

Acknowledgments

This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2002/1 “Science of Intelligence” – project number 390523135.

7. Supplementary

7.1. Video

We encourage readers to watch the accompanying video, which explains the method more intuitively than the still images in the paper.

7.2. Analysis of Variations

Figure 8 shows an example histogram of variation magnitudes per event, using the box scene from EVIMO2 (*scene13_dyn_test_00* in Fig. 5). The order of raw variations is 10^{-5} , while we scale them to $[0, 255]$ for visualization as an MVI.

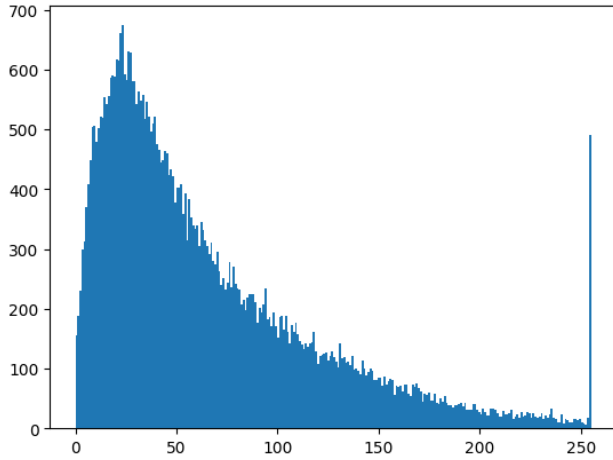


Figure 8. Histogram of the magnitude of variations per event in one sequence example from EVIMO2. The magnitude is scaled to $[0, 255]$ for visualization as an image (i.e., MVI).

8. Sensitivity Analysis

In the datasets collected from real-world environments (Secs. 4.4 and 4.5), significantly more noise is present compared to ideal conditions. The noise yields remnants of first variations being introduced during the separation of estimated motion, and hence, blurs the boundaries of motion segmentation. To address this, one may apply a simple de-noising filter using a Gaussian kernel on the IWE and classify using a certain threshold.

We evaluate the effect of the Gaussian intensity σ and the threshold of the filter on the bounding box accuracy in Fig. 9, using the EVIMO2 dataset. The experimental results show that the IoU consistently remains above 0.7 for the tested parameters (see also Sec. 4.4). We confirm that the proposed method consistently achieves the detection accuracy, even with inversely adjusting σ and the threshold.

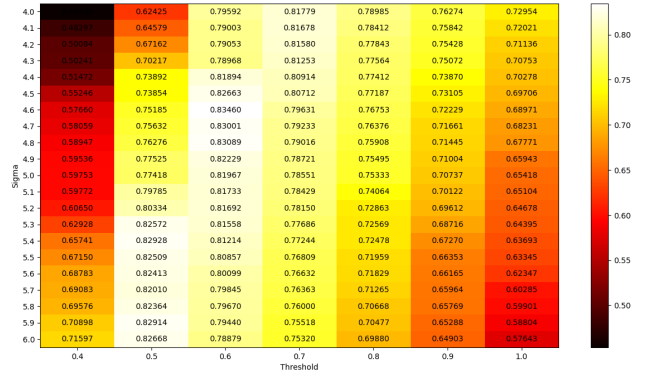


Figure 9. Sensitivity study on the effect of threshold and sigma.