Advanced Engineering Informatics 68 (2025) 103665

Contents lists available at ScienceDirect

Advanced Engineering Informatics

journal homepage: www.elsevier.com/locate/aei

Learning viewpoint control from human-initiated transitions for teleoperation in construction

Sungboo Yoon [a], Moonseo Park [a], Changbum R. Ahn [b,*]

[a] Department of Architecture and Architectural Engineering, Seoul National University, Seoul 08826, Republic of Korea

[b] Department of Architecture and Architectural Engineering, Institute of Construction and Environmental Engineering, Seoul National University, Seoul 08826, Republic of Korea

ARTICLE INFO

Keywords:
Imitation learning
Teleoperation
Viewpoint control
Virtual reality
Construction robotics

ABSTRACT

Visual perception is critical for teleoperation in construction, where optimal visibility directly impacts task performance. Hybrid viewpoint control systems enhance the flexibility of visual perception by adaptively coupling or decoupling the viewpoint from robot movements according to situational demands. However, determining the optimal timing for transitions between these perspectives remains a major challenge, as existing autonomous methods are not directly applicable to hybrid control for construction tasks. In this work, we propose a viewpoint control mode prediction model that autonomously manages transitions during teleoperation with hybrid control. Our learning scheme with a transition-guided weighting method leverages sporadic transition commands from human interactions with the teleoperation system as demonstration data for imitation learning. User evaluation in a virtual reality (VR) environment simulating construction welding tasks shows that our model outperforms the baselines, achieving an 11% improvement over the state-of-the-art behavioral cloning (BC) algorithm and a 19% improvement over the state-of-the-art weighted BC algorithm in replicating human transition behaviors. This work contributes novel insights into the design of visual perception systems for teleoperation in construction, enabling reliable, user-aligned viewpoint transitions.

1. Introduction

Teleoperation plays an important role in construction, including reduced human exposure to risks [1–3], data collection for imitation learning [4–6], and on-the-fly adjustments for autonomous systems [7]. Most teleoperation systems rely on single or multiple cameras that are either fixed in position [6] or mounted in an eye-in-hand configuration [8]. These cameras are typically positioned in a task-agnostic manner, without specific consideration for the visibility requirements of particular tasks [9]. Although several perception systems [10–13] include cameras that allow orientational movements, their movements remain constrained by fixed camera setups [14], often suffering from occlusions in cluttered workspaces or when close-up views are required [9]. This limited visual feedback not only hinders precise and efficient remote task execution but also raises safety concerns in construction [15].

To offer more flexibility to remote perception in teleoperation, recent studies have explored dynamic viewpoint systems that allow real-time adjustments to operator views. Dynamic cameras, using external cameras mounted on unmanned aerial vehicles (UAVs) [16,17] or high-degree-of-freedom (DoF) robotic arms [9,14,18–23], typically operate in either robot-coupled or decoupled modes [24]. In the robot-coupled mode, the operator's viewpoint moves along with the robot, providing a coordinated view that aligns closely with the robot's actions. In the robot-decoupled mode, the viewpoint is entirely independent of the robot's movements, allowing greater flexibility in adjusting the viewpoint as needed. In our previous work [25], we introduced a hybrid viewpoint control system that combines both robot-coupled and decoupled behaviors, enabling operators to adapt their viewpoints during teleoperated construction tasks. Construction welding tasks, featuring spatially dispersed weld seams, provide illustrative examples of such applications. When navigating construction sites, operators can employ a robot-decoupled viewpoint to effectively explore the surrounding environment, helping them avoid obstacles and plan efficient lift paths [26]. When performing precise welding operations, operators can switch to a robot-coupled viewpoint to simplify hand-eye coordination and focus on the task at hand [27].

* Corresponding author at: Department of Architecture and Architectural Engineering, Institute of Construction and Environmental Engineering, Seoul National University, Seoul 08826, Republic of Korea.

E-mail addresses: yoonsb24@snu.ac.kr (S. Yoon), mspark@snu.ac.kr (M. Park), cbahn@snu.ac.kr (C.R. Ahn).

https://doi.org/10.1016/j.aei.2025.103665
Received 27 January 2025; Received in revised form 11 July 2025; Accepted 15 July 2025
Available online 21 July 2025
1474-0346/