

Navion: A Fully Integrated Energy-Efficient Visual-Inertial Odometry Accelerator for Autonomous Navigation of Nano Drones

Amr Suleiman, Zhengdong Zhang, Luca Carlone, Sertac Karaman, and Vivienne Sze

Demo: Trevor Henderson and Diana Wofk

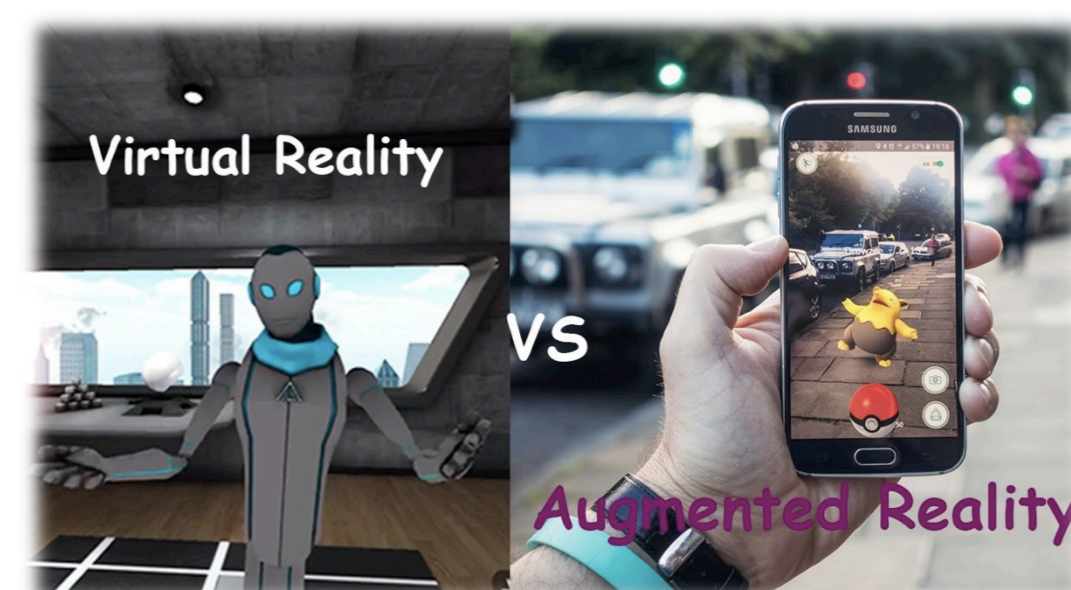


Motivation

Autonomous navigation of *nano drones*



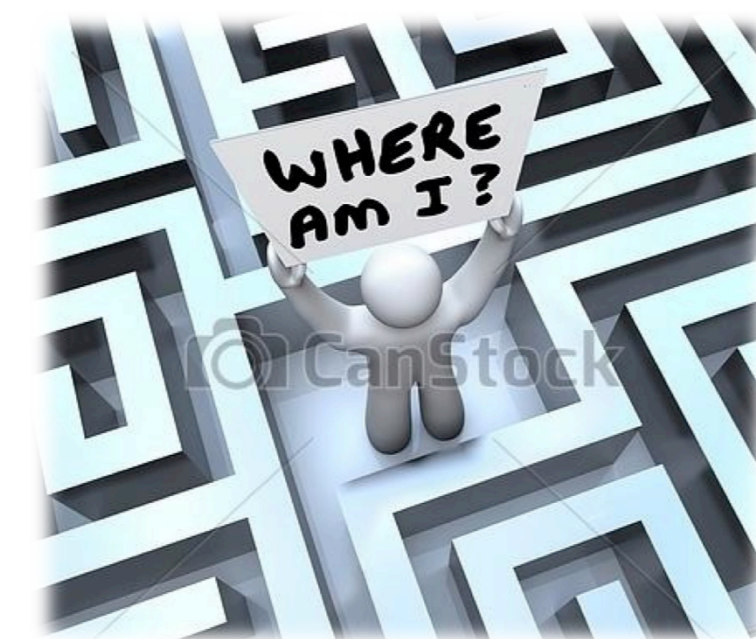
Virtual and augmented reality on *energy-constrained devices*



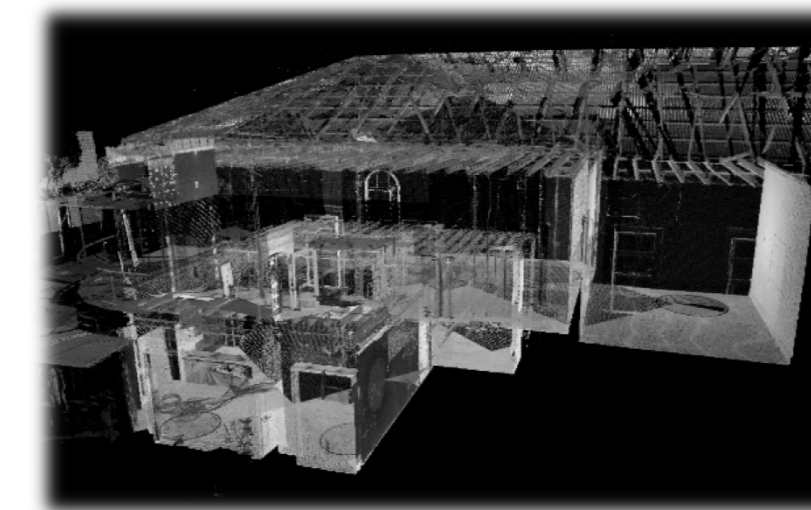
No GPS

Use images and inertial data for *state estimation*

Location (Trajectory)

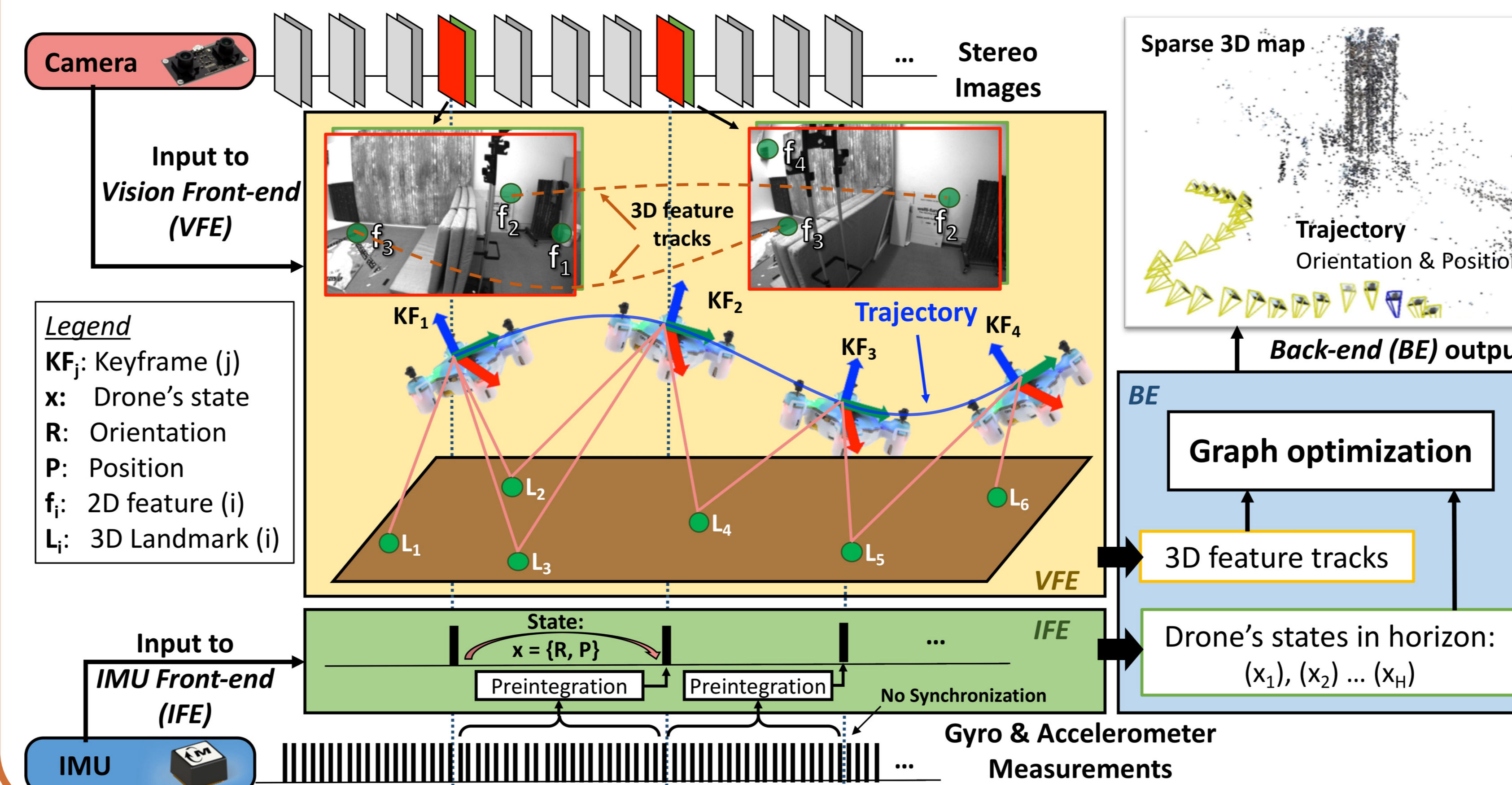


3D Map

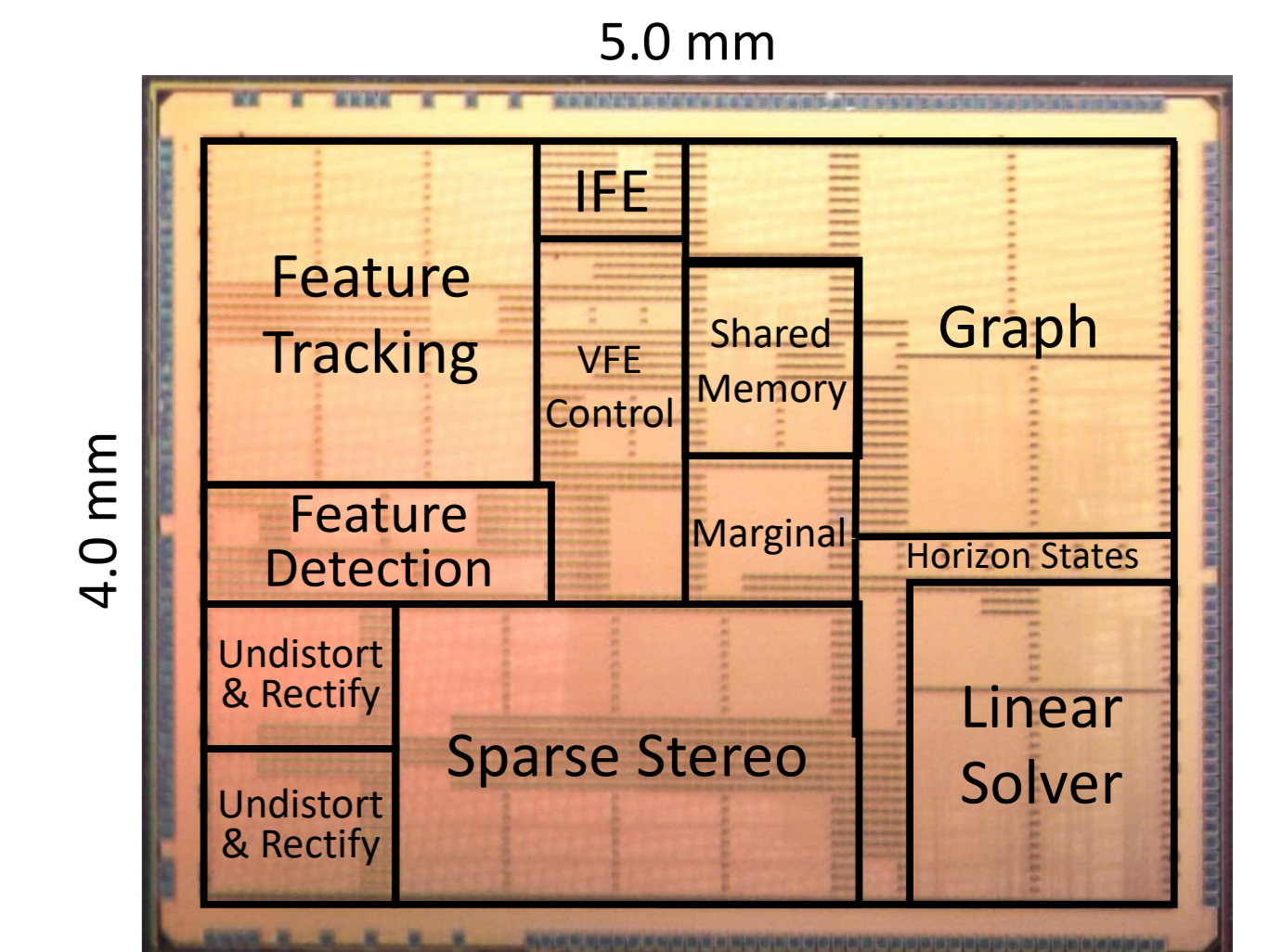


VIO: Visual-Inertial Odometry

VIO is a special instance of Simultaneous Localization and Mapping (SLAM)



Navion Results



Key Challenges

- Large frame memories for feature tracking and stereo computation
- Large graph memory with irregular memory access
- Slow linear solver

Our Approach

- Image compression
 - **4.4x** smaller frame memory
 - **4.9x** less power
- Sparse feature tracks data structure
 - **5.4x** smaller graph memory
- Sparse linear solver
 - **5.2x** smaller linear solver memory
 - **7.2x** speed up

Technology	65nm CMOS
Chip area	4.0 x 5.0 mm ²
Logic gates	2,043 kgates
SRAM	854 KB
Frame size	480x752
Camera	28 – 171 fps
Keyframe	16 – 90 fps
Power	27 – 43.2 mW

Memory size	Baseline	Image Compression	Two-stage Storage	Sparse LS Memory
Frame	3.5 MB	2.4 MB	1.6 MB	854 kB
Graph	-	-	-	-
Linear Solver	-	-	-	-
Misc.	-	-	-	-

Platform	Xeon [2]	ARM [2]	FPGA [2]	ASIC (This work)
Accuracy*	0.20%	0.20%	0.24%	0.28%
Camera max fps	21	5	20	171
Keyframe max fps	13	3	5	90
Power (W)	26.10	2.33	1.46	0.024

*Measured on EuRoc MAV dataset

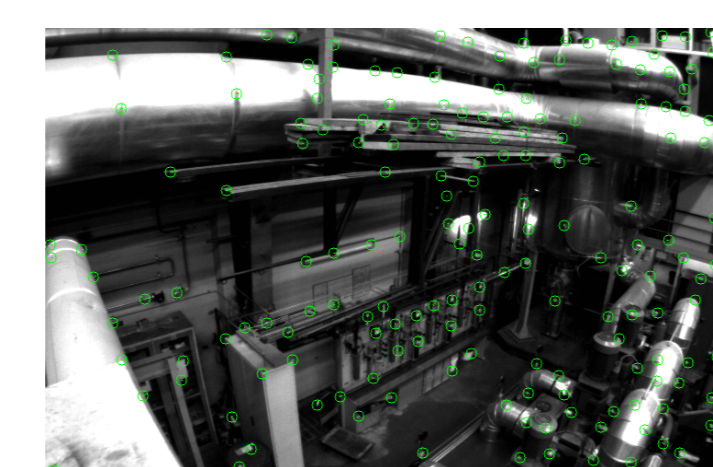
References

- [1] C. Forster, L. Carlone, F. Dellaert, D. Scaramuzza. "On-manifold preintegration theory for fast and accurate visual-inertial navigation," IEEE Trans. Robotics, 2016
- [2] Z. Zhang, A. Suleiman, L. Carlone, V. Sze, S. Karaman, "Visual-Inertial Odometry on Chip: An Algorithm-and-Hardware Co-design Approach," RSS 2017.

Vision Frontend (VFE)

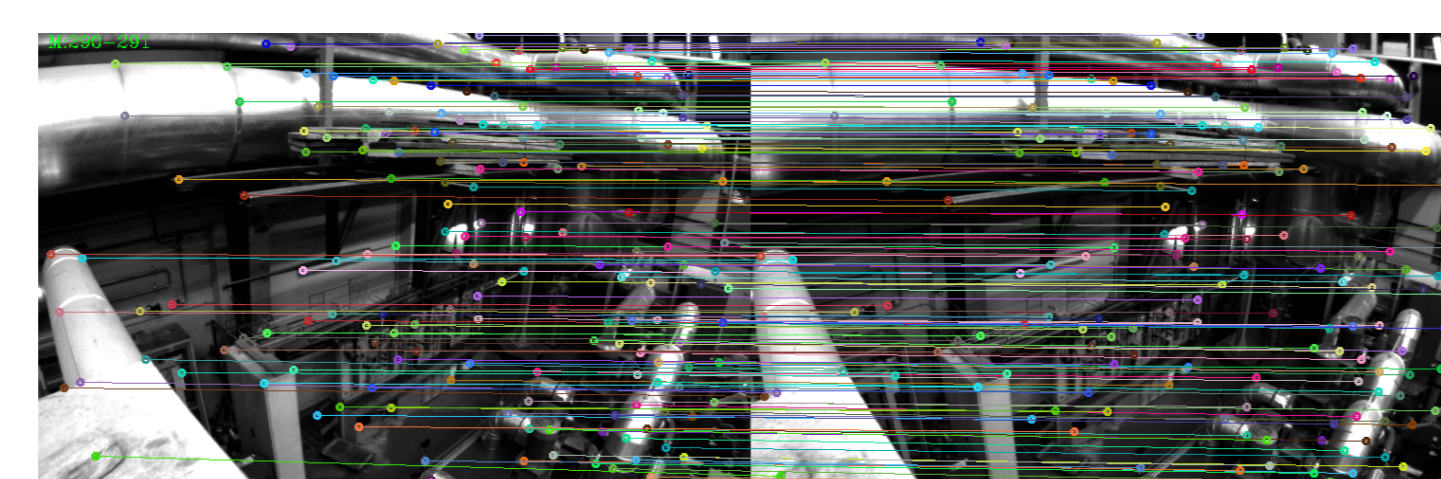
Feature Detection (Identify features)

- Extract Harris corners (up to 200 per frame)
- Search and distribute across the whole frame
- Compute for every keyframe



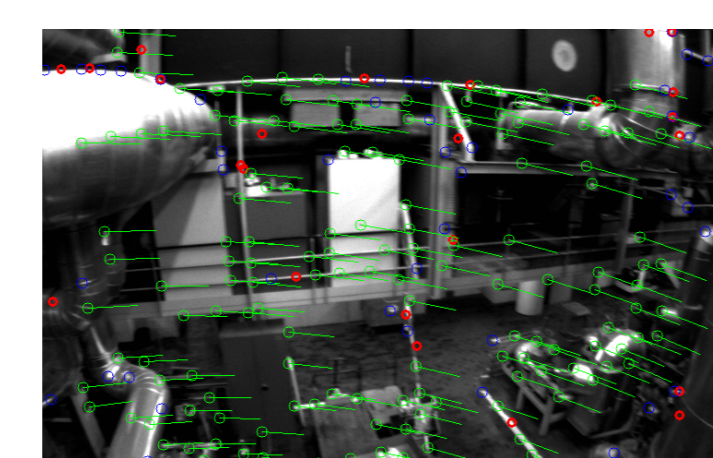
3D Stereo (Estimate depth of features – a.k.a. landmarks)

- Sparse block matching
- Compute for every keyframe



Feature Tracking (Estimate movement of features)

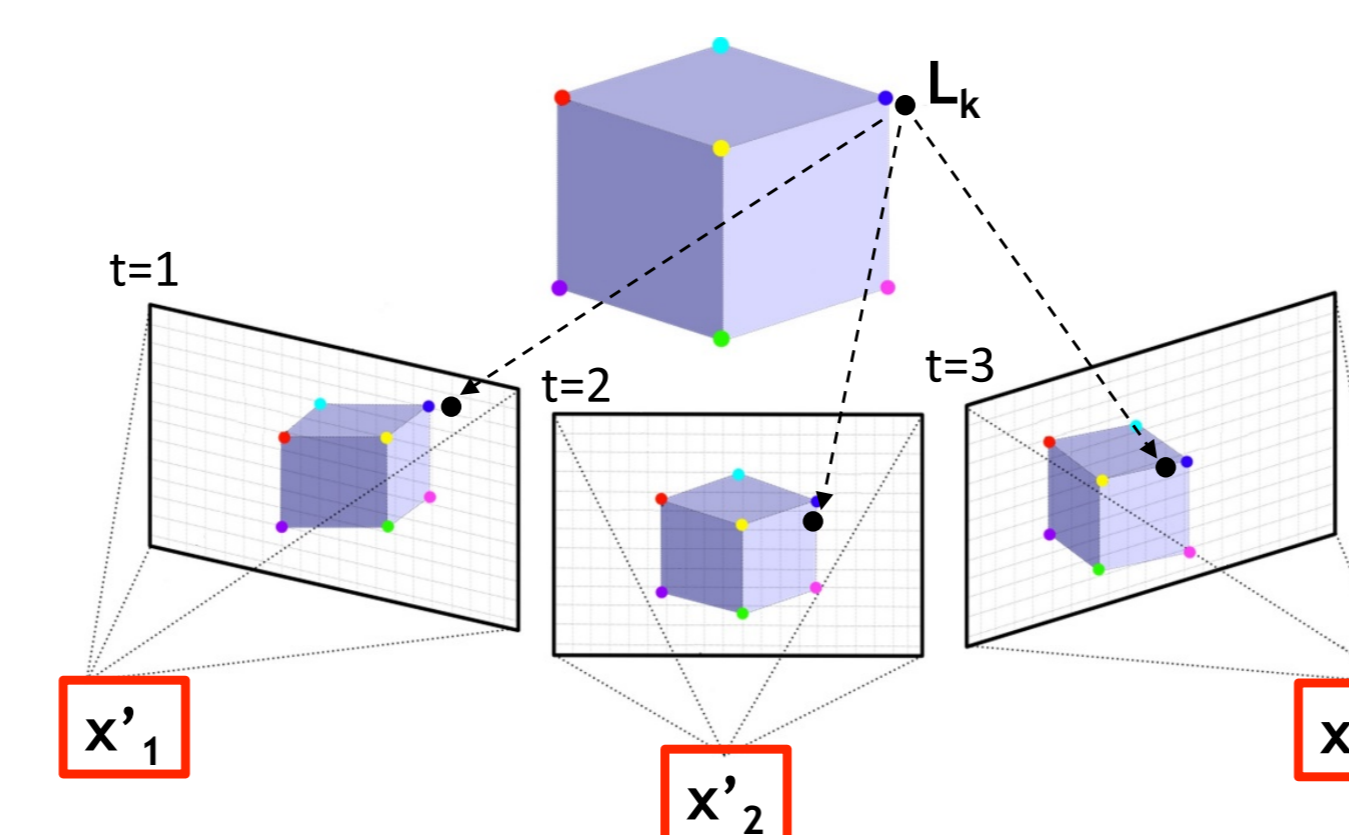
- Lucas-Kanade pyramidal optical flow
- Supports up to 3 pyramid levels
- Compute for all frames (measure movement between keyframes)



Backend (BE)

Minimize inconsistencies between measurements across time

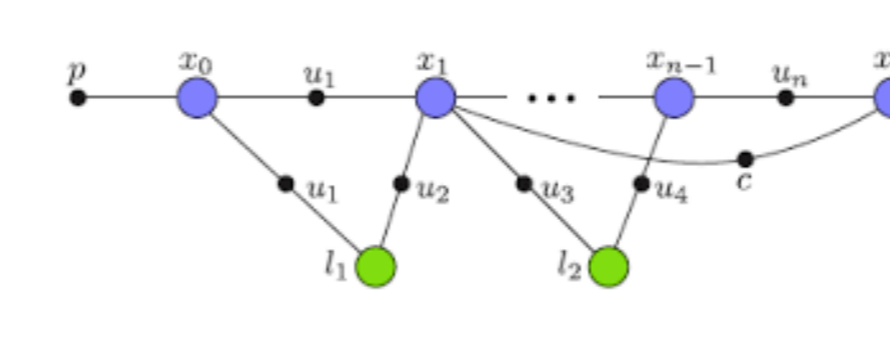
Fuses vision and IMU estimates to refine the final state estimates (x')



Update state estimates (x')

$$\min_x \sum_{(i,j) \in \mathcal{E}_F} \|\tilde{r}_{IMU}(x, \Delta \tilde{R}_{ij}, \Delta \tilde{p}_{ij}, \Delta \tilde{v}_{ij})\|^2 + \sum_{k \in \mathcal{L}} \sum_{i \in \mathcal{E}_k} \|\tilde{r}_{CAM}(x, l_k, u_{ik}^r, u_{ik}^l)\|^2 + \|\tilde{r}_{PRIOR}(x)\|^2$$

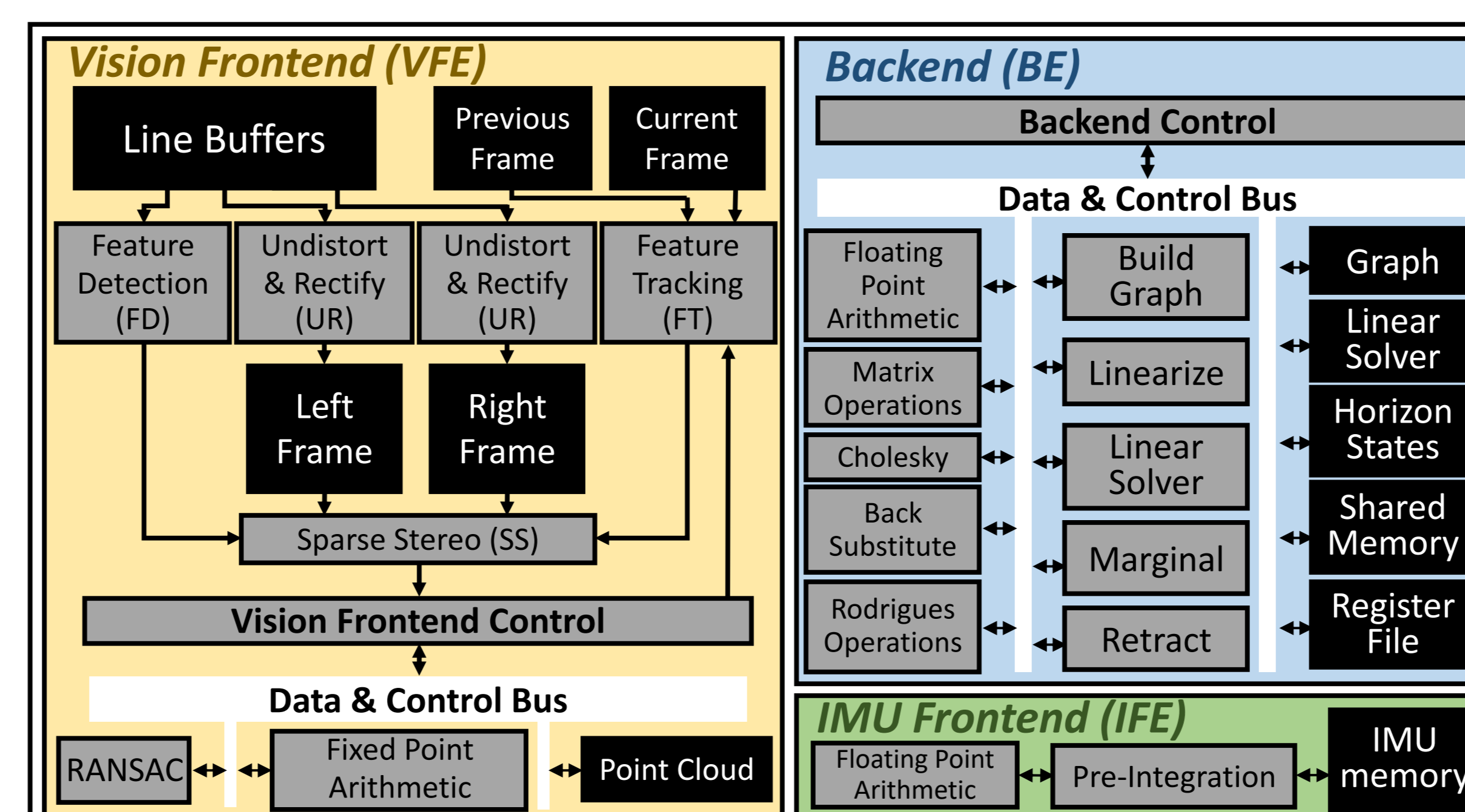
IMU Factors Vision Factors



Over **4000+** factors in optimization

Solve with state-of-the-art factor graph algorithm

Navion Chip Architecture



Fully Integrated System (no off-chip processing or storage)

Sponsors



Want to know more?

Visit navion.mit.edu