マルチタスク学習を用いたシーン理解による複合現実感の向上

研究代表者　馮　起
（理工学術院総合研究所 次席研究員）

１．　研究課題

複合現実感（XR）は，コンピュータで生成された仮想オブジェクトを用いて現実世界を拡張する先端技術であり，近年，学術・産業の両分野において急速に注目を集めている．デジタル世界と物理世界を自然に融合させ，ユーザーに高い没入感と直感的な操作性を提供するためには，仮想空間と現実環境の双方を精緻に理解・統合する技術が不可欠である．本研究では，コンピュータビジョンにおけるマルチタスク学習の枠組みに着目し，シーン理解に関わる複数の関連タスクを単一のモデルで同時に学習させることで，共通する有用な特徴表現を獲得し，XR 技術における認識精度と体験品質の向上を図った．また，知覚的な理解にとどまらず，XR 環境を活用したユーザー支援技術への応用にも取り組んでおり，非利き手のスキル習得を支援するデジタルツイン技術に基づくトレーニングシステムを構築した．

２．　主な研究成果

2.1　単眼 360 度画像からのスケール整合な奥行き推定の実現

従来の単眼画像ベースの奥行き推定技術においてはスケールの曖昧性が大きな課題であり，特に 360 度画像では撮影デバイスや環境条件の差異によってスケール整合性が損なわれ，現実空間の再構成精度が著しく低下する問題があった．そこで本研究では，微分可能な等角投影変換と補助的なスケール推定ネットワークを組み合わせた新たな深層学習フレームワークを提案した．具体的には，360 度画像を一時的に透視投影画像へと変換し，既存の大規模データセットで学習された 60〜100

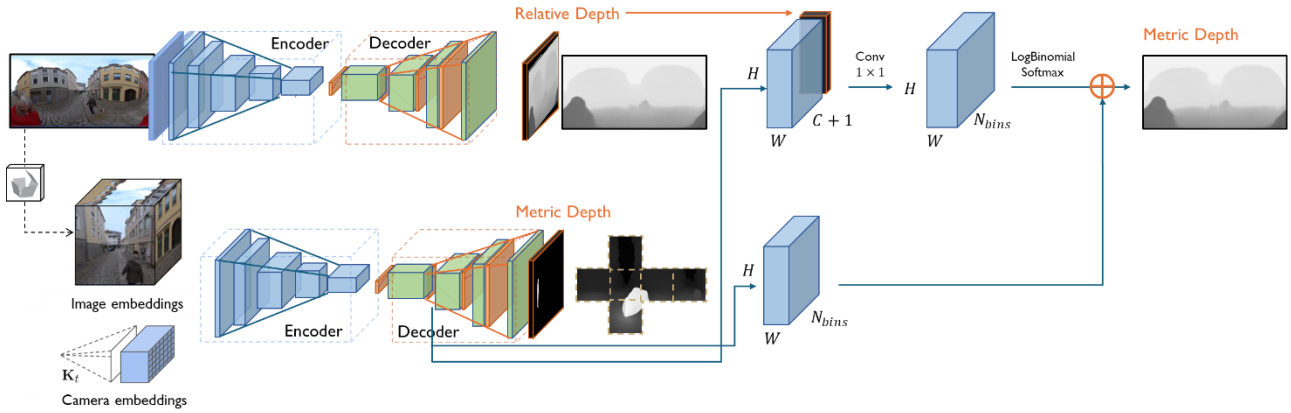

Fig. 1 提案ニューラルネットワーク

度視野角の深度推定モデルを再利用することで，教師データの不足を補いながらスケール整合性の高い奥行き推定を実現している．さらに，透視投影画像から得られるスケール情報を 360 度画像に

転移する補助学習モジュールを導入することで，より一貫性のある奥行きマップの生成を可能とした．この成果は，XR 空間構築やナビゲーション支援をはじめとする多様な応用領域において，現実環境に対する高精度かつ一貫性のある空間理解を支える基盤技術となる．

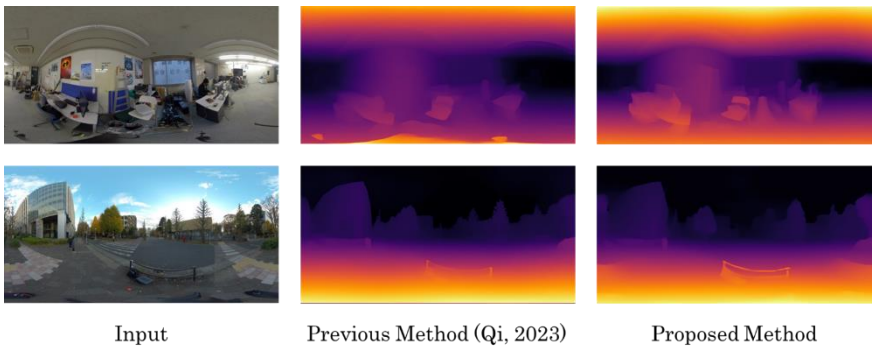

Fig. 2 既存手法との比較結果

2.2 デジタルツインと XR による非利き手スキル習得支援システムの構築

Fig. 3 実験デザイン

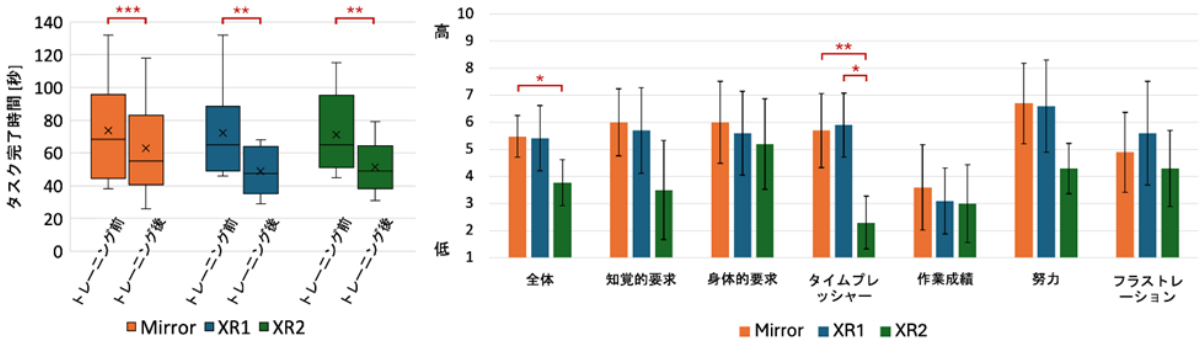

Fig. 4 トレーニングシステムがタスク完了時間（図 3 左）および認知負荷（図 3 右）に与える影響．赤い括弧は各条件間の有意差を表す: *(p < .05), **(p < .01), ***(p < .001)．なお，図 3b のエラーバーは 95 %信頼区間を示す．

非利き手における運動スキルの習得は，日常生活やリハビリテーション，産業訓練において重要な課題である．従来のミラーセラピーは，鏡を用いて利き手の動作を反転表示することで非利き手の補助を行うが，物理的な制約や反転映像による認知的負荷の高さにより，学習効果には限界があった．そこで本研究では，XR 技術を応用し，ユーザーのデジタルツインをリアルタイムに左右反転表示する視覚フィードバックシステムを開発した．ヘッドマウントディスプレイとカメラによるシーン理解を活用し，利き手による熟練動作に基づいた同期型ガイダンス映像を提示することで，直感的な比較と即時のフィードバックを可能とした．30 名の被験者を対象に，箸を用いて豆を容器

間で移動させる課題を通じて本システムの有効性を検証した結果，タスク完了時間が有意に短縮され，主観的評価においても認知的負荷の大幅な低減が確認された．

３． 共同研究者

森島　繁生（理工学術院・教授）　　田中 啓太郎（理工総研・次席研究員）
Hubert Shum（Durham University・教授）　　井上 隆大（応用物理研究科・学生）

４． 研究業績

4.1 学術論文

Iwakata Shogo, Oshima Ryosuke, Tsunashima Hideki, Feng Qi, Kataoka Hirokatsu, & Morishima Shigeo. Viewpoint-Dependent 3D visual Grounding for Mobile Robots. *In 2025 IEEE International Conference on Image Processing*. IEEE. 2025. (Under review.)

Inoue Ryudai, Feng Qi, & Morishima Shigeo. (2025, March). SynchroDexterity: Rapid Non-Dominant Hand Skill Acquisition with Synchronized Guidance in Mixed Reality. *In 2025 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops* (pp. 1342-1343). IEEE. 2025.

Nishizawa Hiroki, Keitaro Tanaka, Asuka Hirata, Shugo Yamaguchi, Qi Feng (*Corresponding author), Masatoshi Hamanaka, & Shigeo Morishima. SyncViolinist: Music-Oriented Violin Motion Generation Based on Bowing and Fingering. *In 2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pp. 5419-5428. IEEE, 2025.

Taichi Higasa, Keitaro Tanaka, Qi Feng, & Shigeo Morishima, Keep Eyes on the Sentence: An Interactive Sentence Simplification System for English Learners Based on Eye Tracking and Large Language Models, *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 2024.

4.2 総説・著書

4.3 招待講演

4.4 受賞・表彰

2024 年度早稲田大学-キオクシア共同研究若手奨励賞受賞

4.5 学会および社会的活動

５． 研究活動の課題と展望

本研究は，最先端のコンピュータビジョン技術を実用的な XR 体験へと橋渡しし，ユーザーにとってより自然で没入感のある体験の実現を目指すものである．近年，高解像度・リアルタイムなカメラアクセスやオンデバイス推論が可能な HMD の登場により，軽量かつ高効率なモデル設計の重要性が一層高まっている．今後は，XR 環境における空間認識とインタラクション体験を高度に統合する CV-XR 融合技術の確立を目指す．