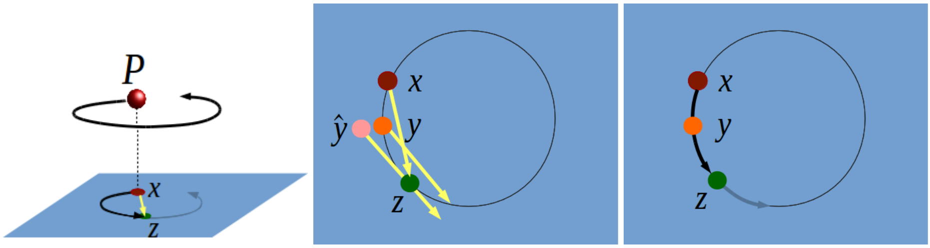
Stereo Video Deblurring
– Supplemental Material –

Anita Sellent[1,2], Carsten Rother[1], and Stefan Roth[2]

[1] Technische Universität Dresden, Germany
[2] Technische Universität Darmstadt, Germany

1 Details on Homography-Based Blur Kernels

One of the key contributions in our stereo video deblurring is to employ 3D scene flow to induce blur kernels based on homographies. As the difference to inducing blur kernels from an optical flow field may seem subtle but increases deblurring performance considerably, we schematically illustrate the difference of the two ways of generating blur kernels in Fig. 10.

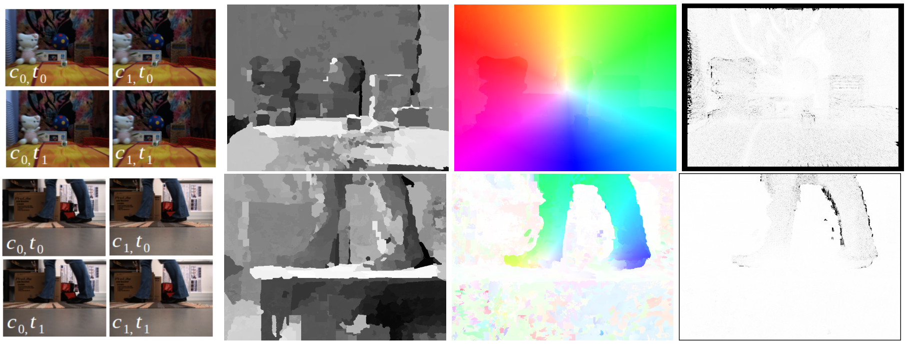


(a) 3D motion

(b) Image plane projection: Blur kernels from linear displacements

(c) Image plane projection: Blur kernels from homographies

Fig. 10. Inducing blur matrices: (a) Assume that 3D point P moves with a constant rigid body motion in 3D, *e. g.* with a yaw motion. The projection of this motion to the image plane (blue) is a circular trajectory. The corresponding 2D ground truth displacement (yellow), however, is a vector in the image plane that connects start and end point of the motion during a time interval. (b) Using optical flow, the blur kernel at a location z is approximated by identifying all pixels that, according to their spatially-variant displacement, pass through z during the time interval. The image content at point x is correctly identified as passing through z. The image content at point y is not identified correctly as its 2D displacement passes z at a distance. Instead, the image content at point $\hat{y}$ is erroneously identified as passing through z, even though $\hat{y}$ has a different distance to the rotation center than z. This results in the distorted kernels shown in Fig. 3b. (c) Assuming 3D points in the vicinity of P to form a plane, we employ 3D homographies to generate blur kernels. The blur kernel at location z is thus formed by the image points whose trajectory, according to the homography, passes through z during the time interval. Consequently, y is correctly identified as passing through z, leading to the kernels shaped as circular arcs shown in Fig. 3a



(a) Input (b) Over-segmentation (c) 2D projection of scene flow (d) Final occlusion weights

Fig. 11. Stereo video deblurring: For two consecutive frames of stereo video (a) we use the scene flow approach of Vogel *et al.* [6] to compute an over-segmentation into planar patches with constant 3D rigid body motion (b). For the real scenes shown, we obtain $1000 - 1500$ segments per frame-pair. For adjacent planes, the projection onto the image plane results in smoothly varying optical flow (c). At object boundaries, discontinuities in the flow coincide with segment boundaries. We use occlusion information from the scene flow to initialize object boundary weights. The weights are refined iteratively to downweight pixels that do not satisfy our image formation model (d)

2 Over-segmentation of Non-Planar Scenes

To enable the use of homographies, we assume that the scene consist of planes. For objects of arbitrary shape, we approximate their geometry with a collection of planar patches. Figure 11b shows the piecewise planar approximations for real scenes. In all our experiments we obtained approximations with $1000 - 1500$ planar patches of different 3D motions for images of size 640×480 pixels.

3 Deblurring Planar Patches

In the main paper we show that homography-based deblurring improves the PSNR for a planar texture undergoing 3 different 3D motions. Here we provide additional results using other textures and types of motions to show that the improvement generalizes. Fig. 12 shows the different planar textures undergoing various 3D motions. The difference images (Figs. 12c and 12e) show that in all cases homography-based deblurring removes subtle errors that optical flow-based kernel approximation accumulates all over the texture. Additionally, some localized error of the initial optical flow estimation is also corrected in our stereo video deblurring (*e. g.* rows 2, 4, 6). We average the performance for each motion across the different textures and summarize the results in Table 4. We observe that scene flow-based stereo video deblurring outperforms all other methods for all motions. In particular, averaged over all textures, it outperforms kernels from 2D ground truth flow also in the case of the 2D '*upward*' motion. This is due to the fact that homographies are invertible and do not suffer from the same rasterization artifacts as kernel generation through forward warping with optical flow.

Table 4. Deblurring planar textures without motion discontinuities: We report the peak signal-to-noise ratio (PSNR) of the deblurred reference frame, the average endpoint error (AEP) of the optical flow, and the average disparity error (ADE). All values are averaged over 8 different planar textures. For all motions the use of scene flow increases deblurring accuracy compared to using 2D displacements. Best results (bold) are obtained with homography based deblurring

blur kernel source	ground truth 2D displacement	initial optical flow [46]		2D projection of scene flow		3D homographies	
	PSNR	AEP	PSNR	AEP	PSNR	ADE	PSNR
forward + roll	32.32	0.10	32.64	0.10	32.73	0.20	**33.57**
forward	34.22	0.10	34.60	0.09	34.83	0.22	**35.40**
upward	26.94	0.37	26.75	0.17	26.88	0.97	**26.95**
forward + yaw	29.40	0.15	29.43	0.14	29.43	0.40	**30.56**
yaw	26.86	3.01	25.37	0.47	26.81	0.28	**27.32**
lateral + pitch	28.26	0.15	28.13	0.20	28.23	0.02	**29.18**

4 Evaluation of the Structural Similarity Index

In the main paper we report the peak-signal-to-noise-ratio (PSNR) between sharp reference images and deblurred images. However, to be *e. g.* consistent with human perception, also other evaluation measures for deblurred images are used. In Table 5 we evaluate the Structural Similarity (SSIM) index [47] between deblurred images and sharp reference images. The evaluation confirms the favorable results of visual inspection and the evaluation with the PSNR: In all but the *'apples'* scene, the use of 3D homographies increases SSIM in comparison to using any form of 2D displacement for blur kernel generation. By suppressing locations with invalid image formation model, the influence of motion boundaries and erroneous scene flow computation can be compensated for and our full algorithm obtains consistently better results than deblurring based on 2D displacements.

Table 5. Deblurring with motion discontinuities: While Tab. 3 evaluates PSNR as error measure, here the Structural Similarity (SSIM) index [47] of deblurred synthetic scenes with motion discontinuities (top) and real scenes with the camera moving on a motorized rail (bottom) is evaluated. Also for this evaluation measure, our homography-based stereo video deblurring with motion boundary weighting (full) clearly outperforms monocular video deblurring with optical-flow induced blur kernels

	initial optical flow [46]	2D projection of scene flow	3D homographies	ours (full)	Kim and Lee [7]
apples	0.881	0.939	0.928	**0.950**	0.928
bunny	0.770	0.797	0.826	**0.834**	0.720
chair	0.817	0.853	0.876	**0.905**	0.835
squares	0.709	0.776	0.810	**0.869**	0.796
triplane	0.818	0.830	0.866	**0.869**	0.693
bottles	0.955	0.953	**0.972**	**0.972**	0.942
office	0.958	0.956	**0.968**	**0.968**	0.935
planar	0.961	0.961	0.975	**0.976**	0.945
toys	0.953	0.954	**0.967**	**0.967**	0.931

(a) Input (b) Using kernels from initial optical flow (c) Difference (b) to reference image (d) Using kernels from homographies (e) Difference (d) to reference image

Fig. 12. Deblurring various planar textures: For a planar texture blurred with 3D rigid body motion (a), deblurring with initial optical flow estimates (b) yields differences (c) to the sharp reference image due to wrongly estimated flow and due to approximated kernels. These errors can be reduced by deblurring with our image formation model that uses scene flow to compute more robust motion and more accurate blur matrices (d), (e)