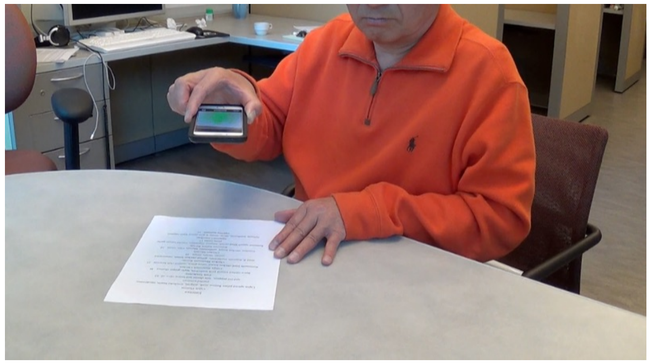
Multi-planar Monocular Reconstruction of Manhattan Indoor Scenes

Seongdo Kim, Roberto Manduchi

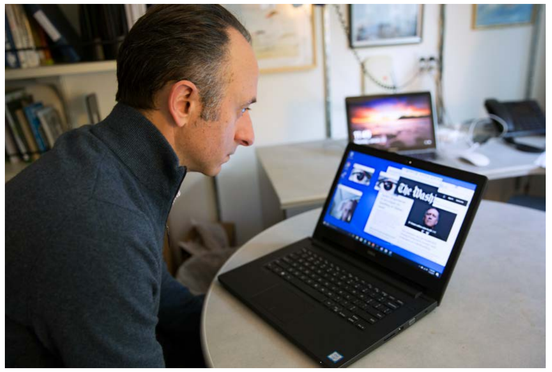
Baskin Engineering UC SANTA CRUZ

Wayfinding by visual odometry

Scene text access



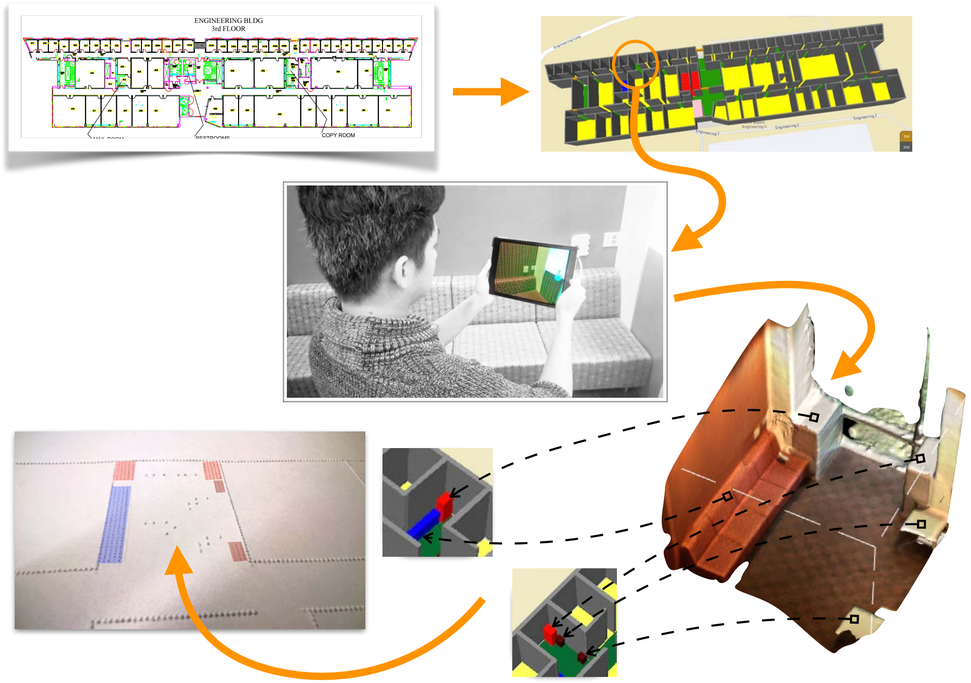
Document access



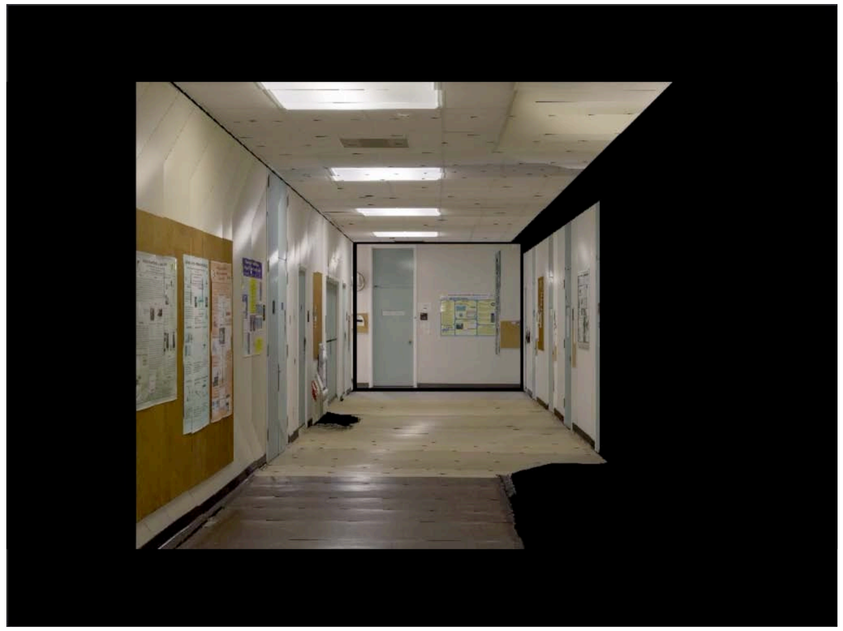
Gaze-contingent screen magnification

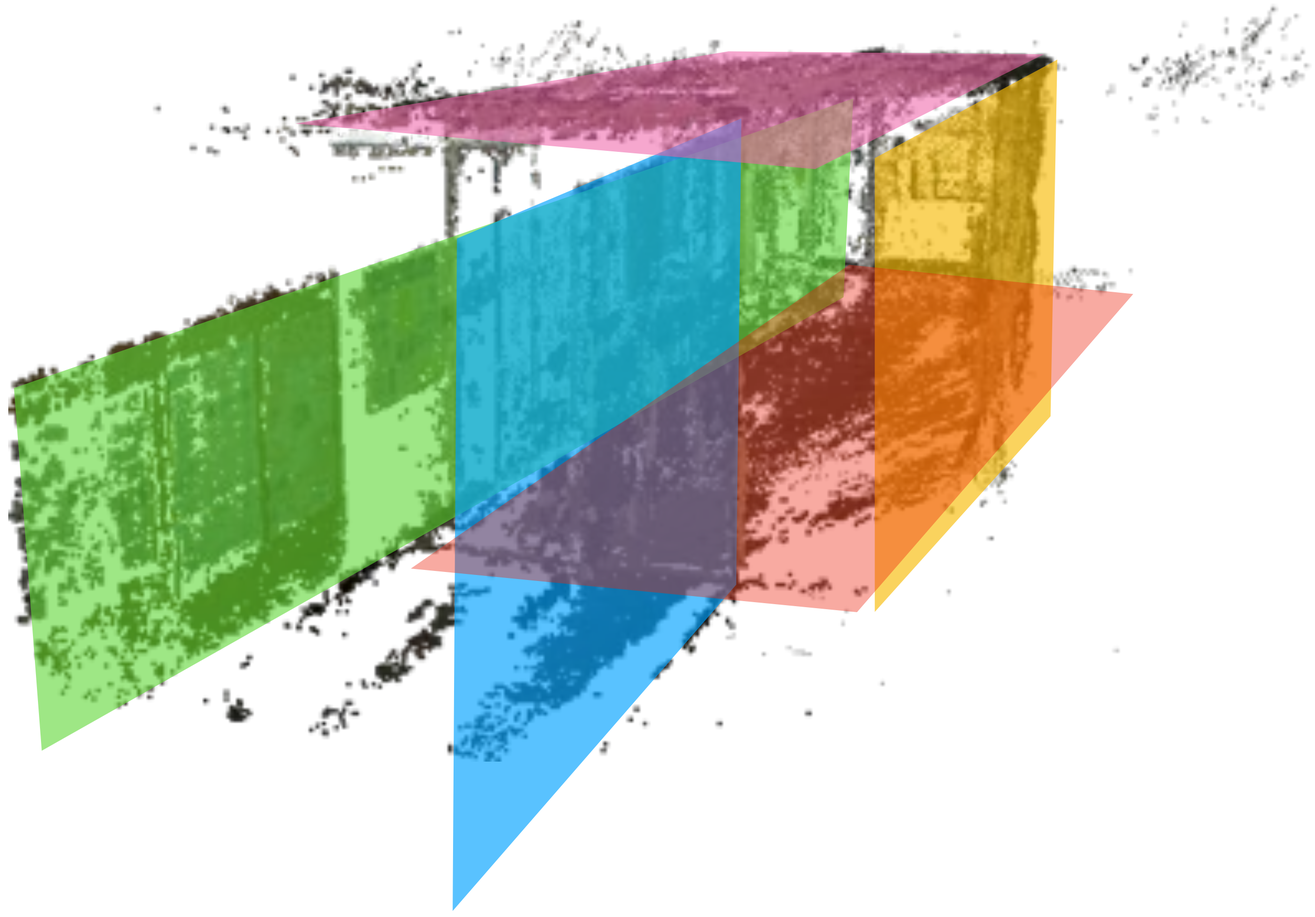
Wayfinding using inertial sensors

Accessible public transit



Tactile conversion of 3-D layouts

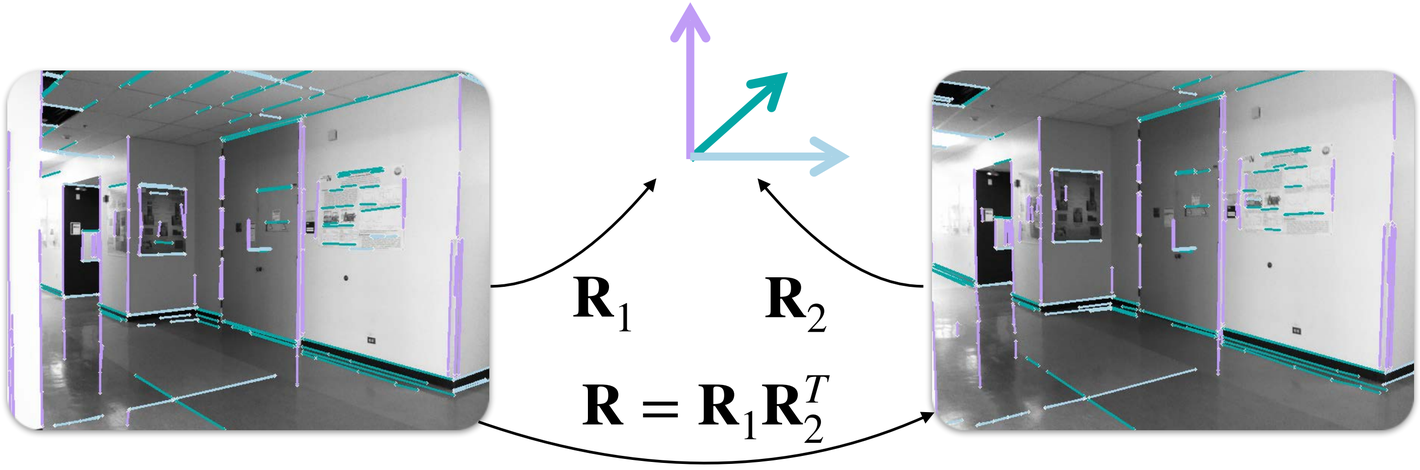


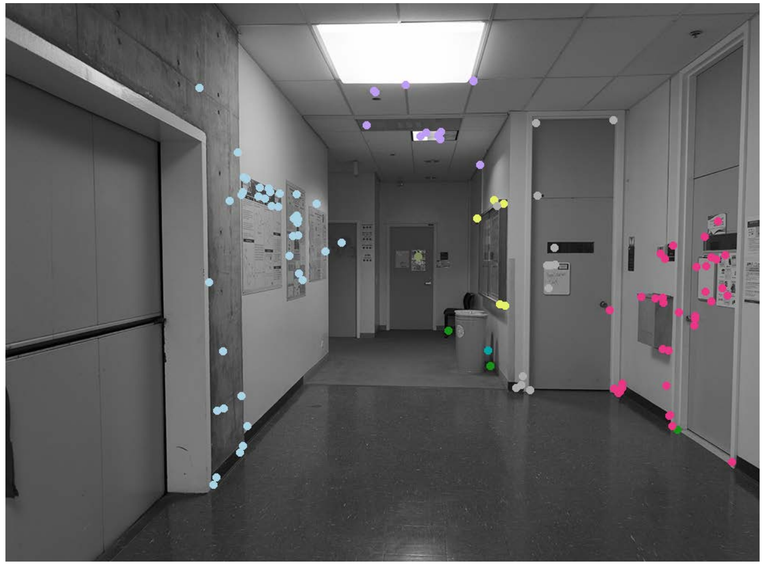


2-Frame Operations

Vanishing points detection

Canonical planes orientation
Relative camera rotation





Ground plane detection

MW-constrained homographic clustering

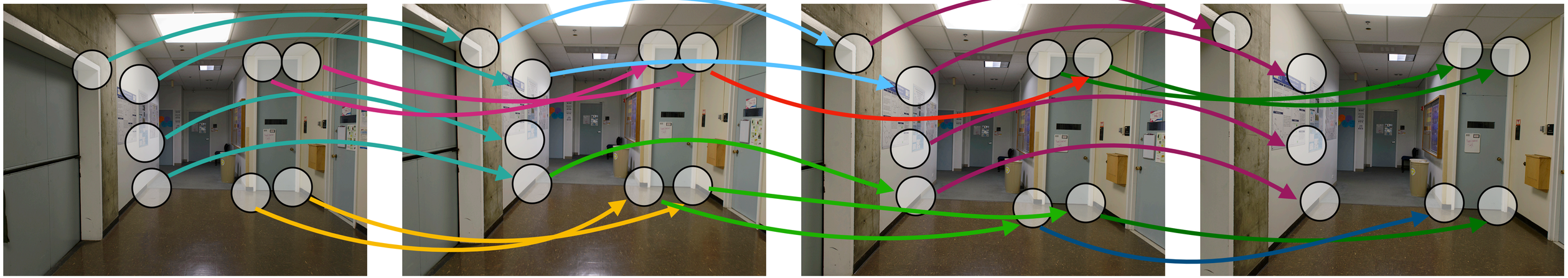
from vanishing points

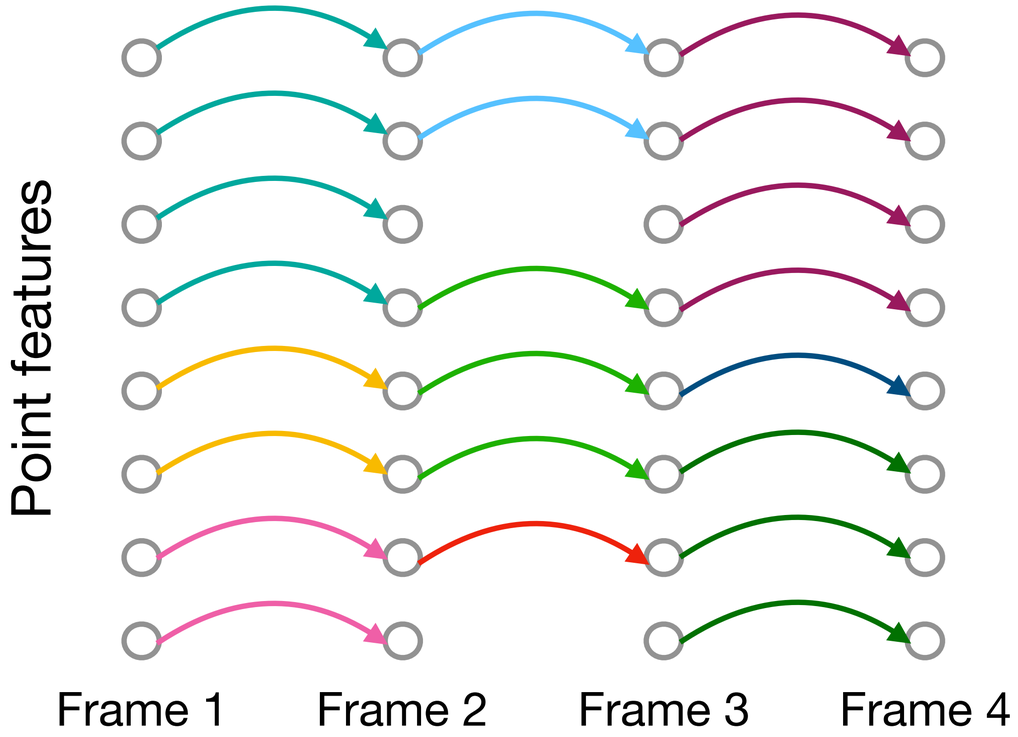
$$\mathbf{H} = \mathbf{K}\left(\mathbf{R} + \frac{\mathbf{t}}{d}\mathbf{n}^T\right)\mathbf{K}^{-1}$$

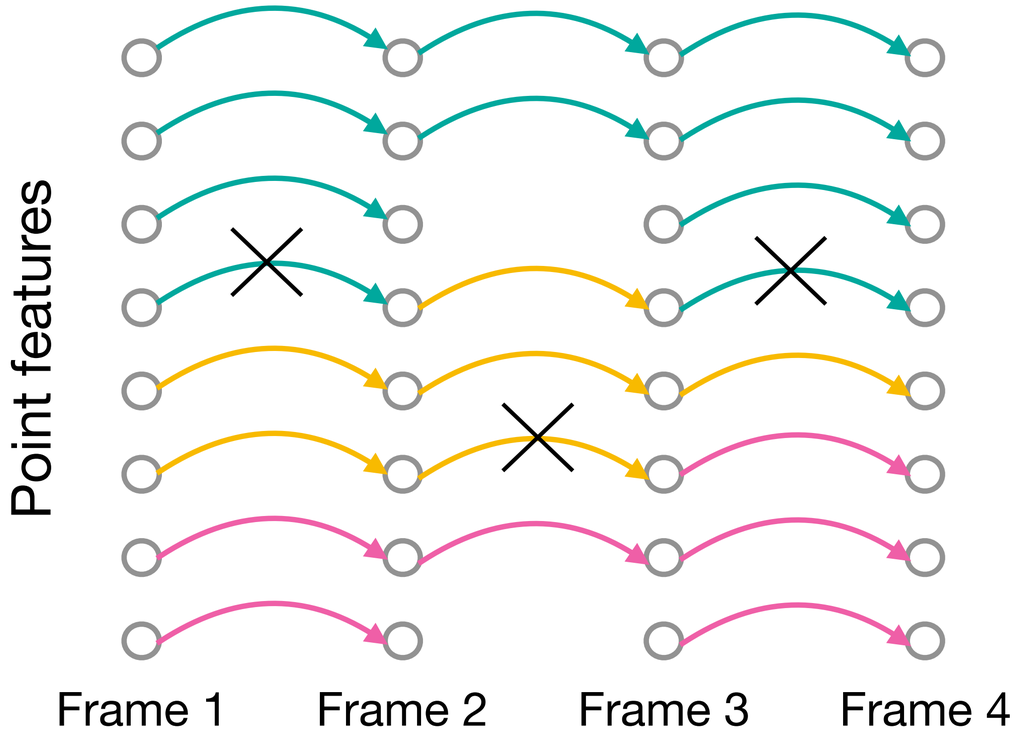
measured

from calibration

$\bar{\mathbf{t}}$ (3 degrees of freedom)

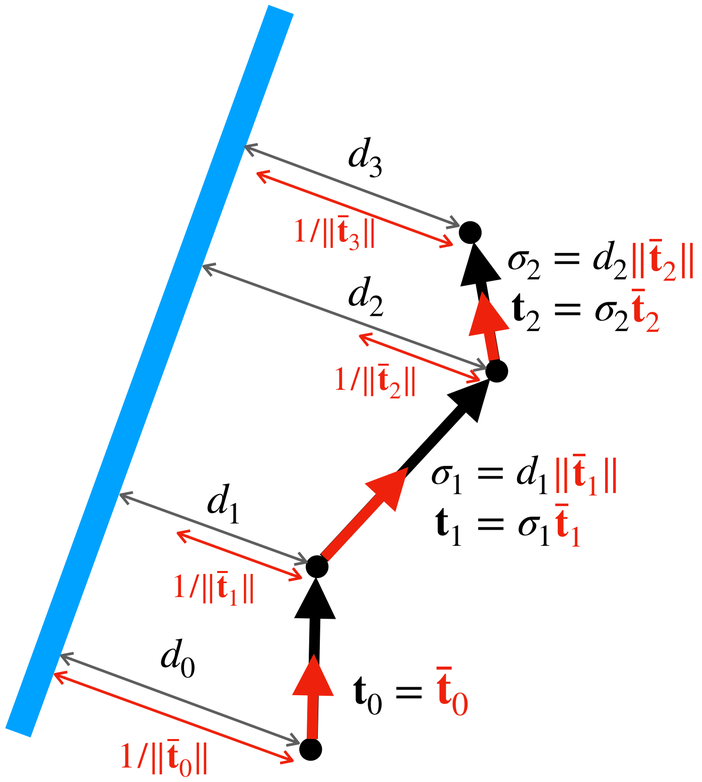



Point features
Frame 1
Frame 2
Frame 3
Frame 4


Point features
Frame 1
Frame 2
Frame 3
Frame 4

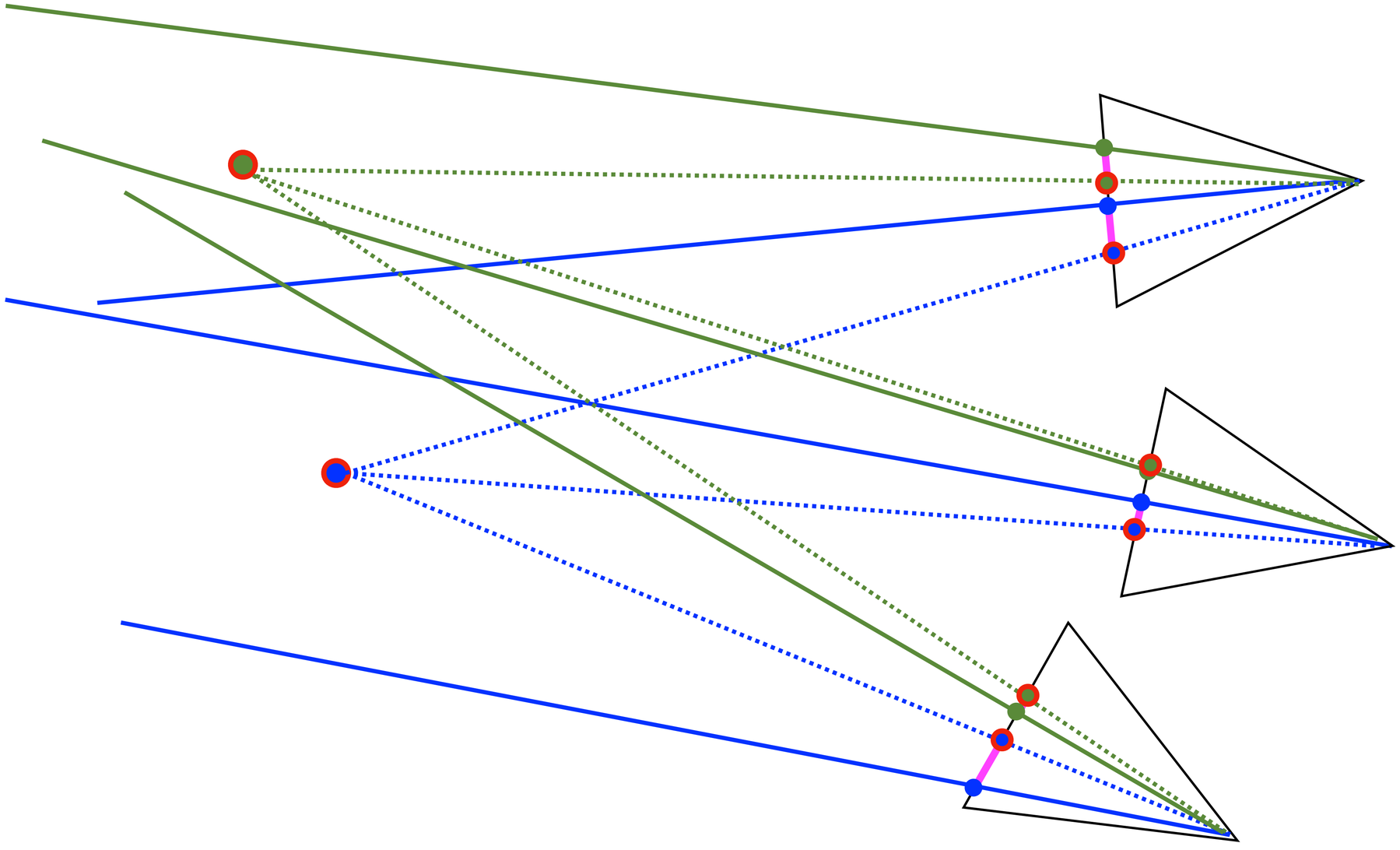
Multi-Frame Operations

Relative scale recovery



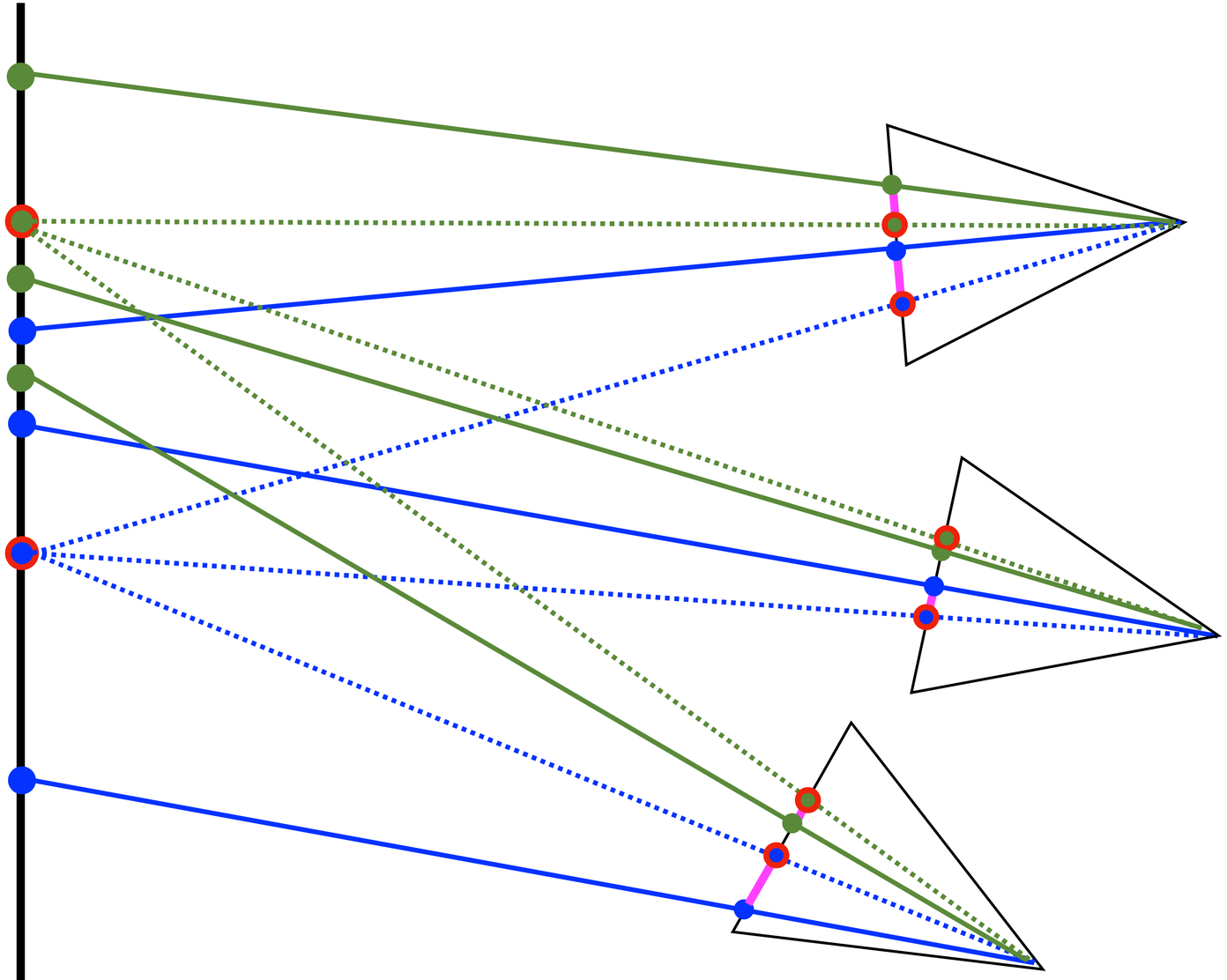
Standard Bundle Adjustment

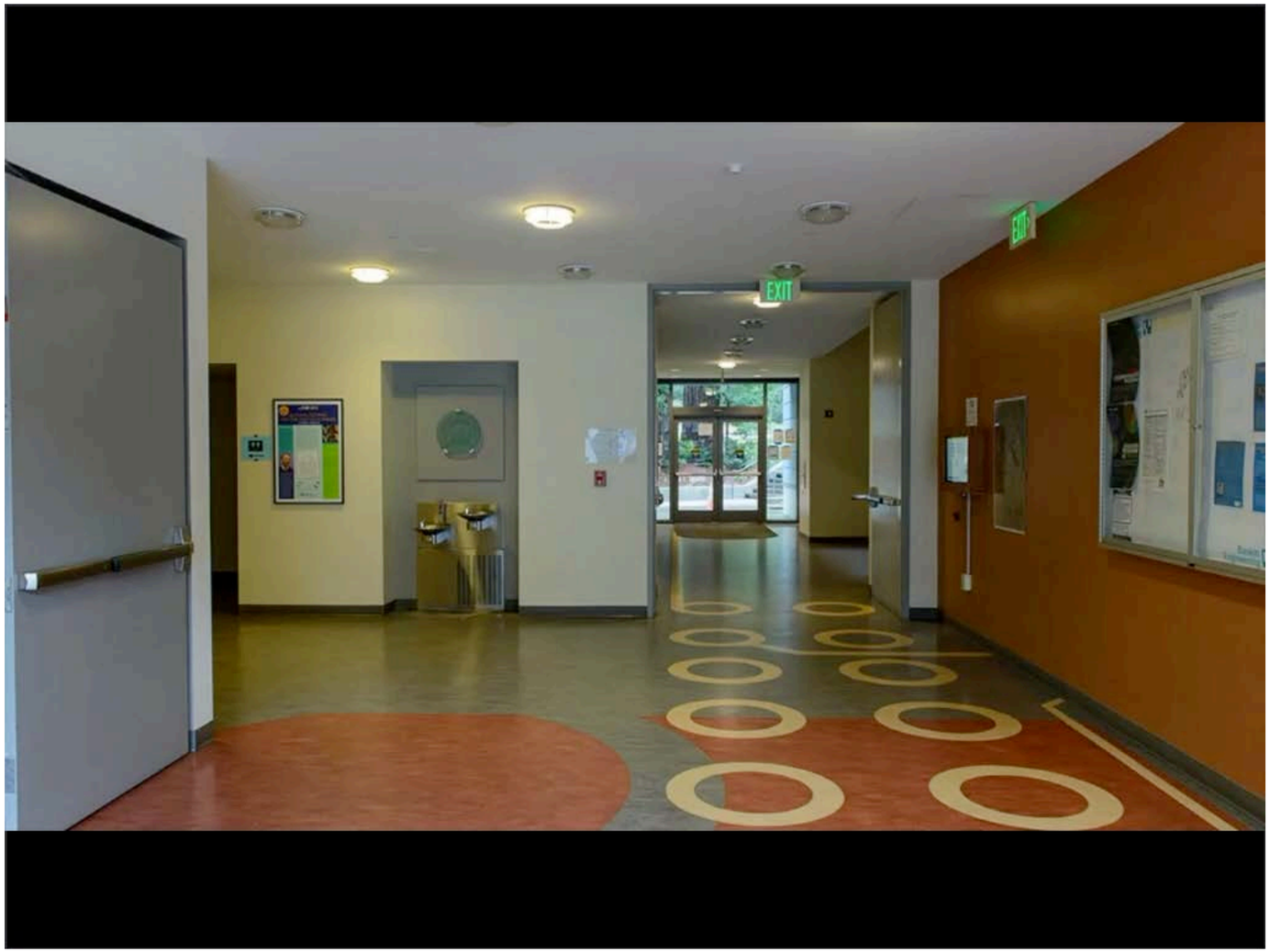
Optimizes for **camera poses** and **point locations**



Plane Bundle Adjustment

Optimizes for **camera poses** and **plane locations**





Processing time per frame

SIFT detection + matching	185 ms
Vanishing points detection	223 ms
CNN ground segmentation	22 ms
Multi-planar clustering	322 ms
Translation vector regression	20 ms
Cluster chain computation	30 ms
Bundle Adjustment (focal length fixed)	135 ms
Bundle Adjustment (focal length optimized)	500 ms
Total (focal length fixed)	**937 ms**
Total (focal length optimized)	**1,302 ms**

Acknowledgments

Awards: BES-0529435
CNS-0709472
PFI-BIC 1632158

Awards: R21 EY017003
R21 EY021643
R21 EY025077
R01 EY029033
R01 EY029260

Award: A16-0495