Ego-Exo Workplace Tasks

Dataset specification

GR-EGO-02 · Physical AI · Available · updated 2026-11-02

1 - Overview

Synchronized first-person and multi-view video of real work in factories, hotels and offices.

Real workers doing their real jobs - small-part assembly and packaging on factory floors, housekeeping in hotel rooms, tidying and cleaning in offices and reception areas - recorded by a head-mounted camera and three fixed cameras at the same time. Sessions are synchronized and calibrated. Annotators describe each action in free text with timestamps, and a pose model estimates body keypoints; raw footage can also be licensed without annotations.

- **Hours of video:** 4,300+
- **Synchronized views:** 4
- **Frame rate:** 30 fps
- **Hours per participant:** <= 10
- **Modalities:** Egocentric video, 3 synchronized exo views, Camera calibration, Action segments, Model-estimated body keypoints

2 - Use cases

- Pre-training vision-language-action (VLA) and humanoid policies on real work
- Ego-exo correspondence and view-invariant representation learning
- Procedure, action and task recognition in workplaces
- Human pose and motion modelling from calibrated multi-view video

3 - Capture

- **Head-mounted camera:** Wide-angle action camera, both hands in view during manipulation
- **Fixed cameras:** 3 static cameras: one full-body view at 2–3 m, one frontal view within 2 m, one on the hands
- **Resolution:** At least 2 MP (1920 × 1080) per camera
- **Frame rate:** 30 fps
- **Synchronization:** Dynamic QR code shown to every camera at the start of each session
- **Calibration:** Calibration-board intrinsics and extrinsics per session

4 - Coverage

- **Manufacturing:** ~ 78% - assembly, packaging, sorting, labelling and inspection
- **Hospitality:** ~ 13% - bed making, cleaning and tidying in hotel rooms and lobbies
- **Offices and public spaces:** Meeting rooms, reception, offices, lounges and dining areas
- **Participants:** Real workers in their own workplaces; at most 10 hours per person
- **Content rules:** No staged scenes or actors; no idle walking; tasks occur naturally
- **Task concentration:** Repetitive bench work dominates; the largest single tasks run 300–450 hours each. Per-task hours on request

5 - Annotation

- **Action segments:** Timestamped free-text descriptions written by annotators in Chinese; task names are not normalized to a fixed taxonomy. English translation on request
- **Body keypoints:** 17-point (COCO) body keypoints per fixed camera, estimated by a pose model with per-point confidence
- **3D pose:** Triangulated from the calibrated fixed cameras; relative scale, metric conversion in preparation
- **Hand keypoints:** Available as a custom add-on
- **Acceptance review:** Per-session checks of sync, camera coverage, image quality and content

6 - Compliance

- **Consent:** Written participant consent covering AI training
- **Privacy:** Face and screen blurring available before delivery
- **Provenance:** Unique session ID with capture metadata and acceptance record
- **Delivery options:** Raw synchronized video with calibration, with or without annotations

7 - Data schema

- `session_id` **(string):** Unique capture session identifier.
- `ego` **(mp4):** Head-mounted video.
- `exo_cam1` … `exo_cam3` **(mp4):** Synchronized fixed-camera videos.
- `calibration` **(yaml):** Per-camera intrinsics and extrinsics.
- `keypoints_2d` **(json):** Per-camera, per-frame 17-point body keypoints with model confidence.
- `keypoints_3d` **(json):** Per-frame triangulated 17-point body keypoints (relative scale).
- `actions` **(json):** Action segments: start, end and description.
- `metadata` **(json):** Scene, task, devices, duration and acceptance status.

8 - Directory layout

```
ego-exo-workplace/            # indicative; final layout confirmed per delivery
|-- manifest.csv
`-- sessions/<scene>/<session_id>/
    |-- ego/
    |-- exo_cam1/
    |-- exo_cam2/
    |-- exo_cam3/
    |-- calibration.yaml
    |-- keypoints_2d.json
    |-- keypoints_3d.json
    |-- actions.json
    `-- metadata.json
```

9 - Evaluation sample

- **Availability:** On request
- **Size:** On request
- **Formats:** MP4, YAML, JSON
- **Contents:** Raw multi-view sessions from manufacturing and hospitality scenes
- **Contents:** Calibration files, action segments and body keypoints

10 - License and contact

- **Graviti Commercial Data License:** Licensed per hour of footage, with rights to train and commercially deploy models. Custom collection to your task list is available.
- **Dataset page:** https://www.graviti.com/datasets/ego-exo-workplace/
- **Sales:** sales@graviti.com