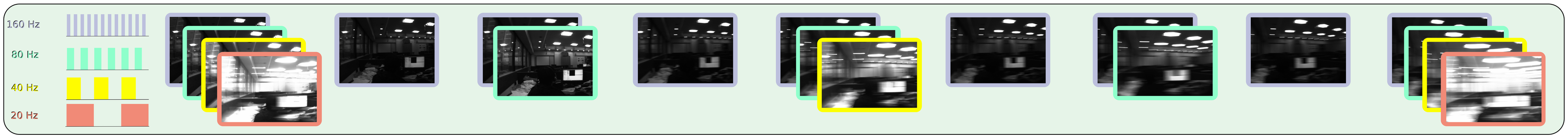
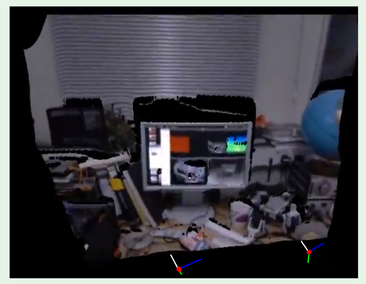
Real Time Camera Tracking: When Is High Frame-Rate Best?

Ankur Handa, Richard A. Newcombe, Adrien Angeli, Andrew J. Davison

What does the same motion look like at different frame-rates? / Which frame-rate is optimal for real-time tracking?

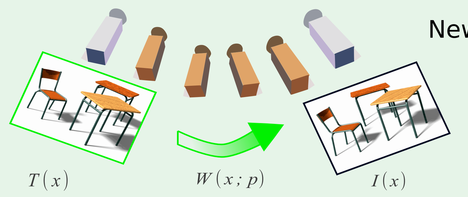


What specific tracking problem are we studying?



Newcombe et al. DTAM, ICCV2011

$T(x)$: Predicted Image
$I(x)$: Current image
p: Camera params
$W(x;p)$: Warp



Real-time 6D single camera tracking in a known 3D rigid scene using iterative whole image alignment.

What questions would we like to answer?

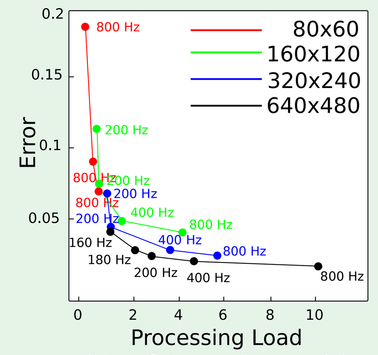
Intuitions

- High frame-rate seems better.
  Image motion is small - tracking is easy.
- Increasing image resolution is desriable.
  More data means better accuracy.

Knowing that, we'd like to find out

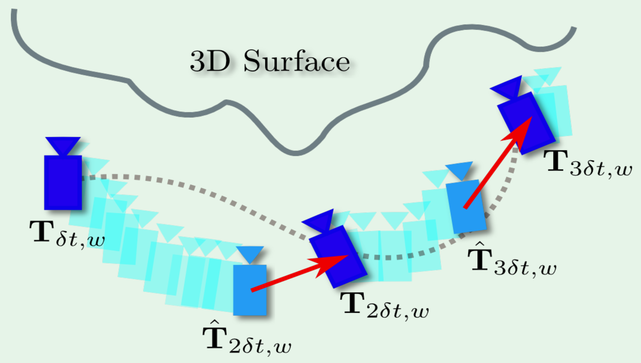
Given a limited processing budget:

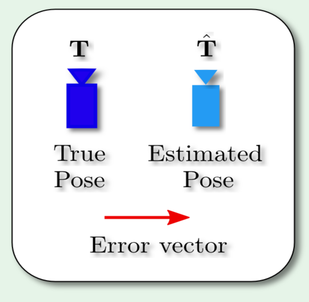
- What is the optimal frame-rate for tracking?
- What is the best image resolution for tracking?



Pareto Front of Operating Points

How do we evaluate a real-time tracker?





- **Error:** $\mathcal{L}_2$ norm of the estimated pose and ground truth pose.
- **Computational Cost:** Cost per frame $\times$ frame-rate.
- **Robustness:** how often a tracker works without gross failures.
  Not analyzed in this paper but plan to work on it in future.

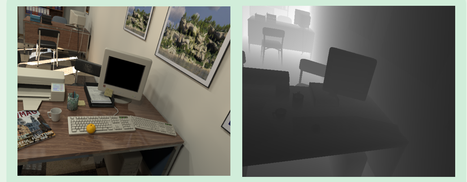
How do we gather experimental data?

Collecting Real Data

- ✗ Not easy to obtain ground truth.
- ✗ Not easy to have repeatable camera trajectory.
- ✗ Cannot obtain resolution other than standard resolutions.
- ✗ Cannot control scene lighting.

Justification for synthetic framework for obtaining data

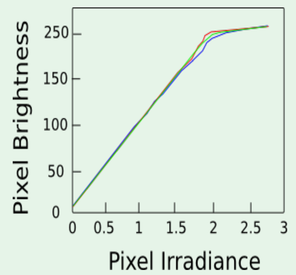
- ✓ Can exercise full control of parameters
- ✓ Perfect ground truth.
- ✓ Repeatable camera motion
- ✓ Controlled scene lighting



POVRay raytraced image and depth-map

How to add photo-realism in synthetic images?

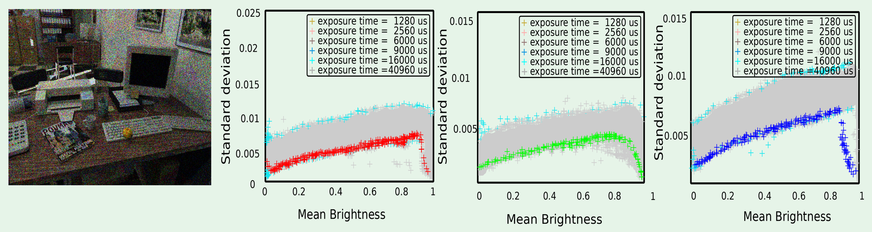
How do we create motion blur?



Calibrating Camera Response Function

- Obtain pixel Irradiance to brightness map of camera.
  Camera Response Function (CRF)
- Use the CRF to map brightness for synthetic images.
  POVRay gives Irradiance.

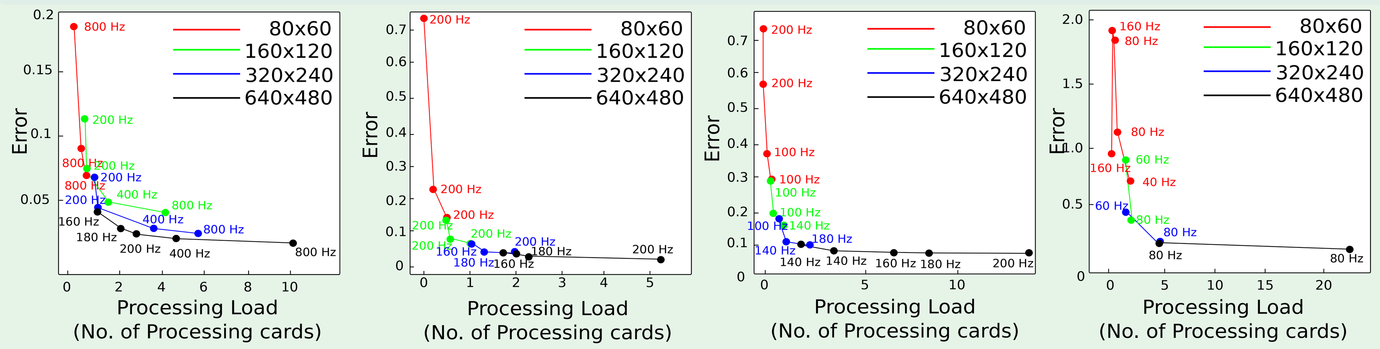
How do we add camera noise?



Calibrating Camera Noise Level Function

- Take multiple shots of static scene with varying shutter times.
- Plot mean brightness of pixel against the standard deviation.
- Fit a parametric function to obtain the noise level for all channels.
- Use this noise in the synthetic images for all frame-rates.

What are our conclusions and results? / What do we interpret from the data?



Perfect Lighting	High Lighting	Moderate Lighting	Low Lighting
No noise and no motion blur	$\alpha = 40$	$\alpha = 10$	$\alpha = 1$

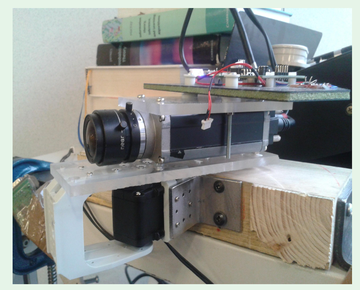
Perfect Lighting: Use the combination of highest possible frame-rate and resolution the budget allows.

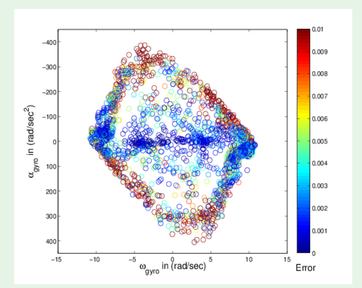
Real Lighting: There is optimal frame-rate that depends on scene lighting.

- High lighting allows high SNR images and therefore optimal frame-rate is high e.g. 200Hz
- Moderate Lighting means optimal frame-rate shifts to a lower value e.g. 100-140Hz
- Low Lighting permits working on only low frame-rates - high frame-rate images are dark and noisy.

Future goals and directions

Real Experiments





- Verify the conclusions from synthetic framework.
- Obtain the robustness curves for all frame-rates.
- Find out answers to more questions e.g.

If camera is moving with a known acceleration and velocity, what is the optimal frame-rate taking accuracy, cost, and now robustness into account?

We are grateful to ERC for the funding (ERC Starting Grant 210346) and our colleagues Margarita Chli, Steven Lovegrove, Hauke Strasdat and Thomas Pock for fruitful discussions and exchange of ideas. We also thank Jaime Vivies Piqueres for his POVRay scene.