Monocular Dense 3D Reconstruction of Construction and Demolition Waste on Conveyor Systems Using Motion Stereo

U.G.U.P. Samarasekara[1], S.A.R.R. Suraweeraarachchi[2], H.M.P.B. Ariyaratne[1], J.A.S.C. Jayasinghe[1], U. Jayasinghe[3], C.S. Bandara[1]*

[1]Department of Civil Engineering, Faculty of Engineering, University of Peradeniya, Peradeniya 20400, Sri Lanka
[2]Department of Mechanical Engineering, Faculty of Engineering, University of Peradeniya, Peradeniya 20400, Sri Lanka
[3]Department of Computer Engineering, Faculty of Engineering, University of Peradeniya, Peradeniya 20400, Sri Lanka
**e19343@eng.pdn.ac.lk*

Accurate volumetric characterisation of Construction and Demolition (C&D) waste is critical for efficient material recovery and automated sorting on conveyor systems. Existing multi-sensor approaches, such as RGB-D cameras, LiDAR, and binocular stereo rigs, are operationally impractical in material recovery facilities due to high hardware costs, complex spatial calibration, and accelerated sensor degradation, while monocular solutions remain entirely unexplored for this material domain. This study aims to develop and evaluate a novel dense 3D reconstruction system for C&D waste from single-camera image sequences, establishing the first empirical benchmark for monocular reconstruction in this domain and enabling metric volume estimation without dual-camera hardware. Exploiting the kinematically constrained, rectilinear trajectory of industrial conveyor belts, temporally displaced monocular frames are treated as synthetic stereo pairs under the Motion Stereo paradigm. Dense disparity maps are computed via a conventional stereo pipeline using Convolutional Neural Networks for both foreground segmentation and learned stereo matching. Metric accuracy is fundamentally coupled to inter-frame baseline estimation; ArUco fiducial markers are deployed on the belt surface to provide drift-constrained displacement measurements, converting disparity maps into metrically consistent point clouds from which object volumes are derived by integrating the reconstructed surface geometry. The proposed system successfully generates dense 3D point clouds from monocular sequences across representative waste categories. Convolutional Neural Network-based segmentation effectively isolates foreground waste objects from the moving belt background. ArUco-constrained baseline tracking substantially reduces metric drift relative to unconstrained estimation. Volumetric estimates derived from the reconstructed point clouds demonstrate consistency with reference measurements, confirming the viability of the Motion Stereo paradigm for this application. This work establishes the first empirical benchmark for monocular dense 3D reconstruction of C&D waste, demonstrating that single-camera systems can deliver metrically accurate volumetric estimates. The approach substantially lowers deployment barriers compared to multi-sensor alternatives, offering a cost-effective pathway toward intelligent, automated material recovery

Keywords: construction and demolition waste, monocular 3D reconstruction, motion stereo, volume estimation, convolutional neural networks

Acknowledgement: This work was supported by the NSF–NSFC Joint Research Program *(Grant No. ICRP/NSF-NSFC/2024/EA&ICT/03).*