

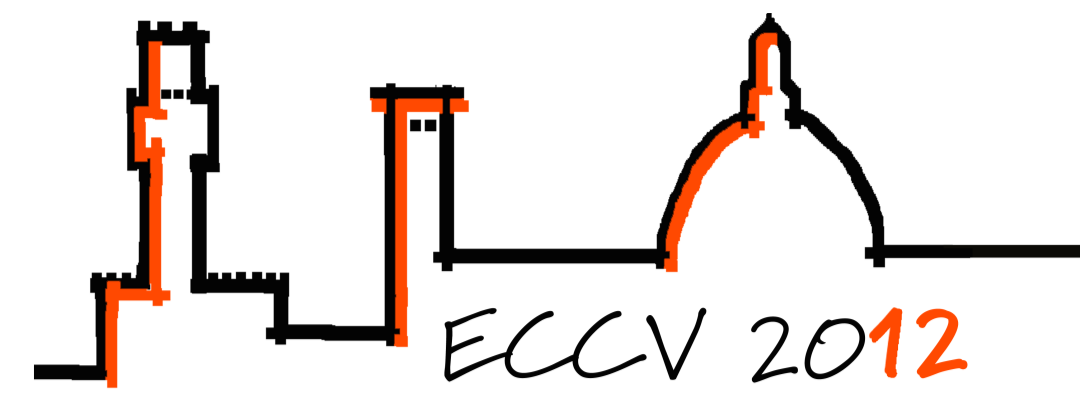


MAX-PLANCK-GESELLSCHAFT

Recording and playback of camera shake: benchmarking blind deconvolution

Rolf Köhler, Michael Hirsch, Betty Mohler, Bernhard Schölkopf and Stefan Harmeling

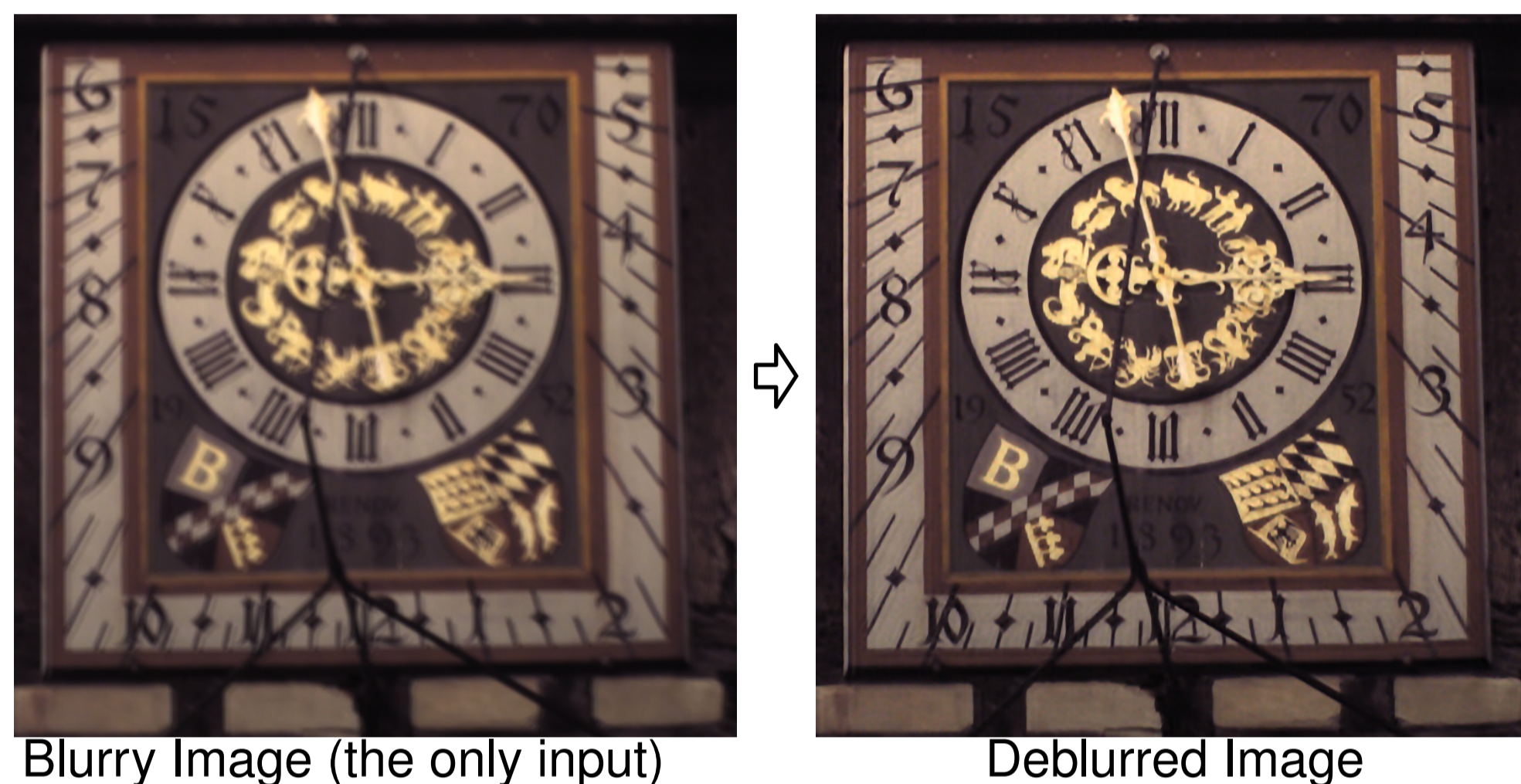
Department of Empirical Inference, Max Planck Institute for Intelligent Systems, Tübingen, Germany



Overview

- ▶ Creation of a benchmark dataset to compare new deblurring algorithms
- ▶ Comparison of 7 state-of-the-art blind deconvolution algorithms
- ▶ Analysing camera shake

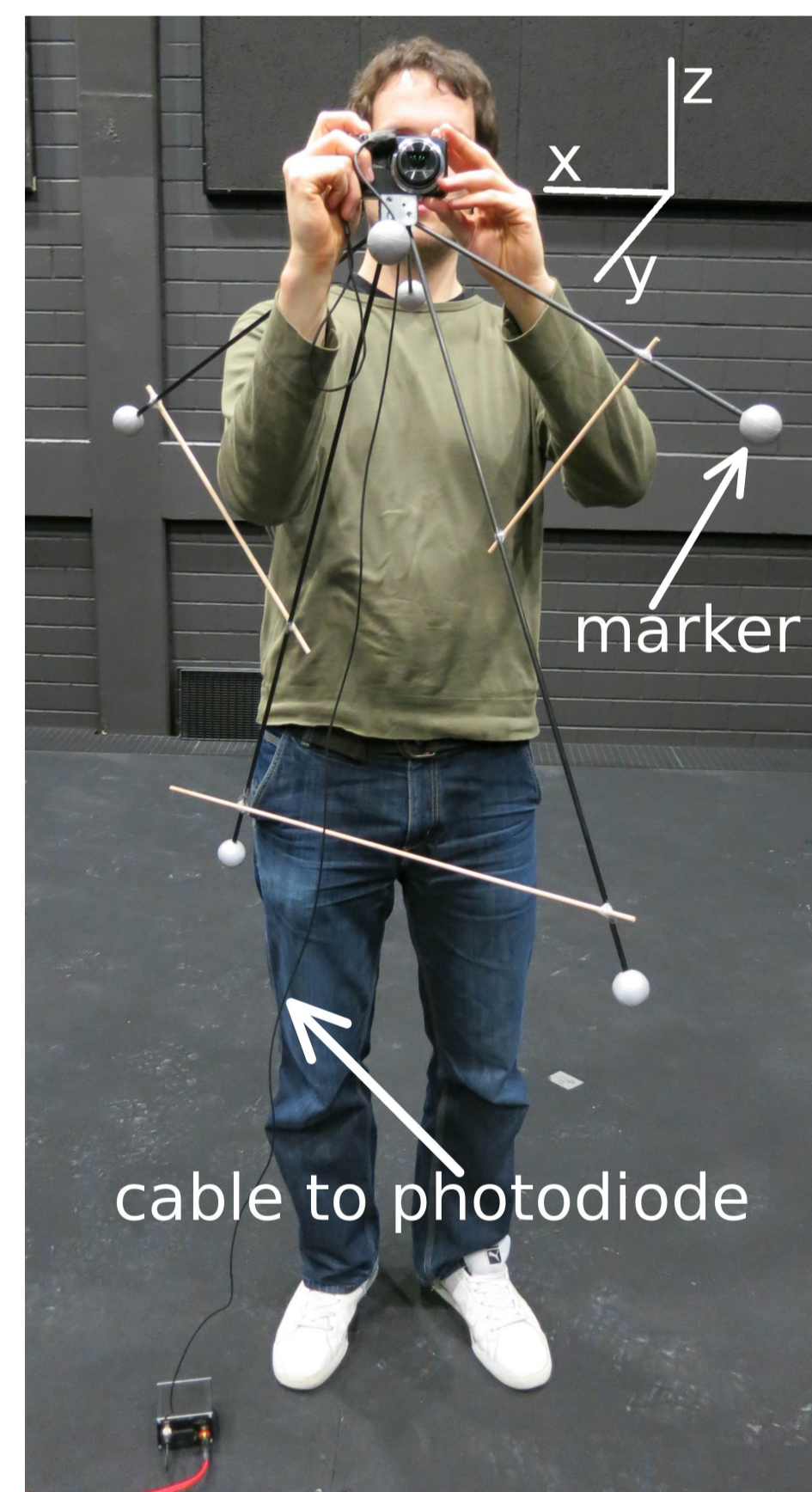
Blind Deconvolution



Blurry Image (the only input)

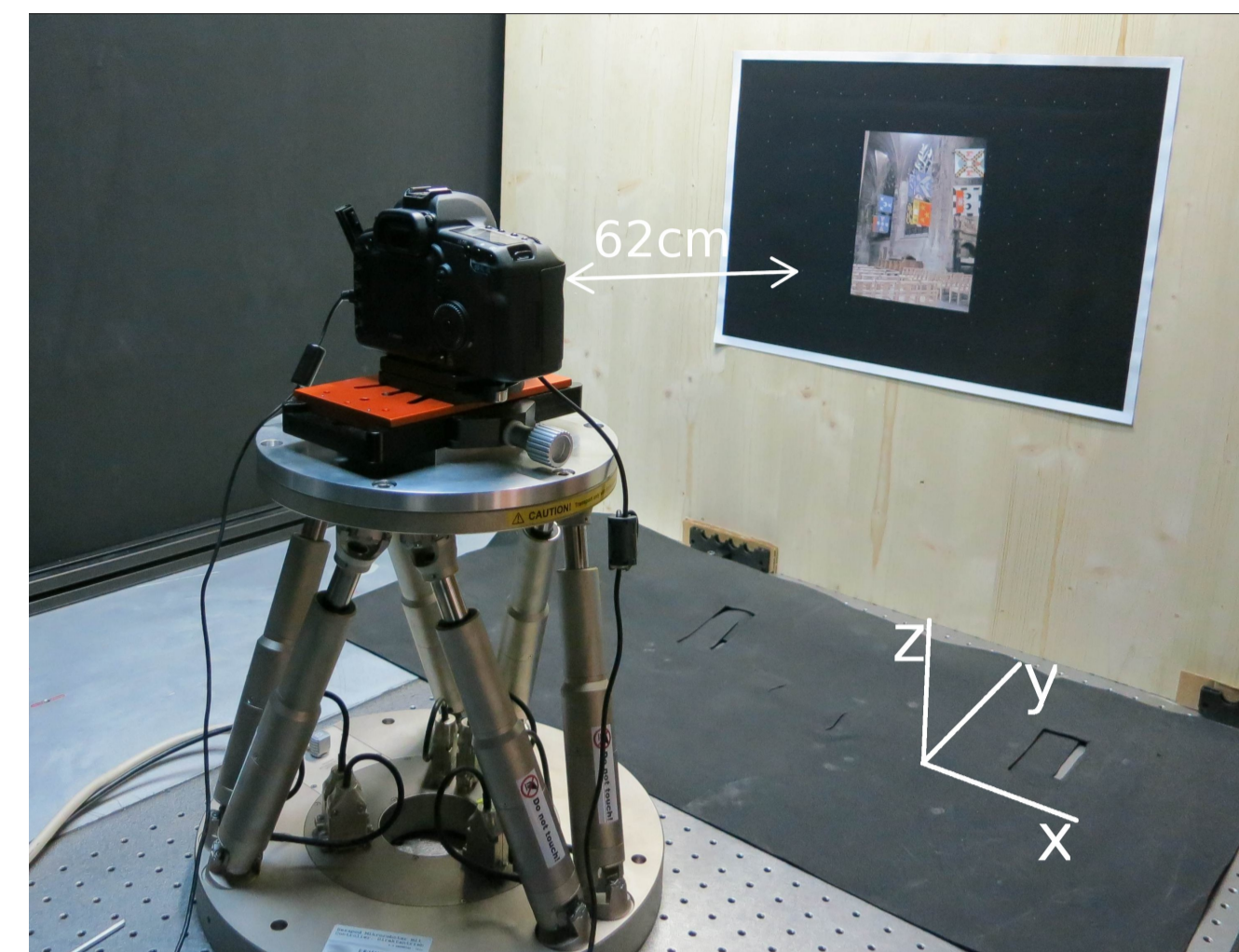
Deblurred Image

Recording Camera Shake



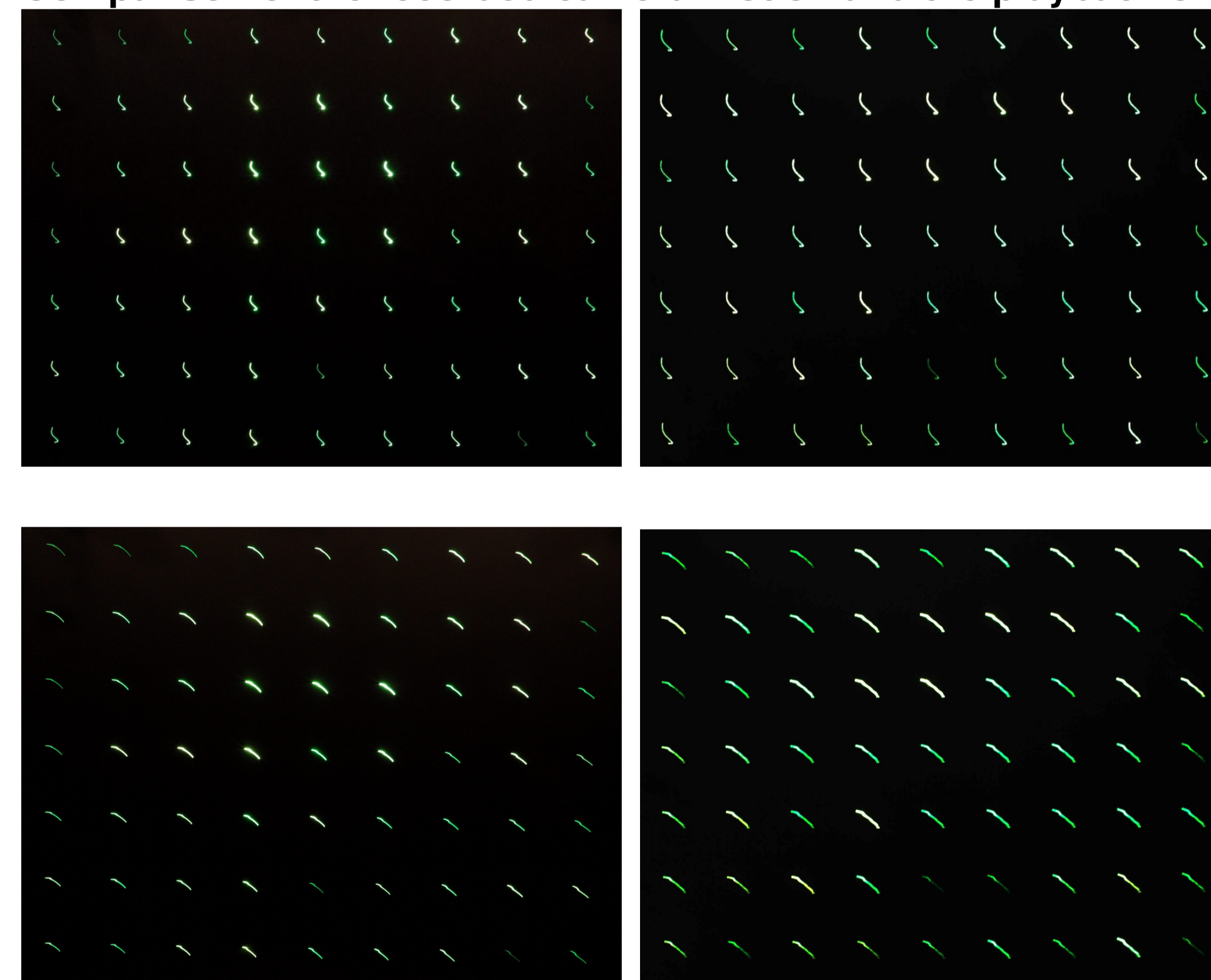
- ▶ camera shake was recorded holding a compact camera (Samsung WB600)
- ▶ exposure time of 1/3 sec
- ▶ recorded with 16 high-speed Vicon MX-13 cameras running at a frame rate of 500 Hz
- ▶ the cameras were calibrated to a cube of roughly 2.5m side length.
- ▶ 6 subjects were recorded, in total 40 recordings.

Playback of Camera Shake on a Hexapod



- ▶ minimum incremental motions of $3\mu\text{m}$ (x and y axis), $1\mu\text{m}$ (z axis) and $5\mu\text{rad}$ (rotations)
- ▶ repeatability $\pm 2\mu\text{m}$ (x and y axis), $\pm 1\mu\text{m}$ (z axis) and $\pm 20\mu\text{rad}$ (rotations).
- ▶ SLR camera (Canon Eos 5D Mark II), ISO 100, aperture f/9.0, exposure time 1sec, taking images in the Canon raw format SRAW2
- ▶ lens: Canon EF 50mm f/1.4

Comparison of the recorded camera motion and the playback of it



Recorded camera motion

Playback of camera motion

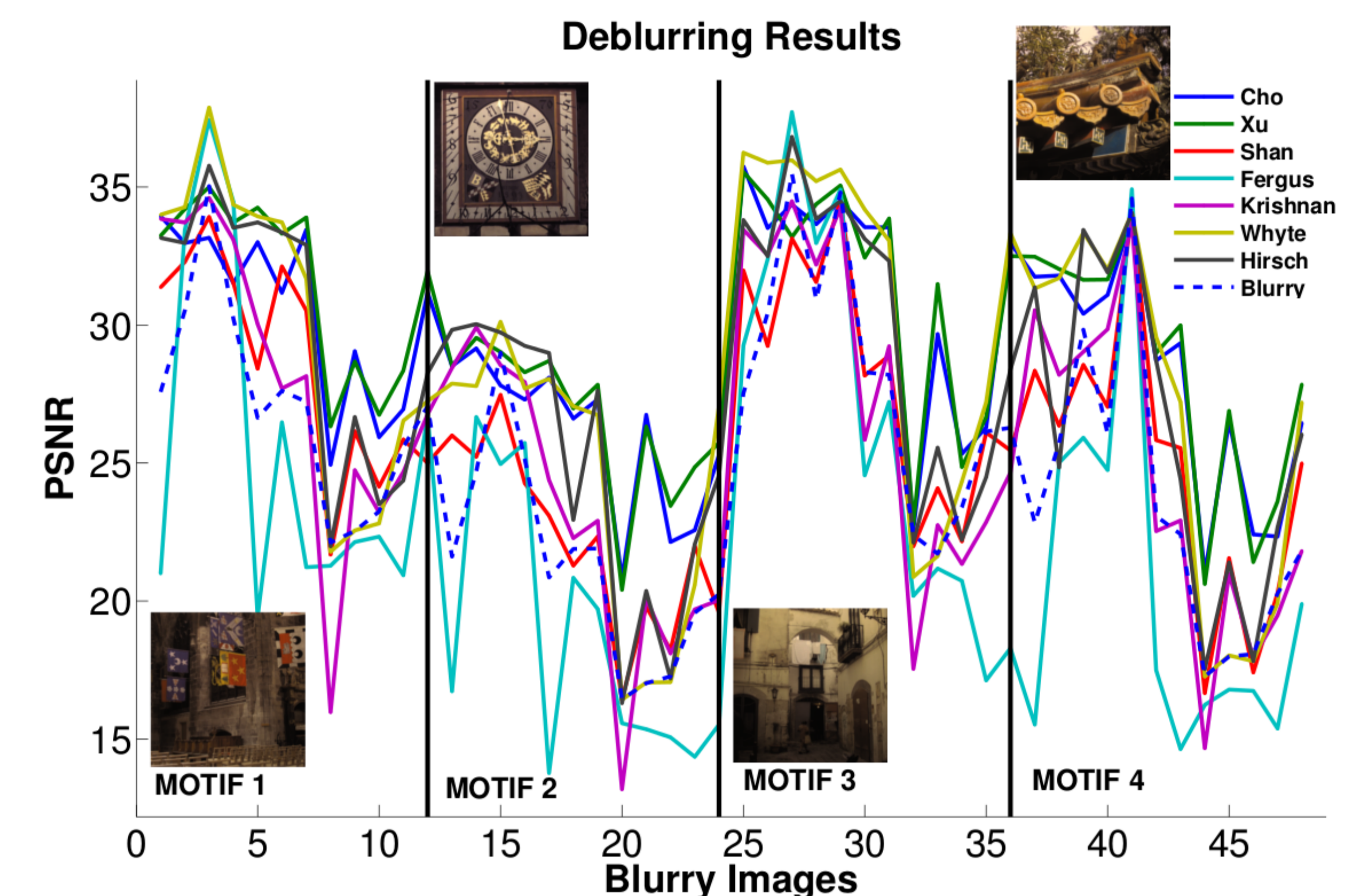
Benchmark dataset

- ▶ 12 different camera shakes (randomly selected 2 of each of the 6 subjects)
- ▶ 4 different motives (ground truth images)
- $(12 * 4) = 48$ blurry images

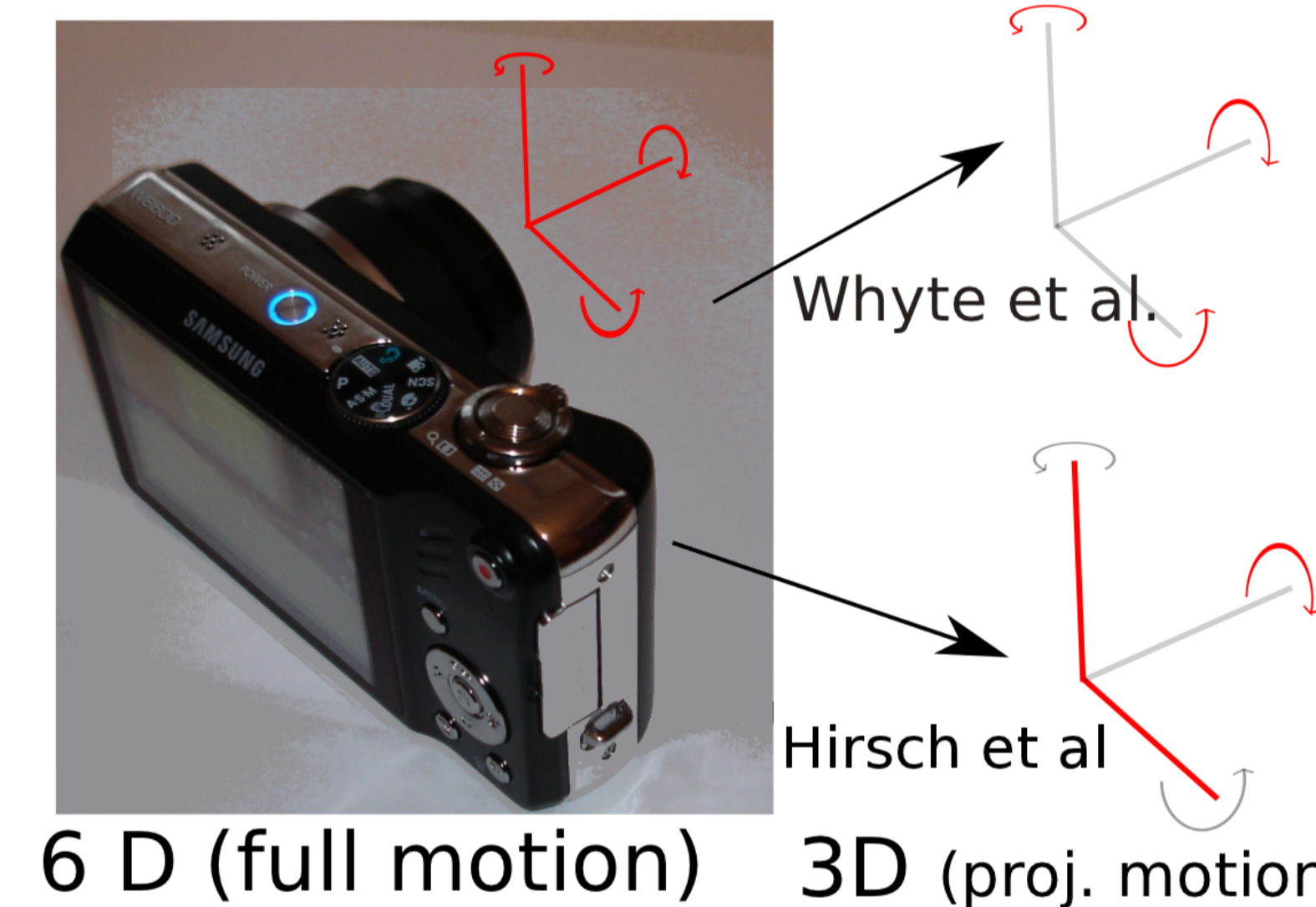


The four original images used in the benchmark.

Results



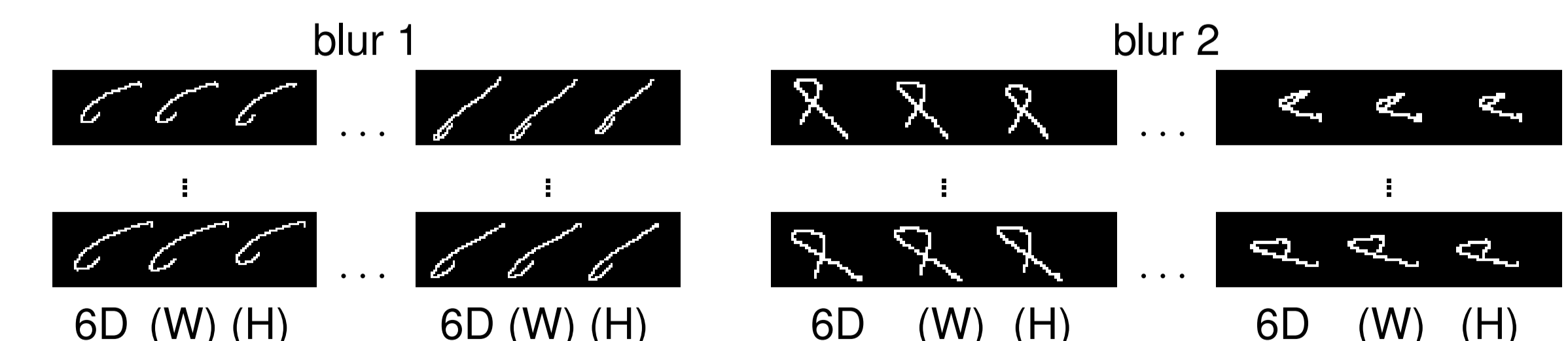
Approximation of 6D camera trajectory by 3D



6 D (full motion) 3D (proj. motion)

Non-uniform Blur models by Whyte [7] and Hirsch [6] approximate the 6D camera trajectory by 3D. We transformed the 6D trajectory to 3D (d is the distance lense ↔ object) according to:

$$(H) p_t = \begin{bmatrix} \theta_x \\ \theta_y \\ \theta_z \\ x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} 0 \\ \theta_y \\ 0 \\ x - d \sin(\theta_z) \\ 0 \\ z + d \sin(\theta_x) \end{bmatrix} \quad (W) p_t = \begin{bmatrix} \theta_x \\ \theta_y \\ \theta_z \\ x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} \theta_x - \arcsin(x/d) \\ \theta_y \\ \theta_z + \arcsin(z/d) \\ 0 \\ 0 \\ 0 \end{bmatrix}$$



Left: 6D trajectory, middle: Whyte, right: Hirsch. Only the four corners of the point grid are mapped. Focal length = 50mm, object distance = 2m.

References

- [1] Fergus, R., Singh, B., Hertzmann, A., Roweis, S.T., Freeman, W.T.: Removing camera shake from a single photograph. In: *ACM Transactions on Graphics (SIGGRAPH)*. (2006)
- [2] Shan, C., Jia, J., Agarwala, A.: High-quality motion deblurring from a single image. In: *ACM Transactions on Graphics (SIGGRAPH)*. (2008)
- [3] Cho, S., Lee, S.: Fast motion deblurring. In: *ACM Transactions on Graphics (SIGGRAPH ASIA)*. (2009)
- [4] Xu, L., Jia, J.: Two-phase kernel estimation for robust motion deblurring. In: *Proceedings of the European Conference on Computer Vision (ECCV)*. (2010)
- [5] Krishnan, D., Tay, T., Fergus, R.: Blind deconvolution using a normalized sparsity measure. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. (2011)
- [6] Hirsch, M., Schuler, C.J., Harmeling, S., Schölkopf, B.: Fast removal of non-uniform camera-shake. In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. (2011)
- [7] Whyte, O., Sivic, J., Zisserman, A.: Deblurring shaken and partially saturated images. In: *Proceedings of the IEEE Workshop on Color and Photometry in Computer Vision, with ICCV 2011*. (2011)
- [8] Guizar-Sicairos, M., Thurman, S.T., Fienup, J.R.: Efficient subpixel image registration algorithms. In: *Optical Letters* 33 (2008) 156-158



<http://webdav.is.mpg.de/pixel/benchmark4camerashake/>

Measuring the deblurring performance

comparing similarity between two images a and b

1. estimate the optimal scaling $\hat{\alpha}$ and translation \hat{T} such that the L2 norm between a and b becomes minimal^a

$$\hat{\alpha}, \hat{T} = \min_{\alpha, T} \|a - T(\alpha b)\|^2$$

2. calculate the peak-signal-to-noise ratio (PSNR)^b as

$$\text{PSNR}(a, b) = 10 \log_{10} \frac{m^2}{\langle \|a_i - \hat{T}(\hat{\alpha} b_i)\|^2 \rangle_i} \quad (1)$$

3. PSNR similarity between an estimated image \hat{u} and the ground truth as the maximum PSNR between \hat{u} and any of the images along the trajectory,

$$\text{SIM} = \max_n \text{PSNR}(u_n^*, \hat{u}). \quad (2)$$

^aWe allow for integer pixel translations only, which we estimate with the Matlab function `fftreg` by [8]
^bwith $\langle \cdot \rangle_i$ denoting an average over pixels and m being the maximal possible intensity value, i.e. $m = 255$ as we work with 8bit encoding.