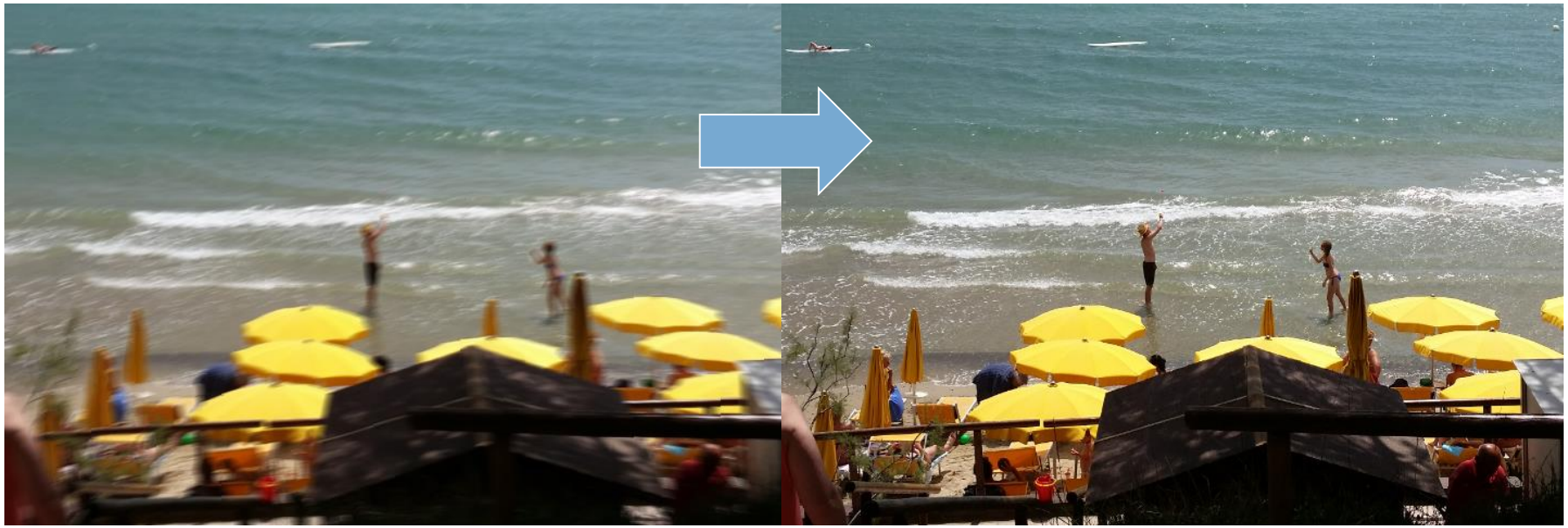
multiframe visual-inertial blur estimation and removal for unmodified smartphones

Gábor Sörös, Severin Münger, Carlo Beltrame, Luc Humair

09.06.2015

WSCG 2015, Plzen, Czech Republic



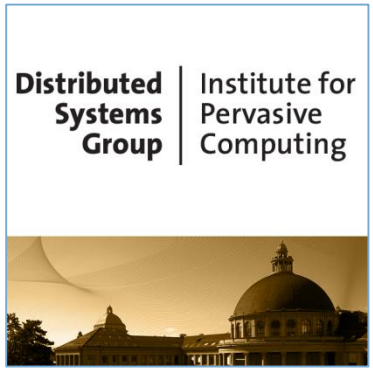
images taken by non-professional photographers are often degraded by motion blur due to object motion or camera motion during the exposure time

typical examples: smartphone/smartglass photography

retaking the photos is often not possible → blur removal needed

target unmodified smartphones (no hardware modifications, no low-level camera control)

basics of blurry image formation

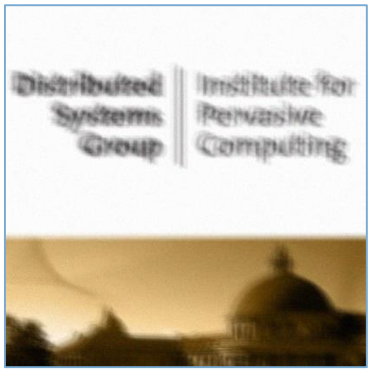


sharp scene

$\boldsymbol{I}$

blurry scene

$\boldsymbol{I} * \boldsymbol{k}$

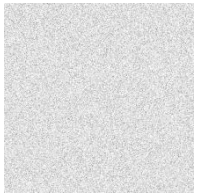


observed image

$\boldsymbol{B} = \boldsymbol{I} * \boldsymbol{k} + \boldsymbol{n}$

convolution with a point spread function (PSF) $\boldsymbol{k}$

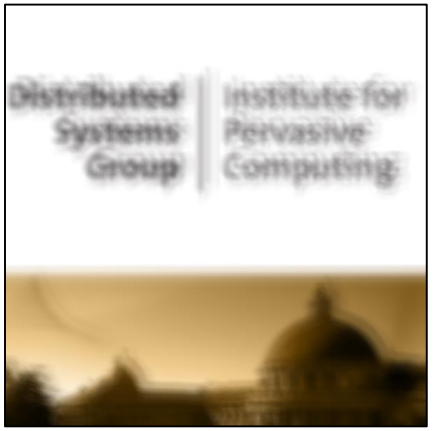
adding camera noise $\boldsymbol{n}$

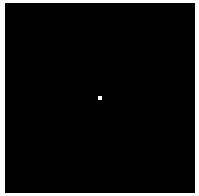


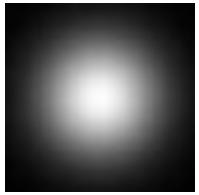
blur removal problems

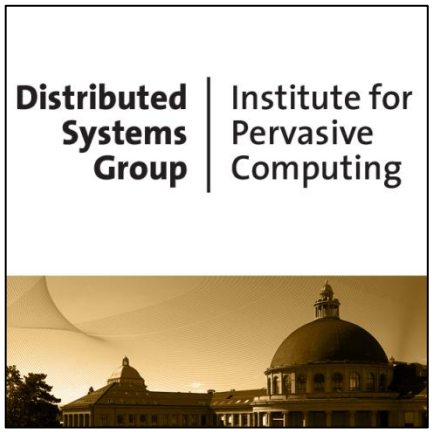
the problem of deblurring an image is ill-posed:

- there are infinite combinations of sharp images and blur functions that result in the same blurry image
- the blur function is usually not known
- even with a known blur function, deblurring is not straightforward









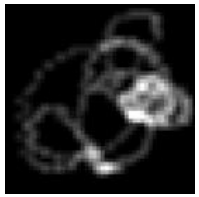


figure inspired by Robert Fergus

deblurring ambiguity

deconvolution:

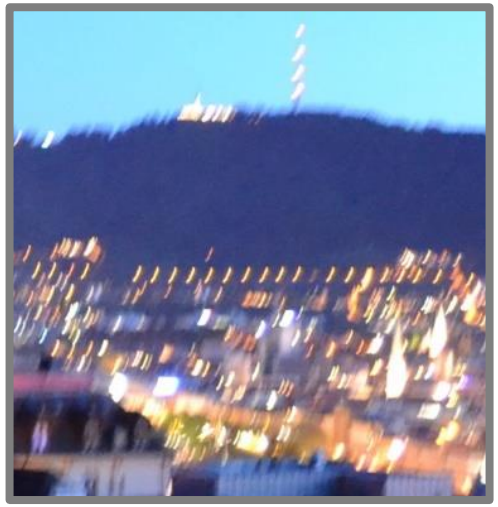
$$\boldsymbol{B} = ? * \boldsymbol{k} + \boldsymbol{n}$$

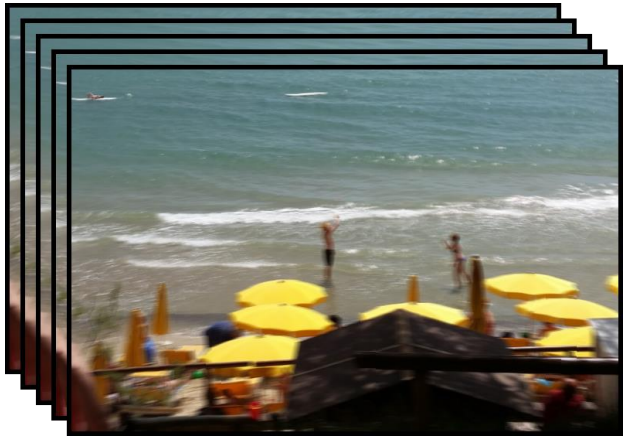
blind deconvolution:

$$\boldsymbol{B} = ? * ? + \boldsymbol{n}$$

additional clues

- the blur is 'encoded' in the image of point light sources
- smartphones have inertial sensors – we can reconstruct the camera motion
- we have multiple images from the camera's video stream





outline

- reconstructing camera motion from sensors
- estimating blur at each part of the image
- restoring individual tiles of the image
- aligning subsequent frames
- restoring tiles with the help of neighboring tiles in time
- advanced issues

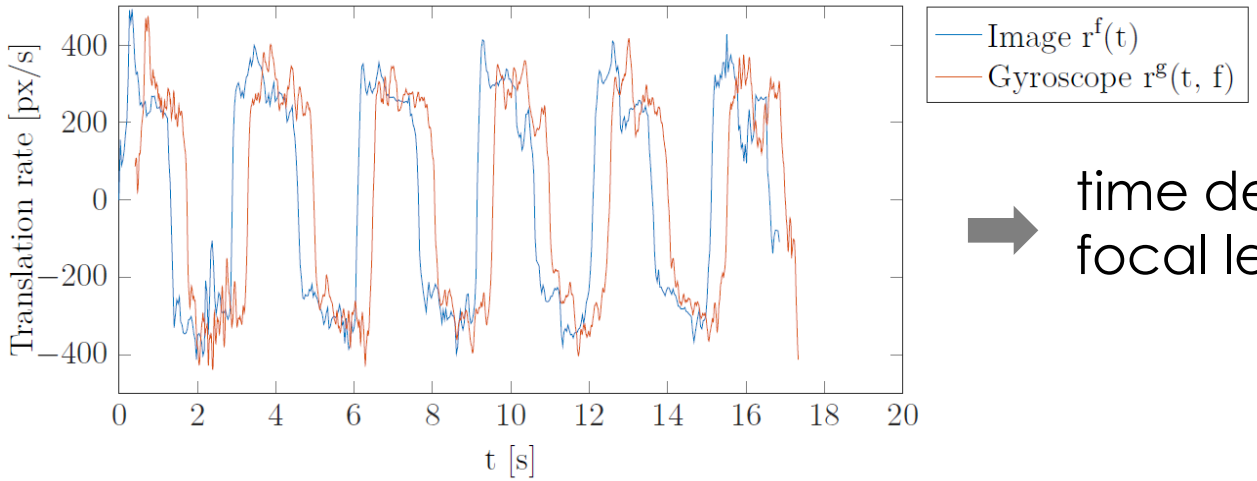
motion sensors

- Accelerometers – linear acceleration
  - gravity compensation difficult
  - double integration amplifies noise
  - translational blur depends on scene depth
- Gyroscopes – rotational velocity
  - rotational blur is dominant in hand shake
  - bias can be neglected in short intervals
  - rotational blur is independent of scene depth

- We use only gyroscopes
- Synchronization with camera required

gyro-camera synchronization

- Previous work
  - hardware modification [Joshi2010, Park2014]
  - phone-specific [Sindelar2014]
- Our current method
  - Extended Kalman Filter [Jia2014] (open source) – initialization?
- Our new method in development
  - pixel translation rates for initialization
  - optimization on coplanarity constraints

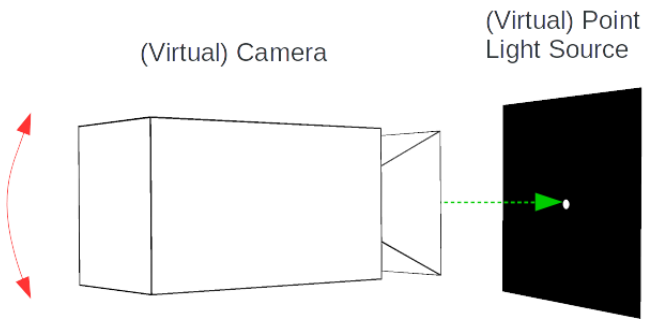


time delay (x match)
focal length (y match)

estimating blur from motion

Kernel rendering

- place a point light source on the image plane
- shake virtual camera by 'replaying' the motion
- super-resolve time by spherical linear interpolation
- blend the rendered dots

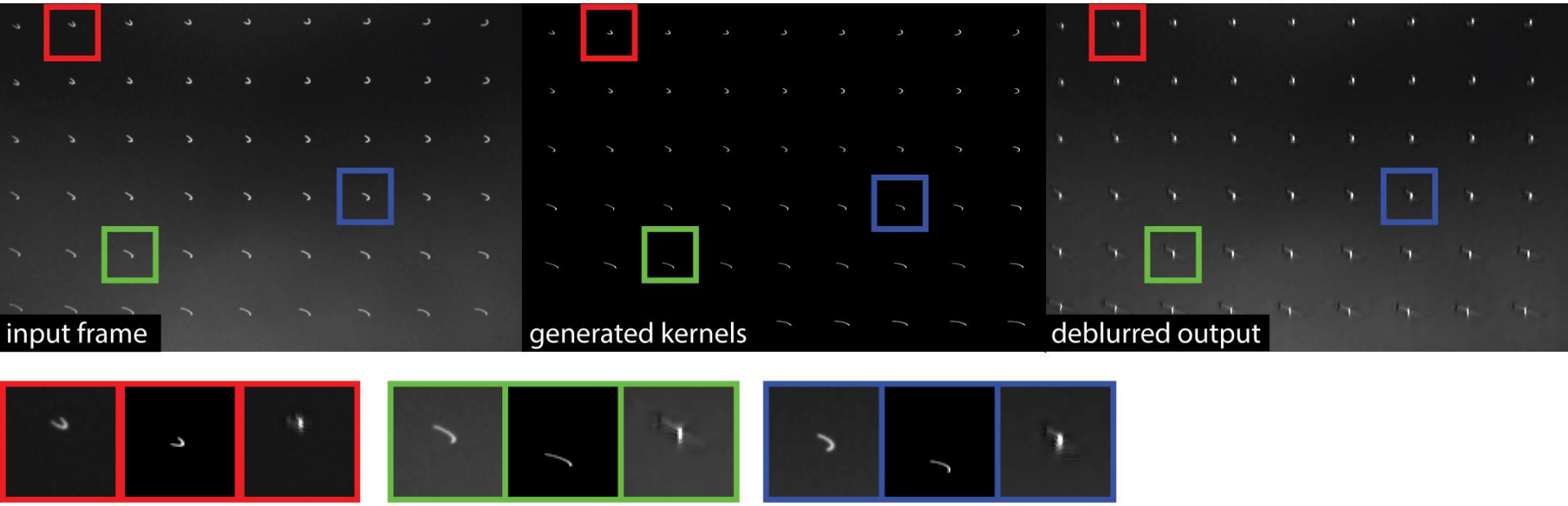


Non-uniform blur (rotations)

- split the image into overlapping regions
- assume uniform blur in each region
- render kernel for each tile

estimating blur from motion (evaluation)

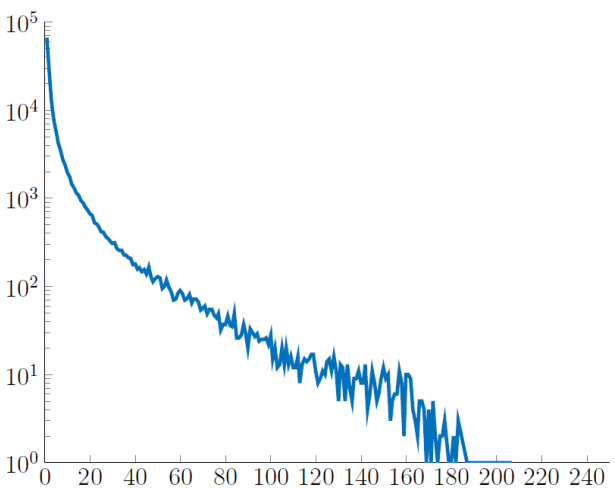
capturing a screen that shows white dots



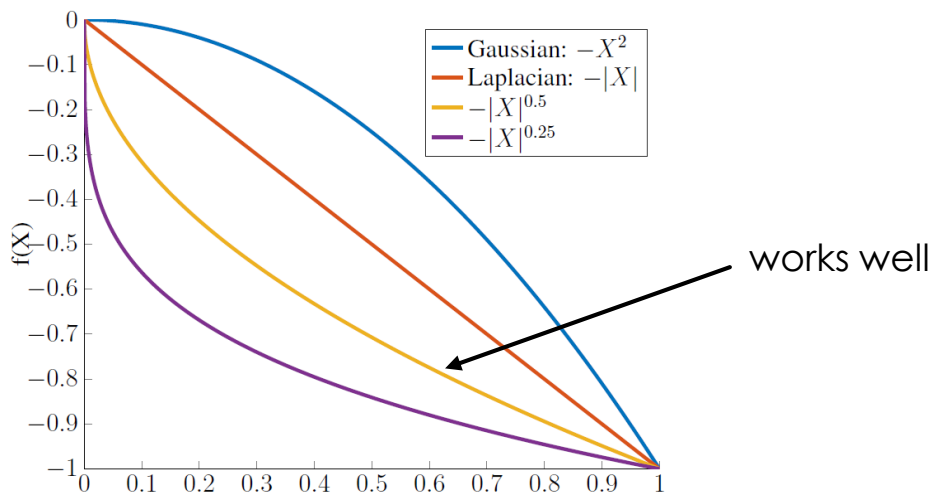
trade-off: number of regions (restoration quality) vs. restoration speed

inverting the blur (non-blind deconvolution)

- even if the blur kernel is known, deconvolution is ill-posed, requires regularization
- natural image statistics for regularization
  → certain distribution of image gradients



log-gradient distribution



parametric models

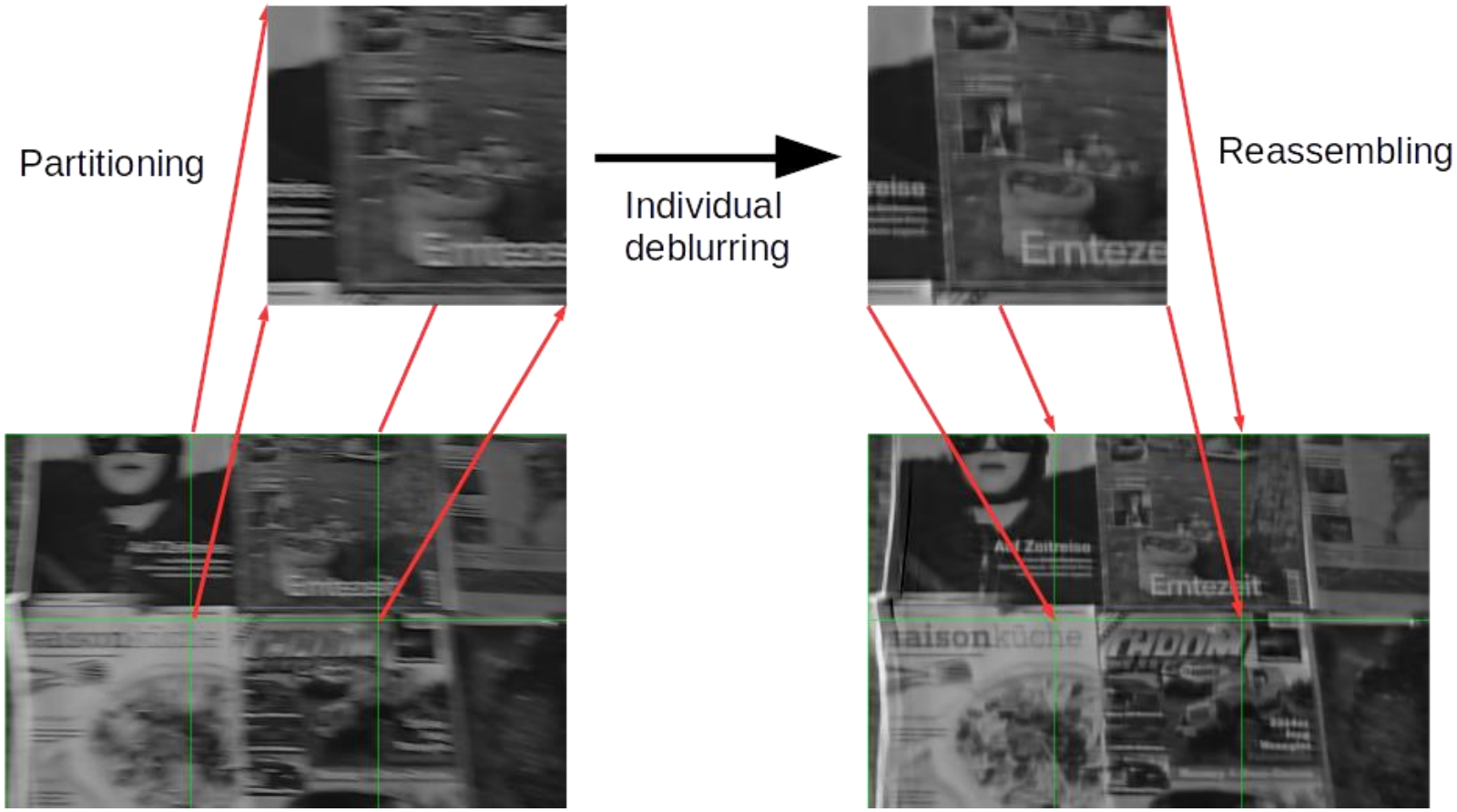
non-blind deconvolution

- algorithm of Krishnan and Fergus [Krishnan2009]
- given the blurry image **B** and the kernel **K**, it solves for the sharp image

$$\mathbf{I} = \underset{\mathbf{I}}{\operatorname{argmin}} \|\mathbf{I} * \mathbf{K} - \mathbf{B}\|_2^2 + \underbrace{\lambda p(\mathbf{I})}_{\text{prior on gradients}}$$

- solution via FFTs and pixel-wise equations
  further details omitted
- fast and good quality (compared to others)
- only uniform blur!

partitioning the image to 'uniformly' blurred tiles



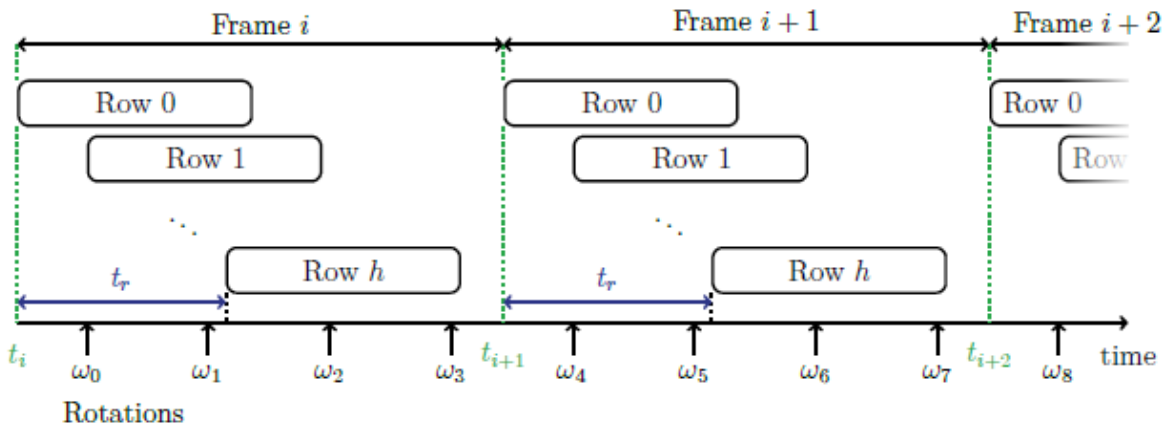
extension to multiple frames

- deblur individual subsequent frames (tiles)
- align the deblurred images
  - extract SURF features [Bay2008]
  - calculate homography map
- deblur the main image (tile) again, but penalize deviations from the helper images (tiles)
- penalty weights of each tile are inversely proportional to the blurriness of that tile
- fast, requires only 2 more FFTs

$$\mathbf{I} = \underset{\mathbf{I}}{\operatorname{argmin}} \|\mathbf{I} * \mathbf{K} - \mathbf{B}\|_2^2 + \lambda p(\mathbf{I}) + \frac{\mu}{\sum \mu_i} \sum_{i=1}^{M} \left(\mu_i \|\mathbf{I} - \mathbf{I}_i\|_2^2\right)$$

rolling shutter distortions

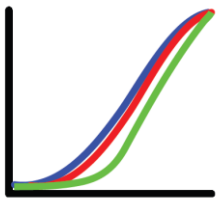
- The smartphone's image sensor is exposed row by row. When the camera undergoes motion, this causes skew distortions in the image



- we warp the input images on the GPU to invert the rolling shutter skew → better image alignment

- we shift the time windows for kernel generation

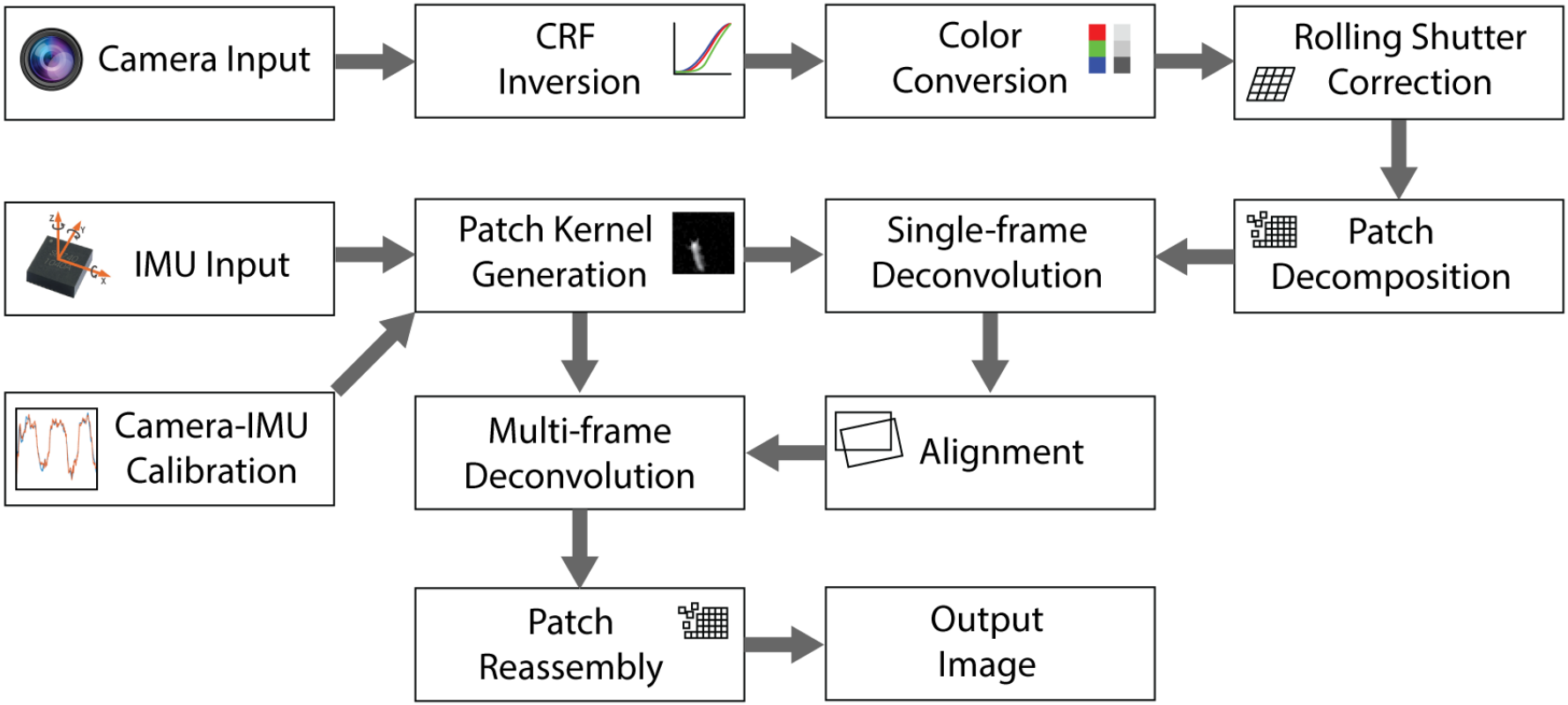
camera response function (CRF)



- our blur model is linear, but
- the camera converts scene intensity to pixel values through a **non-linear** function. This has a significant impact on deblurring [Tai2013]
- the CRF is different for each camera, for each mode
- CRF-estimation algorithms require precise exposure control (not yet available for smartphones)

- we apply a simple gamma curve. Online estimation of the CRF remains an open question. Upcoming smartphones do allow exposure control.

algorithm outline



implementation

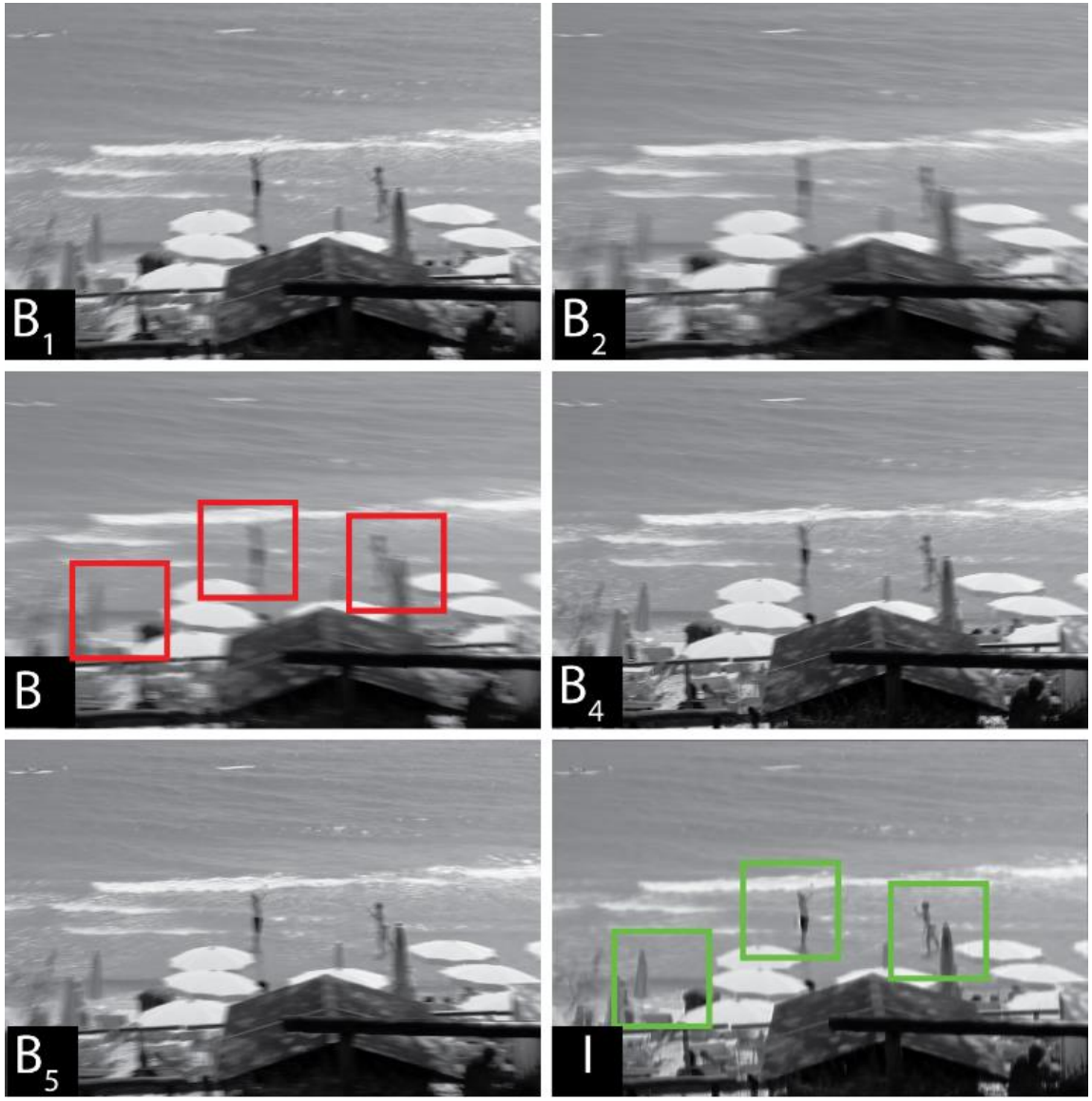
- OpenCV
  cross-platform image processing in C++

- OpenGL ES 2.0
  image warping and color conversions on GPU
  (later also Fourier transforms)

- Android Recorder Application (Google Nexus 4)
  720x480 preview frames @ 30 Hz, gyroscope @ 200 Hz

- Experiments offline on a PC

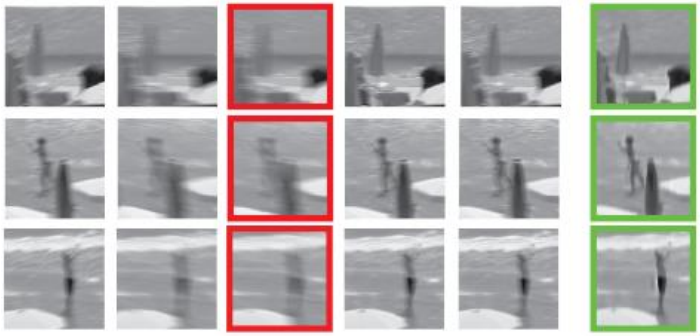
results: removing synthetic blur



synthetic:
- perfect synchronization
- linear CRF
- no RS

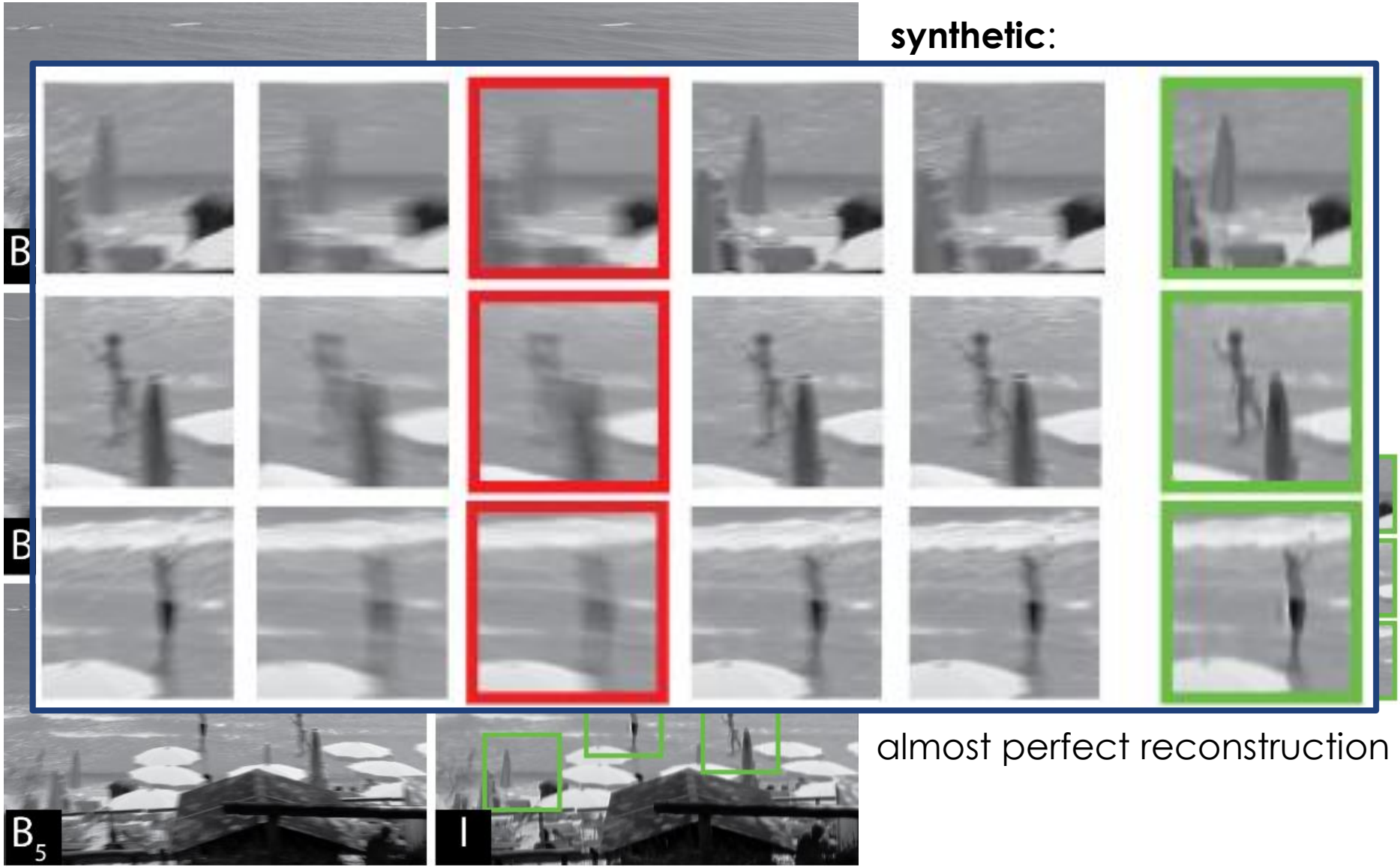
B : main input
$B_{1,2}$ and $B_{4,5}$: helper images
I : output

direct comparison:



almost perfect reconstruction

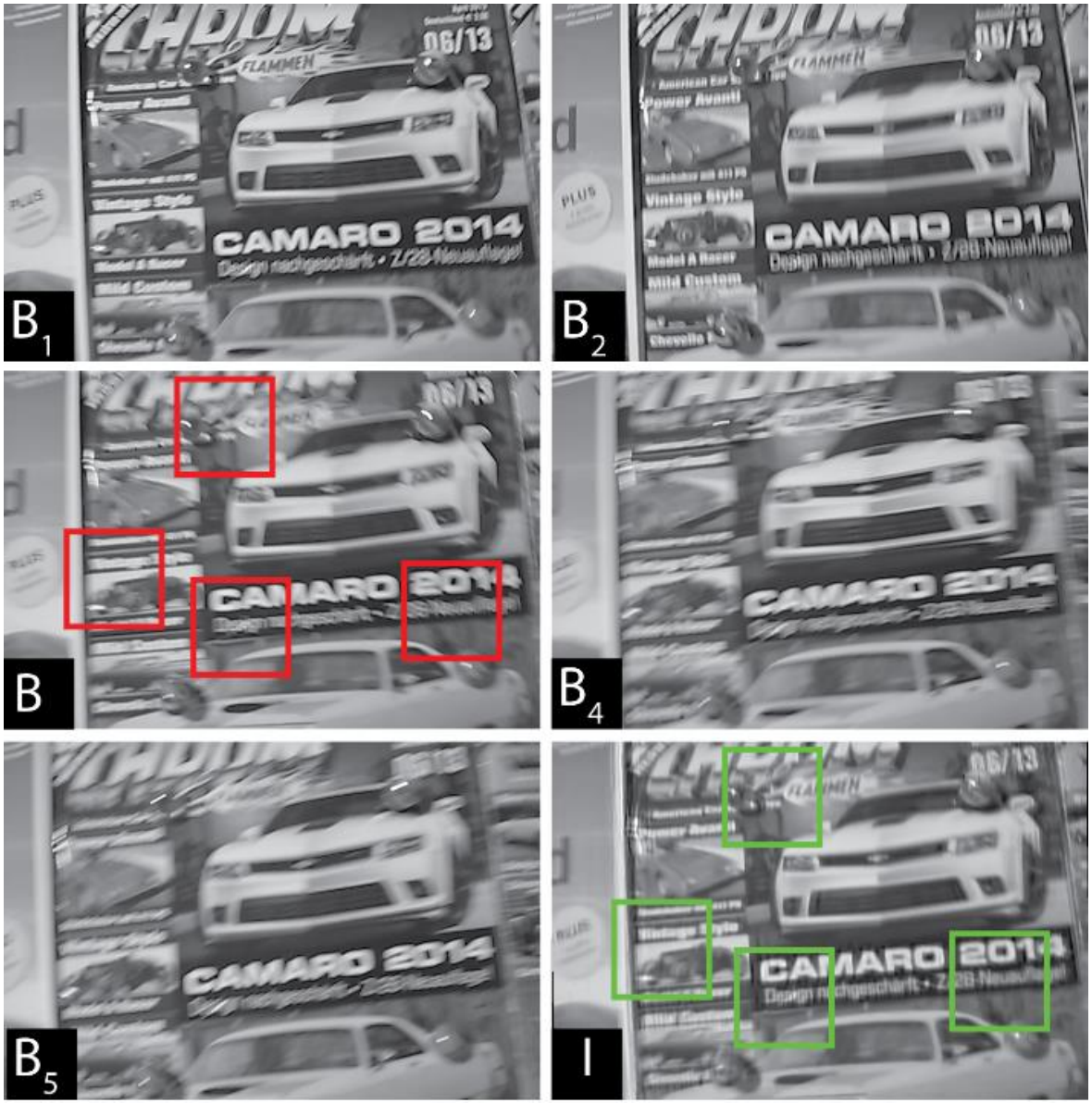
results: removing synthetic blur



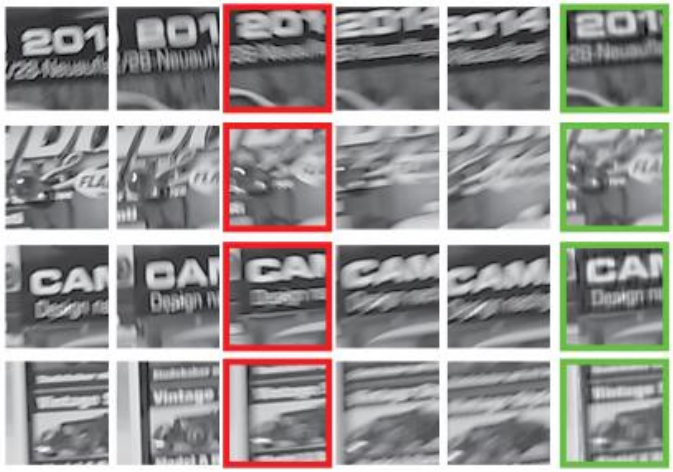
synthetic:

almost perfect reconstruction

results: removing real blur



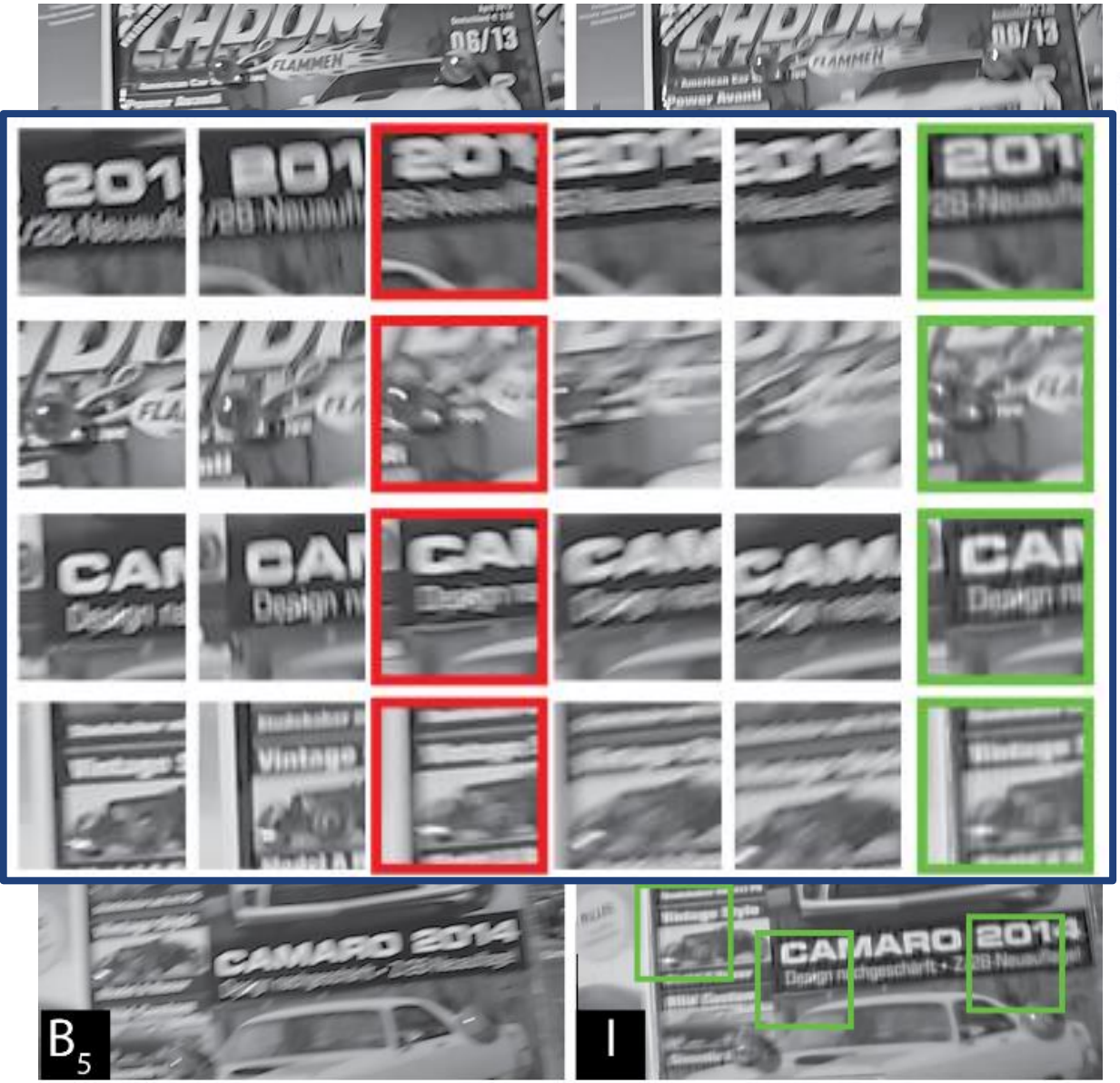
direct comparison:



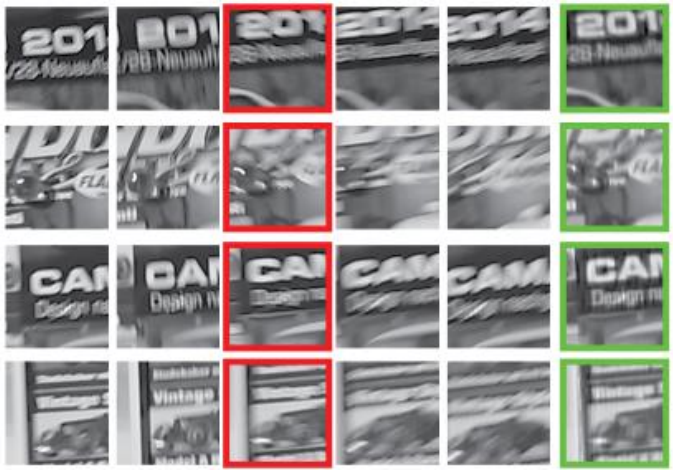
The green outputs are sharper than the red inputs, however, they are sometimes blurrier than a helper image.

Note: helper tiles are not copied, but penalize the reconstruction of red tiles.

results: removing real blur



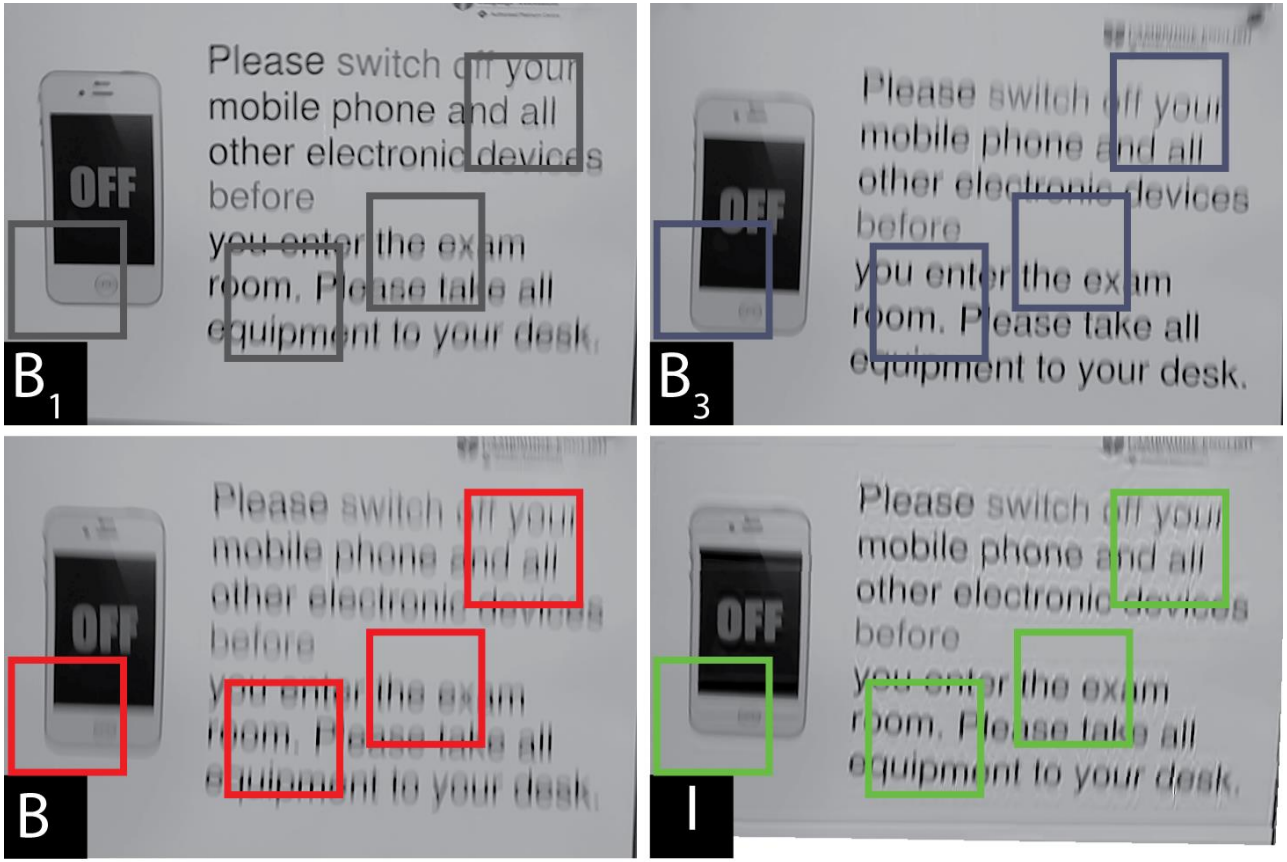
direct comparison:



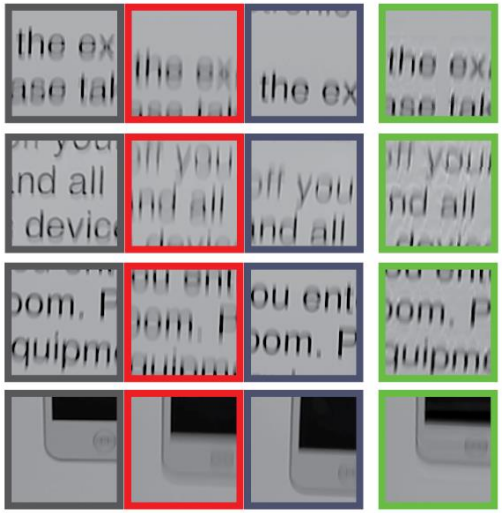
The green outputs are sharper than the red inputs, however, they are sometimes blurrier than a helper image.

Note: helper tiles are not copied, but penalize the reconstruction of red tiles.

results: removing real blur

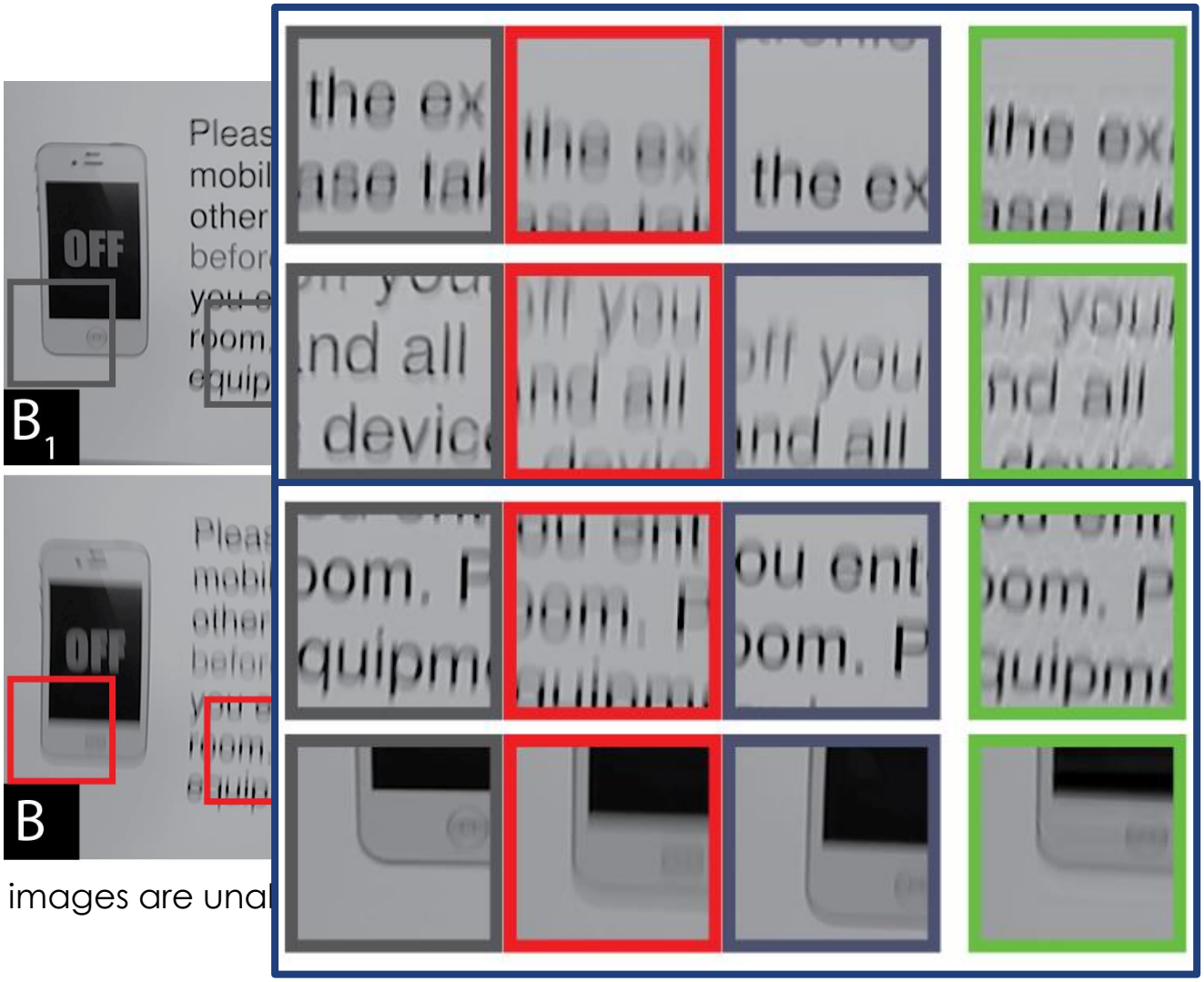


direct comparison:



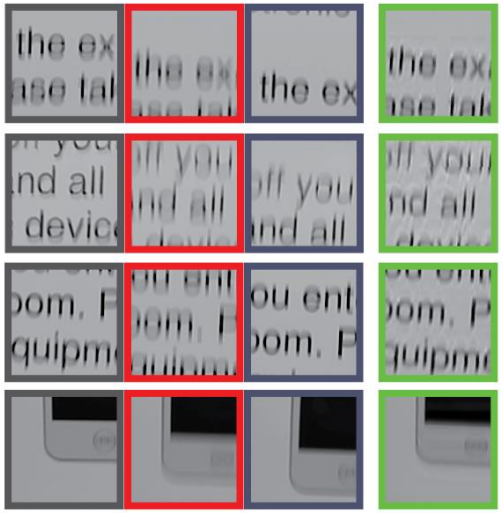
images are unaligned and unrectified for visualization

results: removing real blur



images are unal

direct comparison:



limitations and future work

- the EKF-based gyro-camera synchronization is not robust enough (feature tracking fails under blur)
  → new online synchronization method
- the homography-based image alignment may fail
  → blurry image alignment (blur invariants?)
- translational motion - accelerometers
  → online drift estimation, scene depth estimation
- camera response function
  → online calibration possible?
- optimal weighting strategy in the multiframe setting
  → allow to copy helper tiles (video deblurring)

summary

We described a combined blur removal algorithm

- for *unmodified* smartphones
- using gyroscope measurements and multiple images
- adressing several issues:
  - gyro-camera synchronization
  - blur kernel estimation
  - rolling shutter rectification
  - fast deconvolution with natural image priors
  - multi image alignment
  - speed (runtime order of seconds)

We showed promising qualitative results, and proposed future directions for research

thank you

references

- [Bay2008] H. Bay, T. Tuytelaars, L. van Gool – SURF: Speeded-Up Robust Features, ECCV, 2006
- [Krishnan2009] D. Krishnan, R. Fergus – Fast image deconvolution using hyper-Laplacian priors, NIPS, 2009
- [Joshi2010] N. Joshi, S. B. Kang, C. L. Zitnick, R. Szeliski – Image deblurring using inertial measurement sensors, SIGGRAPH, 2010
- [Tai2013] Y.-W. Tai, X. Chen, S. Kim, S. J. Kim, F. Li, J. Yang, J. Yu, Y. Matsushita, M. Brown – Nonlinear camera response functions and image deblurring: Theoretical analysis and practice, PAMI, 2013
- [Sindelar2014] O. Sindelar, F. Sroubek, P. Milanfar – Space-variant image deblurring on smartphones using inertial sensors, CVPRW, 2014
- [Park2014] S. Park, M. Levoy – Gyro-based multi-image deconvolution for removing handshake blur, CVPR, 2014
- [Jia2014] C. Jia, B. Evans – Online camera-gyroscope autocalibration for cell phones, TIP, 2014