Fast Removal of Non-Uniform Camera Shake

Michael Hirsch, Christian J. Schuler, Stefan Harmeling, and Bernhard Schölkopf

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

MAX-PLANCK-GESELLSCHAFT

ICCV 2011

Problem

Goal is to model and deblur images degraded by real camera shake causing non-uniform blur, i.e. the PSF varies spatially across the image plane, such as shown in the example images below.

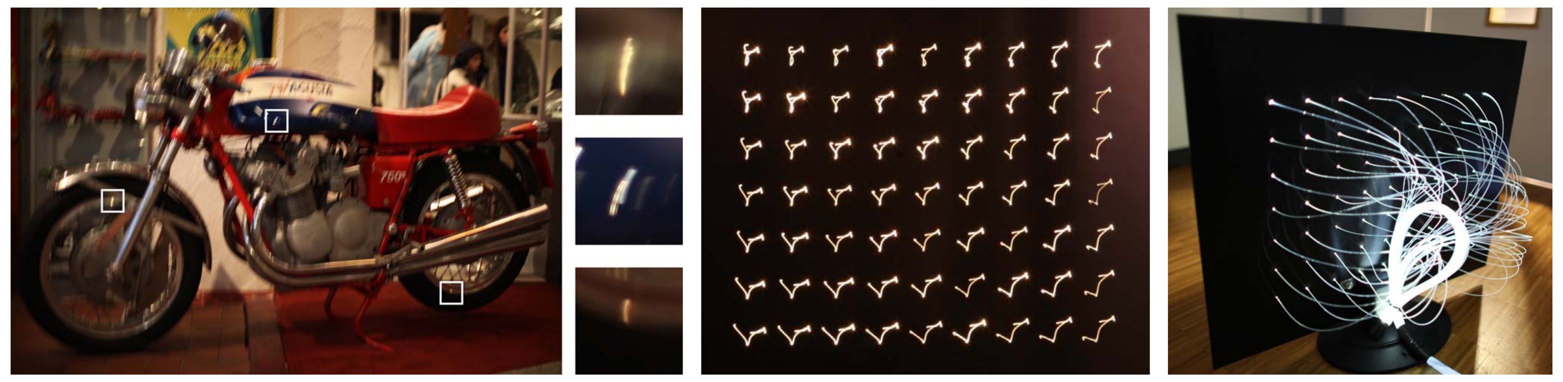
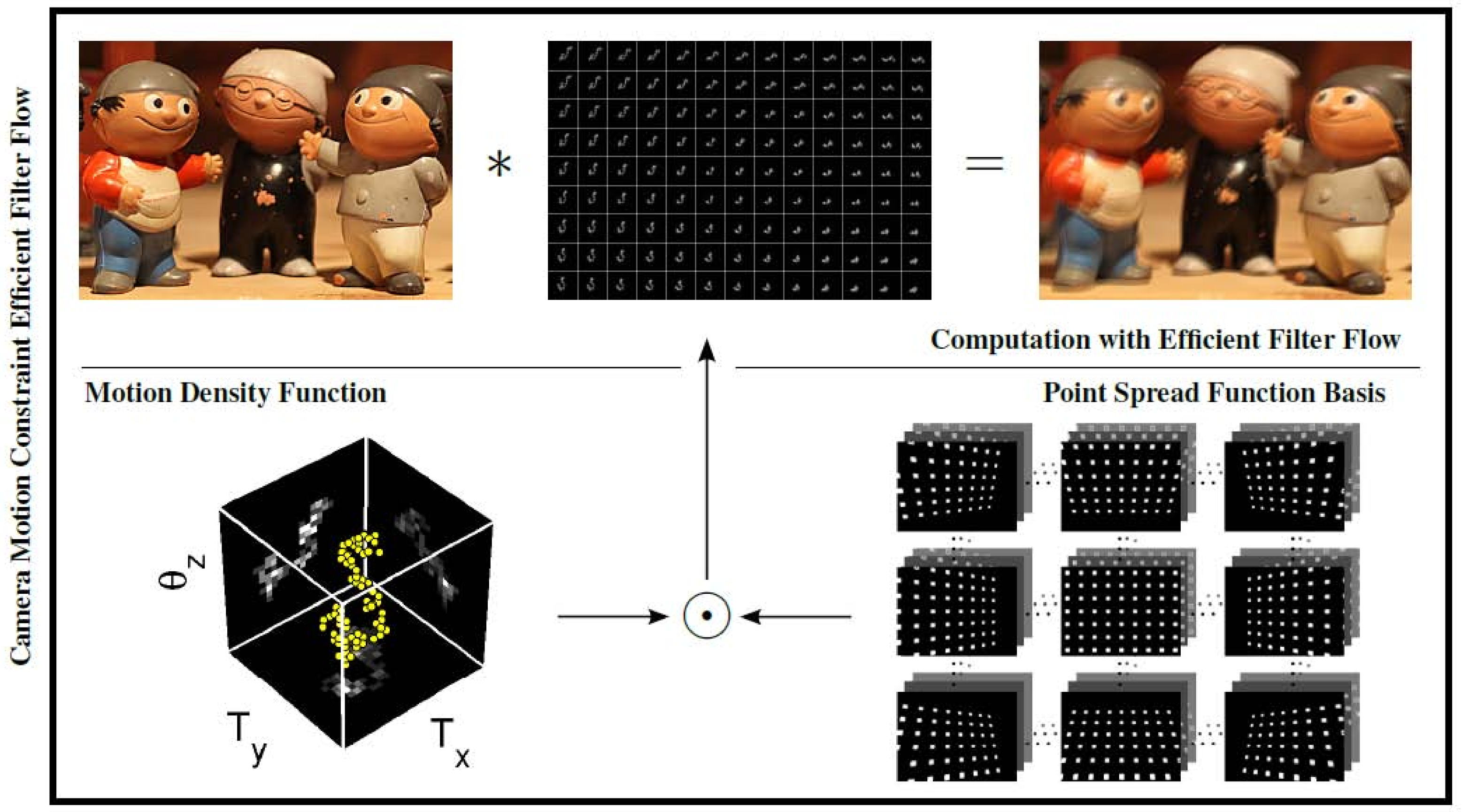


Photo blurred by camera shake — Illustration of PSF variation — Device for PSF sampling

Fast Forward Model



For visualization of the blur parameters (bottom left) we used the plotting function `plot_nonuni_kernel.m` of Whyte et al. (2010).

Efficient Filter Flow approximates a spatially-varying PSF by R local filters $a^{(r)}$:

$$g = \sum_{r=1}^{R} a^{(r)} * \left(w^{(r)} \odot f\right), \tag{1}$$

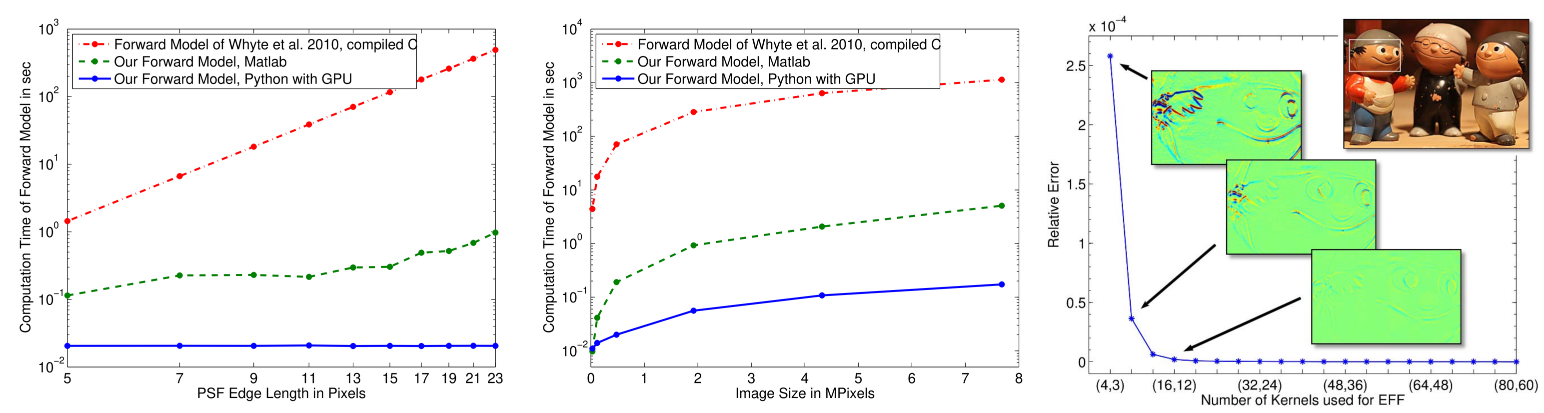
The weighting frames $w^{(r)}$ determine the interpolation scheme between neighbouring filters and ensure a smooth filter flow (cf. Seitz and Baker (2009)) across the image place. Eq. (1) can be computed efficiently by FFTs (Hirsch et al., 2010). To constrain the EFF to motion blur caused by camera shake only, we create a basis from a point grid transformed according to all possible homographies enumerated by the index θ. The $a^{(r)}$ parametrizing the EFF are obtained by a weighted sum of the basis frames:

$$a^{(r)} = \sum_{\theta} \mu_\theta b_\theta^{(r)}, \tag{2}$$

Note that all $b^{(r)}$ can be precomputed. The weighting μ_θ vector corresponds to the time of the camera in a certain pose during exposure. Plugging Eq. (2) into Eq. (1) yields our fast forward model:

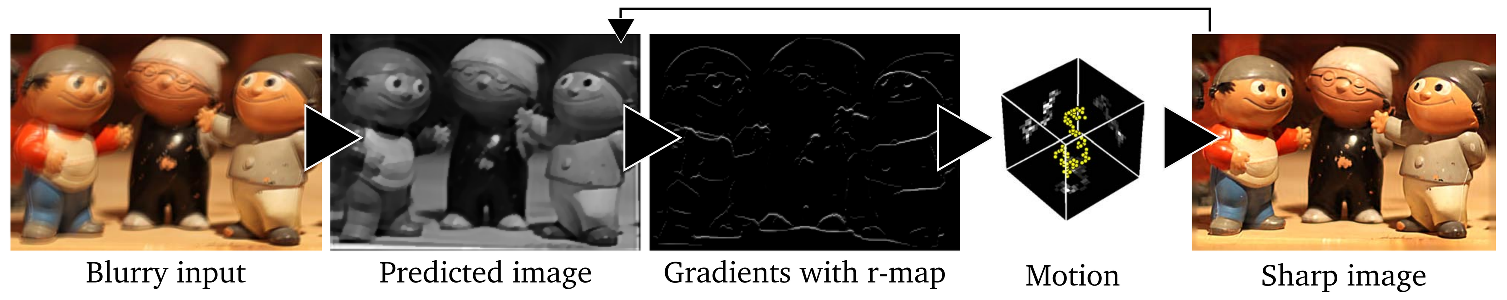
$$g = \mu \diamond f := \sum_r \left(\sum_\theta \mu_\theta b_\theta^{(r)}\right) * \left(w^{(r)} \odot f\right), \tag{3}$$

Note that our model is linear in both blur and image parameters. Hence, there exist matrices M and A such that $g = \mu \diamond f = Mf = A\mu$. Via the EFF we can also obtain fast implementations of the MVMs with M^T and A^T.



Run-time comparison of our forward model with the blurring model of Whyte (2010) and Gupta et al. (2010) as a function of PSF (left panel, image size: 1600×1200) and image size (middle panel, PSF size: 13×13). The right panel shows the accuracy of our approximation of a homographically transformed image (1600×1200 pixels) by the camera motion constrained EFF framework compared to the forward model of Whyte et al. (2010).

Algorithm for Single Image Blind Deconvolution



(i) Blur parameter update step: Initializing f with the blurry image g, the estimation of the camera shake blur parameters μ_θ, is performed by iterating over the following three steps:

- **Predict true image:**
  - remove noise in flat regions of f by edge-preserving bilateral filtering
  - overemphasize edges by shock filtering
  - compute gradient selection mask via *rmap* approach of Xu et al. (2010) to use only informative edges for estimation. In particular, it neglects structures that are smaller in size than the local filters, which could be misleading for the blur parameter estimation.
- **Estimate blur parameters:**
  - update the blur parameters given the blurry image g and the current estimate of the predicted $\tilde{f}$ obtained by bilateral and shock-filtering.
  - for a preconditioning effect use only the gradient images of x
  - enforce smoothness of camera trajectory

$$\|\partial g - m_S \odot \partial(\mu \diamond \tilde{f})\|_2^2 + \lambda\|\mu\|_2^2 + \eta\|\partial\mu\|_2^2, \tag{4}$$

where m_s is a mask (computed by *rmap* approach), that weights gradients according to their information content (see previous step). The regularization constants λ and η balance the likelihood against the prior terms. The above optimization problem is efficiently solved by gradient-based optimization techniques (e.g. lbfgsb or Barzilai-Borwein).

- **Latent image update step:**
  - update the current deblurred image f by solving a least-squares cost function using a smoothness prior on the gradient image via direct deconvolution (see **Direct Deconvolution** section below)

$$\|g - \mu \diamond f\|_2^2 + \alpha\|\partial f\|_2^2 \tag{5}$$

(ii) Non-blind deblurring (following Krishnan and Fergus, 2009): given the EFF parameterized by μ, yield the final image estimate by alternating between the following two steps:

- **Latent variable estimation:** estimate latent variables regularized with a sparsity prior that approximate the gradient of f. This can be efficiently solved with look-up tables as well as analytically, see "w sub-problem" of Krishnan and Fergus (2009) for details.
- **Image estimation step:** update the current deblurred image f by directly solving a least-squares cost function while penalizing the Euclidean norm of the gradient image to the latent variables of the previous step, see "x sub-problem" of Krishnan and Fergus (2009) for details and **Direct Deconvolution** section below.

Direct Deconvolution

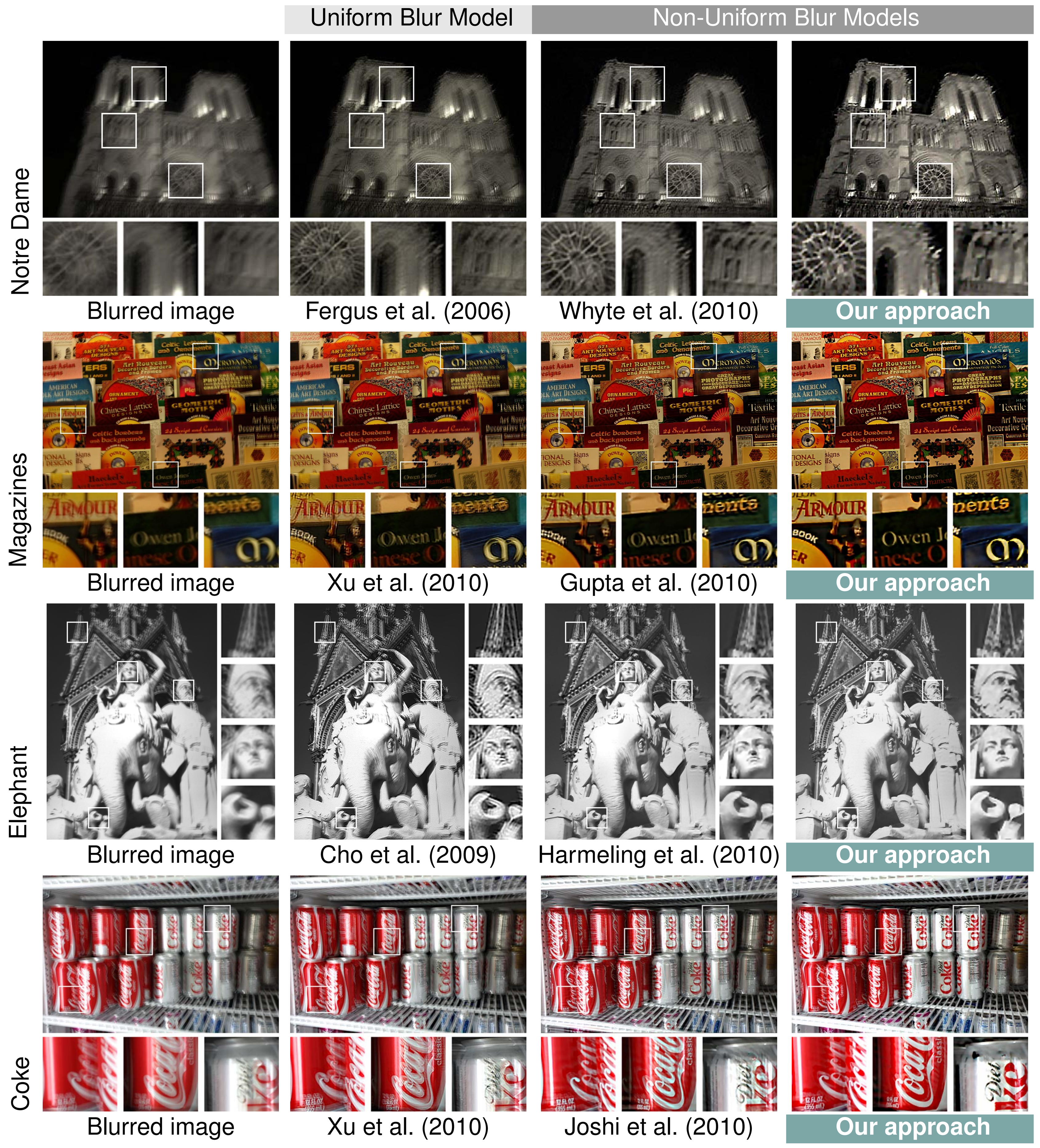
The optimization problem Eq. (5) can be solved directly via an approximate inverse

$$f \approx \mathrm{Diag}(v) \sum_r \mathrm{Diag}(w^{(r)})^{1/2} C_r^\mathsf{T} F^\mathsf{H} \frac{\overline{F Z_a B^{(r)}\mu} \odot (F E_r \,\mathrm{Diag}(w^{(r)})^{1/2} g)}{|F Z_a B^{(r)}\mu|^2 + \frac{1}{2}|F Z_l l|^2} \tag{6}$$

where the term $|FZ_l l|^2$ in the denominator originates from the regularization term in Eq. (5) with $l = [-1, 2, -1]^\mathsf{T}$ corresponding to the discrete Laplace operator. The term $\mathrm{Diag}(v)$ is a corrective weighting term which supresses windowing artifacts.

True image — DD with corrective weighting — DD w/o corrective weighting

Deblurring Results and Comparison



Comparison with state-of-the-art stationary and non-stationary deblurring algorithms on real-world data. Only qualitative comparisons were made. Run-time of our GPU implementation with PyCuda is about 30 seconds on Nvidia C2050 for a 2M pixel image.

Limitations

Saturated pixels — Out-of-focus blur — Objects in motion — Severe blur

References

S. Cho and S. Lee. "Fast Motion Deblurring". *ACM Transactions on Graphics (SIGGRAPH ASIA)*, 2009.
R. Fergus, B. Singh, A. Hertzmann, S.T. Roweis, and W.T. Freeman. "Removing camera shake from a single photograph". *In ACM Transactions on Graphics (SIGGRAPH)*, 2006.
A. Gupta, N. Joshi, L. Zitnick, M. Cohen, and B. Curless. "Single image deblurring using motion density functions". *In Proc. 10th European Conference on Computer Vision (ECCV)*, 2010.
N. Joshi, S. Kang, C. Zitnick, and R. Szeliski. "Image deblurring using inertial measurement sensors". In *ACM Transactions on Graphics (SIGGRAPH)*, 2010.
S. Harmeling, M. Hirsch, and B. Schölkopf. "Space-variant single-image blind deconvolution for removing camera shake". In *Advances in Neural Information Processing Systems (NIPS)*, 2010.
M. Hirsch, S. Sra, B. Schölkopf, S. Harmeling, "Efficient Filter Flow for Multiframe Blind Deconvolution". In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010.
D. Krishnan and R. Fergus. "Fast image deconvolution using hyper-Laplacianpriors". *In Advances in Neural Information Processing Systems (NIPS)*, 2009.
S. Seitz and S. Baker. "Filter Flow". In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009.
Q. Shan, J. Jia, and A. Agarwala. "High-quality motion deblurring from a single image". *ACM Transactions on Graphics (SIGGRAPH)*, 2008.
Stockham, "High-speed convolution and correlation". In *Proceedings of the April 26-28, 1966, Spring joint computer conference*, 1966.
O. Whyte, J. Sivic, A. Zisserman, and J. Ponce. "Non-uniform deblurring for shaken images". In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2010.
L. Xu and J. Jia. Two-phase kernel estimation for robust motion deblurring. In *Proceedings of 10th IEEE European Conference on Computer Vision*, 2010.

http://webdav.is.mpg.de/pixel/fast_removal_of_camera_shake