Monocular Visual Scene Understanding: Understanding Multi-Object Traffic Scenes C. Wojek, S. Walk, S. Roth, K. Schindler, B. Schiele (PAMI 2013)





- A probabilistic 3D scene model for multi-class object detection, object tracking, scene labelling, and 3D geometric relations using monocular video as input
- Extension of Wojek2010ECCV¹ with explicit occlusion reasoning for tracking objects that are partially occluded or that have never been observed to their full extent
- Evaluated on ETH-Loewenplatz, ETH-Linthescher, ETH-PedCross2, MPI-VehicleScenes
- Robust performance due to
 - a strong tracking-by-detection framework with tracklets
 - exploiting 3D scene context by combining multiple cues
- Explicit occlusion reasoning improves results on all sequences.
- Long-term tracking with an HMM does not lead to additional gains.
- Improvement over state-of-the-art object detectors, a stereo-based system, a competing monocular system, basic Kalman filters

¹Monocular 3D Scene Modeling and Inference: Understanding Multi-Object Traffic Scenes, ECCV 2010