

Yan Di

Phone: 017636242558 | Email: shangbuhuan13@gmail.com

WebSite: <https://shangbuhuan13.github.io/>

EDUCATION

Technical University of Munich	Dec 2020 - Feb 2024
Computer Science, Doctor	Munich
Co-supervised by Dr. Federico Tombari and Prof. Nassir Navab	
Tsinghua University	Jun 2017 - Jun 2020
Control theory and engineering Master	Beijing
Supervised by Prof. Xiangyang Ji	
Tsinghua University	Sep 2013 - Jun 2017
Department of Automation Bachelor	Beijing

RESEARCH

My research topic is mainly object pose estimation and its applications in SLAM, robotic grasping, part assembly, shape retrieval and deformation, etc. Here I list some main projects.

- Object pose estimation. Pose estimation refers to estimating the rotation, translation and size of the target object given RGB images or point clouds. I mainly develop geometry-aware methods for pose estimation in different scenes. SO-Pose for instance-level pose estimation in desktop scenes (1-2m). GPV-Pose, RBP-Pose, SSP-Pose are designed for category-level pose estimation in desktop scenes. OPA-3D is proposed to handle traffic scenes (0-80m). U-RED and KP-RED focus on indoor scenes(0-10m).
- Applications of pose estimation. MonoGraspNet employs pose estimation in robotic grasping. U-RED introduces shape retrieval and deformation based on pose estimation for indoor scene understanding.

WORK EXPERIENCES

Google Munich	Aug 2023 - Present
Student researcher	Munich
Large language models + pose estimation + robotics	
SUMSUNG Beijing Research Center	Jun 2016 - Sep 2016
Student researcher	Beijing
SLAM on mobile platforms	

HONORS & AWARDS

Academic Excellence Awards	2015-2016 , 2016-2017
Overall Best Segmentation Method Award, ZebraPoseSAT, ECCV2022 BOP Challenge	2022

PUBLICATIONS

* co-first authors, + corresponding author

Pose Estimation in Different Scenes:

1 *Di, Y.*, Manhardt, F., Wang, G., Ji, X., Navab, N., & Tombari, F. (2021). SO-Pose: Exploiting Self-Occlusion for Direct 6D Pose Estimation. the IEEE/CVF International Conference on Computer Vision (ICCV) 2021.

- 2 **Di, Y***, Zhang, R*, Lou, Z., Manhardt, F., Ji, X., Navab, N., & Tombari, F. (2022). Gpv-pose: Category-level object pose estimation via geometry-guided point-wise voting. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 6781-6791).
- 3 Zhang, R*, **Di, Y***, Lou, Z., Manhardt, F., Tombari, F., & Ji, X. (2022, October). Rbp-pose: Residual bounding box projection for category-level pose estimation. In Computer Vision—ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part I (pp. 655-672). Cham: Springer Nature Switzerland.
- 4 Zhang, R*, **Di, Y***, Manhardt, F., Tombari, F., & Ji, X. (2022, October). SSP-Pose: Symmetry-Aware Shape Prior Deformation for Direct Category-Level Object Pose Estimation. In 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 7452-7459). IEEE.
- 5 Su, Y*, **Di, Y***, Zhai, G., Manhardt, F., Rambach, J., Busam, B., ... & Tombari, F. (2023). OPA-3D: Occlusion-Aware Pixel-Wise Aggregation for Monocular 3D Object Detection. IEEE Robotics and Automation Letters (RAL).
- 6 Zaccaria, M., Manhardt, F., **Di, Y**, Tombari, F., Aleotti, J, Giorgini, M. Self-Supervised Category-level 6D Object Pose Estimation With Optical Flow Consistency (2023). IEEE Robotics and Automation Letters (RAL).

Applications of Pose Estimation:

- 1 **Di, Y***, Zhang, C*, Zhang, R*, Manhardt, F., Su, Y., Rambach, J., Stricker, D., Ji, X., Tombari, F. (2023) U-RED: Unsupervised 3D Shape Retrieval and Deformation for Partial Point Clouds. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)
- 2 Zhai, G., Huang, D., Wu, S. C., Jung, H., **Di, Y***, Manhardt, F., ... & Busam, B. (2022). MonoGraspNet: 6-DoF Grasping with a Single RGB Image. In 2023 IEEE International Conference on Robotics and Automation (ICRA). IEEE.
- 3 Zhang, C*, **Di, Y***, Zhang, R., Zhai, G., Manhardt, F., Tombari, F., & Ji, X. (2023). DDF-HO: Hand-Held Object Reconstruction via Conditional Directed Distance Field. NeurIPS 2023.

3D Reconstruction:

- 1 **Di, Y**, Morimitsu, H., Gao, S., & Ji, X. (2019). Monocular piecewise depth estimation in dynamic scenes by exploiting superpixel relations. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) (pp. 4363-4372).
- 2 **Di, Y**, Morimitsu, H., Lou, Z., & Ji, X. . A unified framework for piecewise semantic reconstruction in dynamic scenes via exploiting superpixel relations. In 2020 IEEE International Conference on Robotics and Automation (ICRA) (pp. 10737-10743).
- 3 Zhang, C., Lou, Z., **Di, Y**, Tombari, F., & Ji, X. (2022). SST: Real-time End-to-end Monocular 3D Reconstruction via Sparse Spatial-Temporal Guidance. ICME2023.
4. Zhai, G., Örneke, E. P., Wu, S. C., **Di, Y***, Tombari, F., Navab, N., & Busam, B. (2023). CommonScenes: Generating Commonsense 3D Indoor Scenes with Scene Graphs. NeurIPS 2023.

Others:

- 1 Zhu, D., Zhai, G., **Di, Y***, Manhardt, F., Berkemeyer, H., Tran, T, Navab, N., Tombari, F., Busam, B. IPCC-TP: Utilizing Incremental Pearson Correlation Coefficient for Joint Multi-Agent Trajectory Prediction (2023), In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)
- 2 Li M, Du Z, Ma X, Gao K, Dong W, **Di Y**, Gao Y. System Design and Monitoring Method of Robot Grinding for Friction Stir Weld Seam. Applied Sciences. 2020; 10(8):2903. <https://doi.org/10.3390/app10082903>