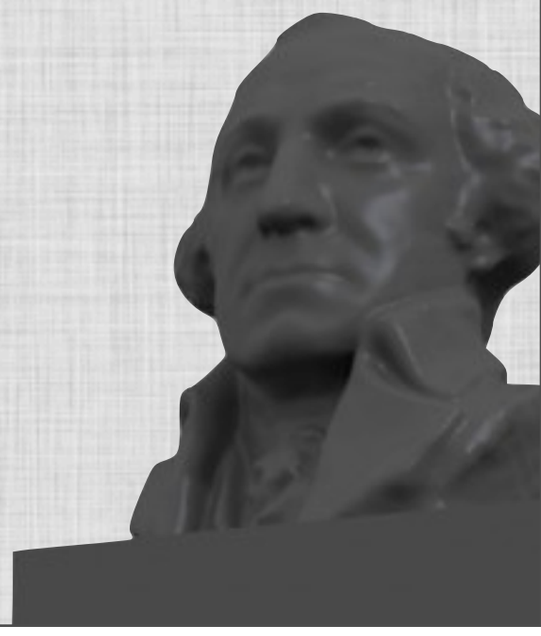
THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON DC

SPLINE FUSION:

A CONTINUOUS-TIME REPRESENTATION FOR VISUAL-INERTIAL FUSION WITH APPLICATION TO ROLLING SHUTTER CAMERAS

Steven Lovegrove, Alonso Patron-Perez
and Gabe Sibley.

Robot Perception Group,
George Washington University.



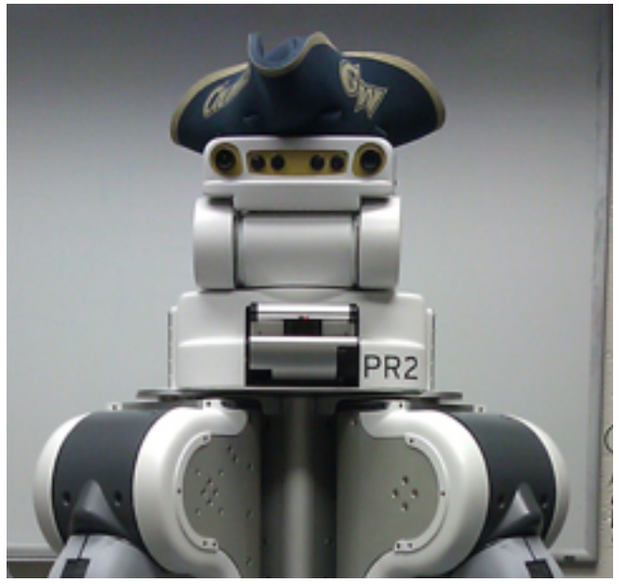
Motivation: Flexible SLAM / Calibration

THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON DC

Visual-Inertial Simultaneous Localization and Mapping

Calibration

Augmented reality
(Image CC Frédéric Bellaiche)



Multi Sensor Mobile Robotics

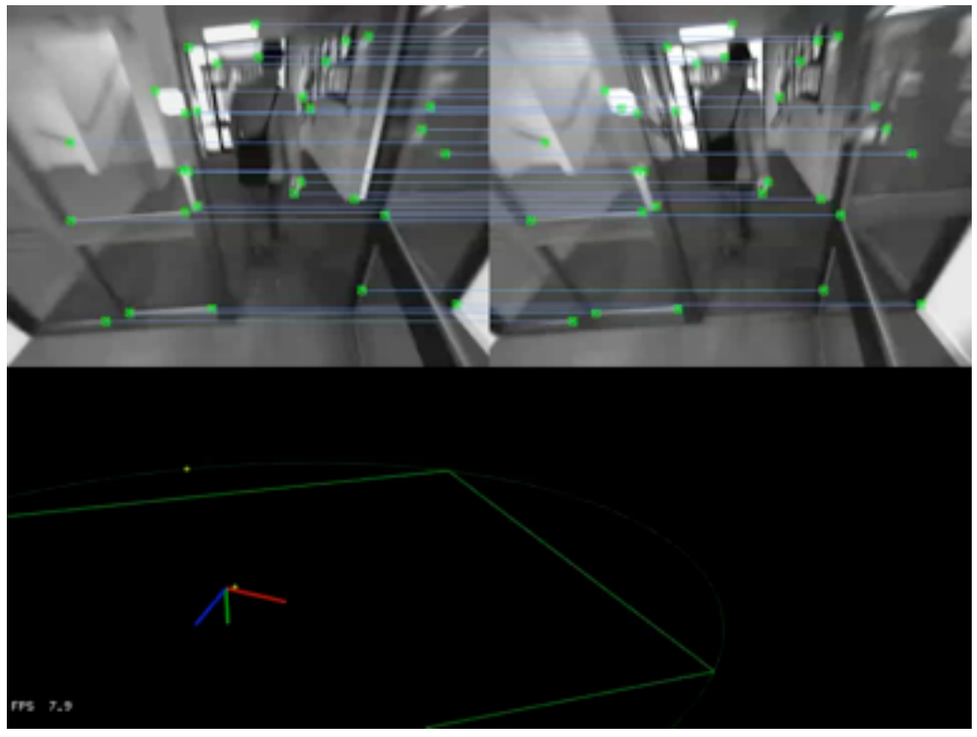
N-Camera

Unsynchronized, high-rate devices

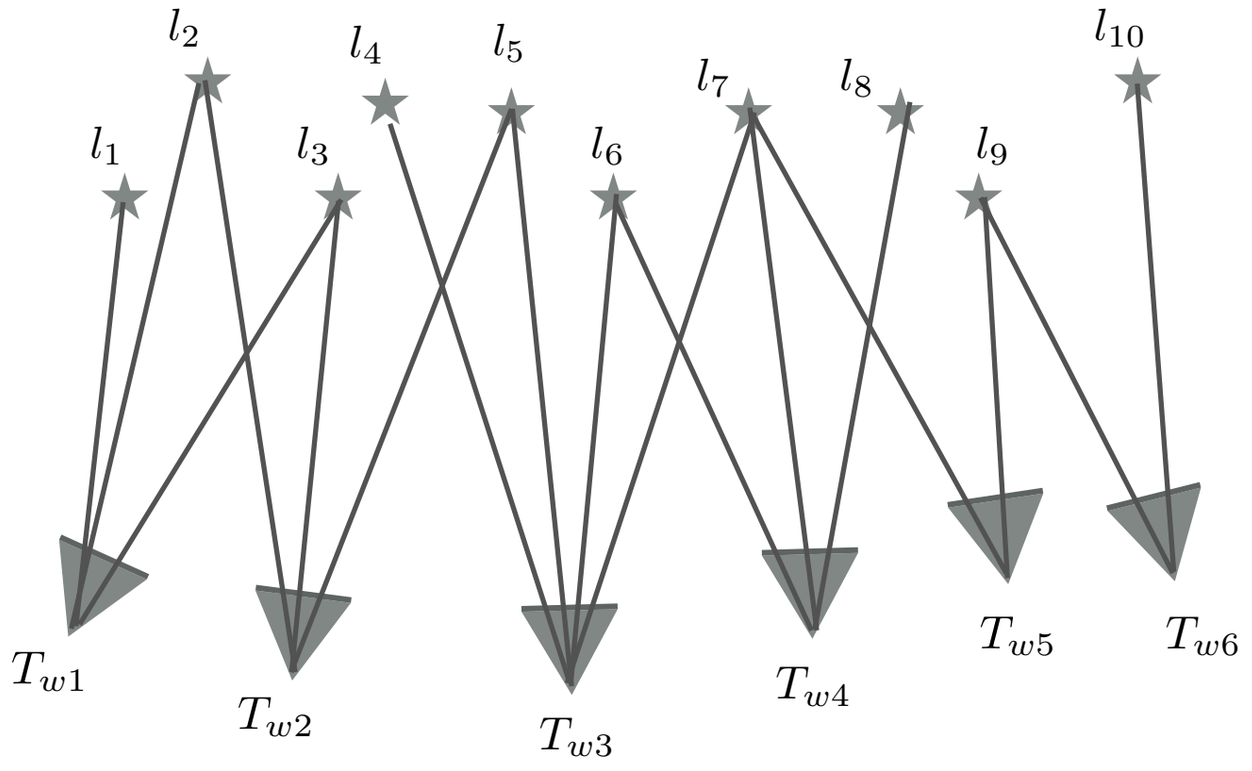
Agile cars; Work of Nima Kievan et al., GWU

Traditional Visual SLAM / SfM

- Minimize landmark reprojection error across images
- Landmark positions and camera poses are estimation parameters
- Least Squares Minimization / EKF / UKF / ...



Planes, Trains and Automobiles , Sibley et al.



What about Rolling Shutter Cameras?

THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON DC

Each line exposed at different instance in time.

Rolling Shutters are common, but usually not modeled.

Can introduce significant bias, particularly during fast motion.

CMOS Sensor Video Frame, Axel1963, Creative Commons

Synthetic Rolling Shutter

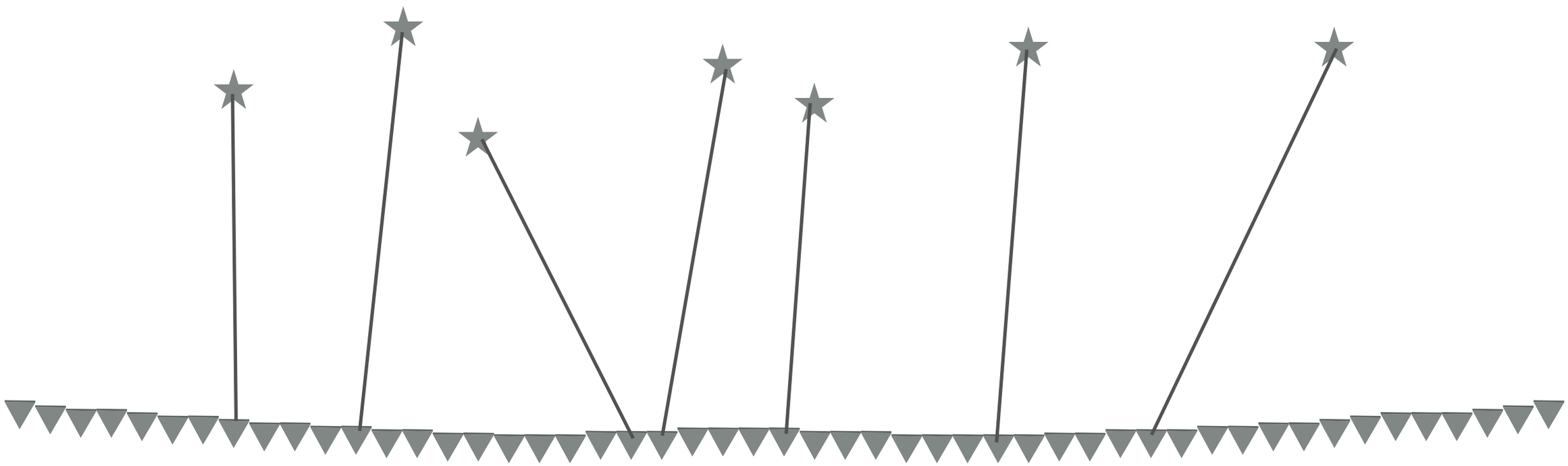
High-Rate / Unsyncrhonized Devices?

THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON DC

Estimating pose for each line / measurement may lead to under-constrained parameters, requiring motion model.

Handling unsynchronized devices within discrete time framework implicitly requires motion model.

Motion models are common in filtering, but not in BA.

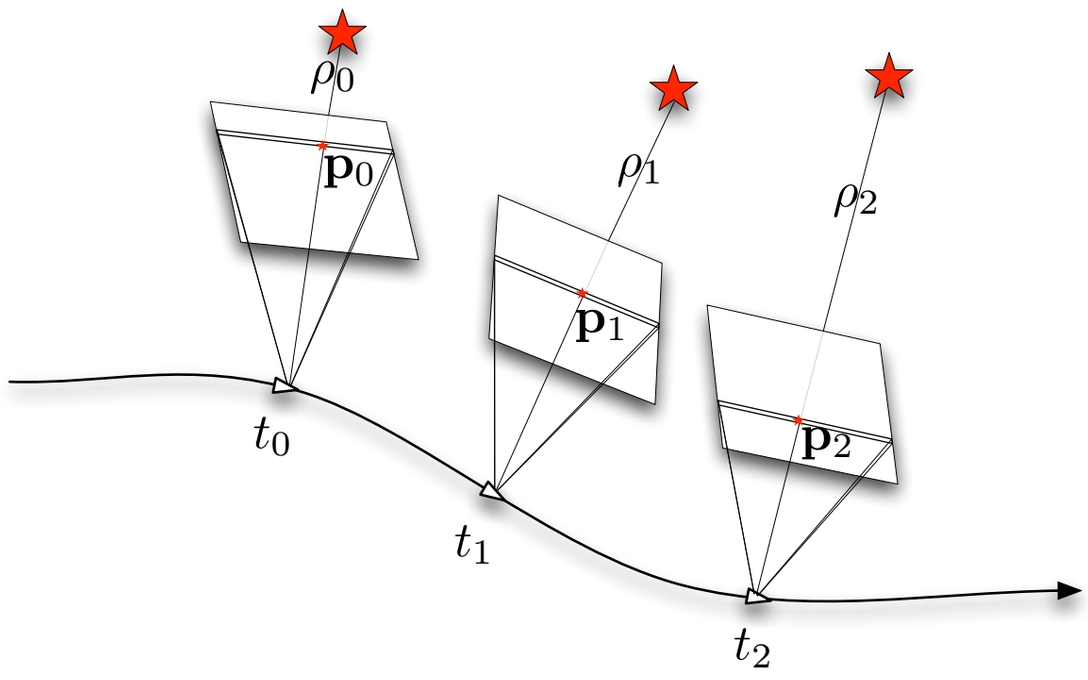


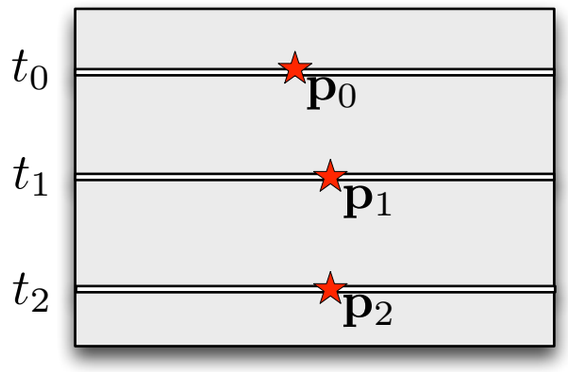
A Solution: Continuous Time SLAM

THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON DC

The pose of the camera / rig / robot can be evaluated at any moment in time, and hence predictions and observed errors can be formed at any instance in time.

Supports high-rate and unsynchronized devices easily.





Previous Research

Rolling Shutter SLAM

1. G. Klein and D. Murray. Parallel tracking and mapping on a camera phone. In *International Symposium on Mixed and Augmented Reality*, 2009.

2. J. Hedborg, P. Forssen, M. Felsberg, and E. Ringaby. Rolling shutter bundle adjustment. In *Conference on Computer Vision and Pattern Recognition*, 2012.

...

Visual-inertial Fusion / Calibration

3. G. Nuetzi, S. Weiss, D. Scaramuzza, and R Siegwart. Fusion of IMU and vision for absolute scale estimation in monocular slam. In *International Conference on Unmanned Aerial Vehicles*, 2010

4. J. Kelly and G. S. Sukhatme. Visual-inertial sensor fusion: Localization, mapping and sensor-to-sensor self-calibration". *International Journal of Robotics Research*, 2010

5. P. Furgale, T. D. Barfoot, and G. Sibley. Continuous-time batch estimation using temporal basis functions. In *International Conference on Robotics and Automation*, 2012

...

↖ Most relevant

What's the Right Representation?

How do you interpolate poses?

We want:

- Local Control - for efficient optimization
- C-2 Continuity - enable us to predict IMU
- Approximation of minimum torque trajectories

We cannot just interpolate in any 6 DoF parametrization, in general we cannot guarantee reasonable trajectories.

Quaternion Interpolation

THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON DC

LERP

Non-uniform angular velocity

SLERP

Constant piecewise angular velocity

SQUAD

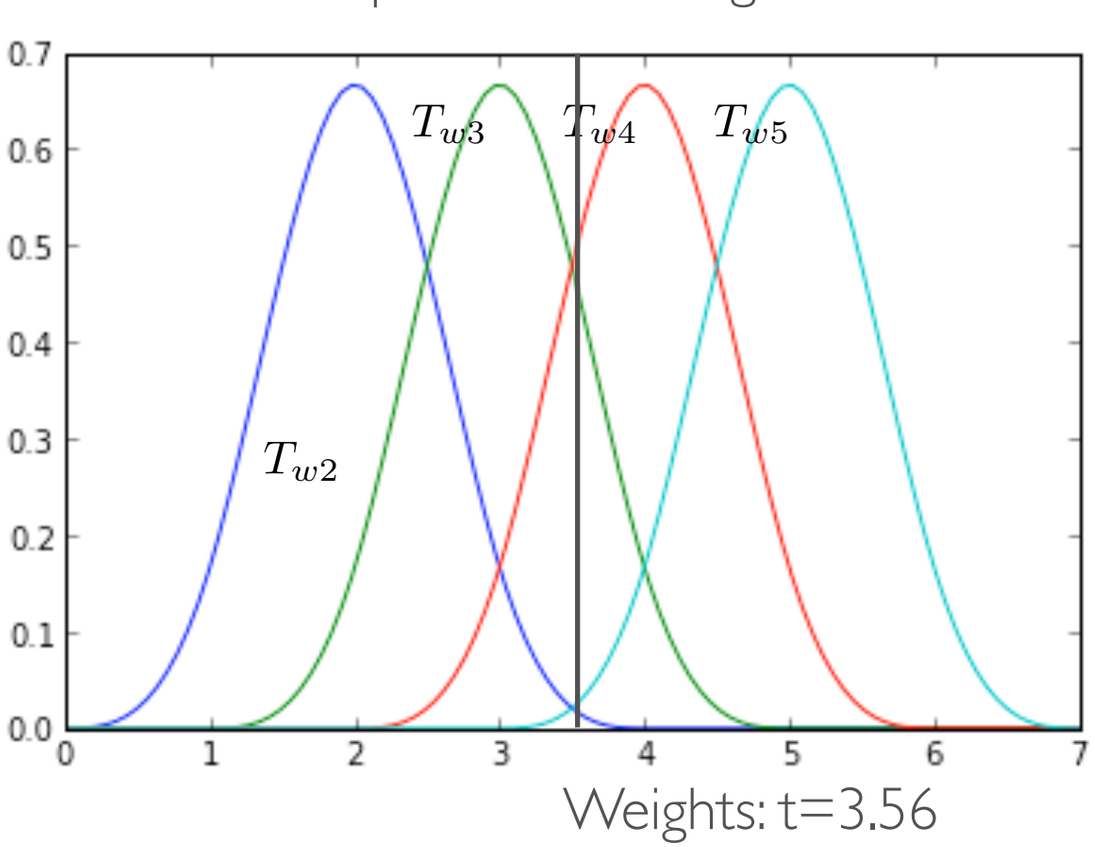
Smooth angular trajectory

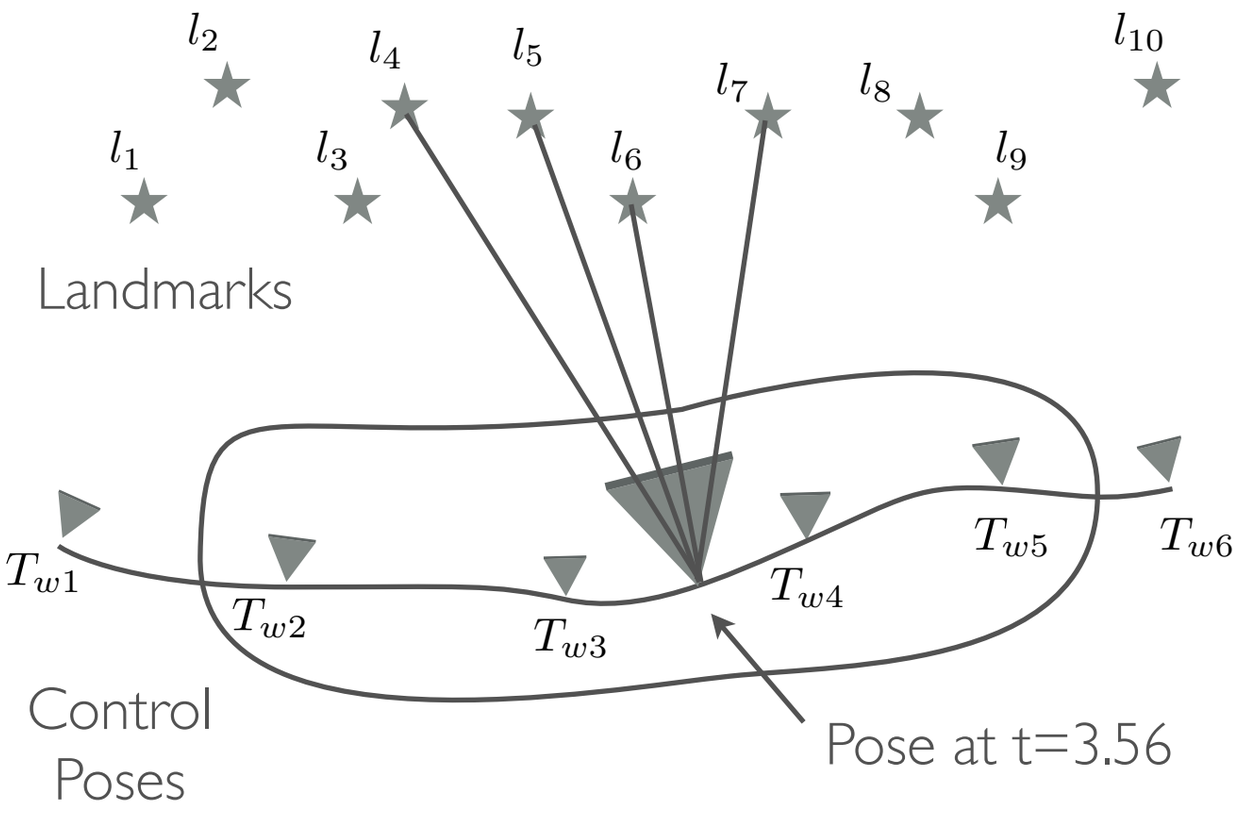
How to generalize SLERP (Shoemake '85) to higher order smoothness? Quaternion Bézier curve or recursive SQUAD not necessarily C-2 continuous and do not have simple closed-form 2nd derivatives (Kim et al. '95).

B-Splines and Control Poses

THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON DC

The pose at time t can be expressed as a weighted average of neighboring control poses expressed using this parametrization.





How do you average poses? Furgale et al. (2012) use cubic B-Spline through Cayley-Gibbs-Rodrigues parametrization.

CUMULATIVE B-SPLINES

THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON DC

Introduced by Kim et al. '95. Applied to Quaternions.

Linear interpolation: $\mathbf{p}_{i,j} \in \mathbb{R}^N, t \in \mathbb{R}$

$$\text{lerp}(\mathbf{p}_i, \mathbf{p}_j, t) = \mathbf{p}_i(1-t) + \mathbf{p}_j t$$ Weighted average form

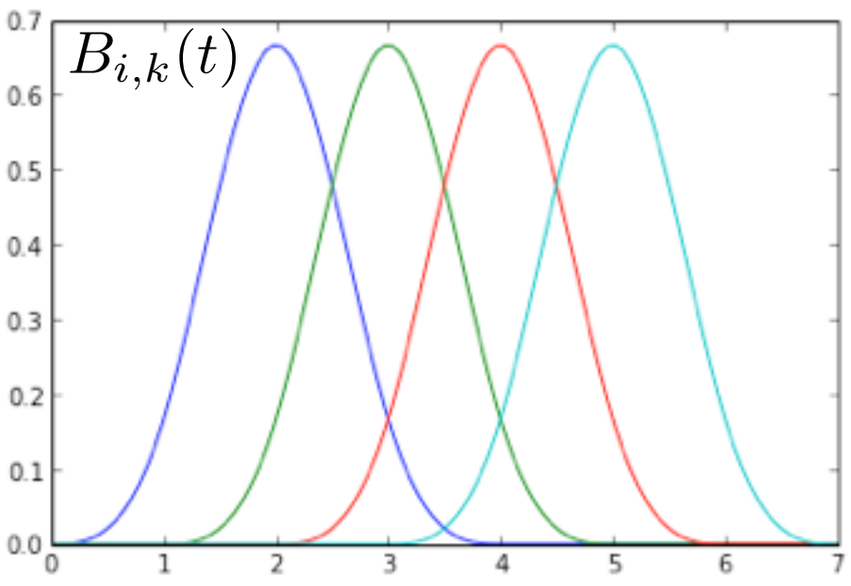
$$= \mathbf{p}_i + \underbrace{(\mathbf{p}_j - \mathbf{p}_i)}_{=\Delta_i} t$$ Cumulative form

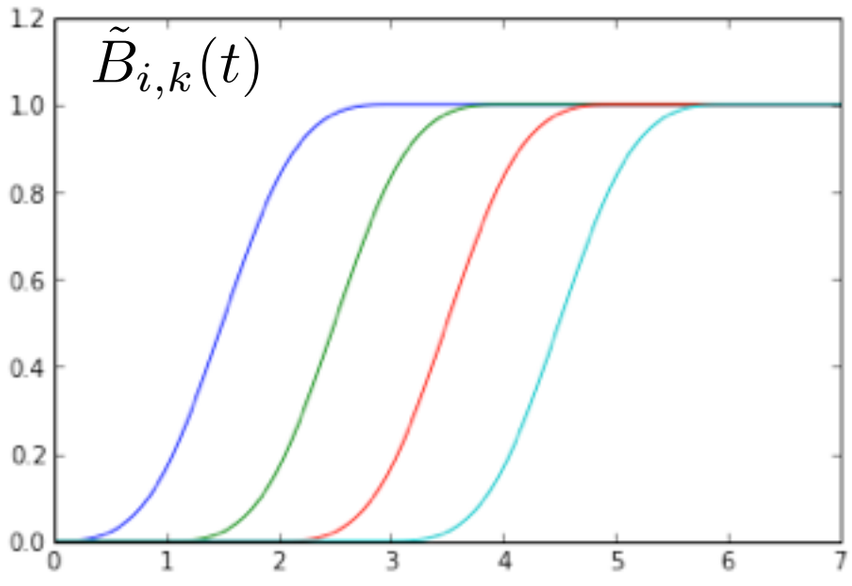
B-Splines:

$$\mathbf{p}(t) = \sum_{i=0}^{n} \mathbf{p}_i B_{i,k}(t)$$ Weighted average form

Cumulative B-Splines:

$$\mathbf{p}(t) = \mathbf{p}_0 \tilde{B}_{0,k}(t) + \sum_{i=1}^{n} (\mathbf{p}_i - \mathbf{p}_{i-1}) \tilde{B}_{i,k}(t)$$ Cumulative form



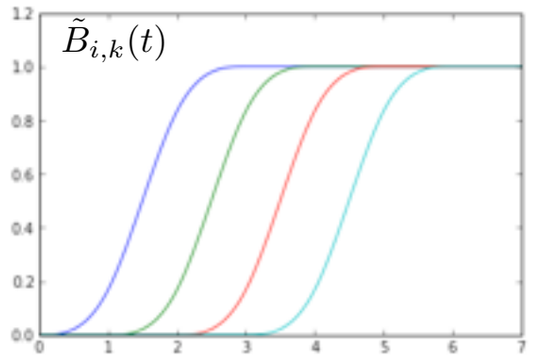


Cumulative B-Splines $\in \mathbb{SE}3$

THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON DC

Cumulative B-Splines:

$$\mathbf{p}(t) = \mathbf{p}_0 \tilde{B}_{0,k}(t) + \sum_{i=1}^{n} (\mathbf{p}_i - \mathbf{p}_{i-1}) \tilde{B}_{i,k}(t)$$



Cumulative B-Splines $\in \mathbb{SE}3$:

$$\mathbf{T}_{w,s}(t) = \exp(\underbrace{\tilde{B}_{0,k}(t) \log(\mathbf{T}_{w,0})}_{= \mathbf{T}_{w,0}}) \prod_{i=1}^{n} \exp(\tilde{B}_{i,k}(t) \underbrace{\log(\mathbf{T}_{w,i-1}^{-1} \mathbf{T}_{w,i})}_{= \Omega_i}),$$

Using the cumulative form, we need only take local differences on the manifold of $\in \mathbb{SE}3$ using log.

These differences $\in \mathfrak{se}3$ represent rotational velocities, and it's these that get blended.

Closed Form 1st and 2nd Derivatives

$$\tilde{\mathbf{B}}(u) = \mathbf{C}\begin{bmatrix}1\\u\\u^2\\u^3\end{bmatrix}, \quad \dot{\tilde{\mathbf{B}}}(u) = \frac{1}{\Delta t}\mathbf{C}\begin{bmatrix}0\\1\\2u\\3u^2\end{bmatrix}, \quad \ddot{\tilde{\mathbf{B}}}(u) = \frac{1}{\Delta t^2}\mathbf{C}\begin{bmatrix}0\\0\\2\\6u\end{bmatrix}, \quad \mathbf{C} = \frac{1}{6}\begin{bmatrix}6 & 0 & 0 & 0\\5 & 3 & -3 & 1\\1 & 3 & 3 & -2\\0 & 0 & 0 & 1\end{bmatrix}$$

$$\mathbf{T}_{w,s}(u) = \mathbf{T}_{w,i-1}\prod_{j=1}^{3}\exp\left(\tilde{\mathbf{B}}(u)_j\,\boldsymbol{\Omega}_{i+j}\right),$$

$$\dot{\mathbf{T}}_{w,s}(u) = \mathbf{T}_{w,i-1}\left(\dot{\mathbf{A}}_0\mathbf{A}_1\mathbf{A}_2 + \mathbf{A}_0\dot{\mathbf{A}}_1\mathbf{A}_2 + \mathbf{A}_0\mathbf{A}_1\dot{\mathbf{A}}_2\right),$$

$$\ddot{\mathbf{T}}_{w,s}(u) = \mathbf{T}_{w,i-1}\begin{pmatrix}\ddot{\mathbf{A}}_0\mathbf{A}_1\mathbf{A}_2 + \mathbf{A}_0\ddot{\mathbf{A}}_1\mathbf{A}_2 + \mathbf{A}_0\mathbf{A}_1\ddot{\mathbf{A}}_2 +\\ 2\left(\dot{\mathbf{A}}_0\dot{\mathbf{A}}_1\mathbf{A}_2 + \dot{\mathbf{A}}_0\mathbf{A}_1\dot{\mathbf{A}}_2 + \mathbf{A}_0\dot{\mathbf{A}}_1\dot{\mathbf{A}}_2\right)\end{pmatrix},$$

$$\mathbf{A}_j = \exp\left(\boldsymbol{\Omega}_{i+j}\tilde{\mathbf{B}}(u)_j\right), \quad \dot{\mathbf{A}}_j = \mathbf{A}_j\boldsymbol{\Omega}_{i+j}\dot{\tilde{\mathbf{B}}}(u)_j,$$

$$\ddot{\mathbf{A}}_j = \dot{\mathbf{A}}_j\boldsymbol{\Omega}_{i+j}\dot{\tilde{\mathbf{B}}}(u)_j + \mathbf{A}_j\boldsymbol{\Omega}_{i+j}\ddot{\tilde{\mathbf{B}}}(u)_j$$

Objective Function

THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON DC

Camera Prediction

$$\mathbf{p}_b = \mathcal{W}(\mathbf{p}_a; \mathbf{T}_{b,a}, \rho) = \pi\Big([\mathbf{K}_b \,|\, \mathbf{0}]\, \mathbf{T}_{b,a} \left[\mathbf{K}_a^{-1} \left[\begin{smallmatrix} \mathbf{p}_a \\ 1 \end{smallmatrix} \right]; \rho \right] \Big)$$

IMU Prediction

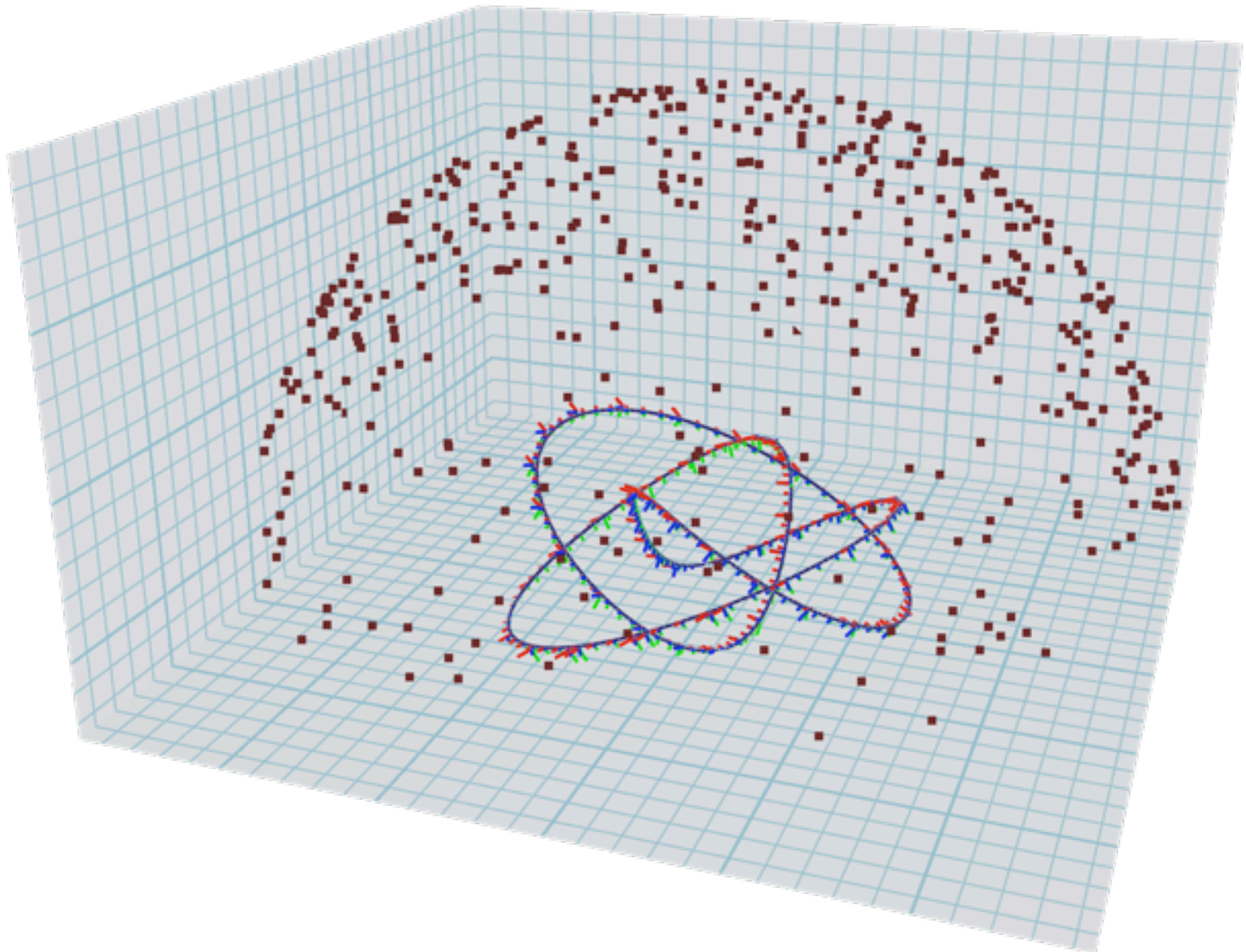
$$\begin{aligned} \text{Gyro}(u) = \quad & \mathbf{R}_{w,s}^{\intercal}(u) \cdot \dot{\mathbf{R}}_{w,s}(u) + \text{bias}, \\ \text{Accel}(u) = \quad & \mathbf{R}_{w,s}^{\intercal}(u) \cdot (\ddot{\mathbf{s}}_w(u) + g_w) + \text{bias} \end{aligned}$$

Objective Function

$$\begin{aligned} E(\boldsymbol{\theta}) = \quad & \sum_{\hat{\mathbf{p}}_m} \Big(\hat{\mathbf{p}}_m - \mathcal{W}(\mathbf{p}_r; \mathbf{T}_{c,s} \mathbf{T}_{w,s}(u_m)^{-1} \mathbf{T}_{w,s}(u_r) \mathbf{T}_{s,c}, \rho) \Big)^2_{\Sigma_p} + \\ & \sum_{\hat{\boldsymbol{\omega}}_m} \Big(\hat{\boldsymbol{\omega}}_m - \text{Gyro}(u_m) \Big)^2_{\Sigma_{\boldsymbol{\omega}}} + \sum_{\hat{\mathbf{a}}_m} \Big(\hat{\mathbf{a}}_m - \text{Accel}(u_m) \Big)^2_{\Sigma_{\mathbf{a}}} \end{aligned}$$

SIMULATED VISUAL-INERTIAL SLAM

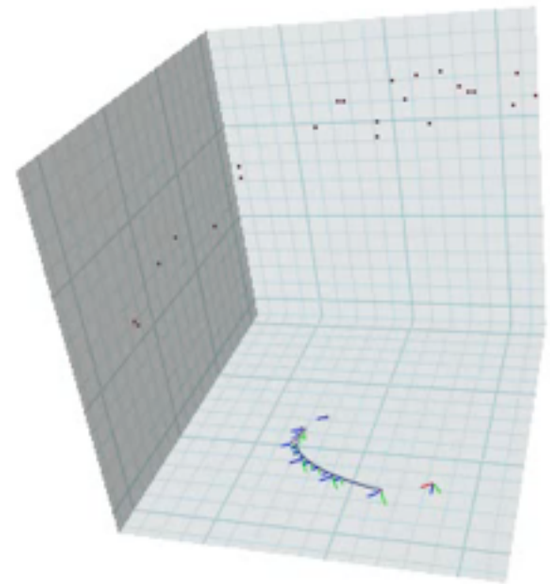
THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON DC



SIMULATED VISUAL-INERTIAL SLAM

THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON DC

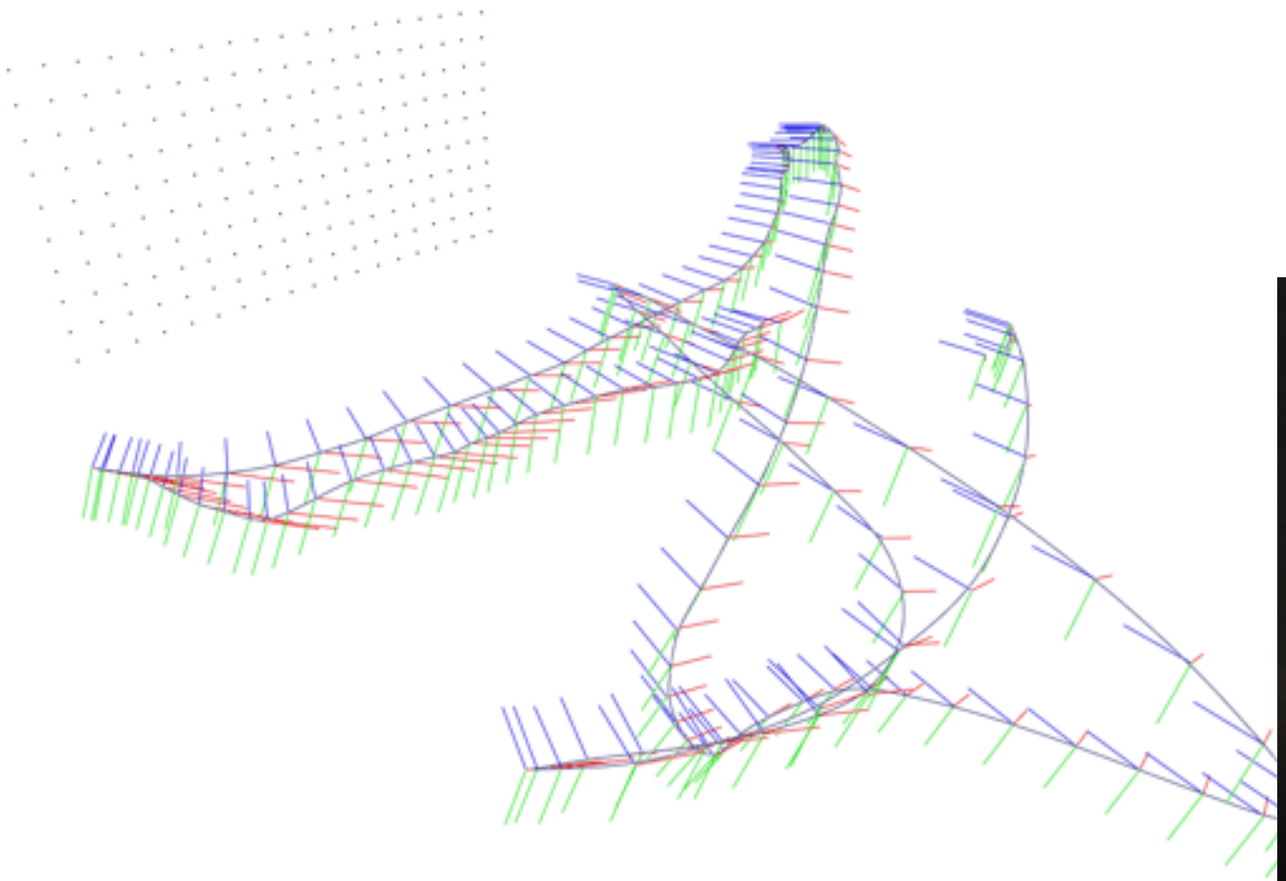
show_time:	show_time: 0



Rolling Shutter and IMU

THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON DC

Recover calibration and scale

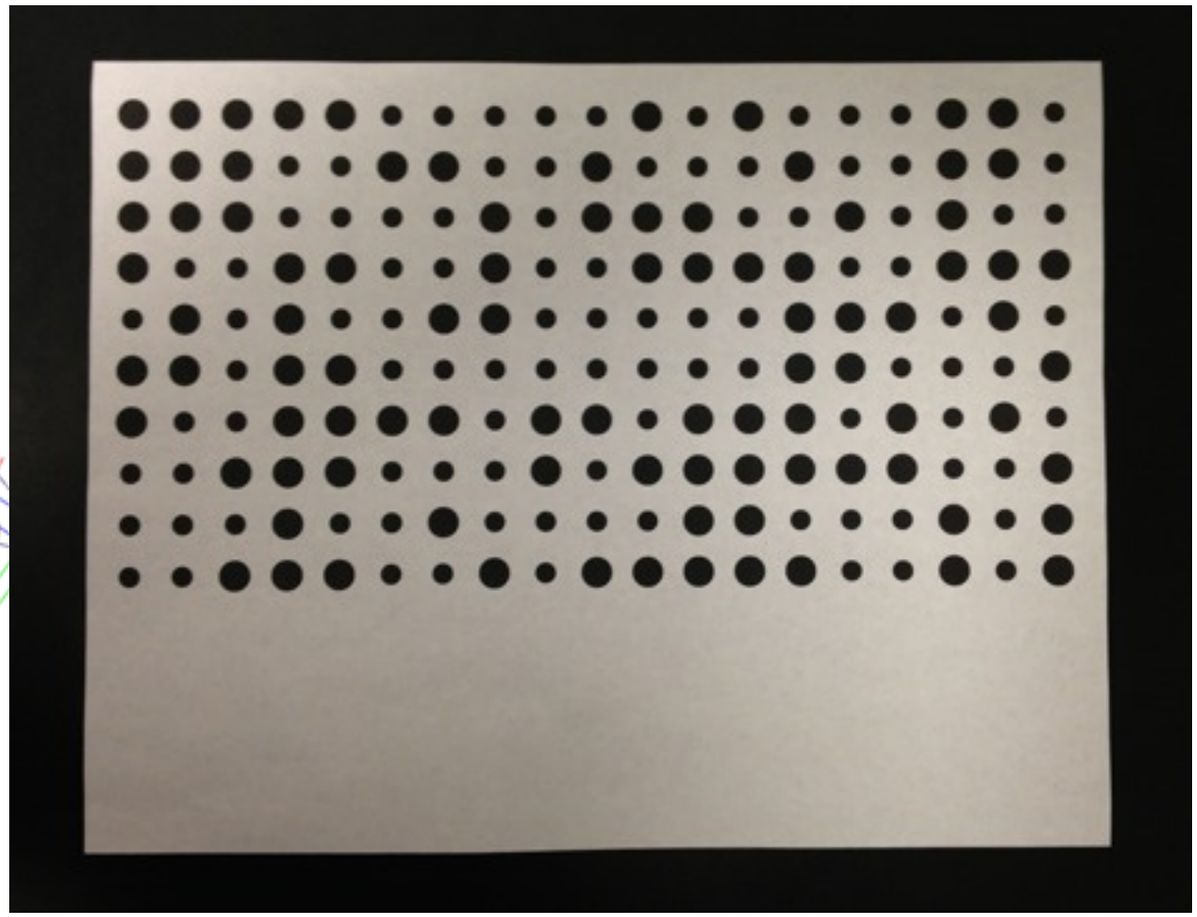


Unknown landmarks
Known data association

True Distance: 25.45cm

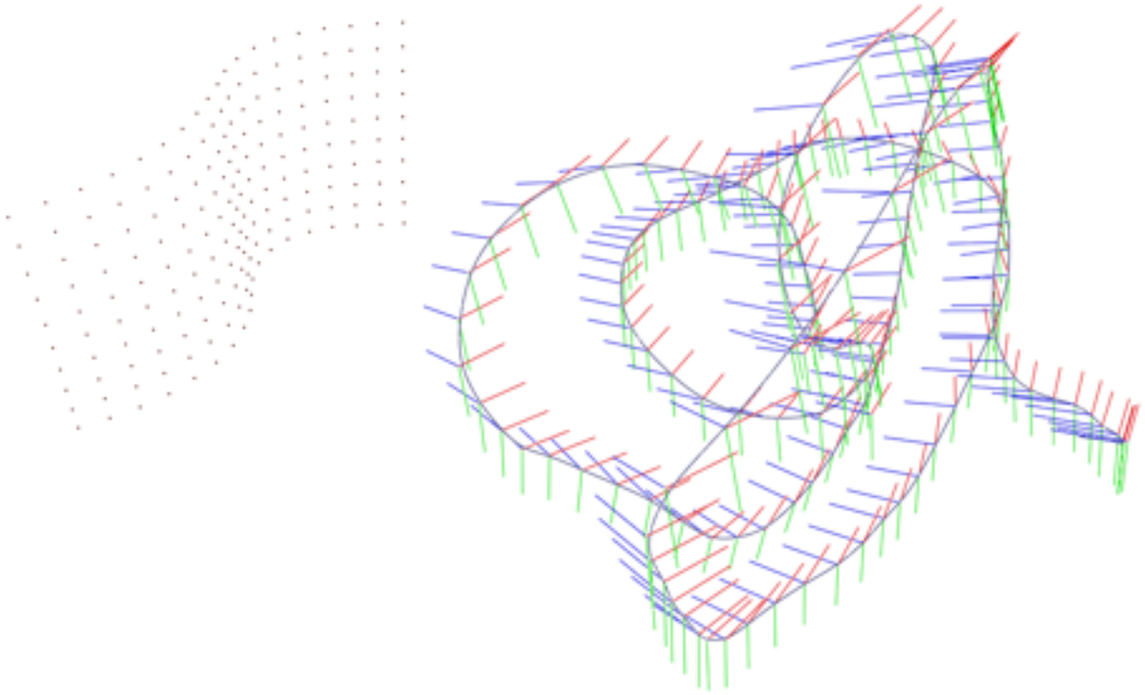
Global Shutter: 25.31cm (0.55% err)

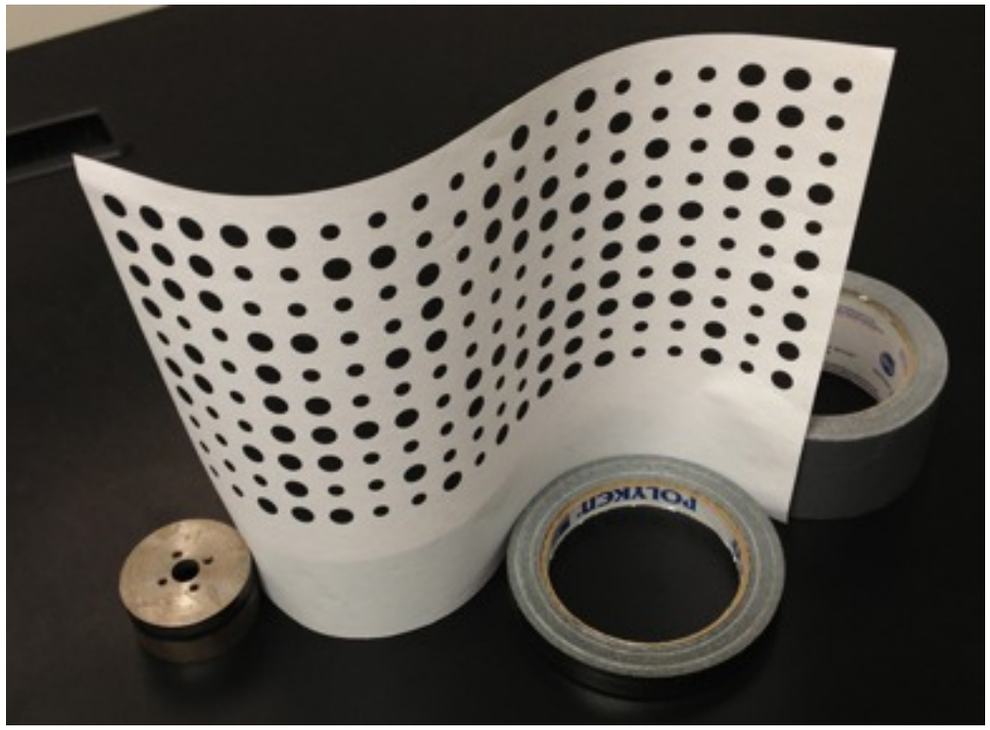
Rolling Shutter: 25.37cm (0.31% err)

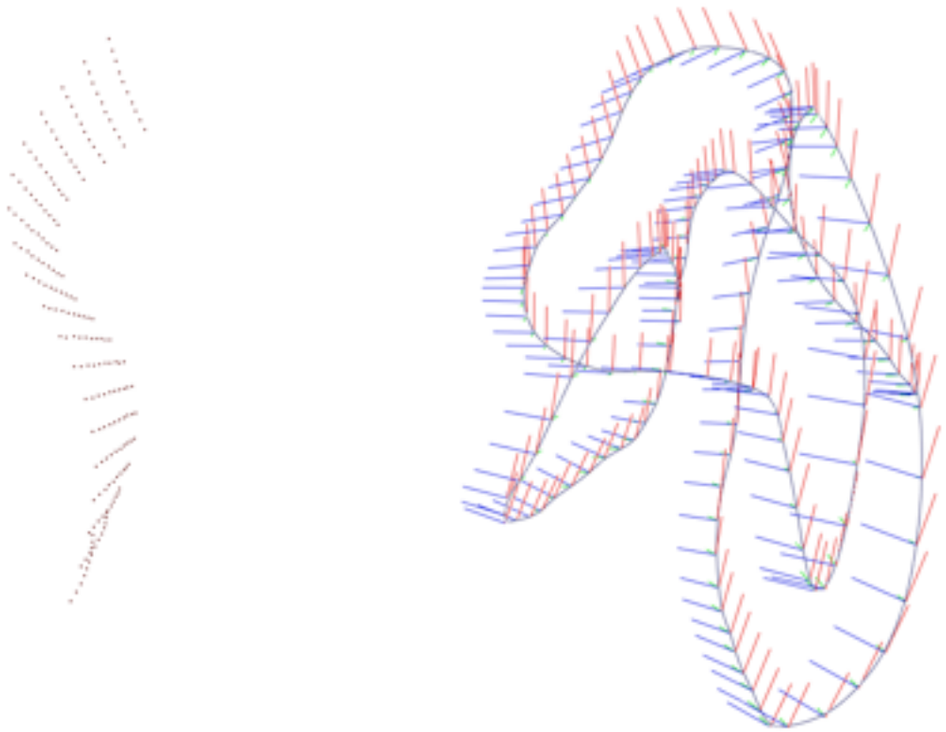


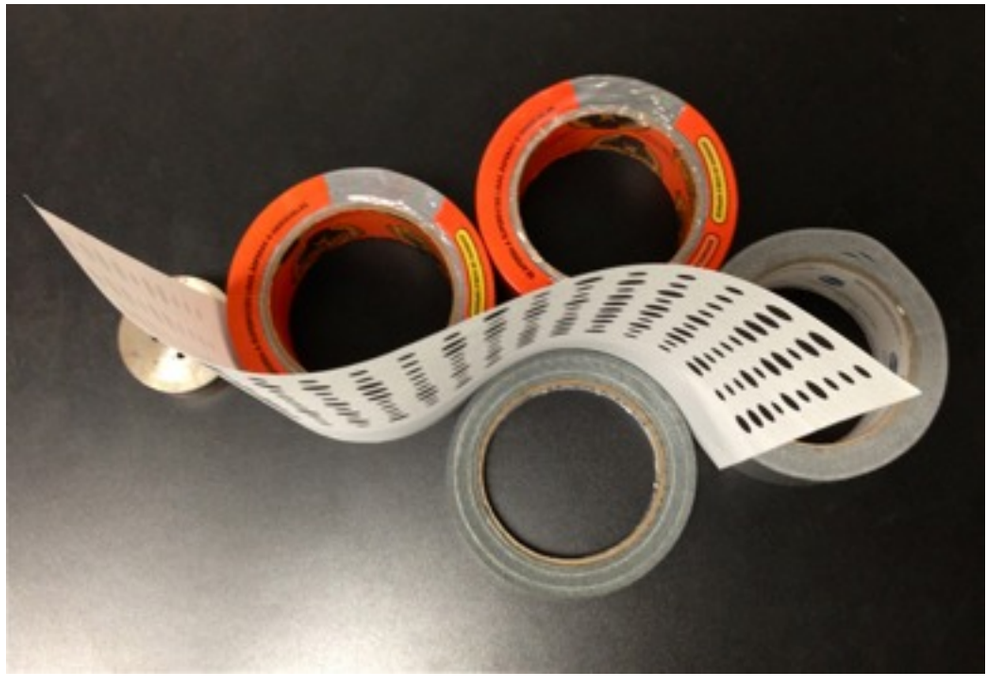
ROLLING SHUTTER AND IMU

THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON DC



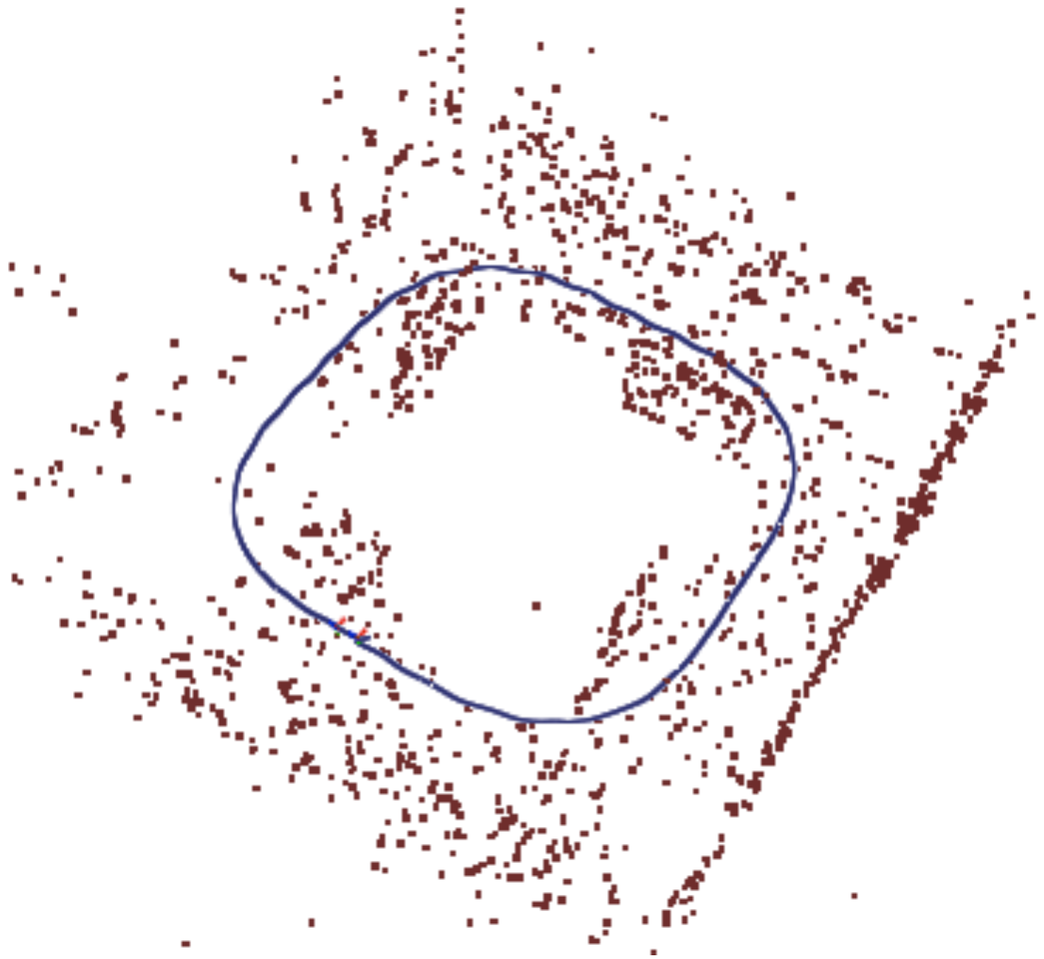






MONOCULAR ROLLING SHUTTER SLAM

THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON DC



Prediction Into Rolling Shutter

Linearization of projection

$$\mathbf{p}_b(t+\Delta t) = \mathcal{W}(\mathbf{p}_a; \mathbf{T}_{b,a}(t), \rho) + \Delta t \frac{d\,\mathcal{W}(\mathbf{p}_a; \mathbf{T}_{b,a}(t), \rho)}{d\,t}$$

Solve for time offset

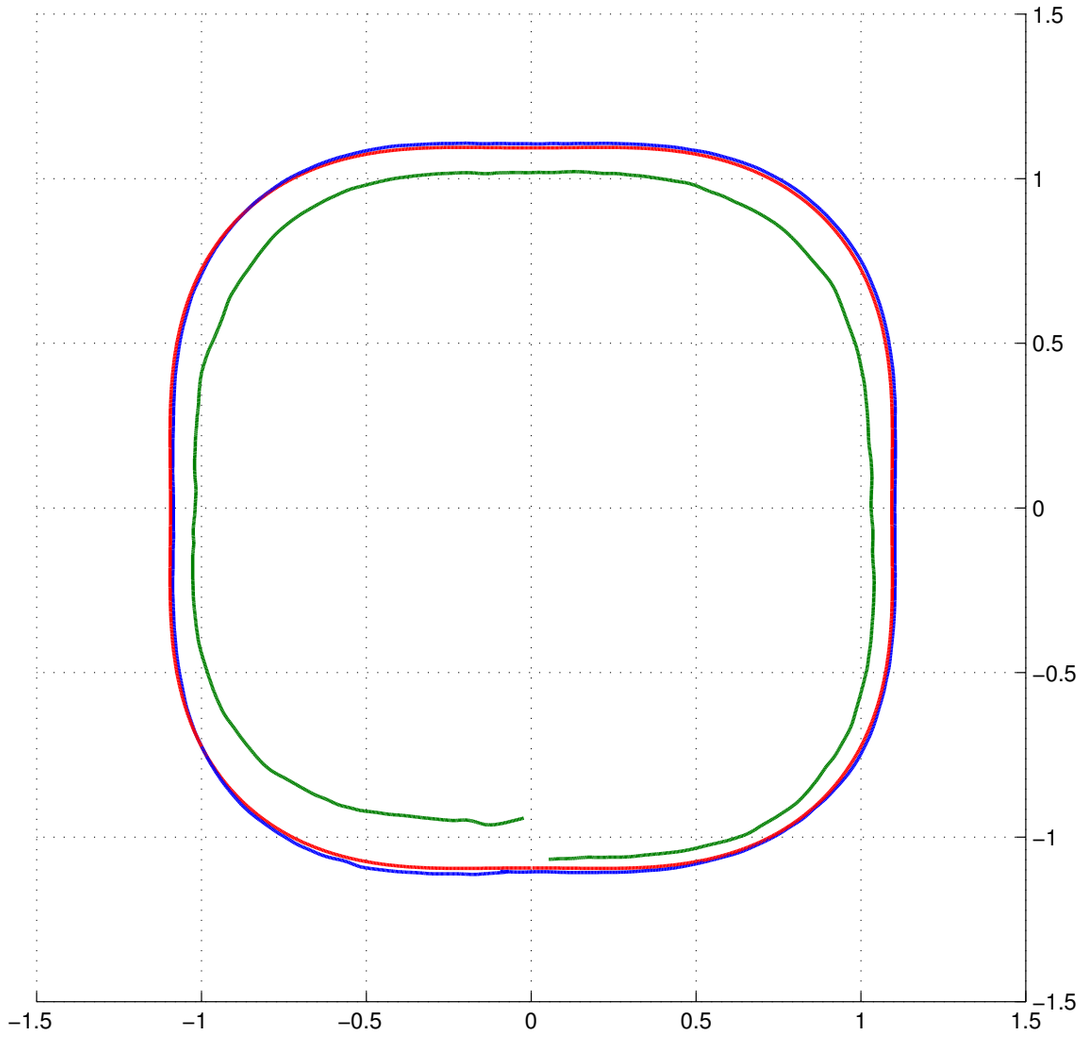
$$y_{b(t+\Delta t)} = \frac{h(t+\Delta t - s)}{e-s}$$

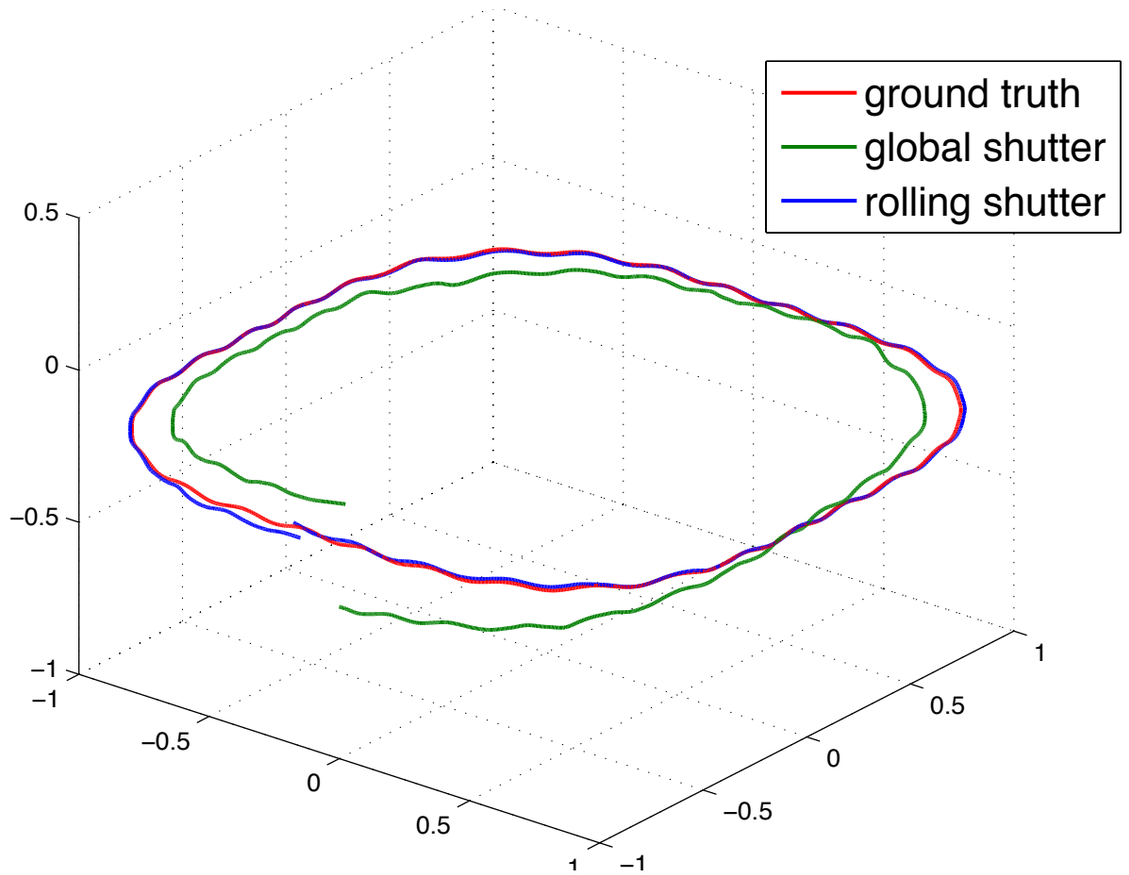
$$\Delta t = -\frac{h.t_0 + s.(y_b(t) - h) - e.y_b(t)}{(s-e)\frac{d\,\mathcal{W}_y(\mathbf{p}_a; \mathbf{T}_{b,a}(t), \rho)}{d\,t} + h}$$

~ 2-3 iterations

Improved Monocular Odometry

THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON DC





SPLINE FUSION

THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON DC

For the future:

- Speed system up - runs at just few FPS right now
- Integrate with better front end
- Improve initialization
- Variable control pose placement

THANKS!
ANY QUESTIONS?