

Motion-from-Blur: 3D Shape and Motion Estimation of Motion-blurred Objects in Videos – Supplementary Material –

Denys Rozumnyi^{1,4} Martin R. Oswald^{1,2} Vittorio Ferrari³ Marc Pollefeys¹

¹Department of Computer Science, ETH Zurich ²University of Amsterdam

³Google Research ⁴Czech Technical University in Prague

{denys.rozumnyi,martin.oswald,marc.pollefeys}@inf.ethz.ch vittoferrari@google.com

1. Evaluation on foreground regions

In the main paper, PSNR and SSIM metrics are calculated over the tightly cropped input frame. However, we can also evaluate on the foreground regions defined by the object mask in the ground-truth, which occupies even less area. In Tab. 1 we now evaluate SfB and our MfB only on the ground-truth foreground region. The improvement made by MfB over SfB is larger but the overall scores are lower than when evaluating on the cropped frame. This is expected since background regions are easier as they are static.

2. Synthetic noise

We run an experiment with additive Gaussian noise, ranging from small to extremely high standard deviation (std). To quantify degradation, in Fig. 1 we measure the same performance metrics and divide them by the original ones without noise (on the Falling Objects and TbD-3D datasets). The method appears robust, *e.g.* when adding noise at 5% of the full RGB range, performance degrades by only 1%-4% depending on the metric on the TbD-3D dataset.

3. Generalization ability

Based on the experiments, we observed that our method can handle various shapes such as a pen, a key, and an Aerobie ring. Since we deform a set of mesh prototypes, our method is not able to generalize to other topologies. Therefore, objects can be quite general once they have the same topology as in the prototype set. Shapes that are too difficult to handle are very spiky objects, *e.g.* a chair with four legs. However, even in such cases, we can add more prototypes, which will cover more object classes and make the method more general. We leave this for future work.

	Falling Objects		TbD-3D Dataset	
	PSNR↑	SSIM↑	PSNR↑	SSIM↑
SfB	21.77	0.597	21.16	0.528
MfB (ours)	23.43	0.649	22.91	0.612

Table 1. Evaluation on foreground regions only.

4. Video processing

The YouTube video, as all other videos, was processed fully automatically. In each frame, we run a state-of-the-art fast moving object detection method, which outputs a bounding box. Outside the bounding box, the input video is not used and is kept unchanged. In the cropped region, we assume that the camera and the background are stationary.

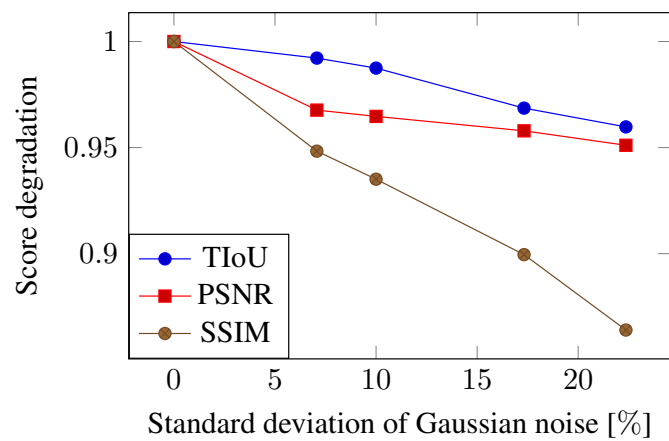
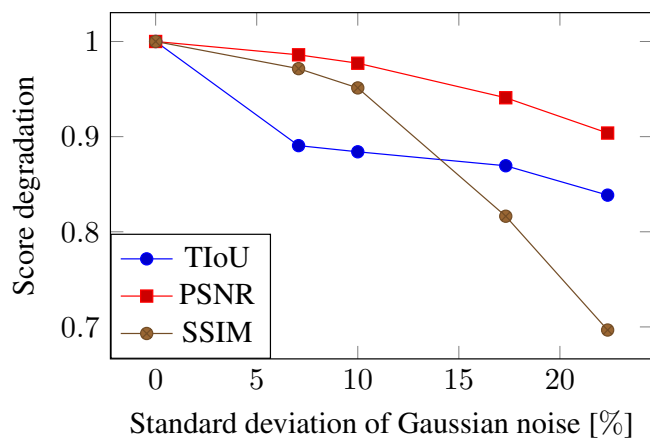


Figure 1. **Score degradation with increasing Gaussian noise.** Left: Falling Objects dataset. Right: TbD-3D dataset.