LooperRobotics®

Looper Data Acquisition Package



Looper Data Acquisition Package is an EGO + UMI wearable multi-view capture system designed for robotics foundation models and embodied AI training data generation. By synchronously recording the head-mounted first-person view and close-range views of both hands, the system turns natural human operations into high-quality robot learning data for annotation, reconstruction, and policy training.

System Components

Module	Recommended Configuration	Function
Head Camera	1 x Insight 9	The headset with Insight 9 weighs 360g in total. It captures Ego-view RGB-D, head motion and 6DoF pose for global scene understanding.
Hand Cameras	2 x Insight 3	Each wrist-mounted Insight 3 module weighs only 158g. The two hand cameras capture UMI close-range hand interactions, complementing the Ego view with fine-grained manipulation data.
Backpack Unit	Jetson Orin NX 128GB NVMe SSD (Standard) V-Mount Battery	The backpack unit weighs approx. 1.8kg and integrates edge computing, local storage, and mobile power. It runs the acquisition program and stores ROS 2 bags, RGB-D, IMU, VIO pose, calibration parameters, and task metadata. Local storage can be upgraded based on recording duration and data volume requirements. The battery supports approx. 150–170 minutes of continuous real-world data capture.
Software Environment	Ubuntu + ROS 2 data acquisition software	Provides unified device connection, recording control, acquisition status management, data directory organization, clip management, and post-processing workflow.

Built on Insight spatial perception cameras, the system uses Insight 9 to capture the EGO global view and head pose, while dual Insight 3 cameras capture UMI close-range hand manipulation details. The system combines Jetson edge computing and local storage to collect, record, and organize multimodal data, including RGB, stereo/depth, IMU, VIO pose, and device calibration information.

Post-Processing Software

Function	Description
Multi-Camera Coordinate Alignment	Automatically performs temporal synchronization and spatial alignment between the EGO head view and UMI hand views, organizing multi-stream RGB-D, IMU, and VIO pose data into a unified reference frame for joint annotation, RGB-D reconstruction, and policy training.
Data Recording & Editing	Supports unified recording and clip-based management of multi-stream data (RGB, depth, IMU, VIO pose). It helps operators quickly remove low-quality segments caused by frame drops, occlusion, or inconsistent actions, improving dataset usability and reducing manual screening costs from the source.
Trajectory Scoring	Leverages onboard VSLAM pose estimation from Insight cameras to automatically evaluate motion smoothness, trajectory continuity, and view completeness. Provides objective quality criteria for data ingestion and helps flag suboptimal samples efficiently.
Trajectory Optimization	Uses high-quality camera pose data to perform smoothing, pose correction, and temporal alignment on jittery, drifting, or low-score trajectories, reducing manual post-processing and making the data more suitable for downstream training.

With a closed-loop "capture → QC → optimize → export" pipeline, the tool quickly turns raw data into high-quality, model-ready training assets.

Core Advantages

EGO + UMI Multi-view Capture

The head view captures the task environment, while the hand views preserve manipulation details, improving the completeness and usability of egocentric operation data.

Onboard VSLAM Output

RGB, depth, IMU, and VIO pose are recorded simultaneously, enabling annotation, reconstruction, and training workflows beyond 2D video.

Optimized for Near-field Manipulation

Insight 3 is ideal for tabletop, arm-level, and close-range fine manipulation within its 0.06m-5m ideal depth range, ensuring every detail of hand manipulation is comprehensively and precisely captured.

Scalable Deployment

The headband, wristbands, Jetson, upgradable NVMe storage, and V-Mount battery backpack form a standardized acquisition kit. With approximately 150–170 minutes of continuous recording, the system supports multi-unit replication, operator training, and standardized data collection workflows.

Post-processing & Quality Control

The post-processing software supports data recording and editing, trajectory scoring, and trajectory optimization, helping teams quickly identify low-quality clips after capture, improve trajectory stability, and convert raw EGO + UMI data into model-ready data for annotation, reconstruction, and training.

Open Interfaces

The system can export RGB-D, IMU, VIO pose, calibration parameters, task metadata, and optimized trajectory data, which can be connected to hand detection, object detection, pose estimation, RGB-D reconstruction, action segmentation, and robot learning data conversion pipelines.

Core Component Specifications





	Insight 9	Insight 3
Product Positioning	Mid- to long-range VSLAM spatial perception camera	Near-field high-precision spatial perception camera
Recommended Role in the Package	Head-mounted main camera / large-workspace view	Left and right hand cameras / near-field manipulation
Baseline	100mm	35mm
Ideal Depth Range	0.3m-30m	0.06m-5m
Depth Accuracy	<2% at 3m	<1% at 0.5m
Depth Output	Resolution up to 544 x 640; frame rate up to 15 fps	Resolution up to 544 x 640;frame rate up to 15 fps
RGB Output	Resolution up to 3840 x 2160; 1920 x 1080 @30fps recommended	Resolution up to 1088 x 1280; 1088 x 1280 @20fps recommended
Weight	176g	103g
Wearable Weight	Head: 360g (headband 184g + camera 176g)	Wrist: 158g (wristband 55g + camera 103g)



Insight 9 sees the big picture, ensuring head motion trajectories and environmental structure are fully captured; while Insight 3 captures the details, ensuring every subtle manipulation at tabletop level is precisely recorded. Together, they form a long-range & near-field, wide-view & fine-detail combination that maximizes the data quality of multi-view capture. After collection, the system can export RGB, depth, IMU, VIO pose, calibration parameters, and task metadata, providing a unified foundation for hand tracking, hand-object interaction analysis, RGB-D reconstruction, behavior annotation, and robot learning training data generation.

To learn more about the Insight Series, please contact us for full product briefs and technical specifications.