良い ・テレビとが違う見方できるのでエンタメ的な活用はありかもしれない ・HMDで視聴してみたい	良

三次元表現については、見た目の改善は必要。演奏指南や見せ方については好意的な感触を得ることができた。

4～6 台のdepth camera (Azure Kinect)により茶道のお手前動作を3次元収録
練習者がAR-HMDにて動作を閲覧することを可能にした。



real scene



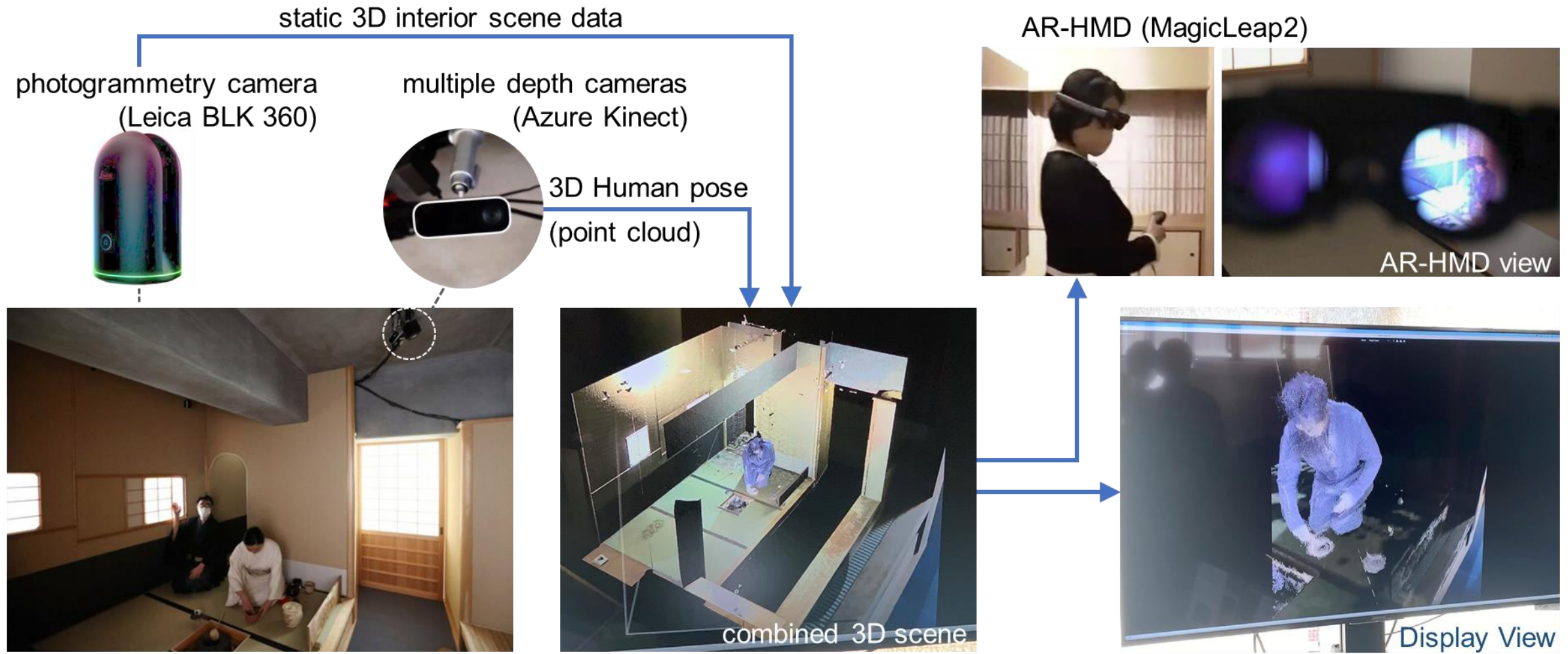
reconstructed scene



view through AR-HMD

協力:ソニーコンピュータサイエンス研究所、裏千家:茶美会文化研究所

[参考] 研究成果展開 茶室技能伝承



協力:ソニーコンピュータサイエンス研究所、裏千家:茶美会文化研究所

教示者と一体「一人称視点」、外部から観測「3人称視点」、背後から観測「背後霊」視点を実現



1st person view



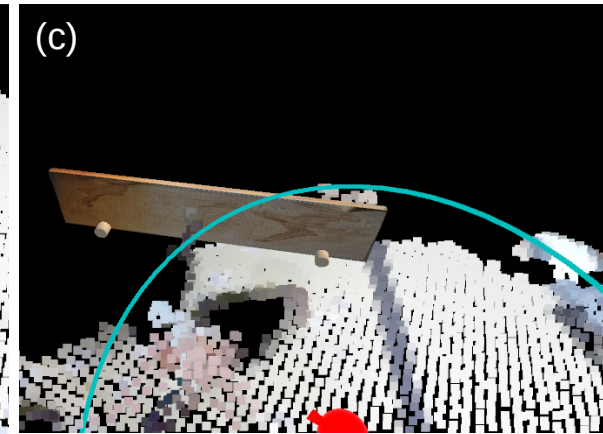
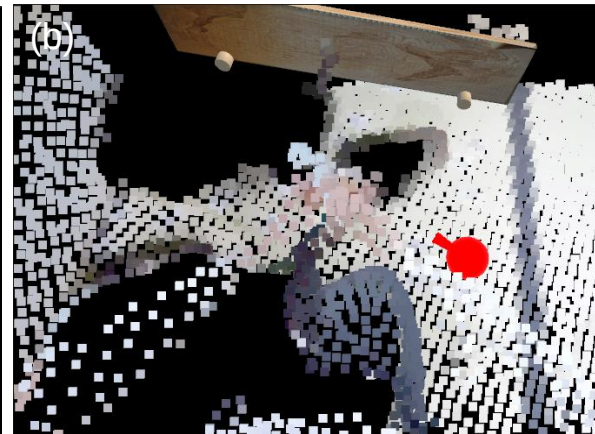
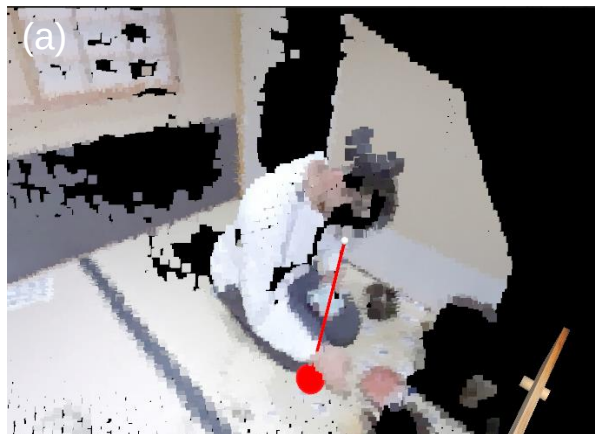
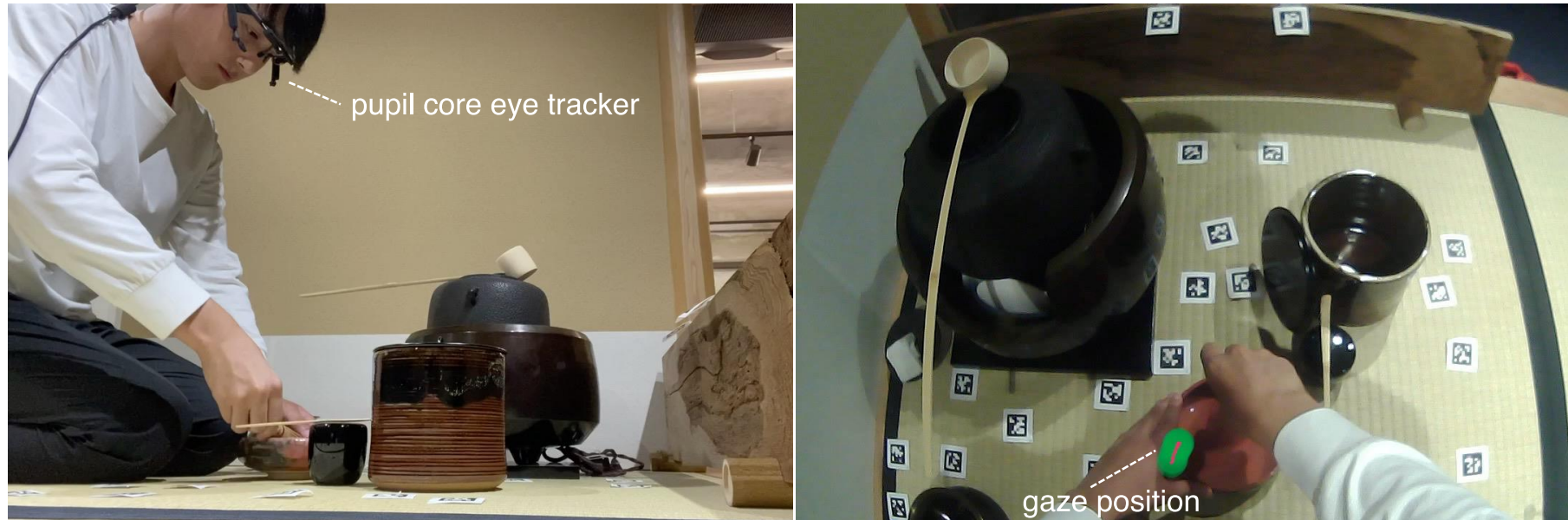
3rd person view



behind view

視線情報の記録、再生

指示者の視線を 3 次元再構築空間で再現実際の動作が開始される前の視線移動が技能理解に有効





実空間にコンピュータ上の空間を重ね合わせることで、
自由な視点から師の作法を学ぶことができる。

[参考] 研究成果展開想定例



遠隔観光



遠隔体験学習



遠隔農業



遠隔生産管理



遠隔外商支援



遠隔宇宙体験

研究グループ全体)

出願区分	実績(2024年2月末) ()内は3月末までの見込み	最終目標
国内出願	3 (3)	2
外国出願	0 (0)	0

2. 外部発表

- 論文発表は総数にて計画数を上回った。
- 一般口頭発表(基調講演複数回を含む)も計画数を上回った。

発表分類	実績(2024年2月末) ()内は3月末までの見込み	最終目標
A.研究論文	0 (0)	2
B.小論文	0 (0)	2
C1.査読付収録論文	4 (8)	1
C2.収録論文	0 (0)	1
D.機関誌論文	0 (0)	1
E.著書等	2 (2)	1
F.学術解説等	0 (0)	1
G.一般口頭発表	13 (13)	2
H.その他の資料	3 (3)	1
I.標準化提案	0 (0)	0
J.標準化採択	0 (0)	0
K.プレスリリース	0 (0)	3
L.報道	0 (0)	—
M.展示会	1 (1)	0
N.受賞	0 (0)	—
O.表彰	0 (0)	—
P.成果の実施	0 (0)	—
Q.その他	0 (0)	0

① 3次元計測技術

- **Depth camera** → **通常カメラ + 深層学習**
NeRF (Neural Radiance Fields), Gaussian Splatting 等の深層学習技術との融合により、通常のカメラでの3次元構築を可能にする。

② 他のモダリティ（視線・触覚情報）との融合

- 遠隔者の視線情報の記録・伝送により意図伝達を円滑化する（一部実施）
- 視線記録可能な1st person head setを開発中

③ 4次元コンテンツのブラウジング手法

- 収録された4D情報(3D+時間)の効率的なブラウジング手法
- 閲覧者の質問に対応（LLMとの融合）

TOPPAN