Visual-Inertial Odometry on Chip: An Algorithm-and-Hardware Co-design Approach – Supplementary Material –

Zhengdong Zhang*, Amr Suleiman*, Luca Carlone, Vivienne Sze, Sertac Karaman
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
Emails: {zhangzd,suleiman,lcarlone,sze,sertac}@mit.edu, Website: http://navion.mit.edu
These authors contributed equally to this work

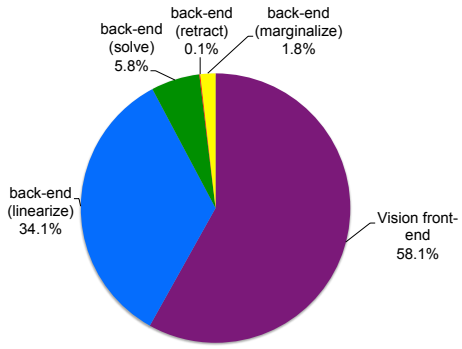

Fig. A1. VIO Complexity breakdown: CPU runtime.

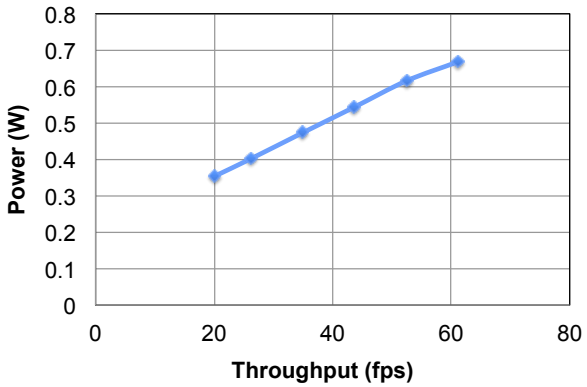

Fig. A2. Throughput versus power trade-off for the front-end, obtained by sweeping the clock frequency.

Platform	Desktop CPU	Embedded CPU	Embedded GPU	FPGA
Processor	Intel Xeon E5 4627v2	ARM Cortex A15	NVidia Maxwell TX1	Xilinx Kintex-7 XC7K355T
# of cores / DSPs	10	4	256	1,440
Frequency (GHz)	3.3	2.1	1.6	any
Memory (MB)	16*	2.5*	2.2 - 2.5*	3.2
Power (W)	20–130	1–8	6.5–15	0.2–32

TABLE A1
RESOURCES FOR DIFFERENT HARDWARE CHOICES. * L2 CACHE.

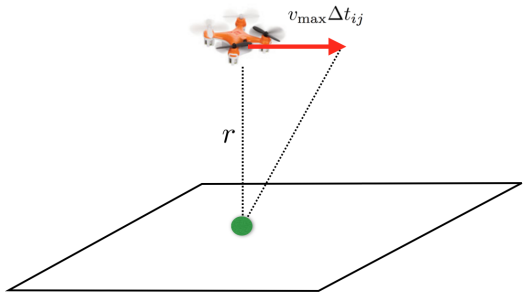

Fig. A3. **Example 1**: keyframe rate design from speed specs. Let us consider a camera with resolution 752×480 and focal lengths of 460 (approximately equal in the x and y direction), and assume a maximum speed of $v_{\max} = 5$m/s among obstacles at a distance of 3m. Then, to maintain a maximum displacement $\Delta u < 160$pixels (i.e., 1/3 of the image height, or equivalently, to be able to track features in 2/3 of the images), we obtain an upper bound on the intra-keyframe time equal to 0.2s (5fps).

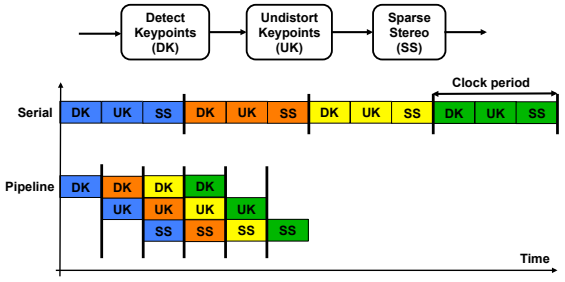

Fig. A4. **Example 2**: pipelined computation of modules in the front-end.

Resource	Front-end	Linear Solver	Linearize	Marginalize	Stereo Factors	IMU & Other Factors	State Estimation	Total (Utilization)
Memory (kB)	1526.3	355.1	0	179.3	85.9	14.7	3.3	2164.6 (67.28%)
Block RAM	378	80	0	41	21.5	4	1	525.5 (73.50%)
Flip Flops	72,301	4,985	67,120	shared	n/a	n/a	n/a	144,406 (32.4%)
LUTs	111,369	15,544	64,912	shared	n/a	n/a	n/a	191,825 (86.17%)
DSP	607	62	102	shared	n/a	n/a	n/a	771 (53.5%)

TABLE A2
RESOURCE UTILIZATION ON KINTEX-7 XC7K355T FPGA

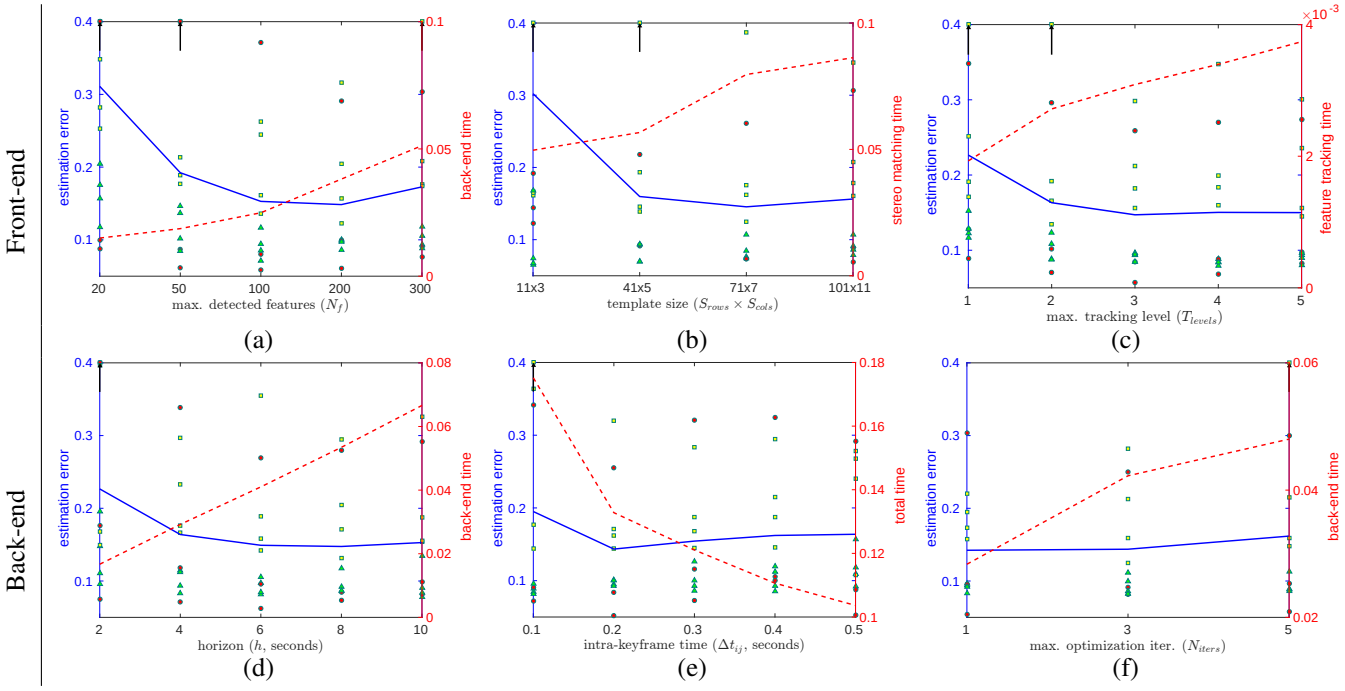

Fig. A5. Algorithmic parameter choice: error-time trade-off for different parameters. Average translation errors are shown as solid blue lines, while time is show as a dashed red line. The figure also shows the details of each datasets as a scatter plot: the EuRoC datasets are classified as "easy", "medium", and "difficult", and, accordingly, we show the average errors for the easy, medium, and difficulty datasets as green triangles, yellow squares, and red circles.