3D Excavator Pose Estimation: Direct Optimization from 2D Pose Using Kinematic Constraints

Leyang Wen,[1] Daeho Kim, Ph.D.,[2] Meiyin Liu, Ph.D.,[3] and SangHyun Lee, Ph.D.[4]

[1]Dynamic Project Management Lab, Department of Civil and Environmental Engineering, University of Michigan, 2350 Hayward St, Ann Arbor, MI 48109; e-mail: wenleyan@umich.edu
[2]Dynamic Project Management Lab, Department of Civil and Environmental Engineering, University of Michigan, 2350 Hayward St, Ann Arbor, MI 48109; e-mail: daeho@umich.edu
[3]Department of Civil and Environmental Engineering, Rutgers University, RWH 420E, 500 Bartholomew Rd, Piscataway, NJ 08854; e-mail: meiyin.liu@rutgers.edu
[4]Dynamic Project Management Lab, Department of Civil and Environmental Engineering, University of Michigan, 2350 Hayward St, Ann Arbor, MI 48109 e-mail: shdpm@umich.edu

ABSTRACT

With the growing prevalence of site surveillance cameras and rapid developments in computer vision techniques, much effort has been made to develop vision-based safety monitoring methods. An excavator is one of the most used pieces of construction equipment, yet its revolving mechanical arm poses major risks for struck-by accidents. Previous works that leverage computer vision techniques have attempted to monitor such risks by capturing a 3D representation of the excavator arm using either stereo vision or deep learning models. However, stereo vision requires rigid camera setups, while deep learning models require vast amounts of annotated 3D excavator pose, which are cumbersome to collect from real excavators. These drawbacks limit the practicality of such methods in actual construction applications. Therefore, we propose an optimization-based algorithm capable of estimating 3D excavator poses using monocular camera images with no dependency on annotated 3D training data. Specifically, the proposed algorithm attempts to find the optimal 3D excavator pose by imposing rigid kinematic constraints (e.g., arm length and bending directions) while also minimizing re-projection joint errors in 2D. The kinematic constraints help to resolve the inherent depth ambiguity from 2D images while the re-projection 2D joint errors serve as the objective function for the optimization process. Tests using synthetically generated datasets showed a mean 3D location error of 0.16m for key excavator joints, demonstrating the capabilities of this proposed optimization-based method. Moreover, the proposed method does not require the demanding preparation works typically required for collecting 3D excavator poses, making it a practical tool to facilitate excavator safety monitoring in actual construction environments.