

VO Methods, datasets & details.

Method	Datasets used	Characteristic	References
① VOLDOR	TUM RGB-D, KITTI	Utilizes externally estimated optical flow fields instead of hand-crafted feature correspondences. Employs a probabilistic model with generalised EM formulation for joint inference of camera motion, pixel depth & motion track confidence	Arxiv.org
② Coded VO Coded VO	ICL NUM	* Introduces custom optics to physically encode metric depth information into imagery addressing the scale ambiguity problem in monocular VO. Achieve state of the Art performance with a known scale.	Arxiv.org
③ TSformer VO	KITTI	Treats monocular vision VO as a video understanding task. Employs a transformer based architecture with spatio-temporal self attention mechanisms to extract features from clips & estimate motions in an end to end manner.	Sensors 2023 Arxiv.org
④ Pseudo Liden for VO	KITTI	Converts depth maps generated ^{by} stereo images into 3D coordinates point clouds (pseudo LIDAR). Utilizes a projection aware dense odometry pipeline to fully exploit the point cloud	Arxiv.org papers with code 2024

Method	Dataset	Characteristics	References
⑤ Attention based Deep-Learning Architecture for Monocular VO	Unspecified	Combines a Convolutional Neural Network (CNN) for image feature extraction with a long short-Term Memory (LSTM) network and a multi-head attention module for video sequence modeling. Designed for real-time monocular visual odometry, particularly in GPS-free drone navigation.	ARxiv.org
⑥ Multimotion Visual Odometry	Oxford Multimotion Dataset (AMP) KITTI	Estimates the full SE(3) trajectory of every motion in the scene, including sensor egomotion, without relying on appearance based information. Extends traditional visual odometry pipelines with multimotion segmentation and tracking techniques.	Robotic esp
⑦ DF-VO	Oxford Robocar, 4 seasons, Internal Singapore Dataset.	Depth & flow for $\frac{1}{2}$ VO (DF-VO) algorithm evaluated for robustness in rainy weather conditions. Performs well for short-range distances (<500m) in Urban Driving scenarios under rain.	ARxiv.org
⑧ Deep Patch Visual Odometry	Unspecified	Utilizes a novel recurrent network architecture designed for tracking image patches across time.	NeurIPS 2023

Method	Dataset	Characteristics	References
Edge VO (9)	TUM datasets	An edge-based visual odometry approach that efficiently selects a small set of edges to improve computational efficiency without sacrificing accuracy.	arXiv 2023
(10) NeRF - VO	Synthetic & Real World datasets	Integrates learning-based sparse visual odometry for low-latency camera tracking with a neural radiance field scene representation for dense reconstruction & novel view synthesis.	arXiv 2023
(11) Brightness-Guided Hybrid Transformer for Visual Odometry	Unspecified	Employs a hybrid transformer architecture guided by brightness information to enhance monocular visual odometry performance.	arXiv 2025
(12) Attenuation-Aware Weighted Optical Flow with Medium Transmission Map for Learning-based VO in Underwater Terrain	unspecified	Integrates principles of underwater optical imaging to manipulate optical flow estimation, improving learning-based monocular VO in underwater environments.	Papers with Code 2024

Tracking Thread

Role: processing of incoming frames to estimate the camera pose & manage keyframes. Here's a breakdown of the

tracking thread's implementation.

- * Extract ORB features
- * System Initialization.
- * Pose estimation
- * Local Map Tracking.
- * Keyframe decision.
- * Relocalization

Main System Components of ORB SLAM 3.

