Autonomous Navigation in Complex Indoor and Outdoor Environments with Micro Aerial Vehicles

Shaojie Shen

Dept. of Electrical and Systems Engineering & GRASP Lab,

University of Pennsylvania

Committee:

- Daniel D. Lee (Chair)
- Vijay Kumar (Advisor)
- Nathan Michael (Advisor)
- Davide Scaramuzza (Member)

Micro Aerial Vehicles

- Low cost, small size (<1m) , adequate payload (1-5kg)
- Superior mobility for indoor and outdoor applications

Transportation

Search and rescue

Aerial photography

Inspection

Law enforcement

Agriculture

Penn

How to Fly a MAV?

- Remote control
  - Requires line of sight and/or communication link
  - Requires skilled pilots
- Inertial navigation
  - Requires aviation grade IMU
  - Heavy and expensive
- GPS-based navigation
  - GPS can be unreliable

Autonomy for MAVs

- Sensing and perception
- State estimation

+

- Control
- Planning

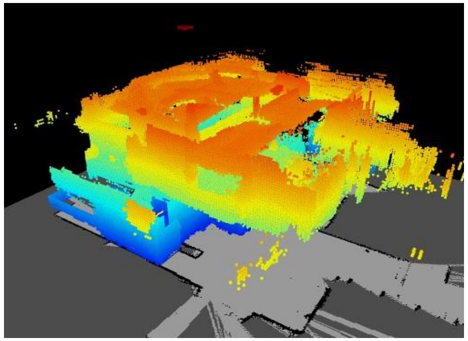
=

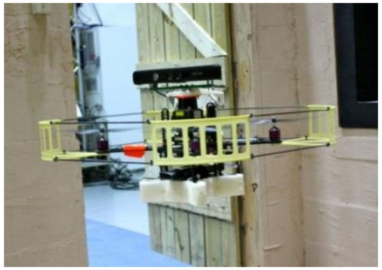
- System integration

Current Research Focus

Build on/Extend Previous Work

Prototype Development



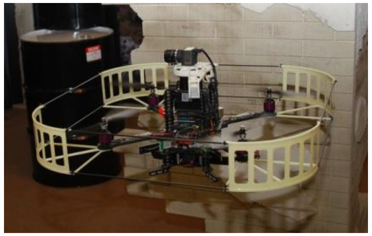


Challenges

- Stabilizing fast vehicle dynamics (5^{th} order system)
  - Requires real-time onboard processing with low latency
  - Requires accurate state estimates
- Limited payload (< 1kg)
  - Limited sensing
  - Limited computation
- Complex environments
  - Unknown environments
  - GPS unreliable or unavailable

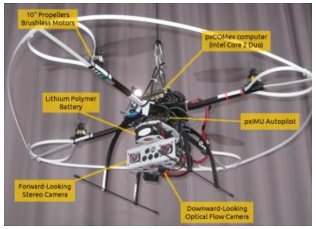
Related Work

- Laser + IMU (Shen, et al. 2011; Bry, et al. 2012)
  - Pros: Computationally very efficient
  - Cons: Requires structured or known environments
- Monocular Camera + IMU (Kottas, et al. 2012; Weiss, et al. 2013)
  - Pros: Low cost, relatively fast processing
  - Cons: Requires good initialization
- Stereo Cameras + IMU (Fraundorfer, et al. 2012)
  - Pros: Directly observable scale
  - Cons: Limited baseline
- RGB-D Sensor + IMU (Huang, et al. 2011)
  - Pros: Depth info directly available
  - Cons: Does not work outdoor
- EKF-Based Multi Sensor Fusion (Lynen, et al. 2013)
  - Pros: Well documented and open source
  - Cons: Considered limited sensor setup

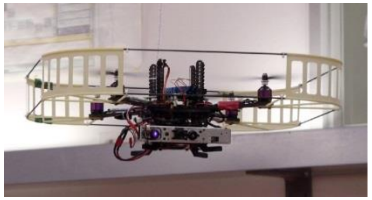


Shen, et al, 2011

Weiss, et al, 2013



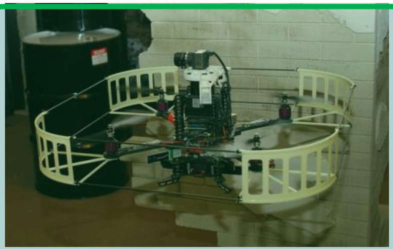
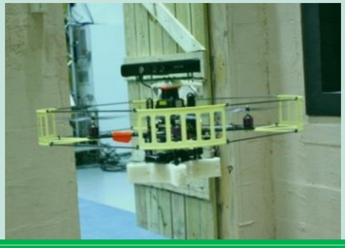
Fraundorfer, et al, 2012



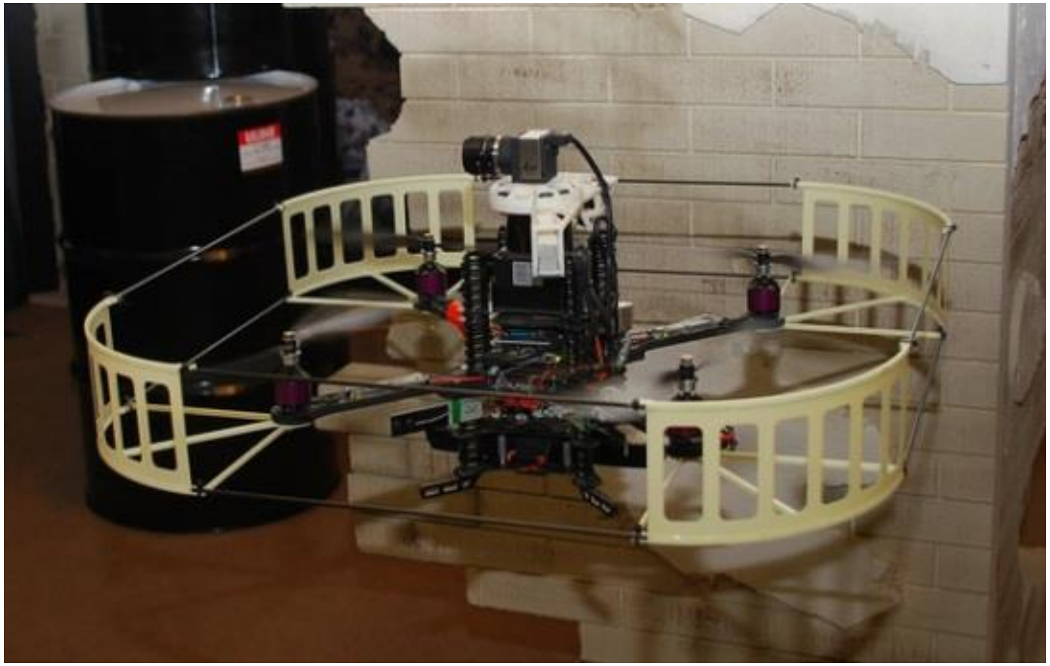
Huang, et al, 2011

State Estimation for MAVs

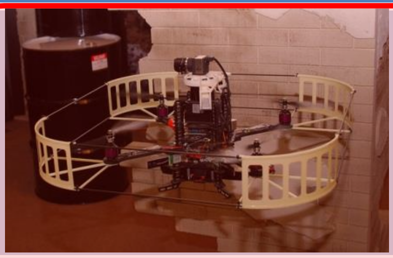
- Power-on-and-go
  – Initialize from an arbitrary unknown state
- Autonomy
  – State estimation in a wide range of environments
- Fault-tolerant
  – Handle failure of one or more onboard sensors
- Fail-safe
  – Recover from total failure of all sensors

Robot	Sensing	Computation	Mass	Environment	Year
	Laser IMU	Intel Atom 1.6GHz	1.7 kg	2.5D indoor	2010-2011
	Laser Kinect IMU	Intel Atom 1.6GHz	1.9 kg	2.5D indoor	2011-2012
	Cameras IMU	Intel Atom 1.6GHz	0.74 kg	3D indoor and limited outdoor	2012-2013
	Laser Cameras GPS IMU	Intel Core i3 1.8GHz	1.9 kg	3D indoor and outdoor	2013-2014
	Camera IMU	Intel Core i3 1.8GHz	1.3 kg	3D indoor and limited outdoor	2014

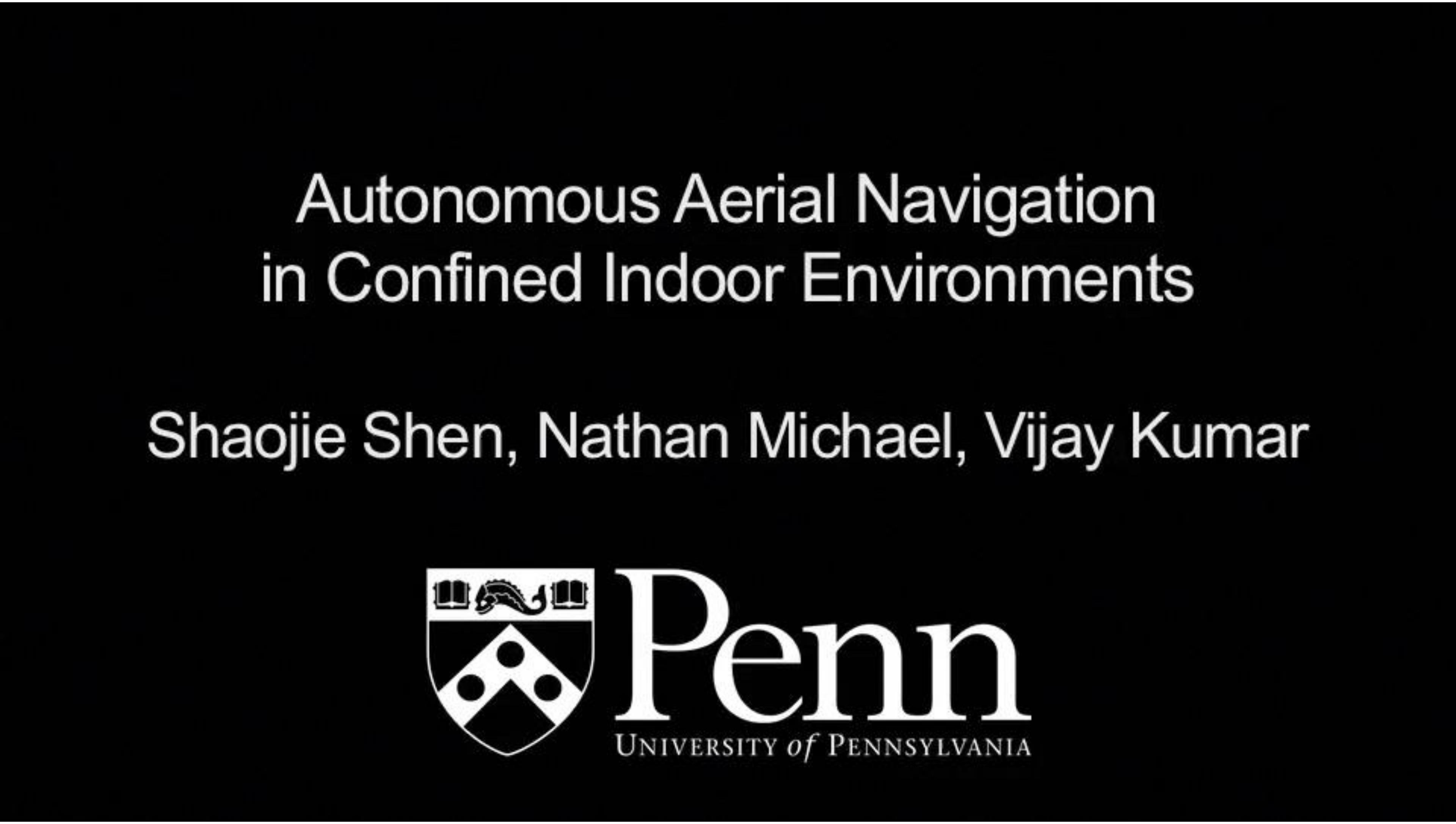
The first self-contained autonomous indoor MAV



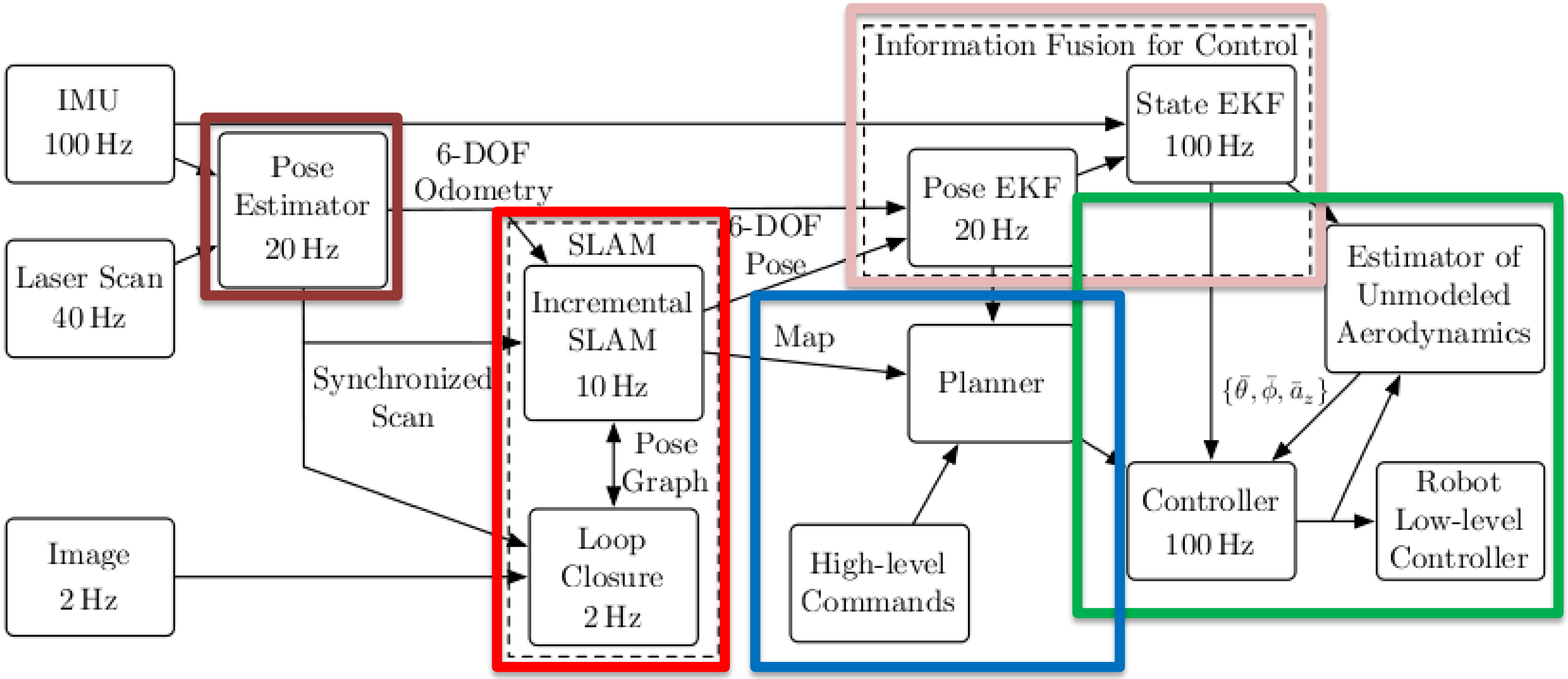
Laser + IMU

Robot	Sensing	Computation	Mass	Environment	Year
	Laser IMU	Intel Atom 1.6GHz	1.7 kg	2.5D indoor	2010-2011
	Laser Kinect IMU	Intel Atom 1.6GHz	1.9 kg	2.5D indoor	2011-2012
	Cameras IMU	Intel Atom 1.6GHz	0.74 kg	3D indoor and limited outdoor	2012-2013
	Laser Cameras GPS IMU	Intel Core i3 1.8GHz	1.9 kg	3D indoor and outdoor	2013-2014
	Camera IMU	Intel Core i3 1.8GHz	1.3 kg	3D indoor and limited outdoor	2014

Laser-Based Autonomous Indoor Navigation

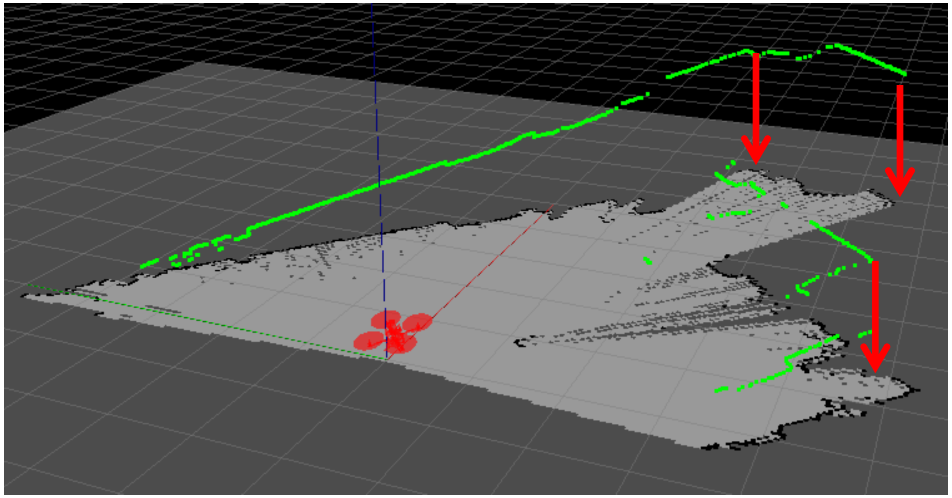


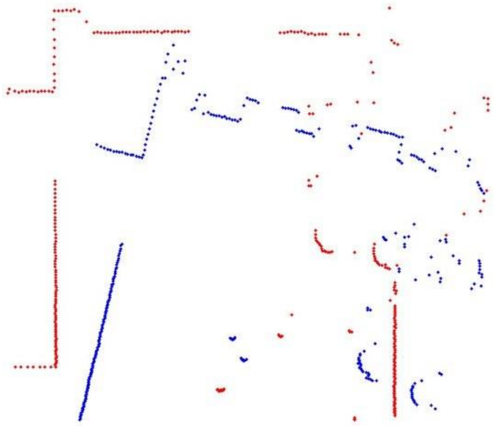
Laser-Based Autonomous Flight



6DOF Pose Estimation with Laser

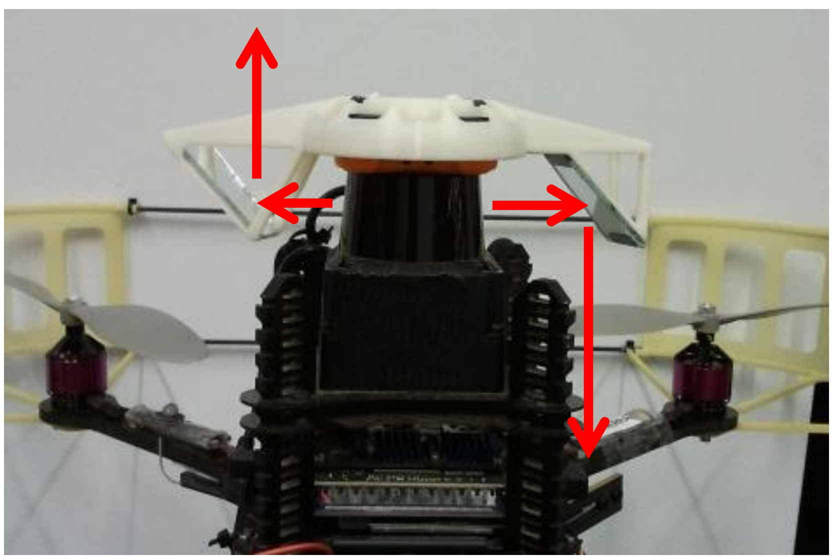
- Laser scans 2D slices of the environment
- 2.5D indoor environment
  - All walls are vertical
  - Given attitude from the IMU, laser scans can be projected onto a common ground plane.
  - 2D scan matching for x, y, and yaw

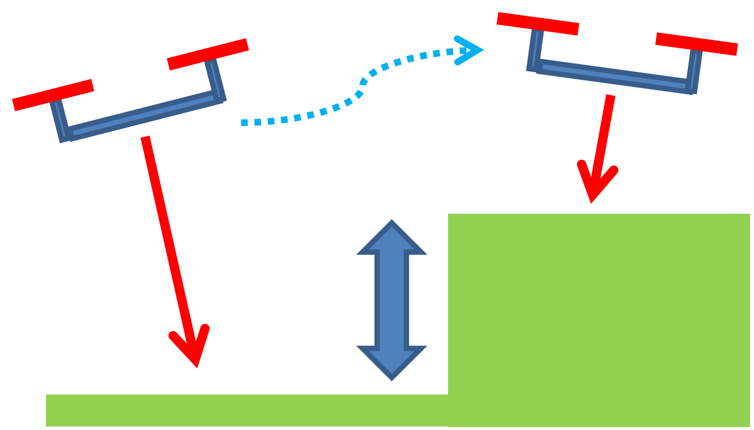




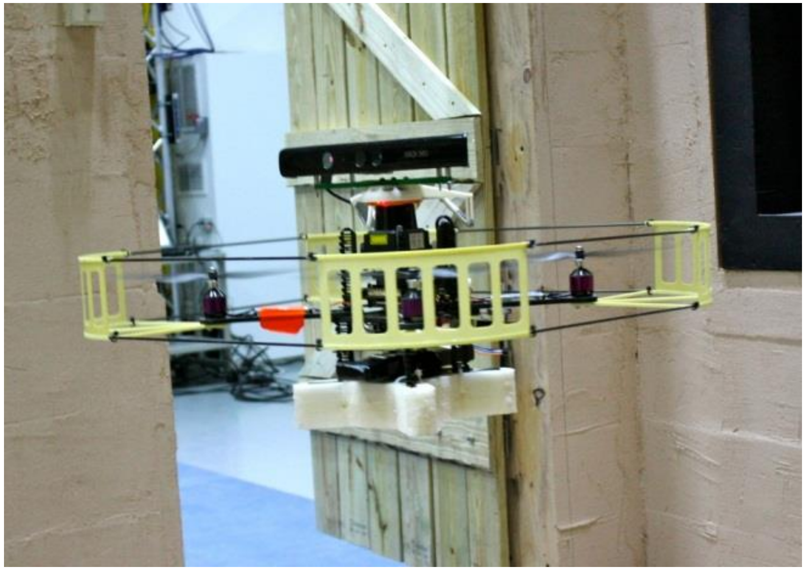
6DOF Pose Estimation with Laser

- Height measurement from redirected laser beams
- Floor detection by fusing IMU and downward laser beams with a Kalman filter





Let the MAV build 3D maps fully autonomously!

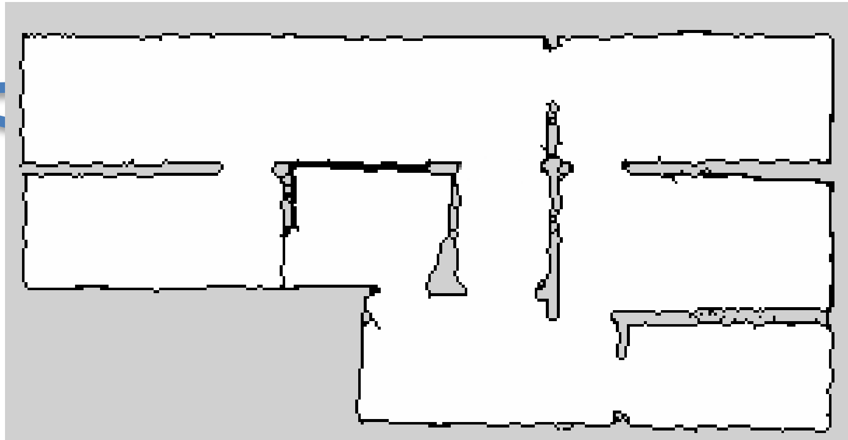


Laser + Kinect + IMU

Exploration – Autonomous Environment Coverage

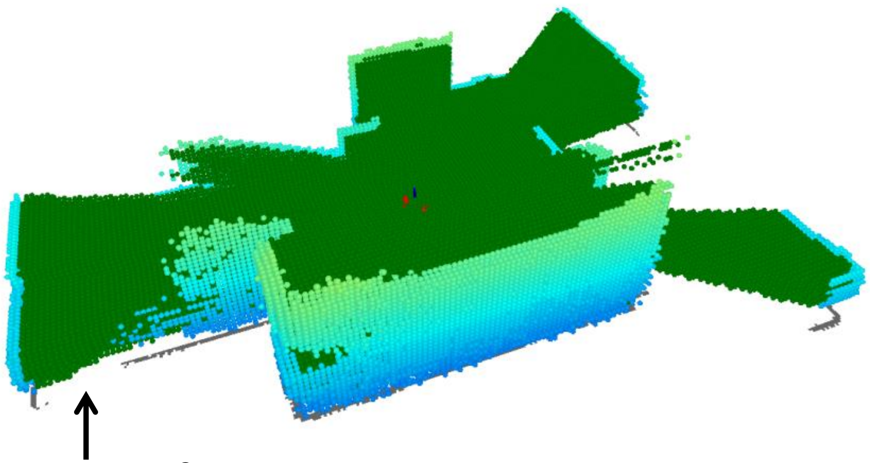
- 2D frontier-based exploration for single- and multi-ground/aerial robot applications
  - Explore boundaries between unoccupied and unknown regions
  - Yamauchi, 1997; Burgard et al., 2005; Fox et al., 2006; Vincent et al., 2008; Bachrach et al., 2011; Pravitra et al., 2011;

Frontier regions

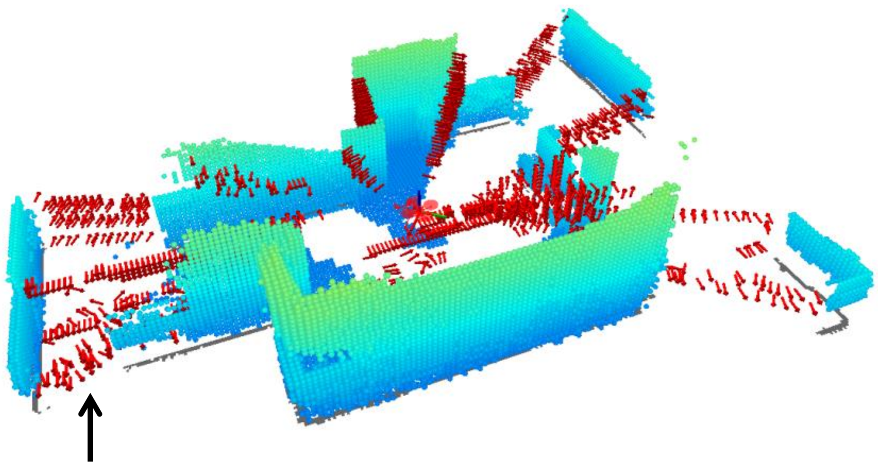


3D Exploration – Challenges

- Challenges to frontier-based approach in 3D
  - Incremental dense free space and frontier regions update
    - Computation and memory demanding
    - Suboptimal exploration behavior due to limited sensing capability



Dense free space



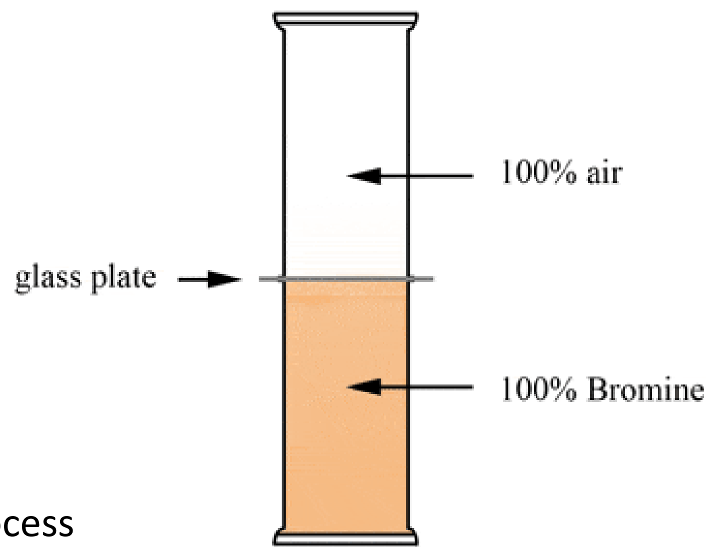
Frontier goals

Stochastic Differential Equation-based Exploration

- Intuition:
  - Gas fills the container – gas "explores" the environment
  - Gas molecule follows a stochastic differential equation (Langevin equation)
    - $\underbrace{\ddot{x}_i(t) = \cancel{-\nabla U(x_i(t))} - \tau^{-1}\dot{x}_i(t)}_{\text{Deterministic Newtonian dynamics}} + \underbrace{\sqrt{2\tau^{-1}k_b T\eta(t)}}_{\text{Stochastic diffusion term}}$ ***Ideal Gas***

Deterministic Newtonian dynamics

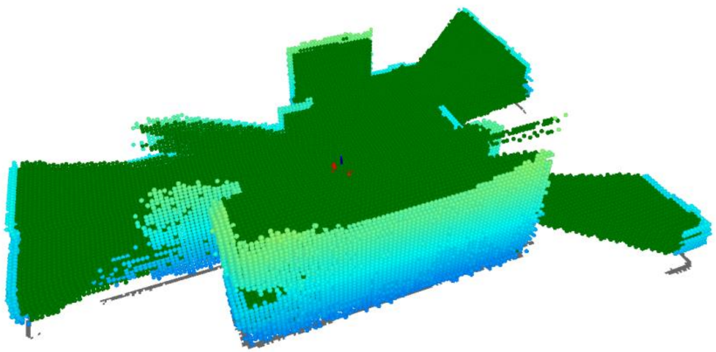
Stochastic diffusion term



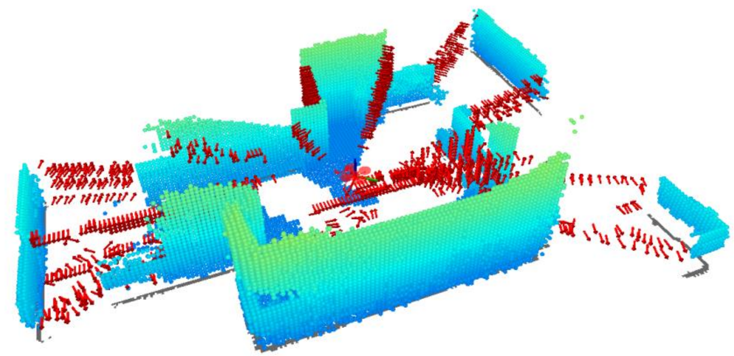
τ: Time constant
k_b: Boltzmann constant
T: Temperature
$\eta(t)$: δ-correlated stationary zero-mean Gaussian process

* Source: http://www.micromountain.com/sci_diagrams/sci_app/sci_app_pages/gasdiff_anim.htm

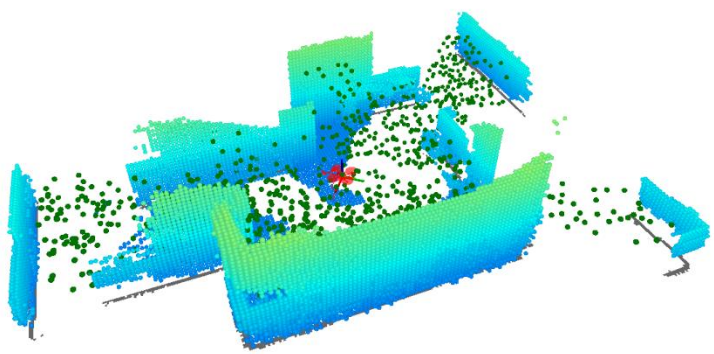
Stochastic Differential Equation-based Exploration



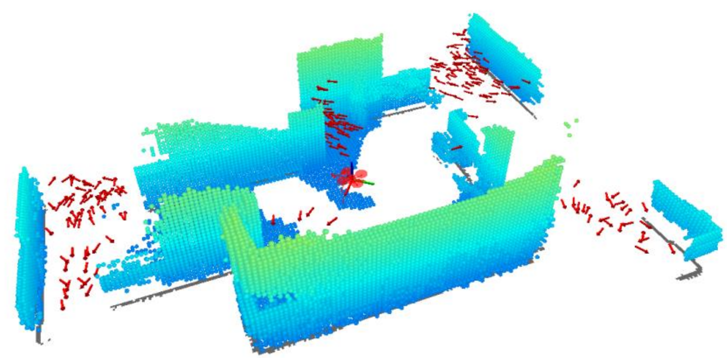
Dense free space



Frontier goals



Sparse free space



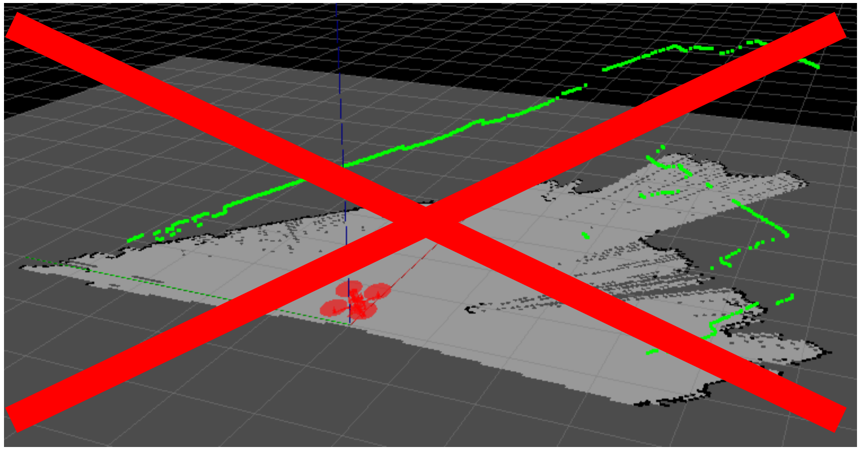
SDEE goals

Autonomous Indoor 3D Exploration with a Micro-Aerial Vehicle

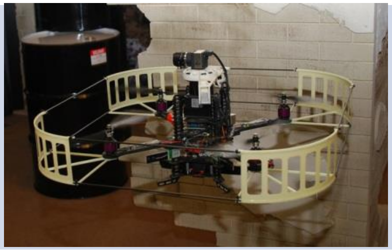
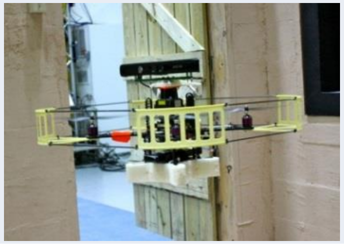
Shaojie Shen, Nathan Michael, Vijay Kumar

Penn
UNIVERSITY of PENNSYLVANIA

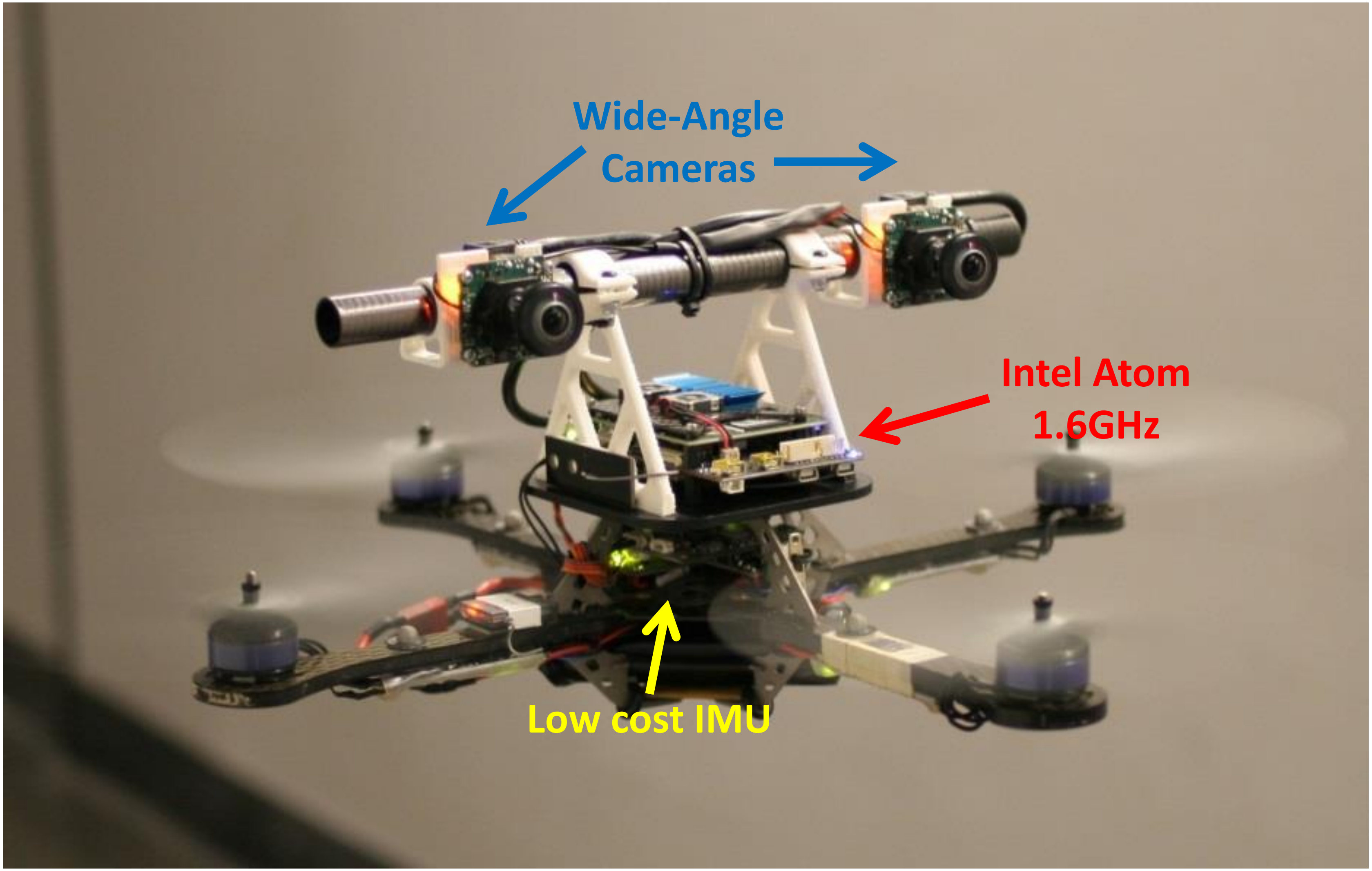
Relax the 2.5D assumption, and fly both indoor *and* outdoor



MAV with heterogeneous sensors

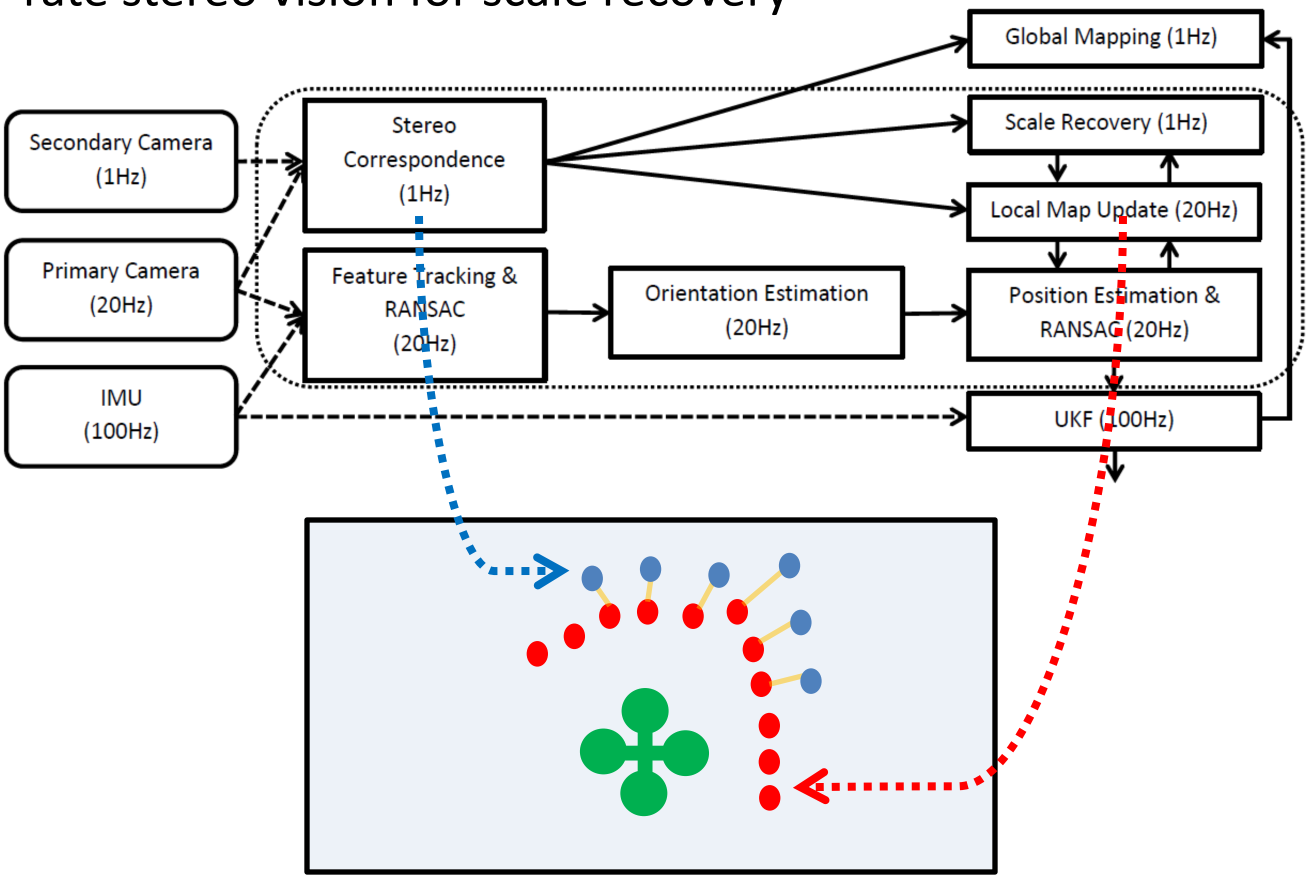
Robot	Sensing	Computation	Mass	Environment	Year
	Laser IMU	Intel Atom 1.6GHz	1.7 kg	2.5D indoor	2010-2011
	Laser Kinect IMU	Intel Atom 1.6GHz	1.9 kg	2.5D indoor	2011-2012
	Cameras IMU	Intel Atom 1.6GHz	0.74 kg	3D indoor and limited outdoor	2012-2013
	Laser Cameras GPS IMU	Intel Core i3 1.8GHz	1.9 kg	3D indoor and outdoor	2013-2014
	Camera IMU	Intel Core i3 1.8GHz	1.3 kg	3D indoor and limited outdoor	2014

Vision-Based Autonomous Flight

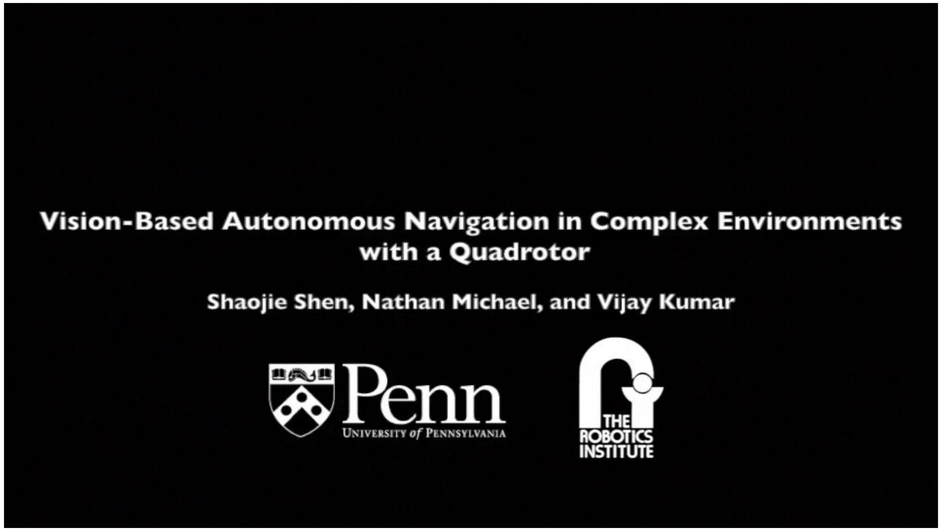


Vision-Based Autonomous Flight

- Decoupled system design for efficient pose estimation
  - Linear rotation, position, and map estimation
- Low-rate stereo vision for scale recovery



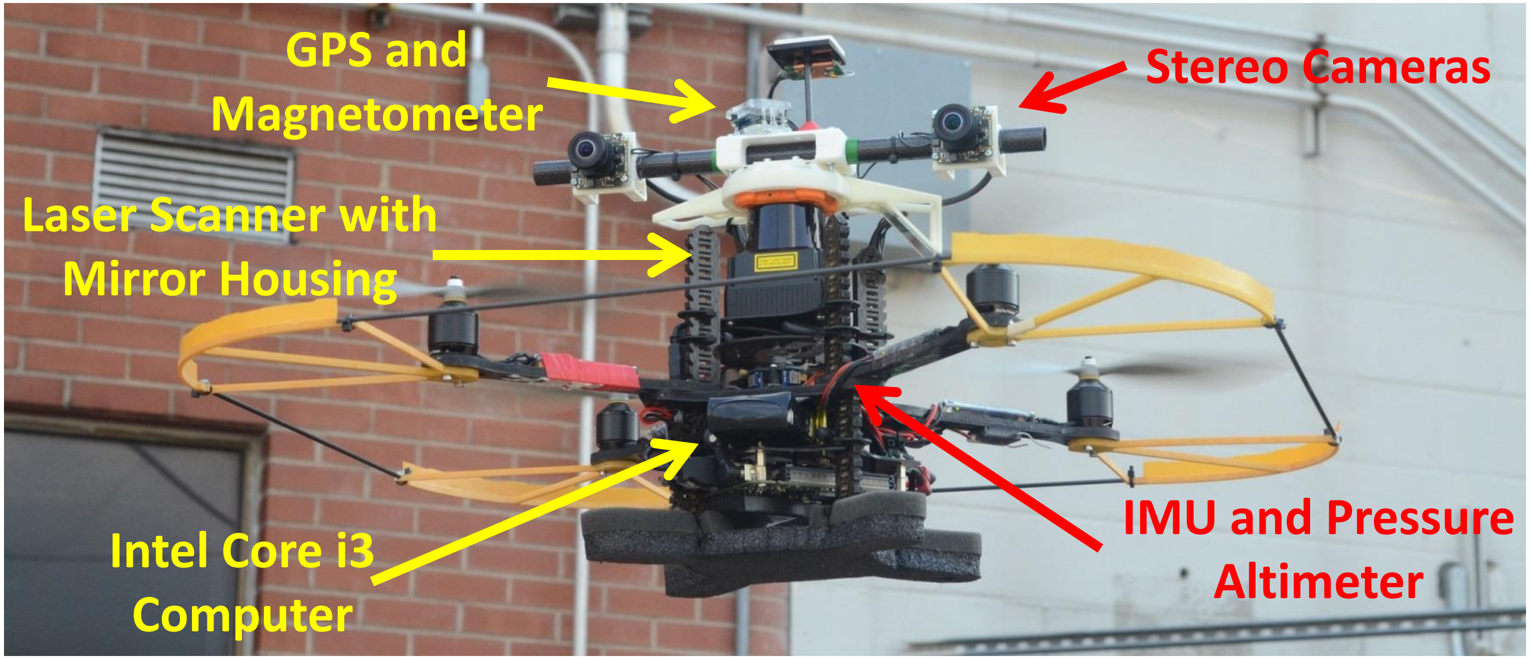
Indoor



Indoor and Outdoor SLAM

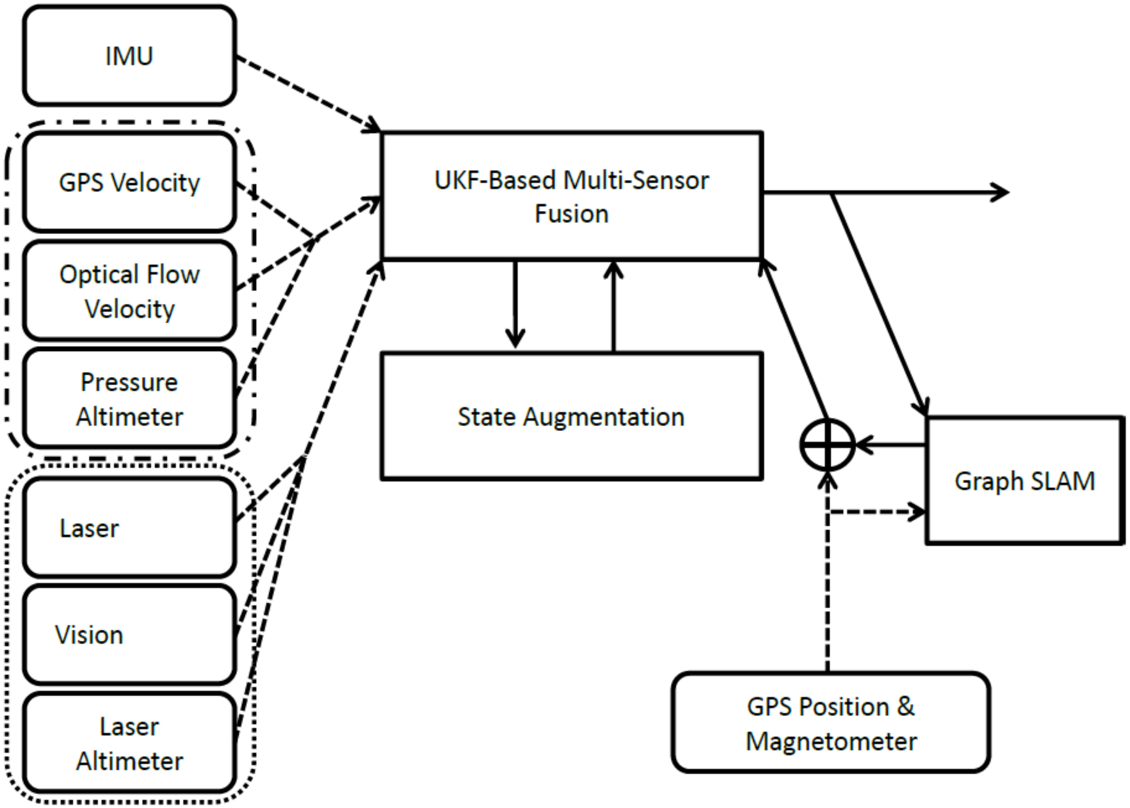
Multi-Sensor Fusion for MAVs

- Overview:
  - A modular Unscented Kalman Filter (UKF) for fusing heterogeneous sensors
  - Rapid sensor reconfiguration with minimum coding and calculation
  - Handling of GPS measurements to ensure smoothness



Multi-Sensor Fusion for MAVs

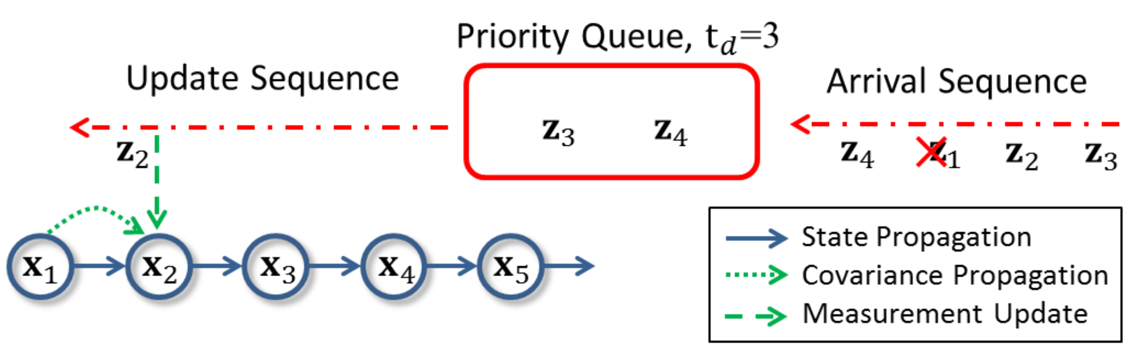
- Add/remove heterogeneous sensors with minimum coding and calculation (no computation of Jacobian required)



Multi-Sensor Fusion for MAVs

- Modular derivative-free Unscented Kalman filter
  - Main state: $\mathbf{x} = [\underbrace{\mathbf{p}^w}_{\text{Position}}, \underbrace{\mathbf{\Phi}^w}_{\text{Orientation}}, \underbrace{\dot{\mathbf{p}}^b}_{\text{Velocity}}, \underbrace{\mathbf{b}_a, \mathbf{b}_\omega, b_z}_{\text{Bias}}]$
  - IMU-based state propagation model: $\mathbf{x}_{t+1} = f(\mathbf{x}_t, \overbrace{\mathbf{a}_t}^{\text{Accelerometer}}, \overbrace{\mathbf{w}_t}^{\text{Gyroscope}}, \overbrace{\mathbf{n}_t}^{\text{Noise}})$

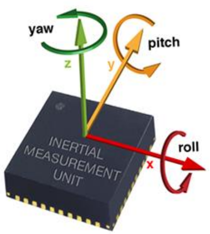
- Handles delayed and out-of-sequence measurements using fixed-lag priority queue:



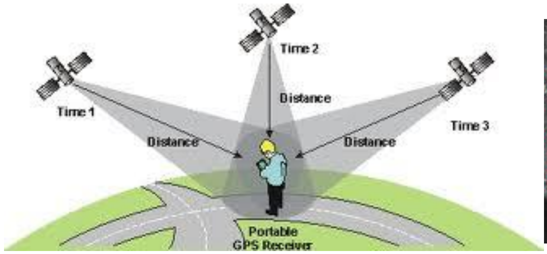
Penn

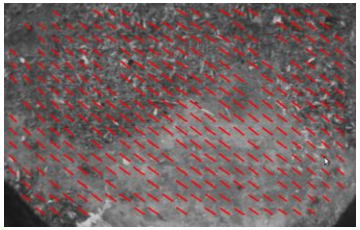
Heterogeneous Sensor Measurements

- Proprioceptive sensor
  - Low cost MEMS IMU

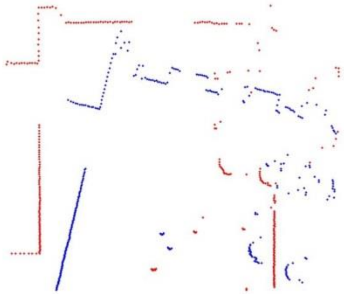


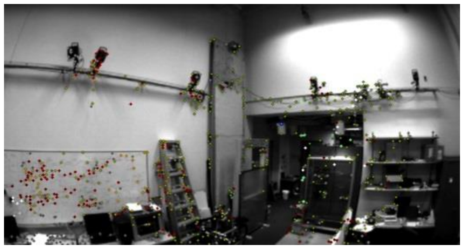
- Absolute measurements
  - GPS and magnetometer
  - Pressure altimeter
  - Optical flow velocity sensor
  - $\mathbf{z}_t = h(\mathbf{x}_t) + \mathbf{n}_t$





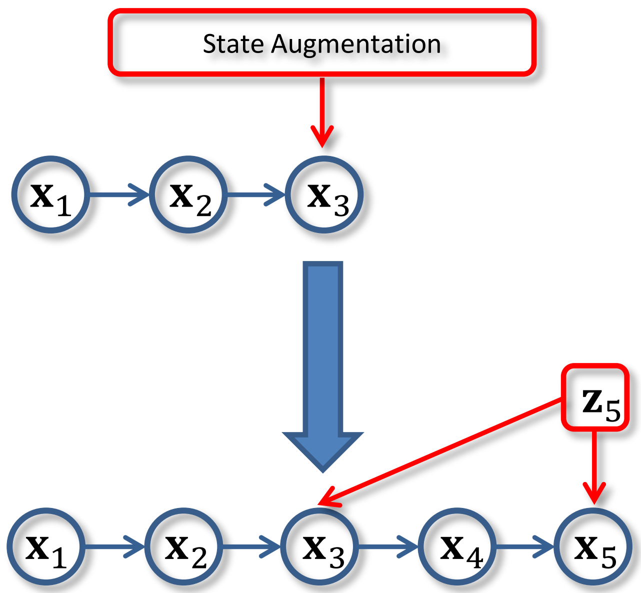
- Relative measurements
  - Laser scan matching – 3DOF Pose
  - Visual odometry – 6DOF Pose
  - Laser altimeter
  - $\mathbf{z}_t = h(\mathbf{x}_t, \mathbf{x}_{t-k}) + \mathbf{n}_t$





Heterogeneous Sensor Measurements

- Kalman filter requires all measurements to be related to the current state only
- State augmentation (Roumeliotis and Burdick, 2002)
  - May augment arbitrary copies of states depending on the availability relative sensors



$$[\mathbf{x}_3] \quad [\mathbf{P}_{33}] \longrightarrow \begin{bmatrix} \mathbf{x}_3 \\ \mathbf{x}_3 \end{bmatrix} \quad \begin{bmatrix} \mathbf{P}_{33} & \mathbf{P}_{33} \\ \mathbf{P}_{33} & \mathbf{P}_{33} \end{bmatrix}$$

Main

Augmented

$$\begin{bmatrix} \bar{\mathbf{x}}_5 \\ \mathbf{x}_3 \end{bmatrix} \quad \begin{bmatrix} \bar{\mathbf{P}}_{55} & \bar{\mathbf{P}}_{35} \\ \bar{\mathbf{P}}_{53} & \mathbf{P}_{33} \end{bmatrix} \longrightarrow \begin{bmatrix} \mathbf{x}_5 \\ \mathbf{x}_3 \end{bmatrix} \quad \begin{bmatrix} \mathbf{P}_{55} & \mathbf{P}_{35} \\ \mathbf{P}_{53} & \mathbf{P}_{33} \end{bmatrix}$$

Handling Ill-conditioned covariance matrix due to state augmentation

- Review of unscented transform:
  - Gaussian approximation of a nonlinear transform of a Gaussian random variable:

$$\mathrm{x} \sim \mathcal{N}(\hat{\mathrm{x}}, \mathrm{P}^{\mathrm{xx}})$$
$$\mathrm{y} = g(\mathrm{x})$$

$$\mathcal{X} = \left[\hat{\mathrm{x}},\ \hat{\mathrm{x}} \pm \boxed{\left(\sqrt{(n+\lambda)\mathrm{P}^{\mathrm{xx}}}\right)}\right] \quad \text{for } i = 1, \ldots, n$$
$$\mathcal{Y}_i = g(\mathcal{X}_i),$$

Cholesky decomposition

$$\hat{\mathrm{y}} = \sum_{i=0}^{2n} w_i^m \mathcal{Y}_i$$
$$\mathrm{P}^{\mathrm{yy}} = \sum_{i=0}^{2n} w_i^c (\mathcal{Y}_i - \hat{\mathrm{y}})(\mathcal{Y}_i - \hat{\mathrm{y}})^{\mathrm{T}}$$
$$\mathrm{P}^{\mathrm{yx}} = \sum_{i=0}^{2n} w_i^c (\mathcal{Y}_i - \hat{\mathrm{y}})(\mathcal{X}_i - \hat{\mathrm{x}})^{\mathrm{T}}$$

Gaussian approximation

- Ill-conditioned covariance matrix due to state augmentation
  - Not full rank; Positive semi-definite
  - Cholesky decomposition not unique $[\mathbf{P}_{33}] \rightarrow \begin{bmatrix} \mathbf{P}_{33} & \mathbf{P}_{33} \\ \mathbf{P}_{33} & \mathbf{P}_{33} \end{bmatrix} \rightarrow \sqrt{\begin{bmatrix} \mathbf{P}_{33} & \mathbf{P}_{33} \\ \mathbf{P}_{33} & \mathbf{P}_{33} \end{bmatrix}} = ?$

Handling Ill-conditioned covariance matrix due to state augmentation

- Unscented transform as statistical linearization (Lefebvre, et al, 2002):
  - Linear approximation: $\mathbf{y} = g(\mathbf{x}) \longrightarrow \mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b} + \mathbf{e}$
  - Minimize linearization error: $\min_{\mathbf{A},\mathbf{b}} \sum_{i=0}^{2n} w_i(\mathcal{Y}_i - \mathbf{A}\mathcal{X}_i - \mathbf{b})(\mathcal{Y}_i - \mathbf{A}\mathcal{X}_i - \mathbf{b})^{\mathrm{T}}$
  - Optimal linearization: $\mathbf{A} = \mathbf{P}^{\mathbf{yx}}{\mathbf{P}^{\mathbf{xx}}}^{-1}, \quad \mathbf{b} = \hat{\mathbf{y}} - \mathbf{A}\hat{\mathbf{x}}$

Handling Ill-conditioned covariance matrix due to state augmentation

- Statistical linearization and state propagation
  - Cholesky decomposition of only main states
  - Unique decomposition

$$\check{\mathbf{x}}_{t|t} = \begin{bmatrix} \hat{\mathbf{x}}_{t|t} \\ \hat{\mathbf{x}}_{\mathcal{I}_{t|t}} \end{bmatrix}, \quad \check{\mathbf{P}}_{t|t} = \begin{bmatrix} \mathbf{P}^{\mathbf{xx}}_{t|t} & \mathbf{P}^{\mathbf{x}\mathbf{x}_{\mathcal{I}}}_{t|t} \\ \mathbf{P}^{\mathbf{x}_{\mathcal{I}}\mathbf{x}}_{t|t} & \mathbf{P}^{\mathbf{x}_{\mathcal{I}}\mathbf{x}_{\mathcal{I}}}_{t|t} \end{bmatrix}$$

$$\check{\mathbf{x}}_{t+1|t} = f(\check{\mathbf{x}}_{t|t}, \mathbf{u}_t, \mathbf{v}_t)$$

w.r.t. Main States

w.r.t. Augmented States

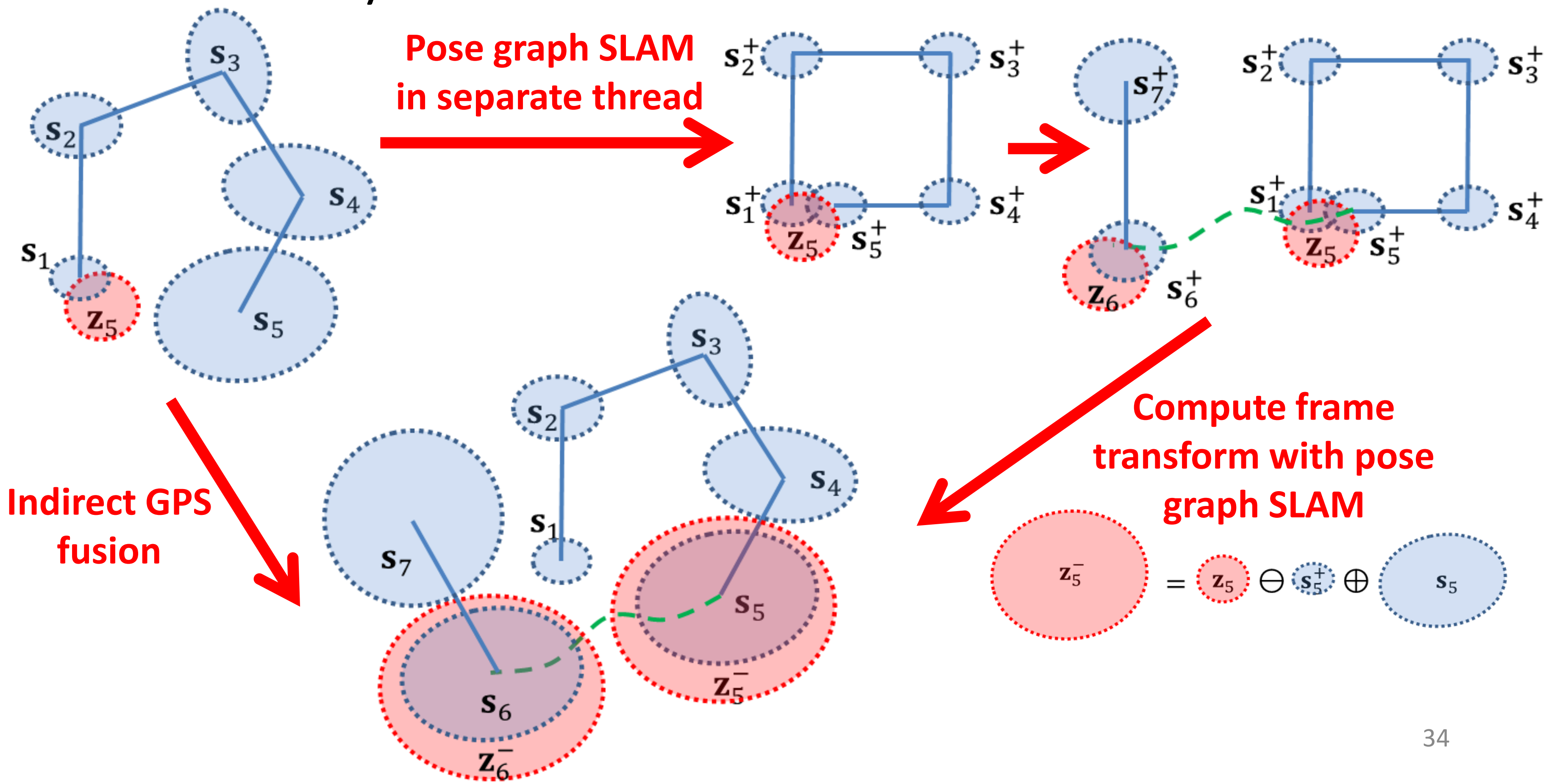
$$= \begin{bmatrix} \mathbf{F}_t & 0 \\ 0 & \mathbf{I}_{|\mathcal{I}|} \end{bmatrix} \check{\mathbf{x}}_{t|t} + \begin{bmatrix} \mathbf{J}_t & \mathbf{G}_t \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u}_t \\ \mathbf{v}_t \end{bmatrix} + \mathbf{b}_t + \mathbf{e}_t,$$

$$\check{\mathbf{x}}_{t+1|t} = \begin{bmatrix} \hat{\mathbf{x}}_{t+1|t} \\ \hat{\mathbf{x}}_{\mathcal{I}_{t|t}} \end{bmatrix}, \quad \check{\mathbf{P}}_{t+1|t} = \begin{bmatrix} \mathbf{P}^{\mathbf{xx}}_{t+1|t} & \mathbf{F}_t \mathbf{P}^{\mathbf{x}\mathbf{x}_{\mathcal{I}}}_{t|t} \\ \mathbf{P}^{\mathbf{x}_{\mathcal{I}}\mathbf{x}}_{t|t} \mathbf{F}_t^T & \mathbf{P}^{\mathbf{x}_{\mathcal{I}}\mathbf{x}_{\mathcal{I}}}_{t|t} \end{bmatrix}$$

- Measurement update the same as standard EKF

Smooth State Estimates

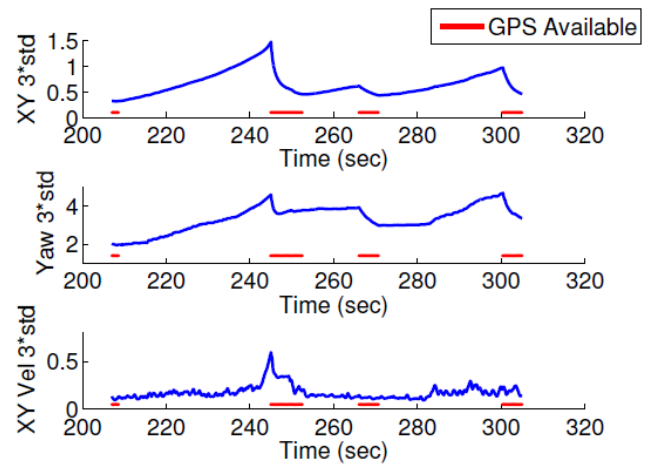
- Sudden GPS availability causes discontinuities
- Alternate way of GPS fusion via SLAM and frame transform



Autonomous Navigation in Industrial Complex

Results

Trajectory with reference map

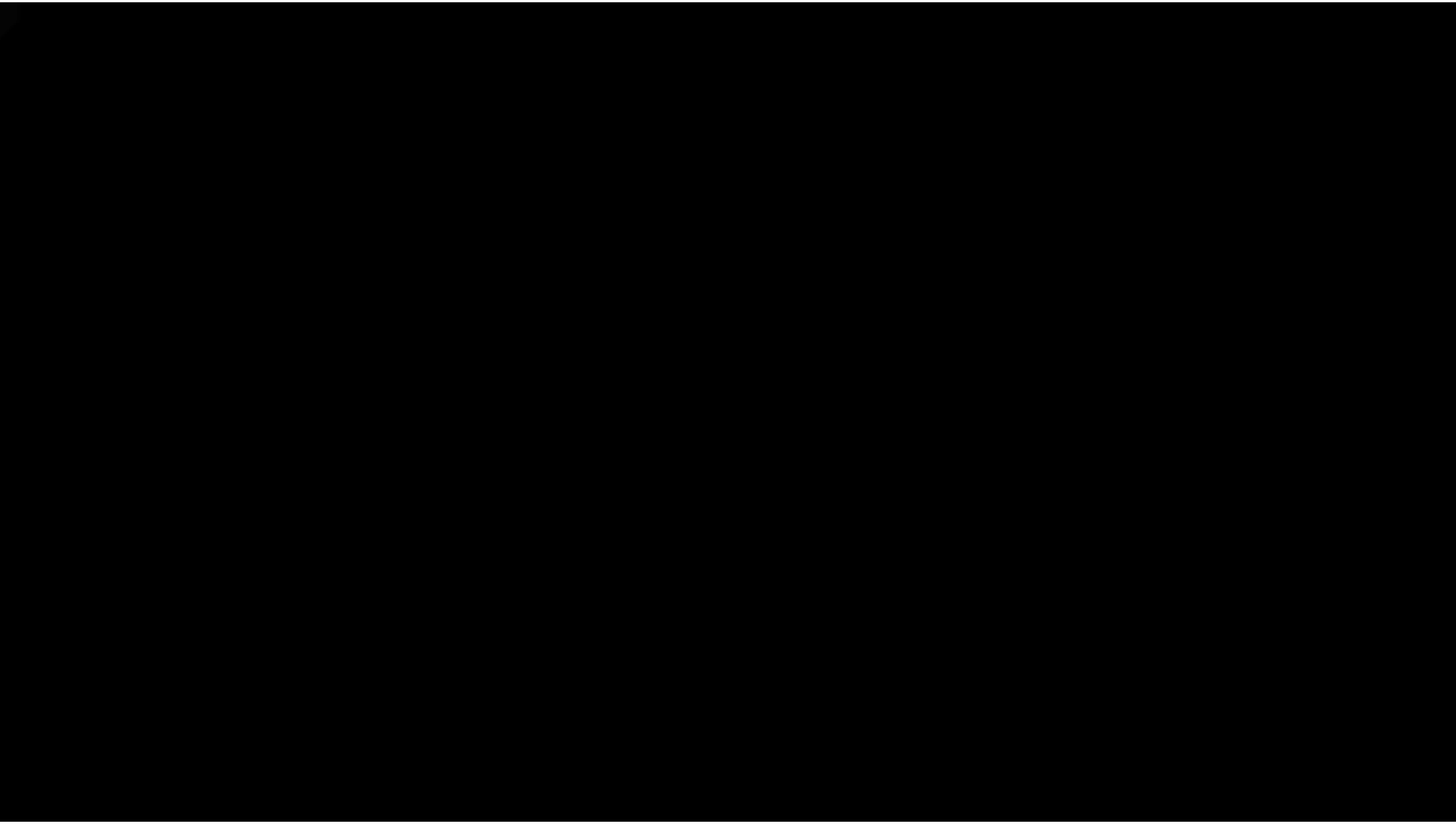


Covariance

Results – Dam Inspection

- Emergency flooding gates at the Carter’s Dam, GA, USA

Carter’s Dam Autonomous Flight & Mapping



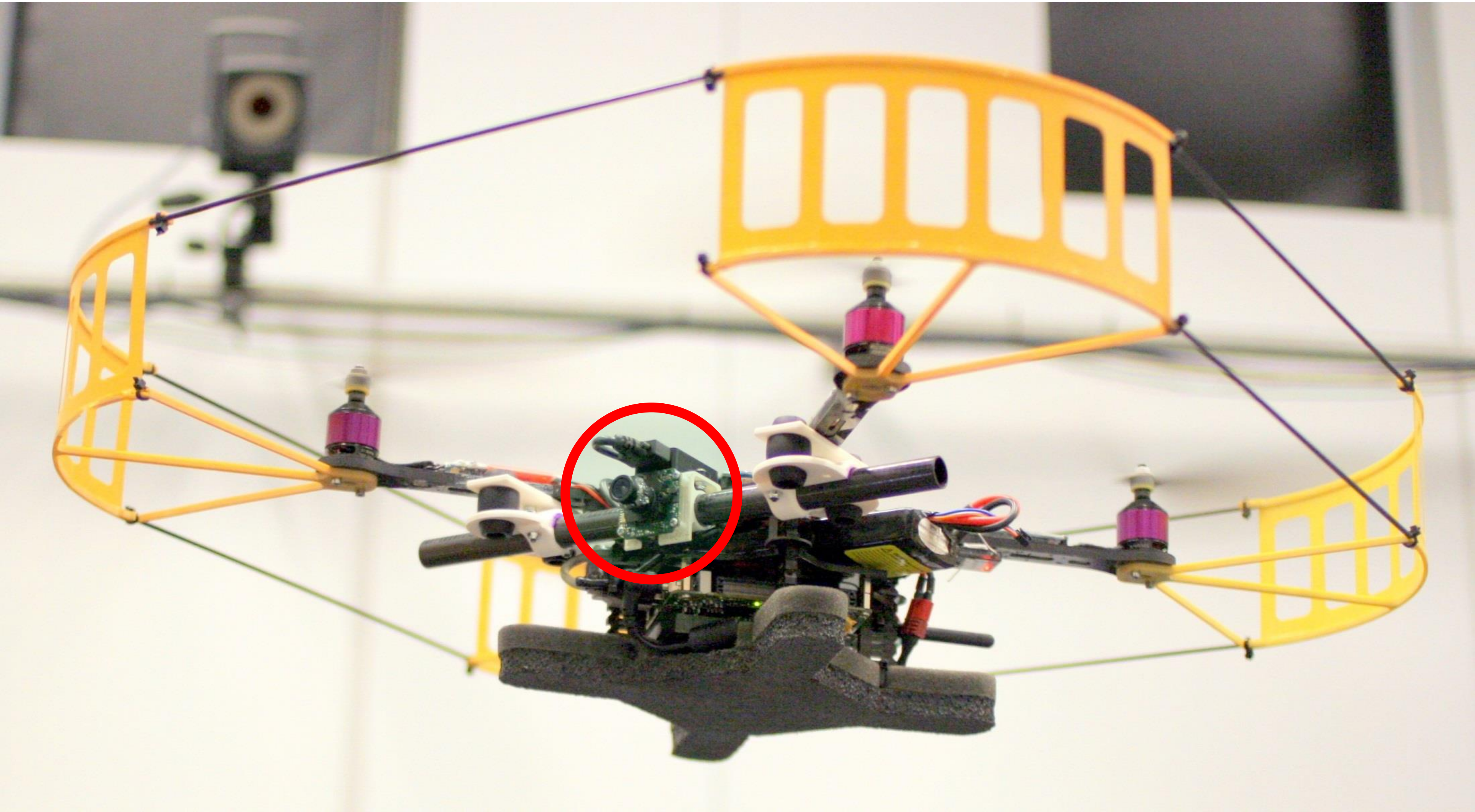
What's Left? – Initialization and Failure Recovery

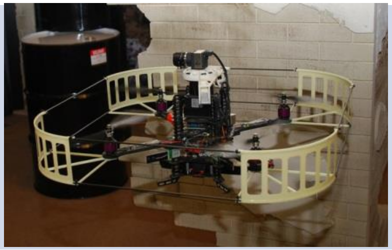
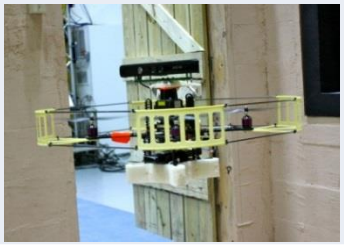
- Power-on-and-go ?
  - Initialize from an arbitrary unknown state
- Autonomy
  - State estimation in a wide range of environments
- Fault-tolerant
  - Handle failure of one or more onboard sensors
- Fail-safe ?
  - Recover from total failure of all sensors

Assumptions for Initialization of Nonlinear Estimators

- Full state observability – Initialize with the first measurement
  - GPS-based navigation
- Known initial conditions
  - Takeoff from stationary condition
- Approximations
  - Velocity from numerical differentiation of poses
  - Attitude from accelerometer
  - Visual scale from barometer/prior depth knowledge
- What if all assumptions are invalid?

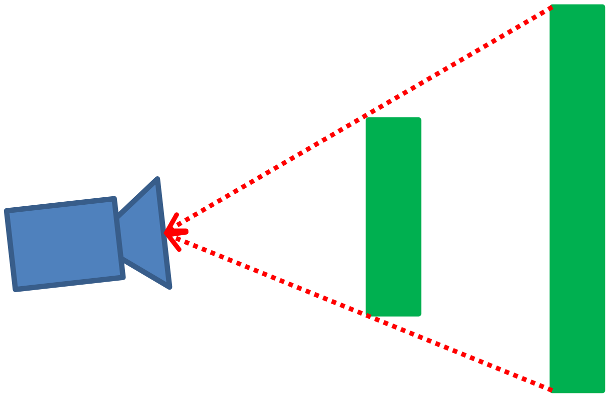
Monocular Visual-Inertial Systems



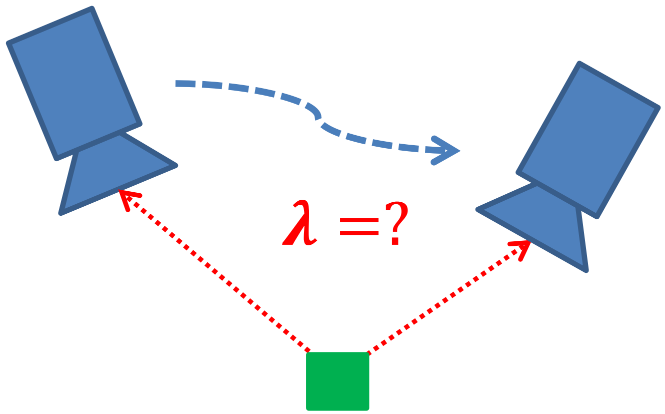
Robot	Sensing	Computation	Mass	Environment	Year
	Laser IMU	Intel Atom 1.6GHz	1.7 kg	2.5D indoor	2010-2011
	Laser Kinect IMU	Intel Atom 1.6GHz	1.9 kg	2.5D indoor	2011-2012
	Cameras IMU	Intel Atom 1.6GHz	0.74 kg	3D indoor and limited outdoor	2012-2013
	Laser Cameras GPS IMU	Intel Core i3 1.8GHz	1.9 kg	3D indoor and outdoor	2013-2014
	Camera IMU	Intel Core i3 1.8GHz	1.3 kg	3D indoor and limited outdoor	2014

Challenges

- Scale ambiguity

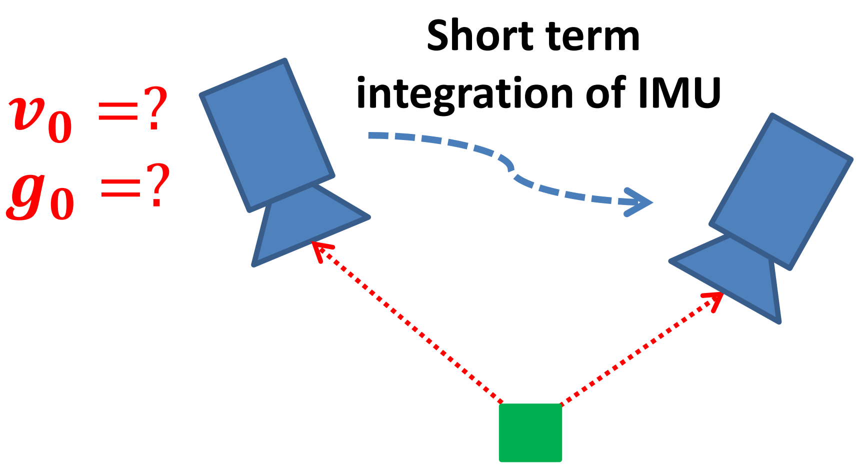


- Up-to-scale motion estimation and 3D reconstruction (Structure from Motion)



Challenges

- With IMU, scale is observable
- But…
  - Requires initial velocity and attitude (gravity)
  - Highly nonlinear system – requires initial values to converge



Can we operate without initialization?

Or…

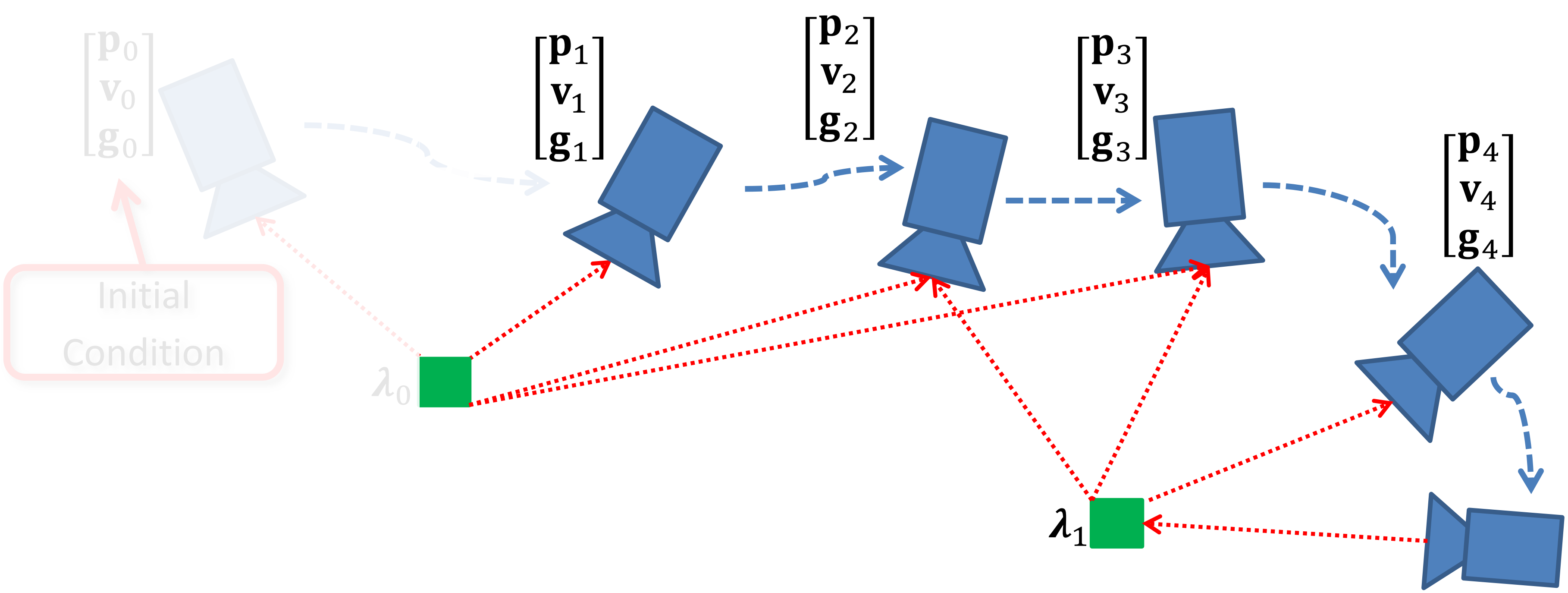
Can we launch by throwing it into the air?

Related Work

- State-of-the-art solutions to monocular VINS
  - Filtering-based (Jones, et al. 2011; Weiss, et al. 2012; Li, et al. 2013)
  - Nonlinear optimization-based (Indelman, et al. 2013; Leutenegger, et al. 2013)
  - Requires good initial conditions, unable to recover from failure
- Closed-form solutions (Lippiello, et al. 2013; Martinelli, 2013)
  - Suboptimal solution, poor performance with noisy IMU measurements
- Initialization-free state estimation (Lupton, et al, 2012)
  - Reduces nonlinearity via frame transform
- Handling degenerate motion in monocular VINS (Kottas, et al. 2013)
  - Explicit handling of hover motion, may lead to pessimistic covariance

Linear Sliding Window Monocular Visual-Inertial Estimator

- Estimates position, velocity, gravity, and feature depth
- Linear formulation enables recovery of initial condition
- Marginalizes selected poses to bound computation cost

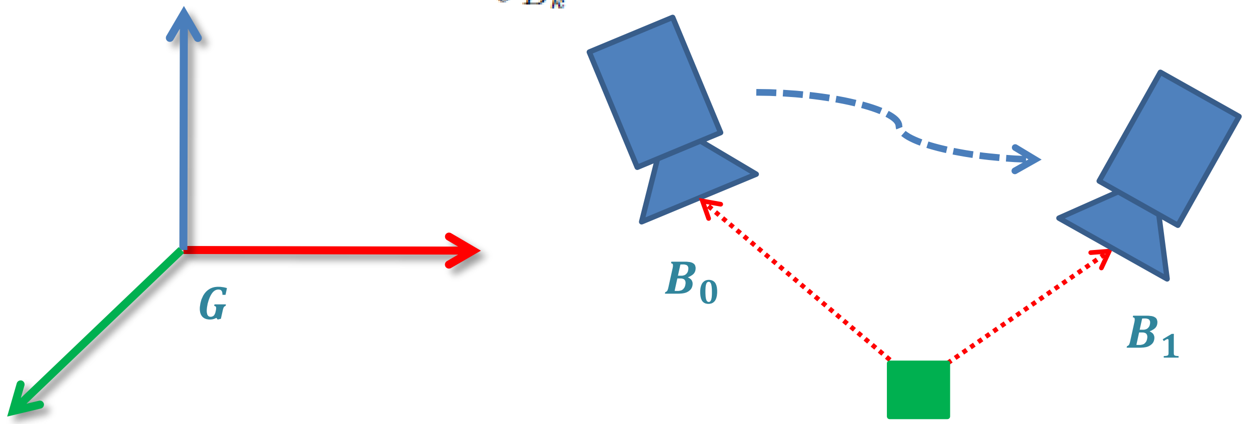


IMU Model

- IMU integration in global frame
  - IMU has higher rate than camera
  - Nonlinearity from global rotation
  - Requires global rotation at the time of integration

$$\mathbf{p}^G_{B_{k+1}} = \mathbf{p}^G_{B_k} + \mathbf{v}^G_{B_k}\Delta t + \iint_{B_k}^{B_{k+1}} \left(\mathbf{R}^G_B \mathbf{a}^B - \mathbf{g}^G\right) dt^2$$

$$\mathbf{v}^G_{B_{k+1}} = \mathbf{v}^G_{B_k} + \int_{B_k}^{B_{k+1}} \left(\mathbf{R}^G_B \mathbf{a}^B - \mathbf{g}^G\right) dt$$



IMU Model

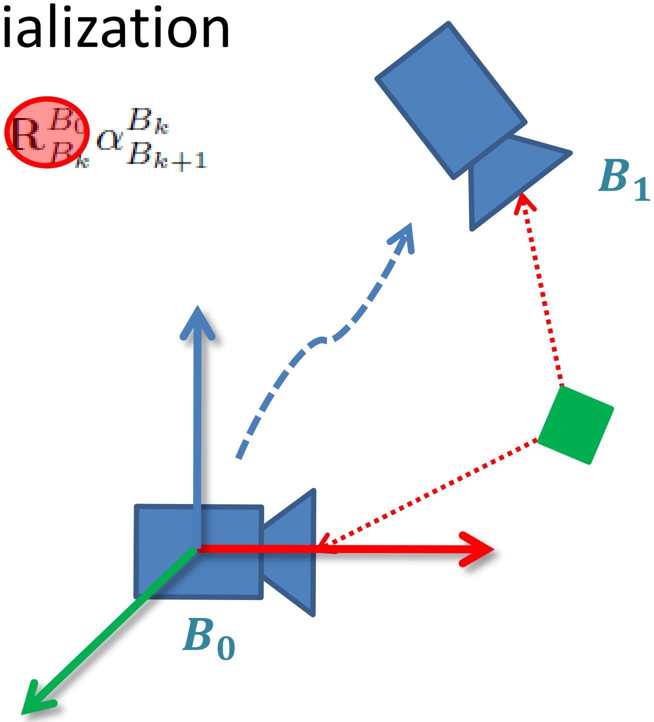
- IMU integration in the body frame of the first pose
  - Nonlinearity from relative rotation only
  - Linear update equations for position, velocity, and gravity
  - IMU Integration without initialization

$$\mathbf{p}^{B_0}_{B_{k+1}} = \mathbf{p}^{B_0}_{B_k} + \mathbf{v}^{B_0}_{B_k}\Delta t - \mathbf{g}^{B_0}\Delta t^2/2 + \mathbf{R}^{B_0}_{B_k}\alpha^{B_k}_{B_{k+1}}$$

$$\mathbf{v}^{B_0}_{B_{k+1}} = \mathbf{v}^{B_0}_{B_k} - \mathbf{g}^{B_0}\Delta t + \mathbf{R}^{B_0}_{B_k}\beta^{B_k}_{B_{k+1}}$$

$$\alpha^{B_k}_{B_{k+1}} = \iint_{B_k}^{B_{k+1}} \mathbf{R}^{B_k}_{B}\mathbf{a}^{B} dt^2$$

$$\beta^{B_k}_{B_{k+1}} = \int_{B_k}^{B_{k+1}} \mathbf{R}^{B_k}_{B}\mathbf{a}^{B} dt$$

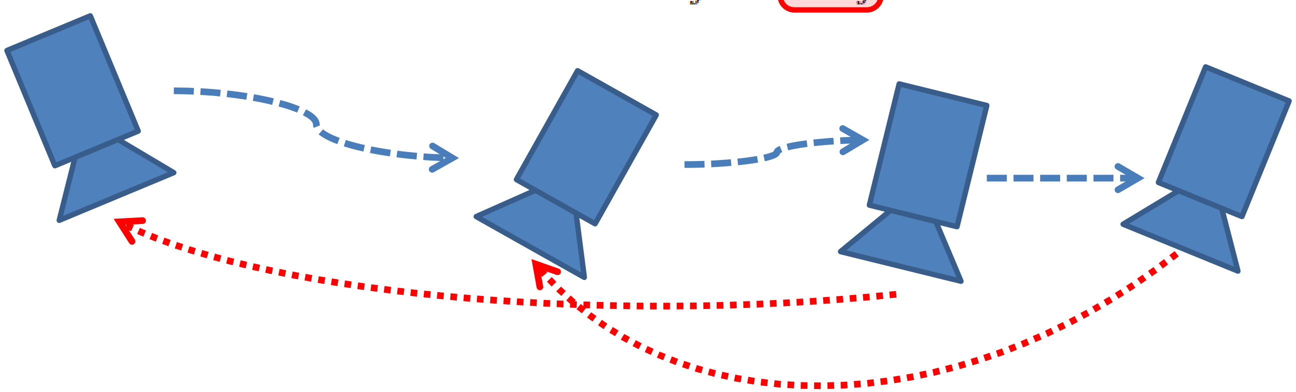


Linear Rotation Estimation

- Relative rotation constraints
  - Short term gyroscope integration
  - Epipolar constraints

Relative rotation

$$\mathbf{R}_{B_0}^{B_0} = \mathbf{I}_3, \quad \mathbf{R}_{B_j}^{B_0} = \boxed{\hat{\mathbf{R}}_{B_j}^{B_i}} \mathbf{R}_{B_i}^{B_0}$$



- Rotation estimation by relaxing orthonormality constraints

$$\left[\mathbf{I}_3, \; -\hat{\mathbf{R}}_{B_j}^{B_i}\right] \boxed{\begin{bmatrix}\mathbf{r}_i^k \\ \mathbf{r}_j^k\end{bmatrix}} = 0 \quad k = 1, 2, 3$$

Columns of rotation matrices

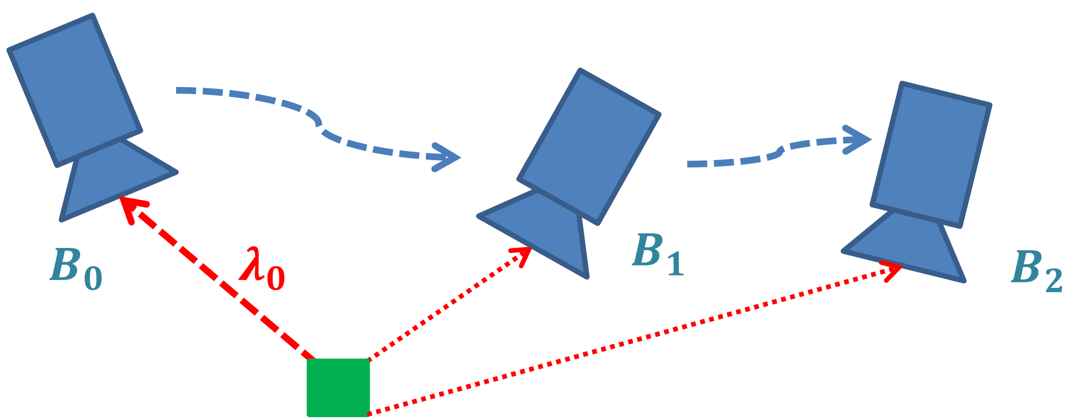
Camera Model

- Linear in position and feature depth
- Nonlinear in feature observation

$$\mathbf{0} = \begin{bmatrix} -1 & 0 & u_l^{B_j} \\ 0 & -1 & v_l^{B_j} \end{bmatrix} \mathbf{R}_{B_0}^{B_j} \left(\mathbf{p}_{B_i}^{B_0} - \mathbf{p}_{B_j}^{B_0} + \lambda_l \mathbf{R}_{B_i}^{B_0} \begin{bmatrix} u_l^{B_i} \\ v_l^{B_i} \\ 1 \end{bmatrix} \right) = \mathbf{H}_l^{B_j} \mathcal{X}$$

  - Unknown scaling factor in observation covariance

$$\mathbf{z}_l^{B_j} \sim \mathcal{N}\left(\mathbf{H}_l^{B_j} \mathcal{X}, {\lambda_l^{B_j}}^2 \bar{\mathbf{P}}_l^{B_j} \right)$$



Linear Sliding Window Monocular Visual-Inertial Estimator

- Linear system
  - Prior is not needed
  - Initial condition recoverable
  - Recoverable from failure

$$\min_{\mathcal{X}} \left\{ \underbrace{\cancel{\left(\mathbf{b}_p - \boldsymbol{\Lambda}_p \mathcal{X}\right)}}_{\text{Prior}} + \underbrace{\sum_{k\in\mathcal{D}} \left\| \hat{\mathbf{z}}^{B_k}_{B_{k+1}} - \mathbf{H}^{B_k}_{B_{k+1}}\mathcal{X}\right\|^2_{\mathbf{P}^{B_k}_{B_{k+1}}}}_{\text{IMU Constraints}} + \underbrace{\sum_{(l,j)\in\mathcal{C}} \left\|\hat{\mathbf{z}}^{B_j}_l - \mathbf{H}^{B_j}_l \mathcal{X}\right\|^2_{\mathbf{P}^{B_j}_l}}_{\text{Camera Constraints}} \right\}$$

$$\left(\cancel{\boldsymbol{\Lambda}_p} + \boldsymbol{\Lambda}_{imu} + \boldsymbol{\Lambda}_{cam}\right)\mathcal{X} = \left(\cancel{\mathbf{b}_p} + \mathbf{b}_{imu} + \mathbf{b}_{cam}\right)$$

Tightly-Coupled Nonlinear Sliding Window Optimization

- Nonlinear optimization based on linear initialization
  - Optimize position, velocity, rotation, and feature depth simultaneously:
  - Rotation error modeled on the tangent space of rotation manifold

$$\mathcal{X} = \left[\mathbf{x}_0^0,\ \mathbf{x}_1^0,\ \cdots\ \mathbf{x}_N^0,\ \lambda_0,\ \lambda_1,\ \cdots\ \lambda_M\right]$$

$$\mathbf{x}_k^0 = \left[\mathbf{p}_k^0,\ \mathbf{v}_k^k,\ \mathbf{q}_k^0\right] \quad \text{for } k = 1, \ldots, N$$

$$\mathbf{q} \approx \hat{\mathbf{q}} \otimes \begin{bmatrix} \frac{1}{2}\delta\boldsymbol{\theta} \\ 1 \end{bmatrix}$$

  - Iteratively minimize residuals from all sensors

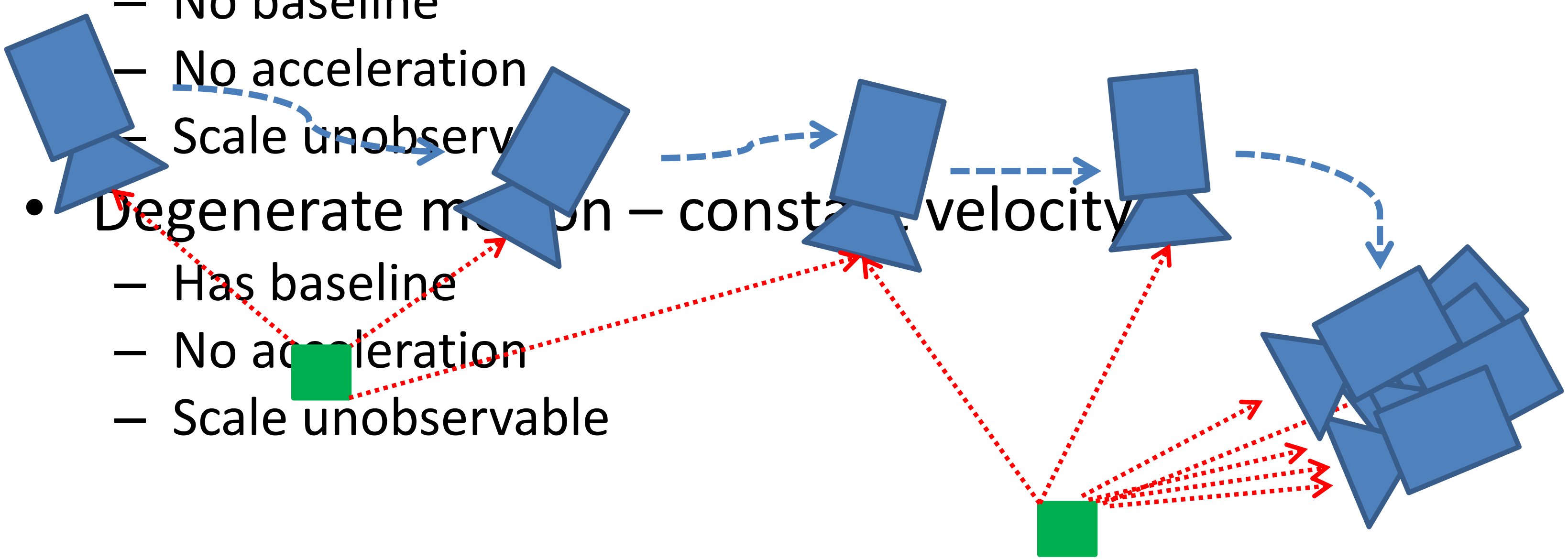
$$\min_{\mathcal{X}} \left\{ \boxed{(\mathbf{b}_p - \mathbf{\Lambda}_p \mathcal{X})} + \boxed{\sum_{k \in \mathcal{D}} \left\| r_{\mathcal{D}}(\hat{\mathbf{z}}_{k+1}^k,\ \mathcal{X}) \right\|_{\mathbf{P}_{k+1}^k}^2} + \boxed{\sum_{(l,j) \in \mathcal{C}} \left\| r_{\mathcal{C}}(\hat{\mathbf{z}}_l^j,\ \mathcal{X}) \right\|_{\mathbf{P}_l^j}^2} \right\}$$

Linearization

$$\min_{\delta\mathcal{X}} \left\{ \underbrace{\boxed{\left(\mathbf{b}_p - \mathbf{\Lambda}_p \hat{\mathcal{X}}\right)}}_{\text{Prior}} + \underbrace{\boxed{\sum_{k \in \mathcal{D}} \left\| r_{\mathcal{D}}(\hat{\mathbf{z}}_{k+1}^k,\ \hat{\mathcal{X}}) + \mathbf{H}_{k+1}^k \delta\mathcal{X} \right\|_{\mathbf{P}_{k+1}^k}^2}}_{\text{IMU residual}} + \underbrace{\boxed{\sum_{(l,j) \in \mathcal{C}} \left\| r_{\mathcal{C}}(\hat{\mathbf{z}}_l^j,\ \hat{\mathcal{X}}) + \mathbf{H}_l^j \delta\mathcal{X} \right\|_{\mathbf{P}_l^j}^2}}_{\text{Reprojection error}} \right\}$$

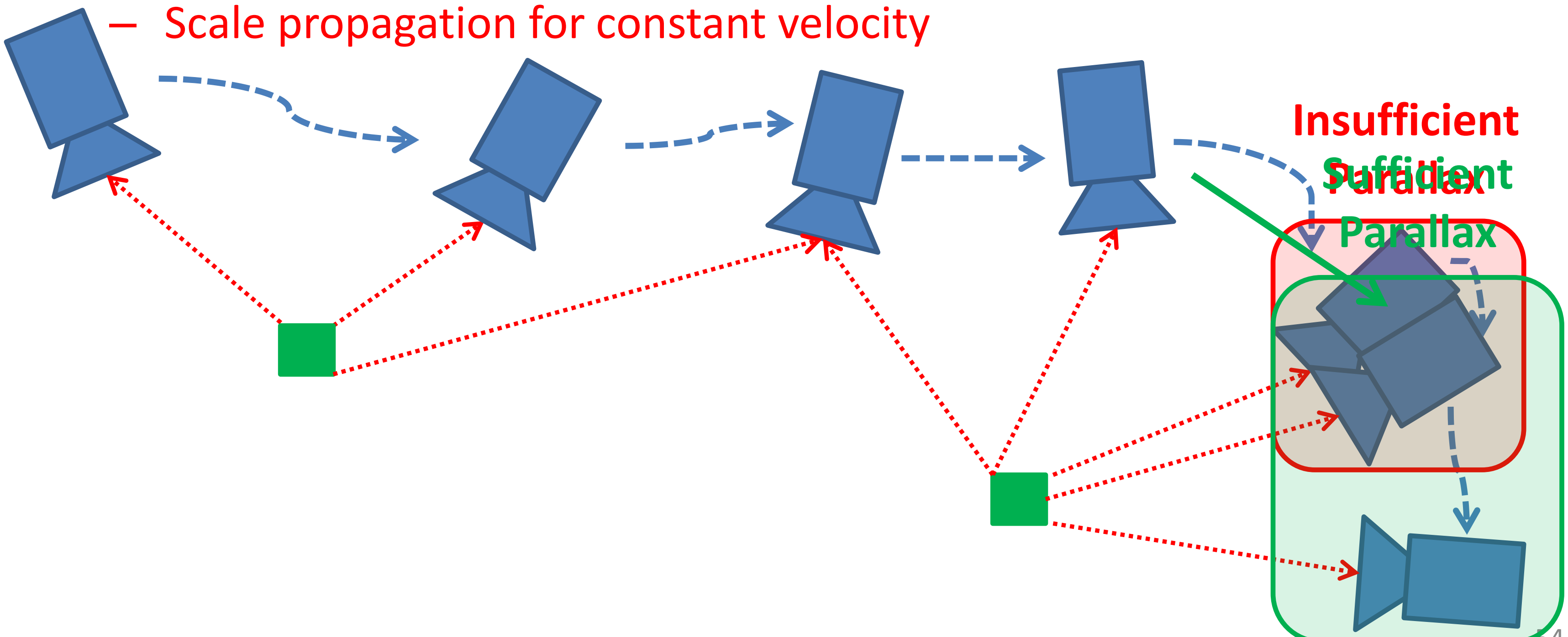
Marginalization

- Remove old poses to bound computation complexity
  - Convert removed measurements into prior
- General motion:
  - Linear acceleration is required for scale observability
- Degenerate motion – hover
  - No baseline
  - No acceleration
  - Scale unobservable
- Degenerate motion – constant velocity
  - Has baseline
  - No acceleration
  - Scale unobservable

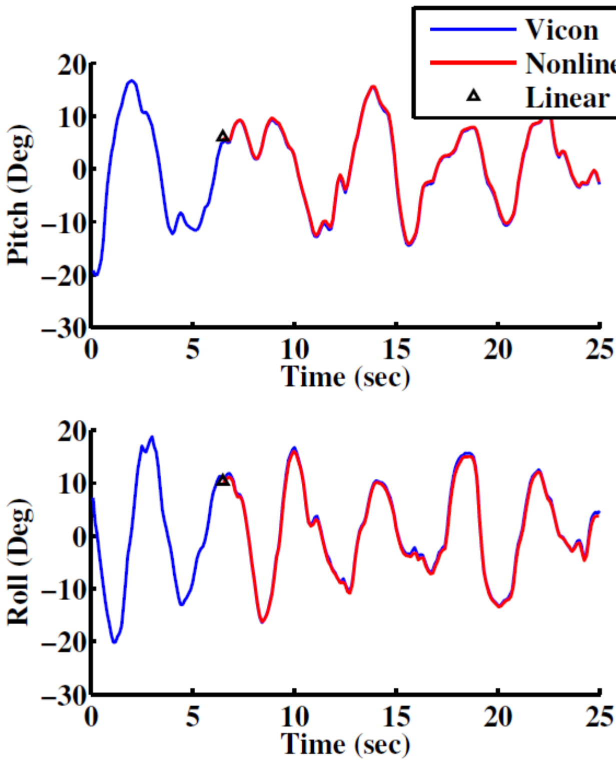


Marginalization

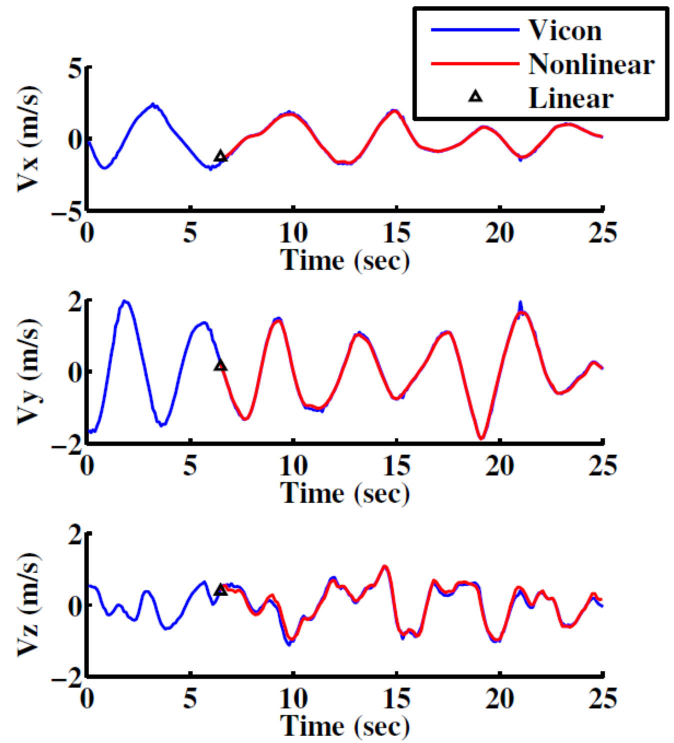
- Two-way marginalization
  - Preserve acceleration and baseline within the sliding window
  - Marginalize either recent or old pose based on parallax heuristic
  - Scale observable for hover
  - Scale propagation for constant velocity



Results – On-the-Fly Initialization



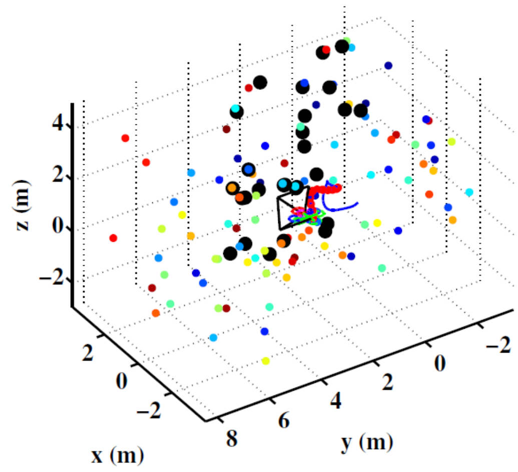
Attitude



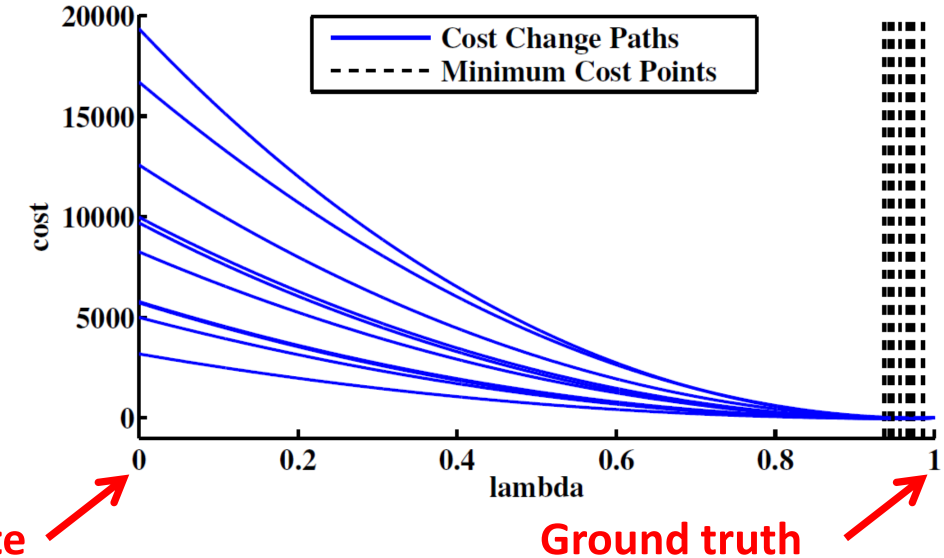
Velocity

Results – Simulation

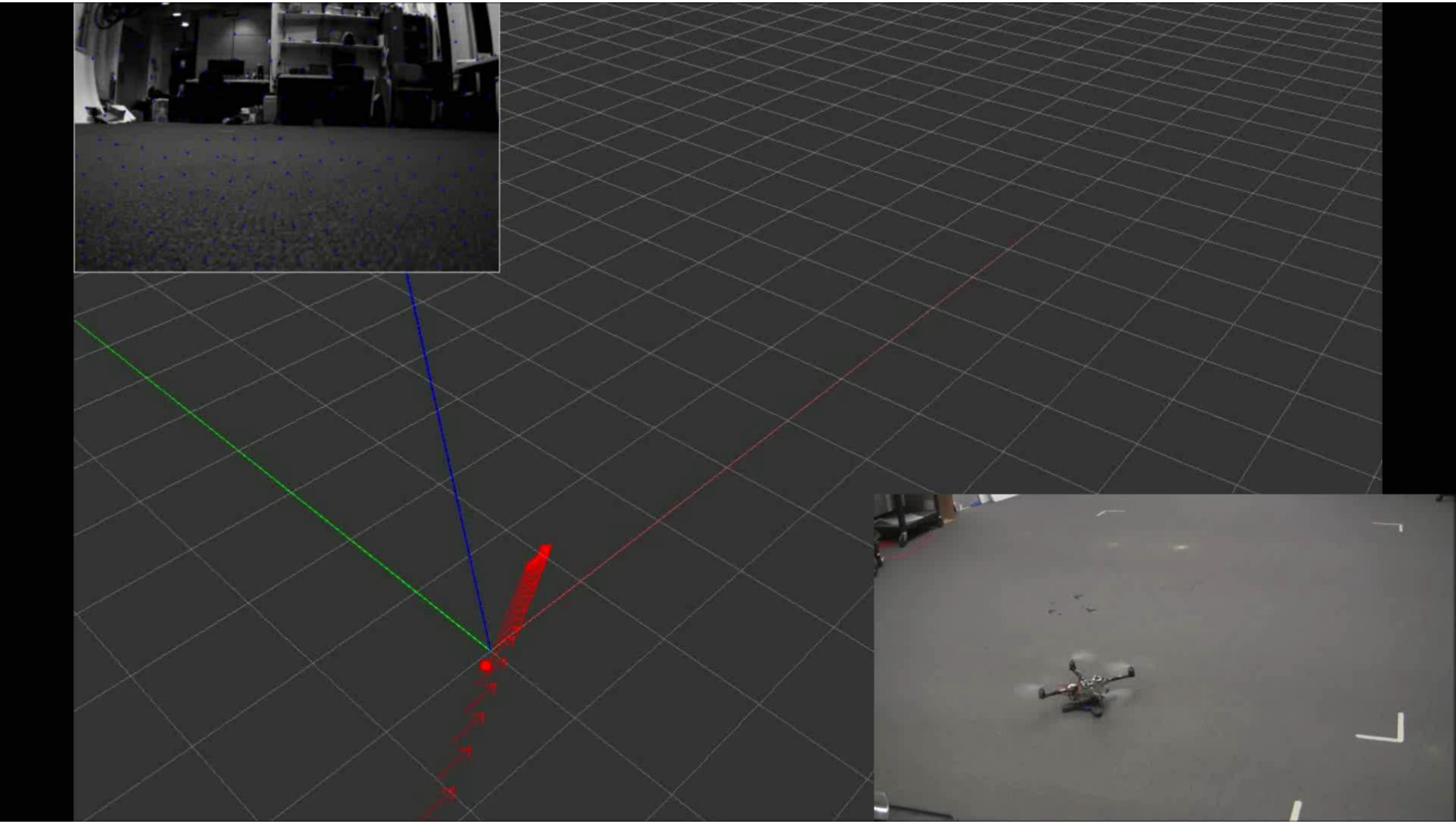
- Simulation environment with sensor noise and MAV dynamics



- The linear estimator well approximates the ground truth



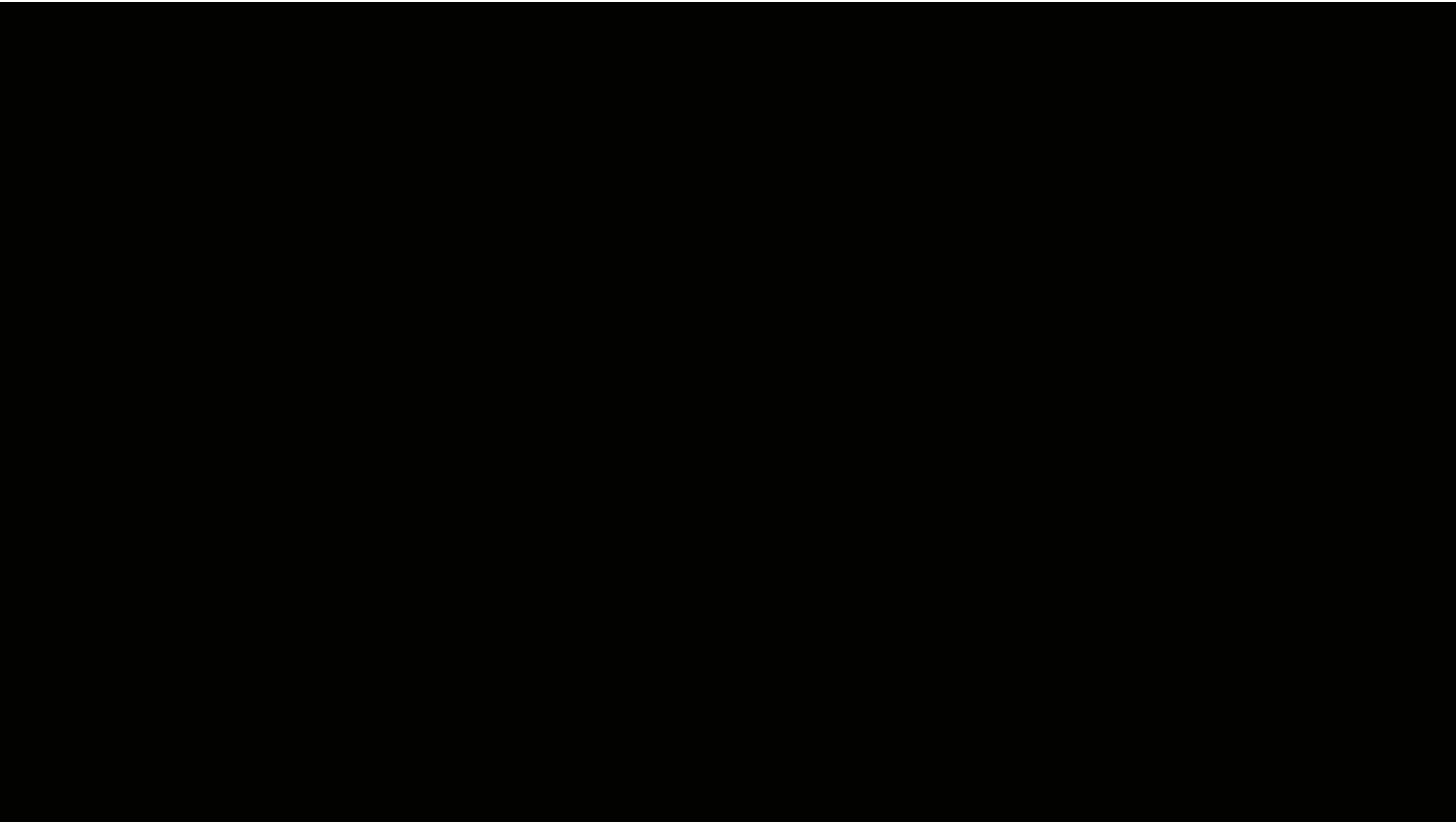
Fast Trajectory Tracking



Max Velocity: 2m/s, Max Attitude: 30 degrees

Drift: [Position: < 0.1m], [Yaw: < 1 degree]

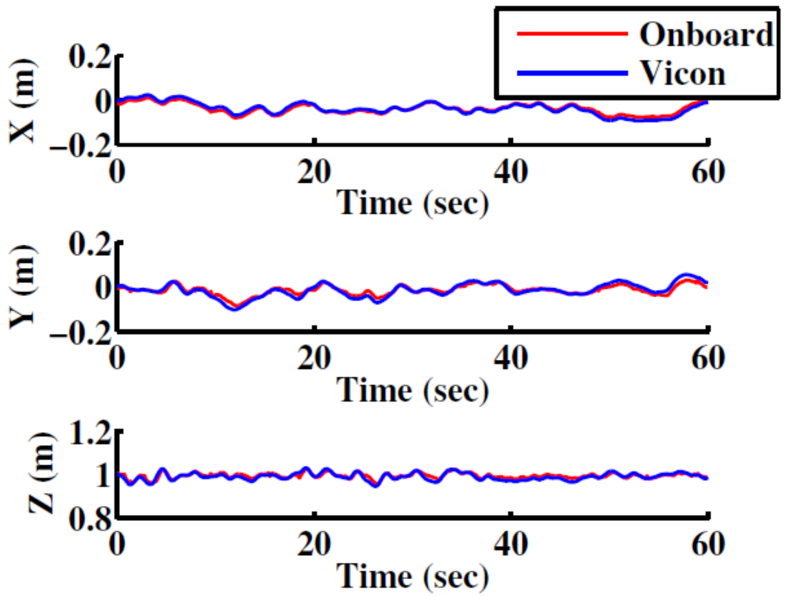
Indoor Hand Carrying



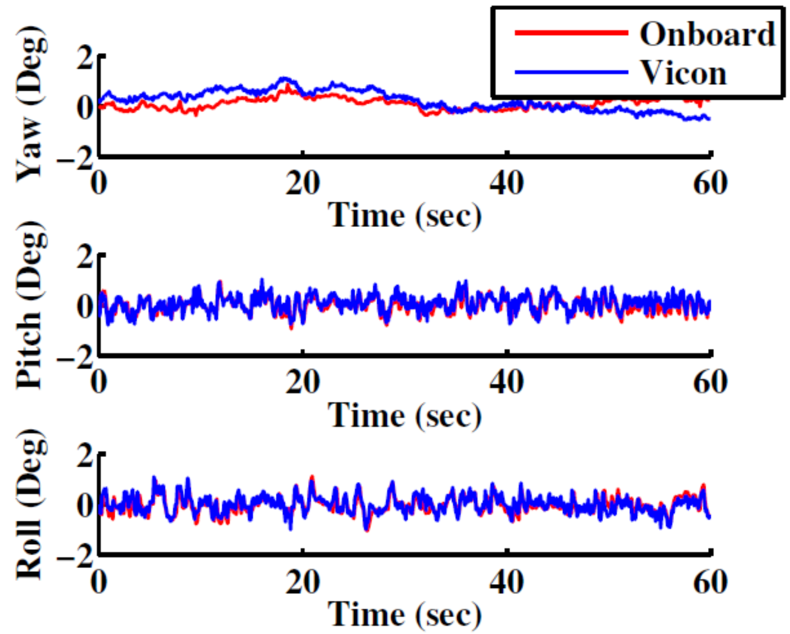
Drift: [XY: 0.15m], [Z: 0.04m] [Yaw: 0.6 degrees]

Results - Hover

- No drift in 6-DOF pose
- Position estimation StdDev: [0.0099, 0.0124, 0.0161] meters
- Hover StdDev: [0.0282, 0.0307, 0.0161] meters



Position



Orientation

Conclusion

- Autonomous navigation in complex indoor and outdoor environments with micro aerial vehicles using a variety of sensors
- State estimation for autonomous MAVs
  - On-the-fly initialization
  - Autonomy
  - Fault-tolerant
  - Fail-safe
- Fully integrated systems with onboard processing
- Publications:
  - Journal: 3
  - Conference: 10

Future Work

- Sensing and perception
- State estimation

+

- Control
- Planning

=

- System integration

- High level environment understanding
- Estimation-aware control
- Planning in partially known environments
- Real world deployments

Thank You!

Questions?