



nVIDIA®

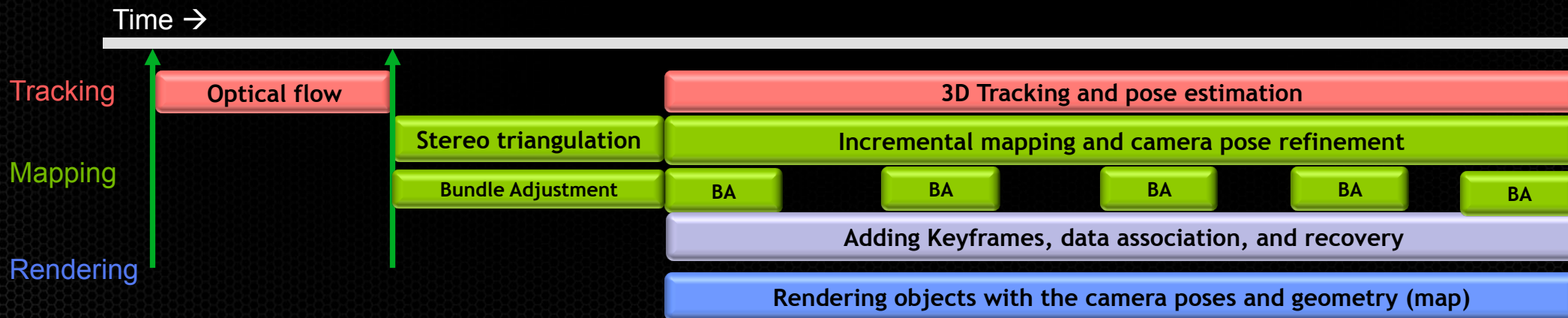
SLAM@NVIDIA

Kari Pulli | Senior Director of Research

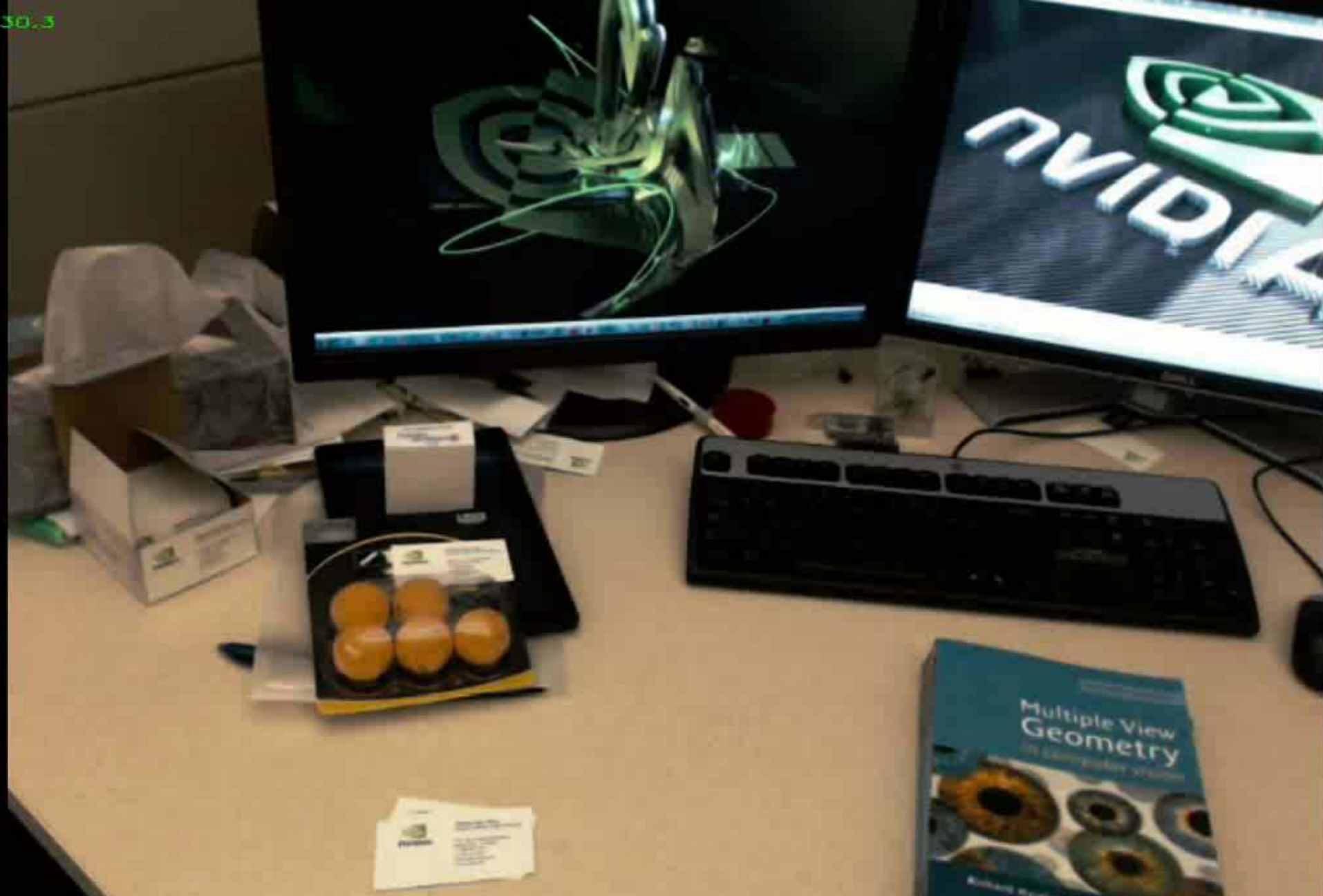
Overview

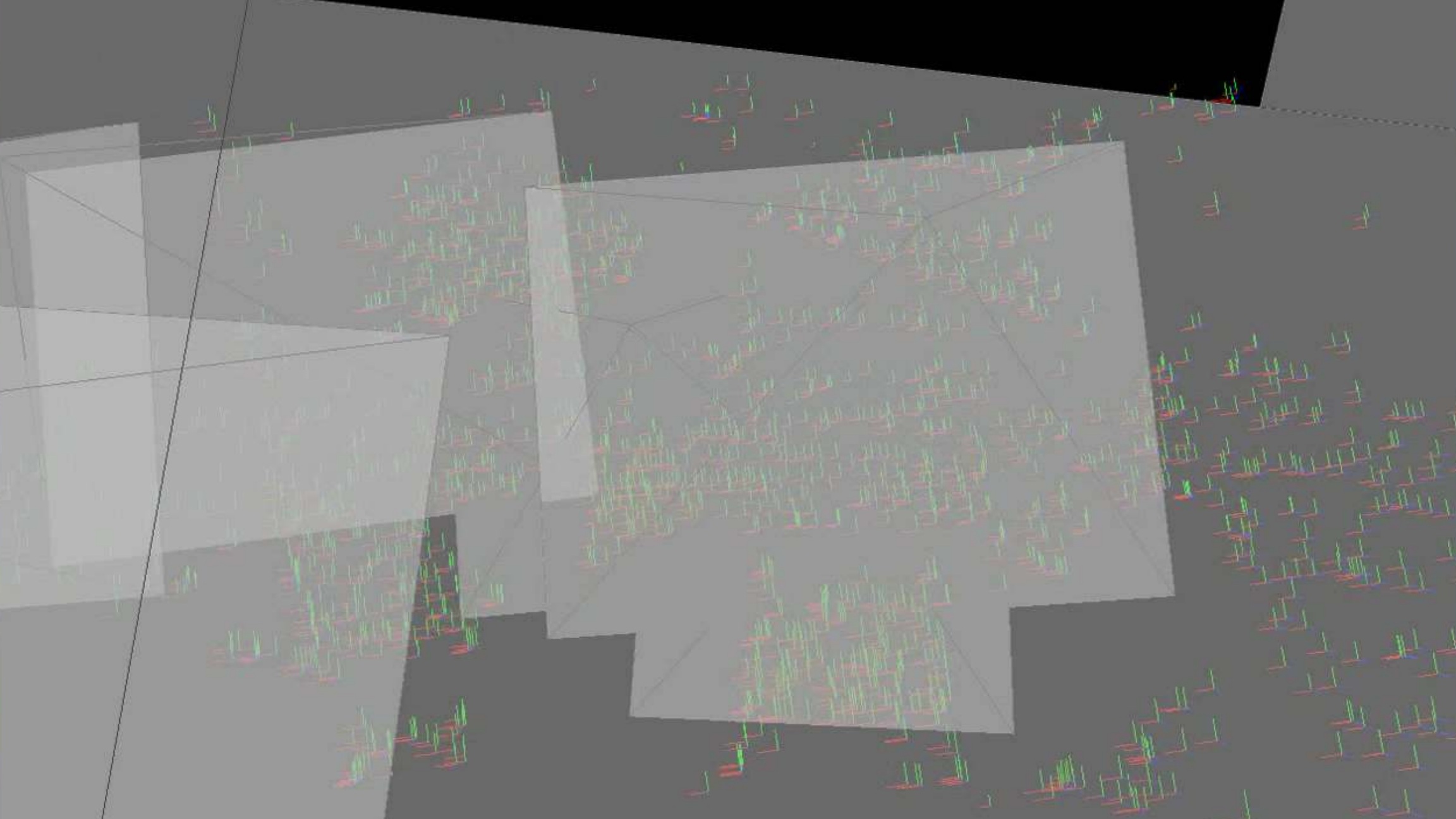
- Keyframe-based SLAM
- 3D rendering for Augmented Reality
- Problems with traditional keyframe-based SLAM
- Solution: Deferred Triangulation SLAM

KeyFrame-based SLAM

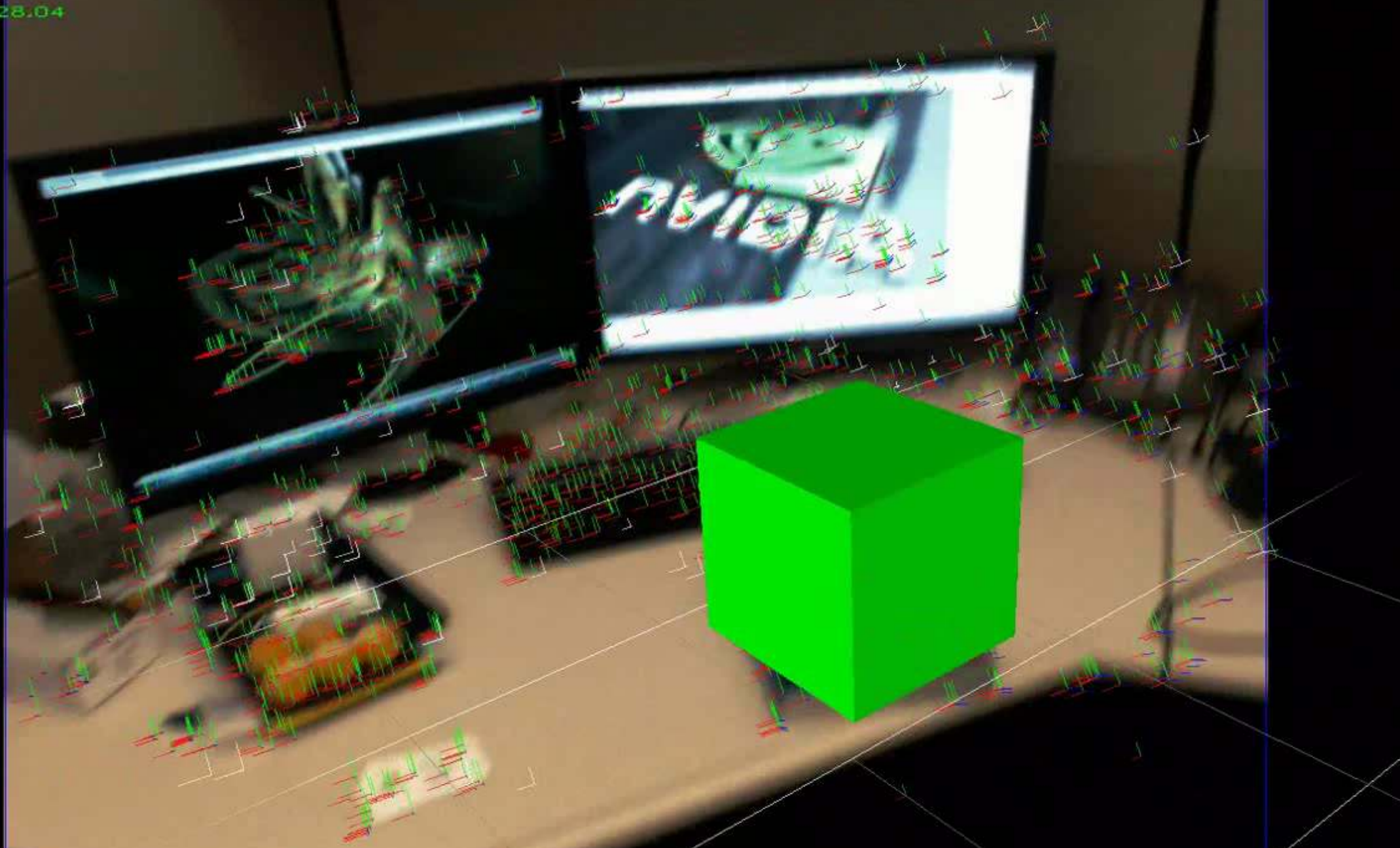


FPS: 30.3

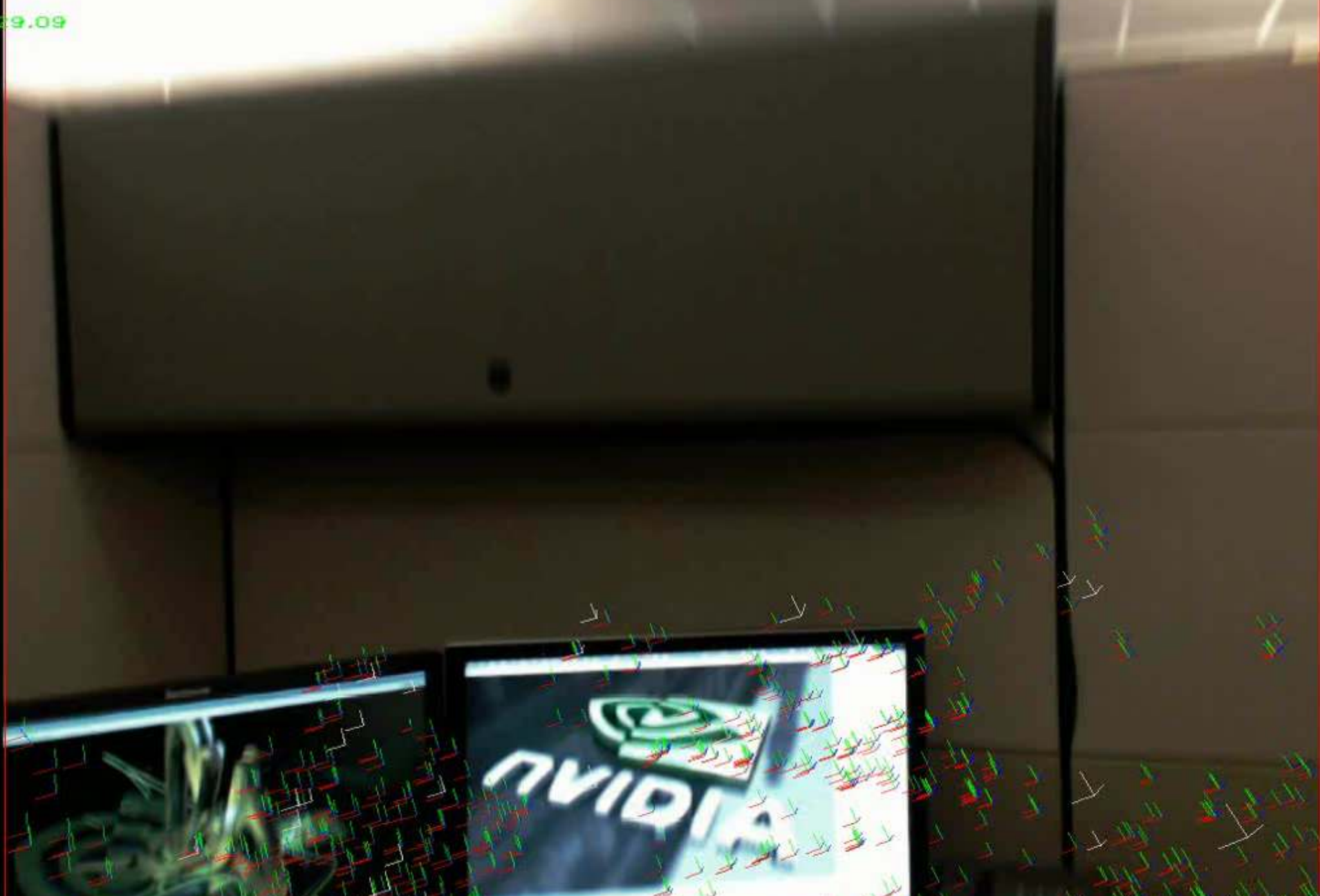




FPS: 28.04



FPS: 29.09





**SimpleClothing - 8 - Male / Trenchcoat 800 - GPU,
localspace
FPS = 29**



**SimpleClothing - 8 - Male / Trenchcoat 800 - CPU,
localspace
FPS = 29**



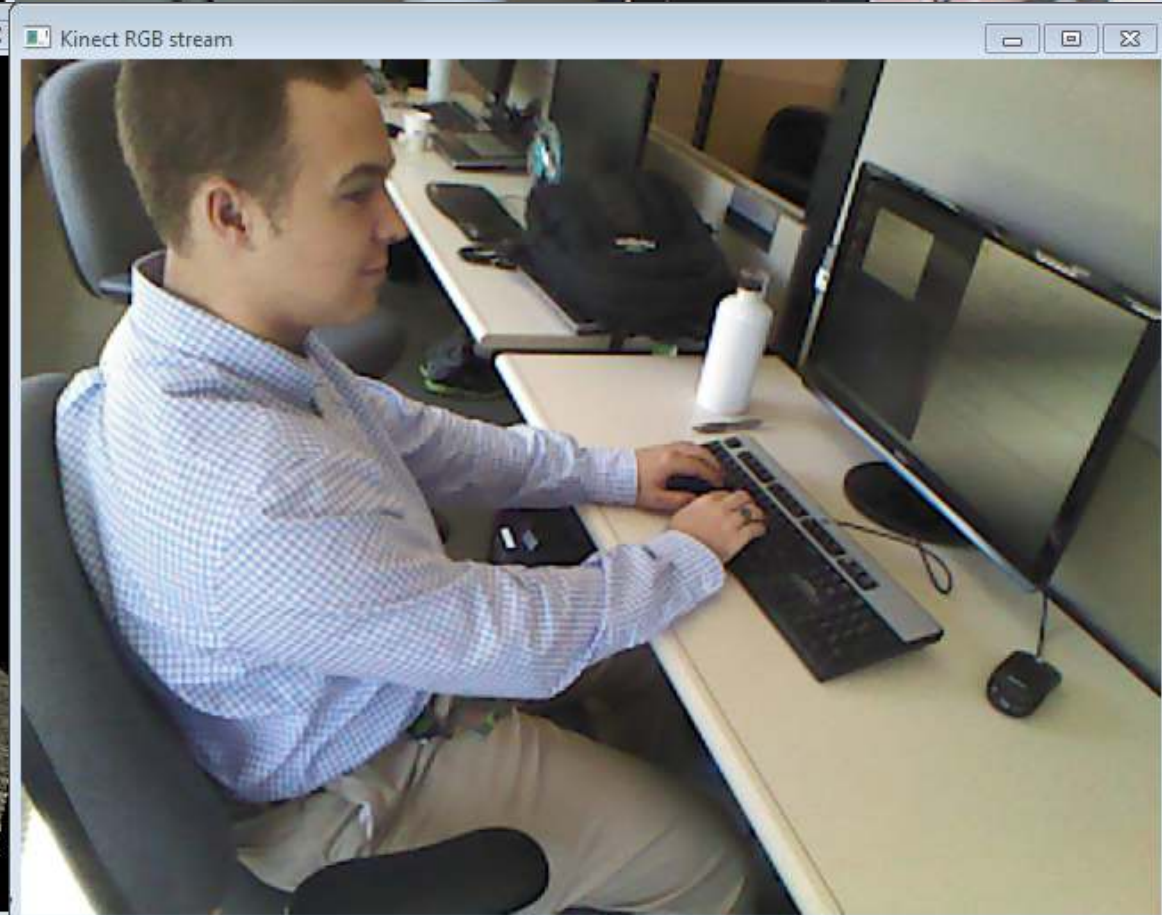
FPS: 15.72

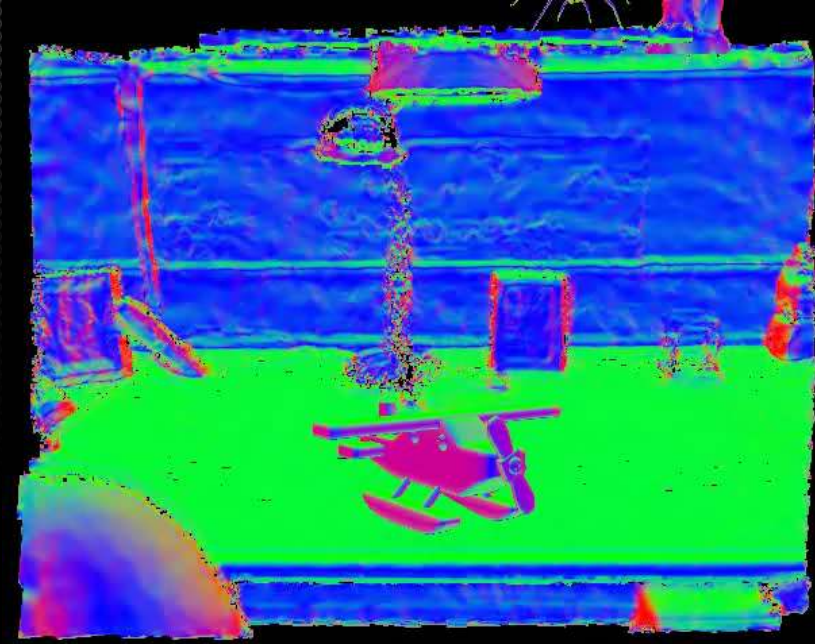
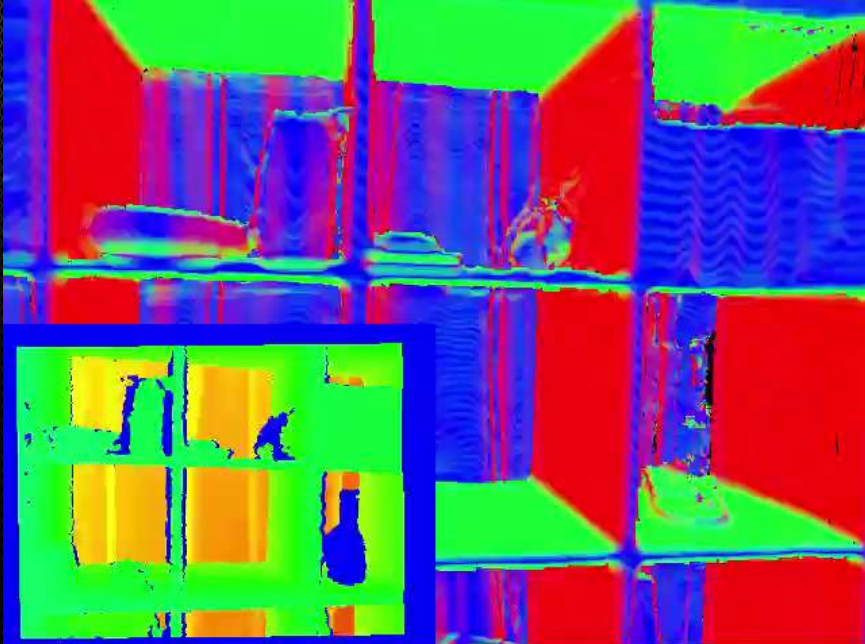


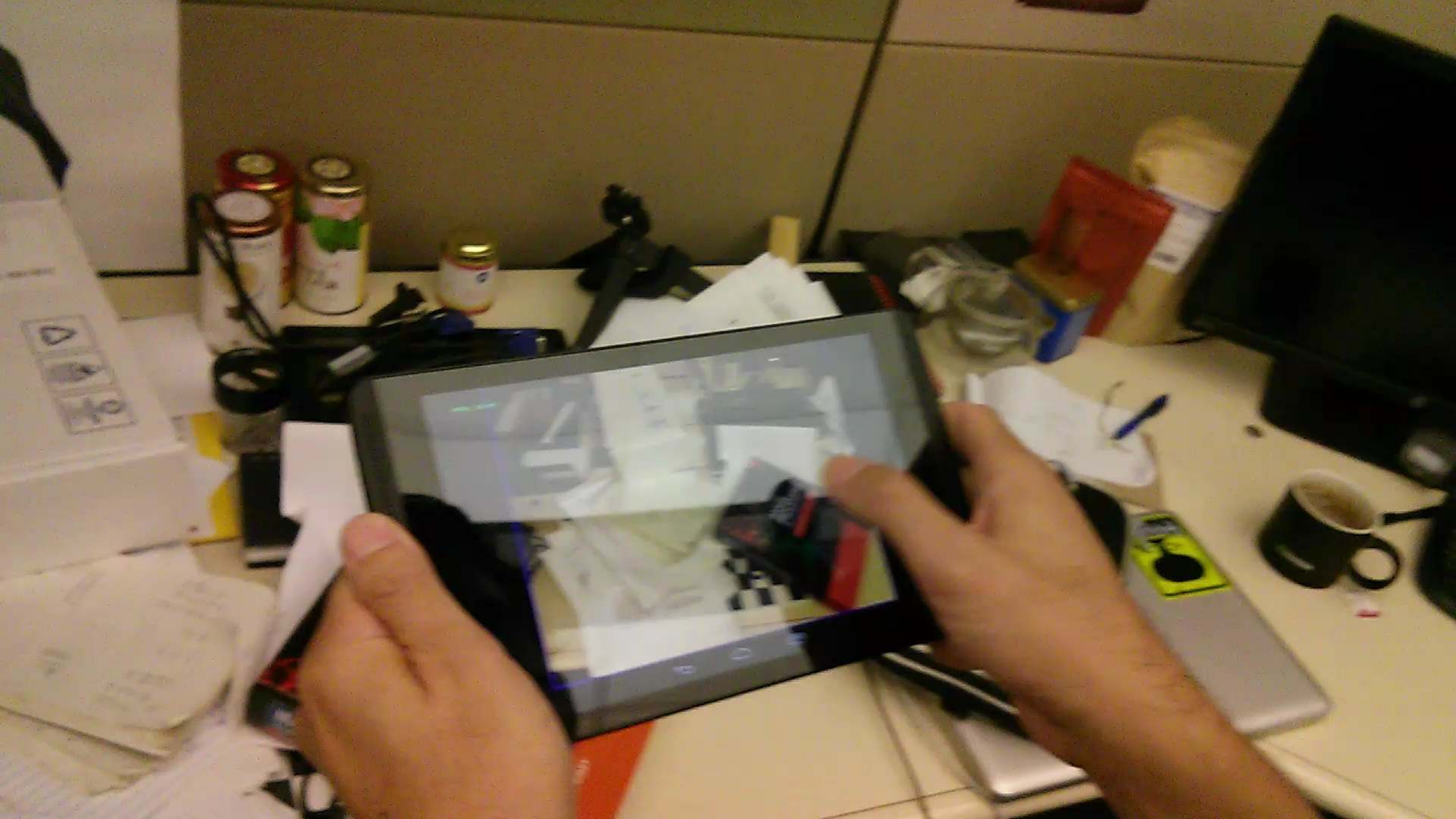
DTAM











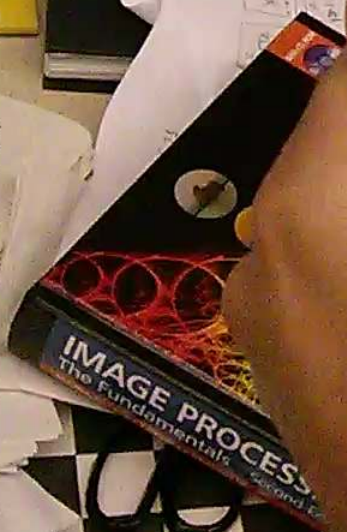
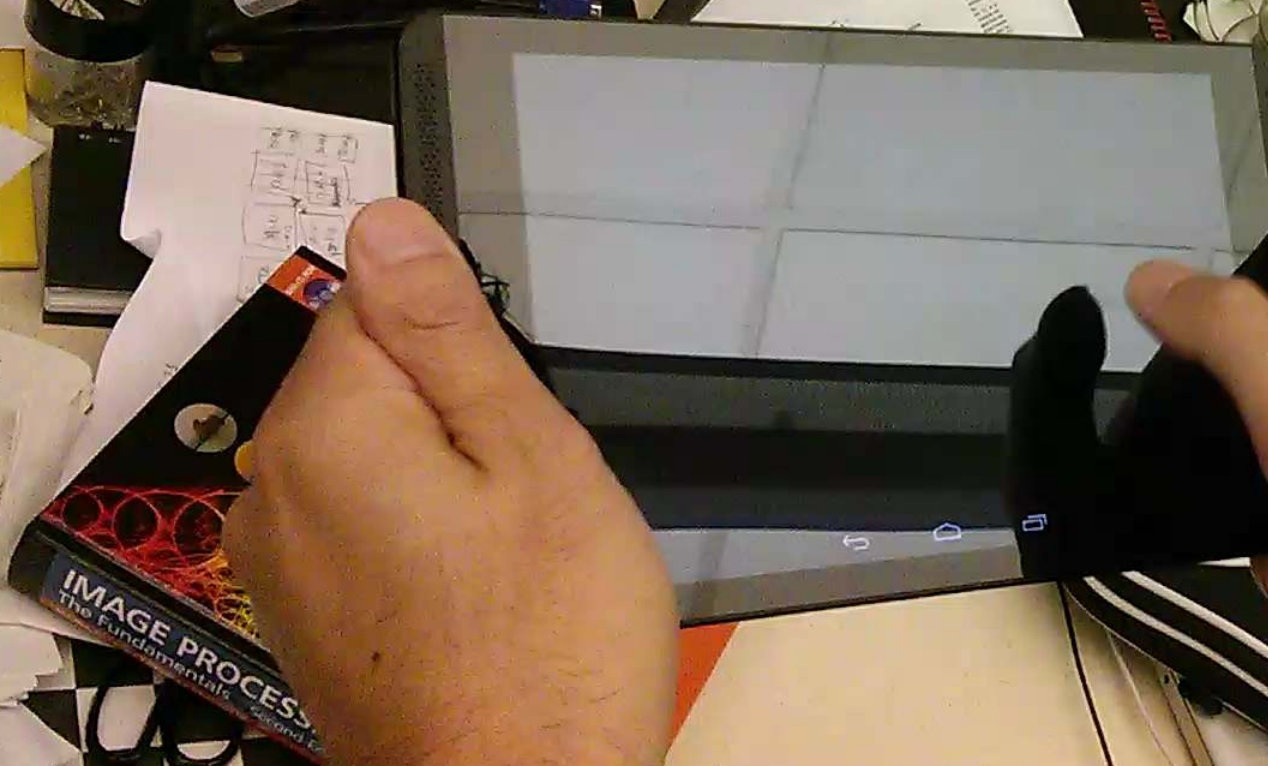




IMAGE PROCESSING
The Fundamentals
Rafael G. Baraniuk, et al.

This is how

DT-SLAM: Deferred Triangulation for Robust SLAM

Daniel Herrera C.[†], Kihwan Kim[‡], Juho Kannala[†], Kari Pulli[‡], and Janne Heikkilä[†]

[†]University of Oulu

[‡]NVIDIA Research

3DV 2014



How to deal with the rotation?

- Deferred triangulation



0.5x Speed for visualization

● Deferred 2D points

● Triangulated 3D points

How to deal with the rotation?

- Deferred triangulation
- Jointly (2D/3D) constrain a pose

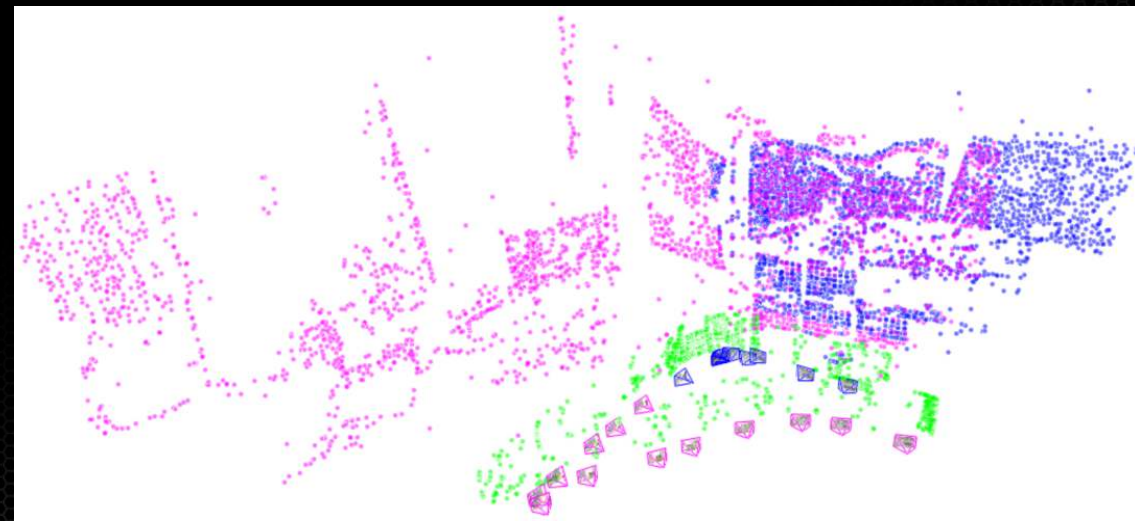
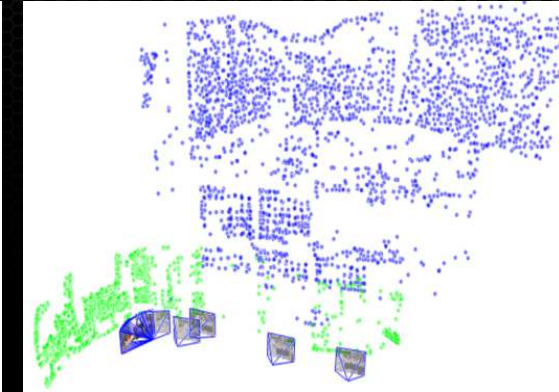
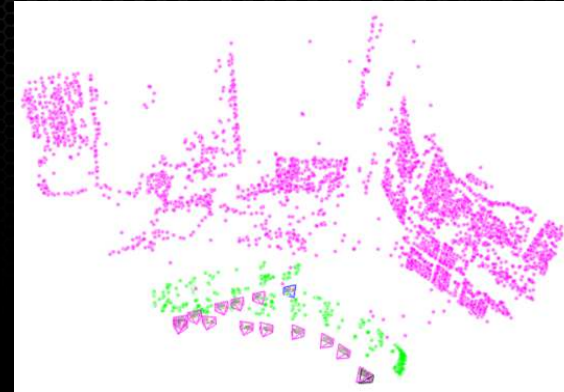


0.5x Speed for visualization

- Deferred 2D points
- Triangulated 3D points

How to overcome the rotation?

- Deferred triangulation
- Jointly (2D/3D) constrain a pose
- Region merging



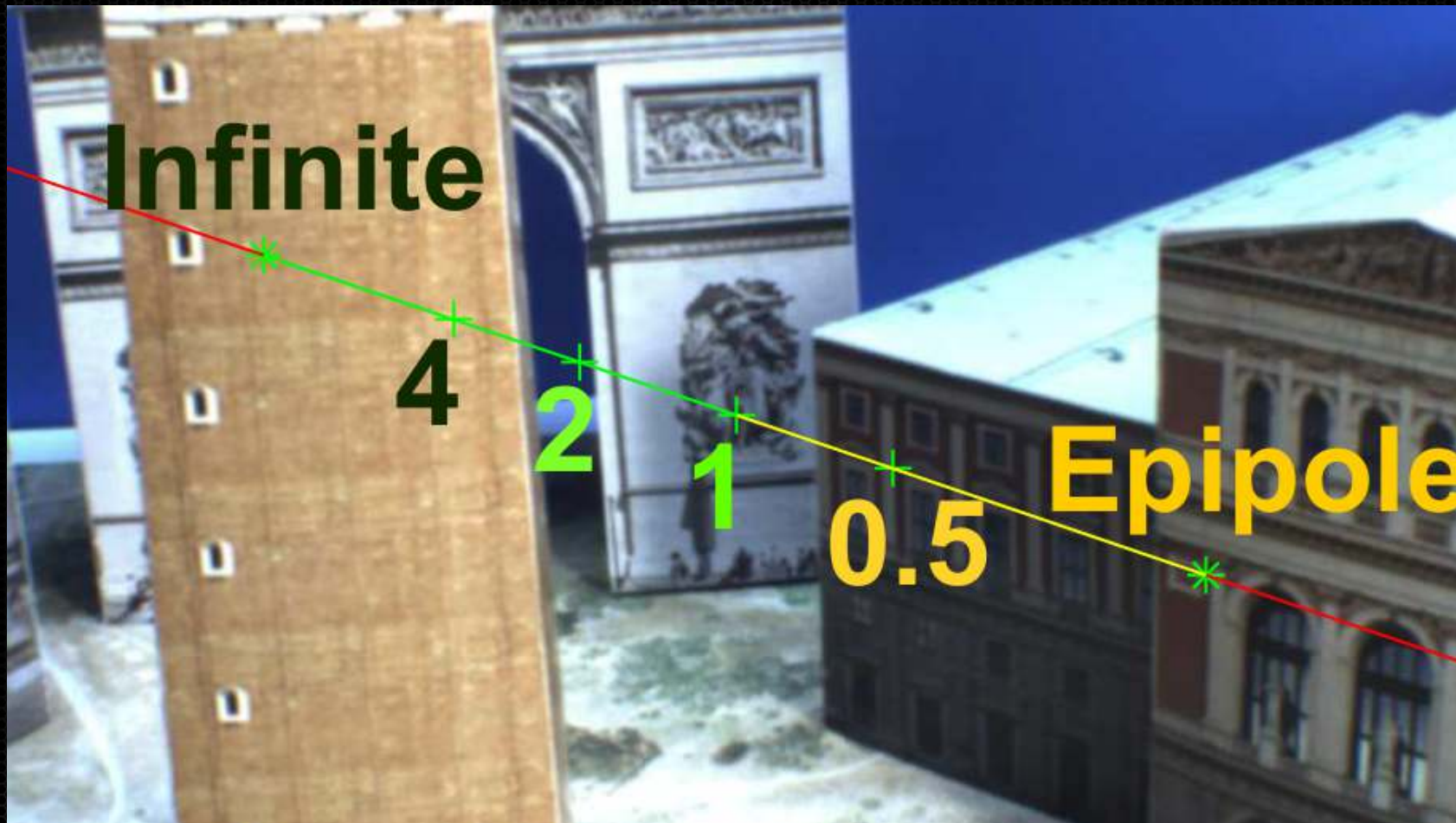
Pose estimation

$$E_{3D} = \|\phi([\mathbf{R}_k | \mathbf{t}_k] \tilde{\mathbf{x}}) - \mathbf{m}_k\|^2.$$

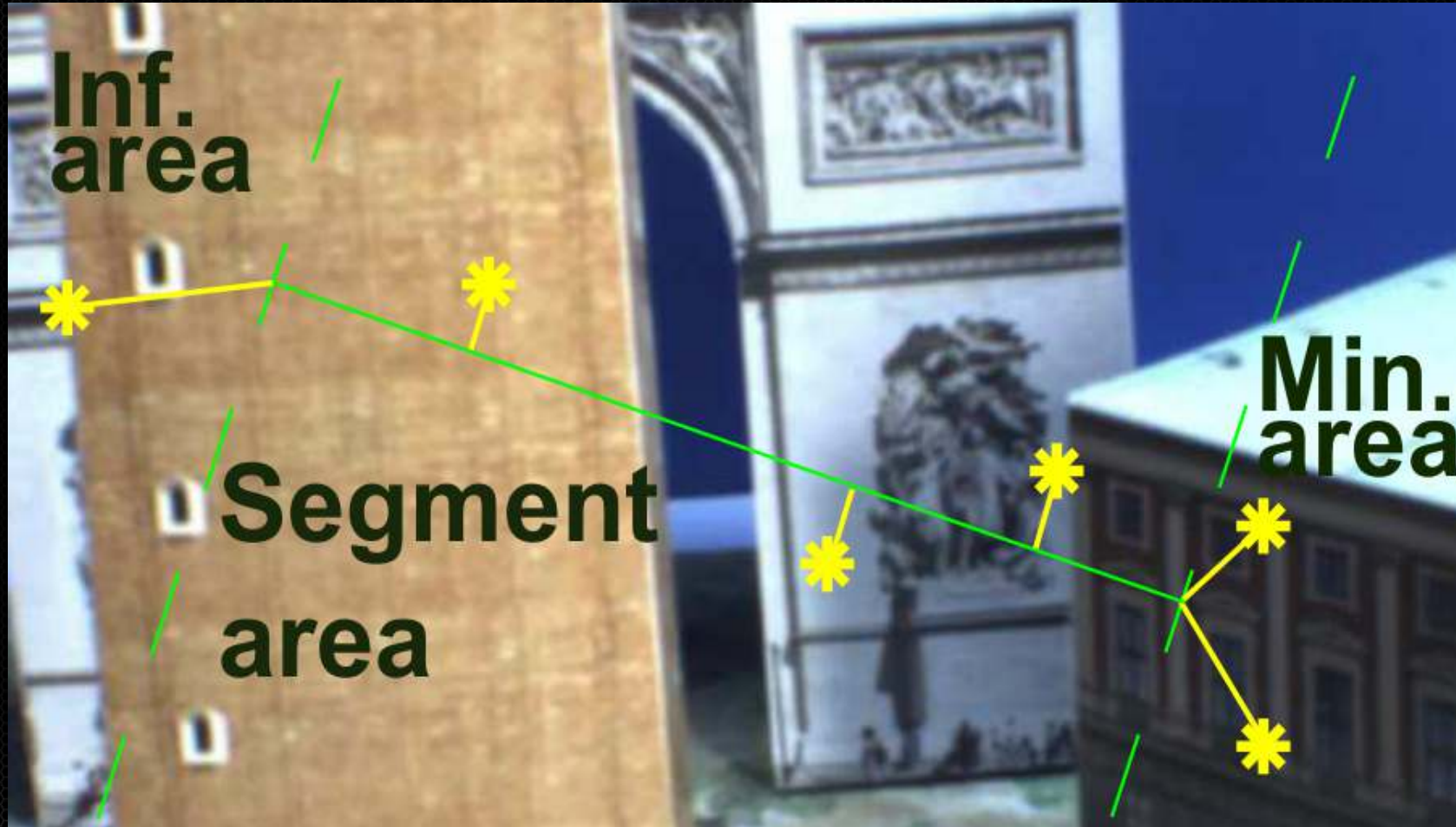


$$\arg \min_{\mathbf{R}_k, \mathbf{t}_k} \sum_i \rho(E_{3D,i})$$

Epipolar segment



Epipolar segment



Pose estimation

$$E_{3D} = \|\phi([\mathbf{R}_k | \mathbf{t}_k] \tilde{\mathbf{x}}) - \mathbf{m}_k\|^2.$$



$$\arg \min_{\mathbf{R}_k, \mathbf{t}_k} \sum_i \rho(E_{3D,i}) + \sum_j \rho(E_{2D,j})$$

Bundle Adjustment

$$\arg \min_{\mathcal{R}, \mathcal{T}, \mathcal{X}} \sum_{k \rightarrow K} \left(\sum_{i \rightarrow M} \rho(E_{3D,k,i}) + \sum_{j \rightarrow N} \rho(E_{2D,k,j}) \right)$$

Draw AR Off

View Map Off

Spacebar

Reset

Menu:

ESOKU
L S vk M T R T P L S
2 3 31 1 2 3 4 5 6 7
9 10 32 8 9 10 11 12 13 14
16 17 33 15 16 17 18 19 20 21
23 24 34 22 23 24 25 26 27 28
30 31 35 29 30 31

MARRAKU
L S vk M T K T P L S

SOULOKU
vk M T R T P L S
2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

Translate the camera slowly sideways, and press spacebar again to perform stereo init.

FFmpeg
Frame 0



ELOKUU

T	P	L	S	vi	M	T	K	T	P	L	S
1	2	3		31	1	2	3	4	5	6	7
7	8	9	10	32	8	9	10	11	12	13	14
14	15	16	17	33	15	16	17	18	19	20	21
21	22	23	24	34	22	23	24	25	26	27	28
28	29	30	31	35	29	30	31				

MARRASKUU

T	P	L	S	vi	M	T	K	T	P	L	S
1	2			44	1	2	3	4	5	6	
6	7	8	9	45	7	8	9	10	11	12	13

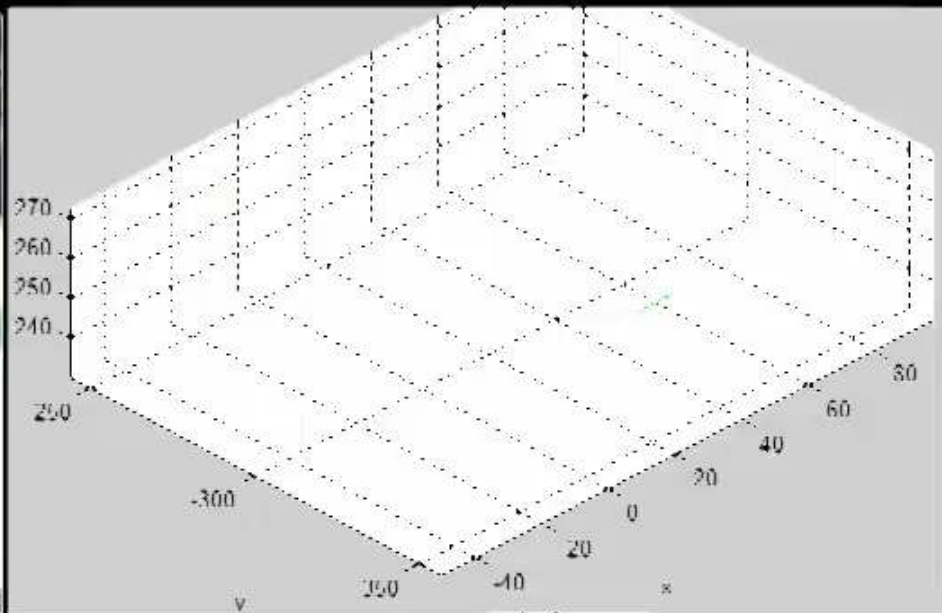
JOUKOKUU

T	P	L	S	vi	M	T	K	T	P	L	S
1	2	3	4	46							
4	5	6	7	47	5	6	7	8	9	10	11

Regions: 1, Keyframes: 1, Features (2D: 2060, 3D : 0)

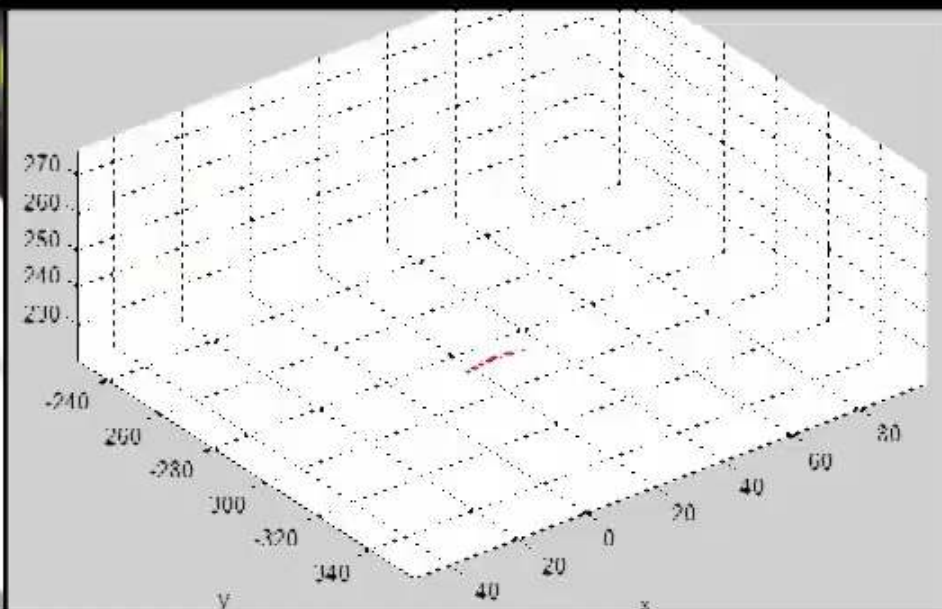
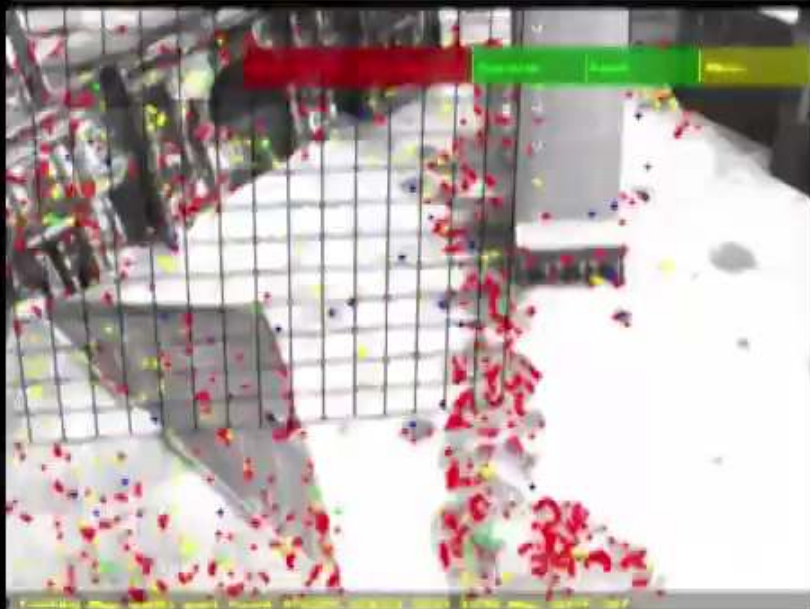


DT-SLAM



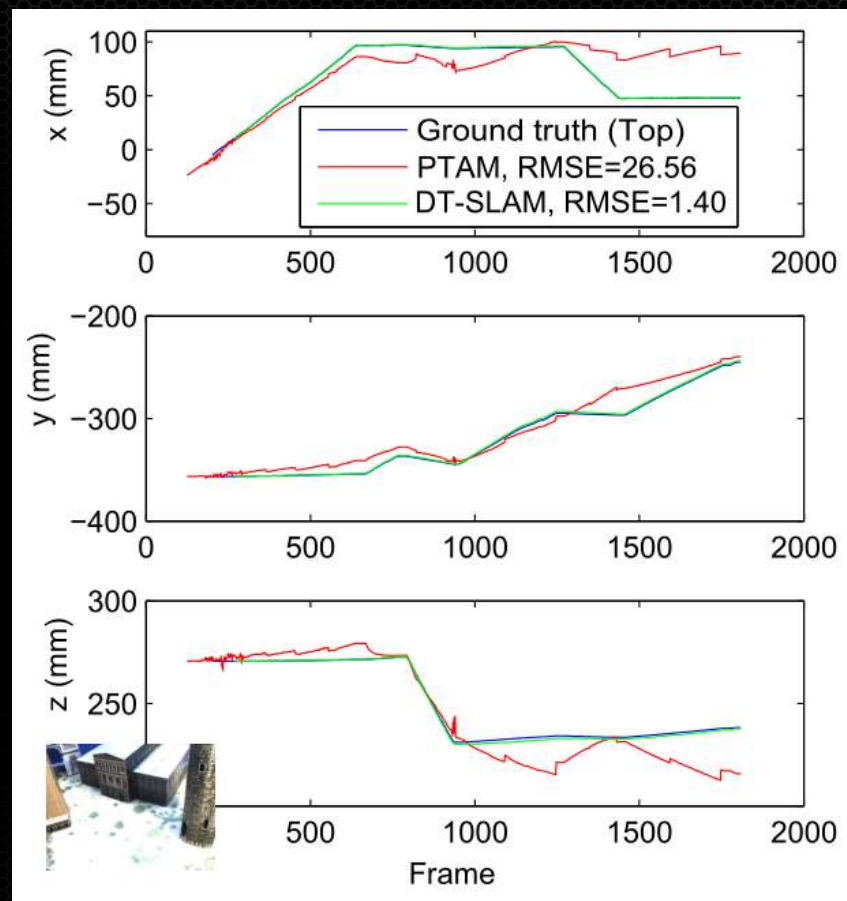
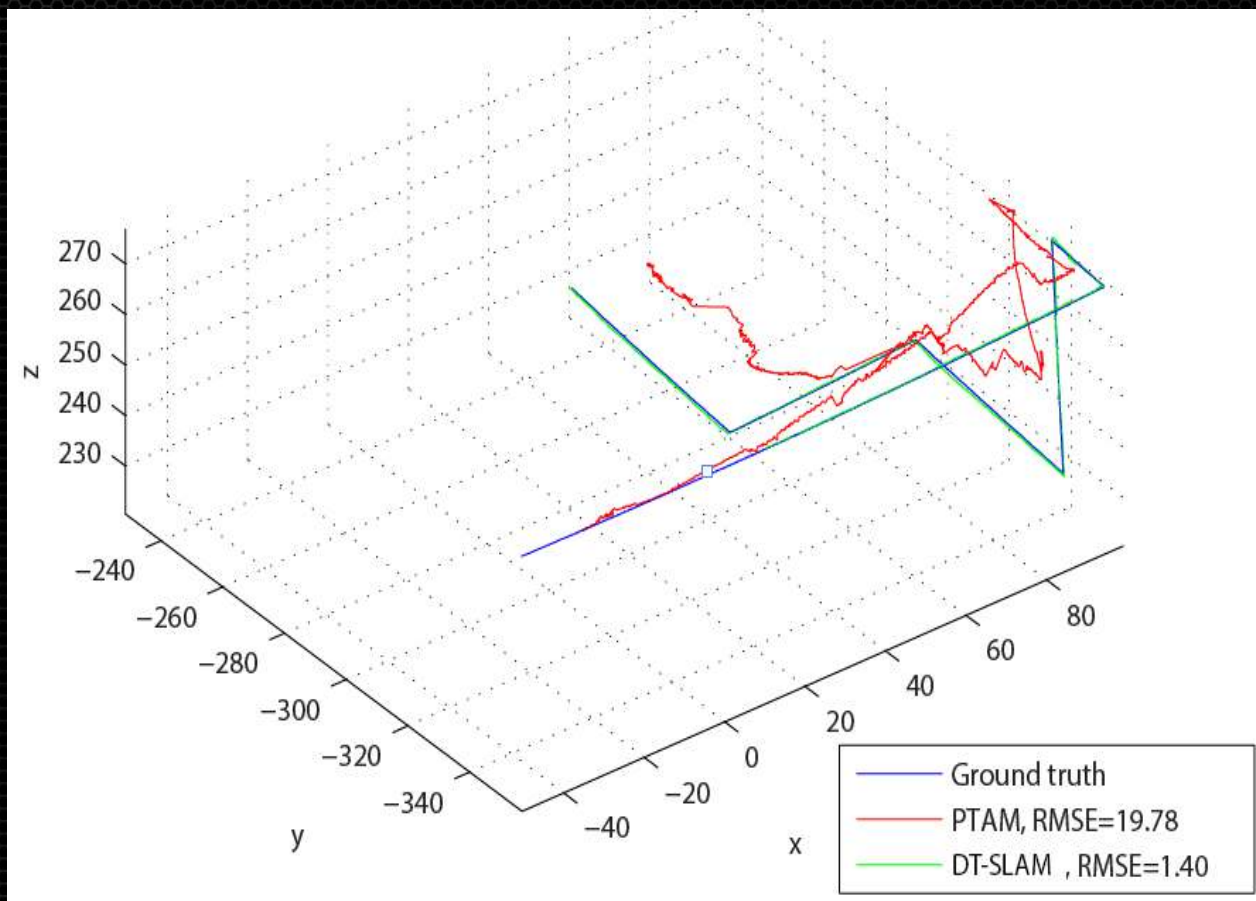
Trajectory
DT-SLAM

PTAM
Klein et al.



Trajectory
PTAM

x2 speed



Comparison with Hybrid SLAM and PTAM

Results from Hybrid SLAM and PTAM taken directly from Pirchheim et al.
The footage contains captions from the original video

DT-SLAM
(Ours)



Hybrid SLAM
(Pirchheim et al.)



PTAM
(Klein et al.)



Another scene reconstruction example



Hand-held camera scene in City of Sights dataset

x2 speed

Summary

- Keyframe-based SLAM is efficient
 - and can run in real time on mobile devices
- But it has problems
 - A separate initialization phase is annoying
 - Breaking with pure rotations is a critical failure
- Both can be addressed by
 - tracking first in 2D
 - deferring triangulation until there is enough baseline between the keyframes
- Bonus: we plan to open source the implementation