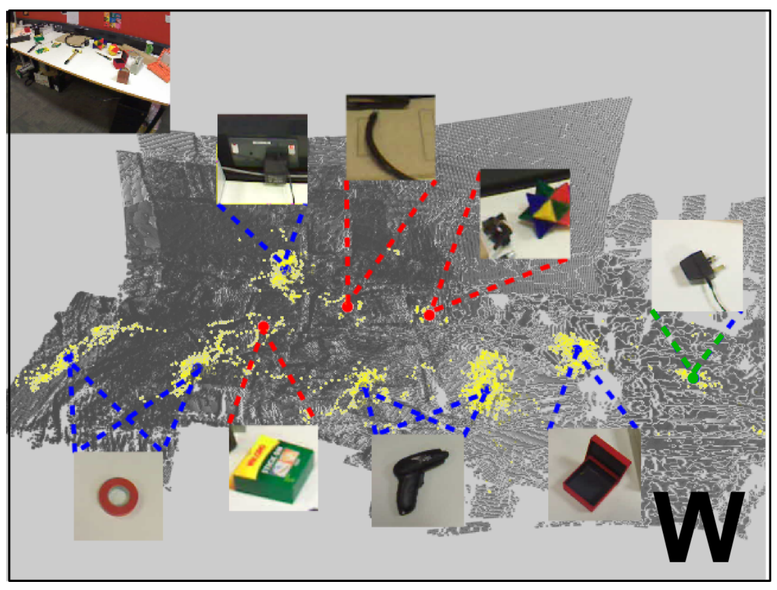
You Do, I Learn: Multi-User Egocentric Online System for Unsupervised Assistance on Object Usage

Dima Damen, Teesid Leelasawassuk, Osian Haines, Andrew Calway, Walterio Mayol-Cuevas

Overview

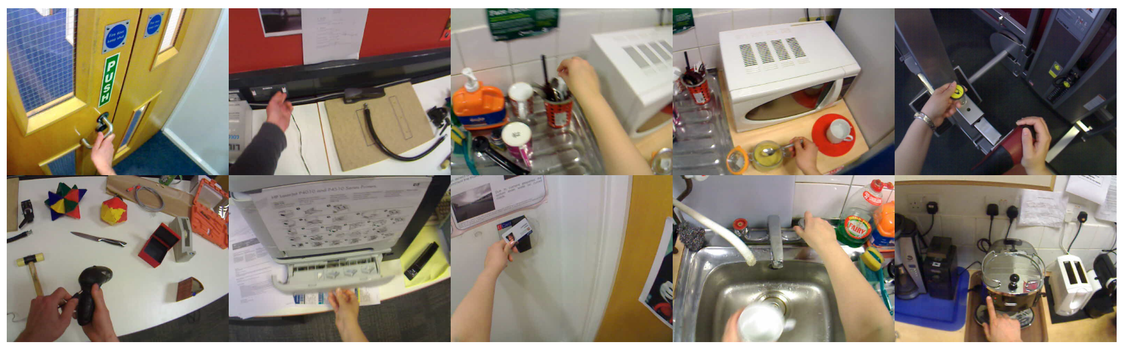
- Online fully unsupervised
- Using egocentric video from multiple users performing tasks
- Discover task-relevant objects
- Extract guidance videos on how objects are used
- Distinguish multiple modes of interaction with the same object
- Provide guidance triggered by user's gaze/attention
- Assistive mode demonstration available on Google Glass

Learning Mode



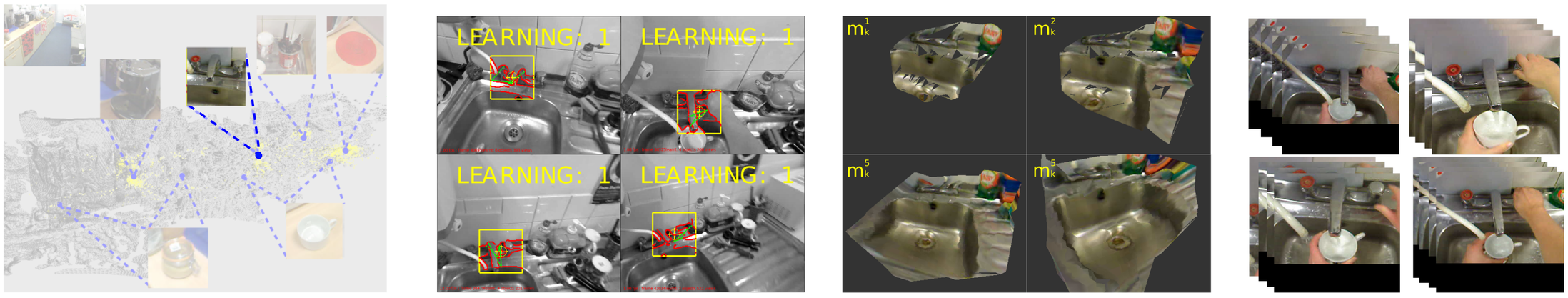
Dataset

Bristol Egocentric Object Interactions Dataset (BEOID)

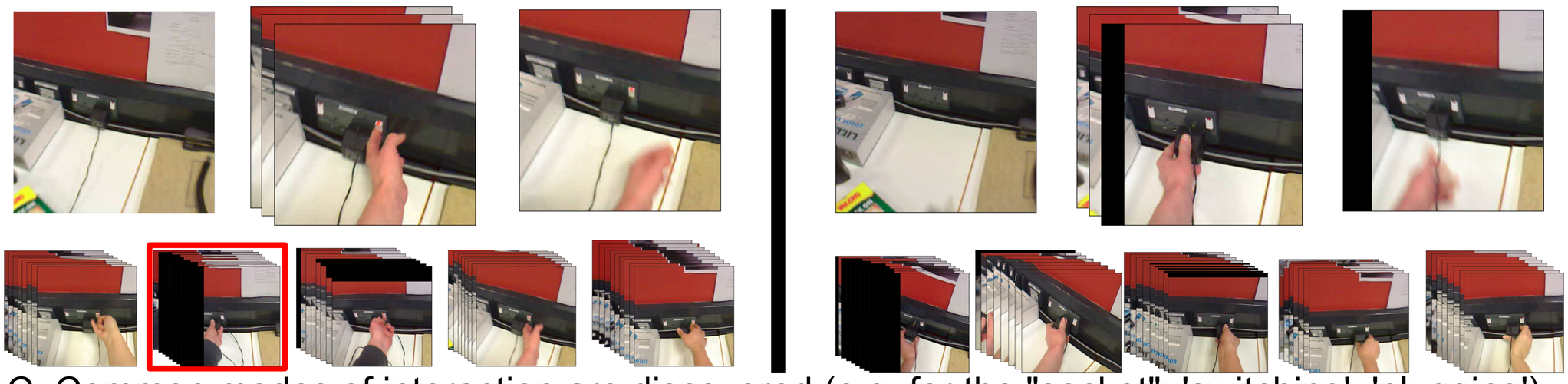


Five operators, six locations: kitchen, workspace, laser printer, corridor with a locked door, cardiac gym and weight-lifting machine with synchronised gaze, recorded using ASL Mobile Eye XG

A. Static and moveable objects are discoverd by clustering appearance, position and attention.



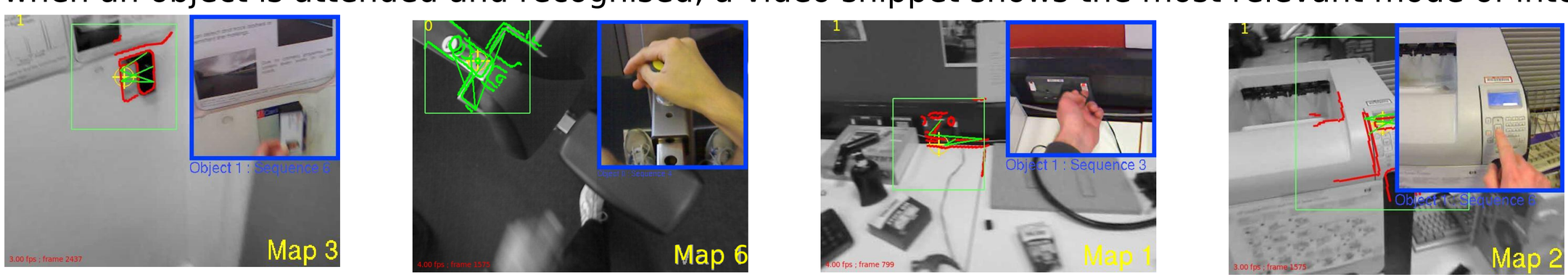
B. Multiple users enable learning multiple views, building 3D model and collecting guiding video snippets.



C. Common modes of interaction are discovered (e.g. for the "socket": 'switching', 'plugging')

Assistive Mode

when an object is attended and recognised, a video snippet shows the most relevant mode of interaction.



References

- Damen, Dima and Leelasawassuk, Teesid and Haines, Osian and Calway, Andrew and Mayol-Cuevas, Walterio (2014). **You-Do, I-Learn: Discovering Task Relevant Objects and their Modes of Interaction from Multi-User Egocentric Video**. British Machine Vision Conference (BMVC).
- Damen, Dima and Haines, Osian and Leelasawassuk, Teesid and Calway, Andrew and Mayol-Cuevas, Walterio (2014). **Multi-user egocentric Online System for Unsupervised Assistance on Object Usage**. ECCV Workshop on Assistive Computer Vision and Robotics (ACVR).