3D Reconstruction with Computer Vision
Meeting 11: Street View Structure from Motion

CS 378 Fall 2014, UT Austin, Bryan Klingner, 2 October

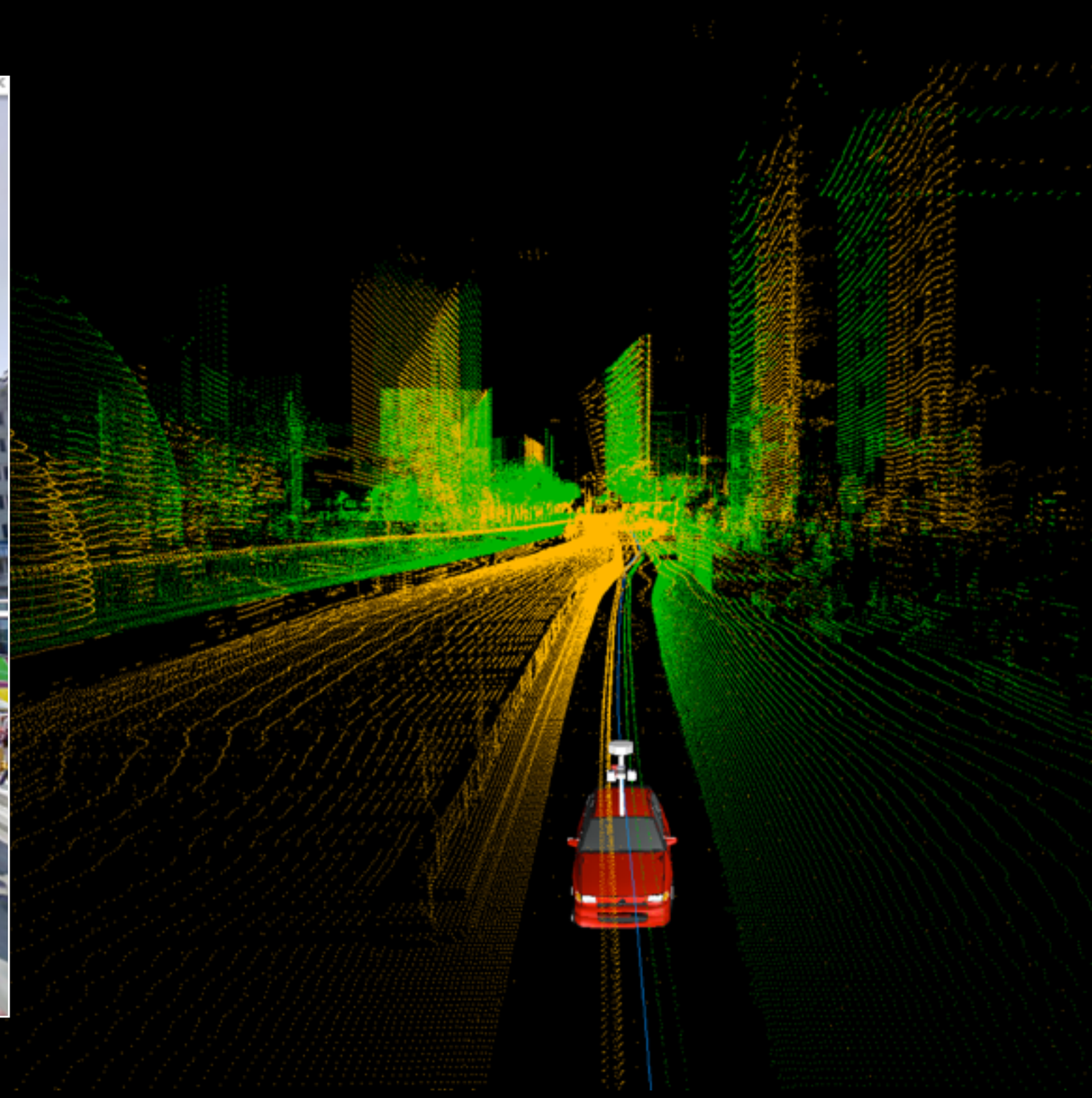
Our goal: Perfect Pose

San Francisco (Financial District)

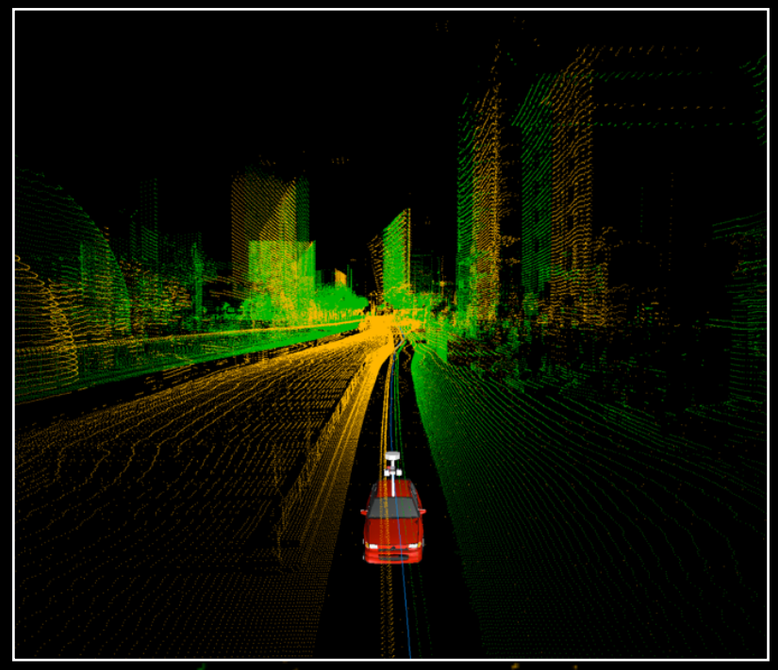
Our goal: Perfect Pose

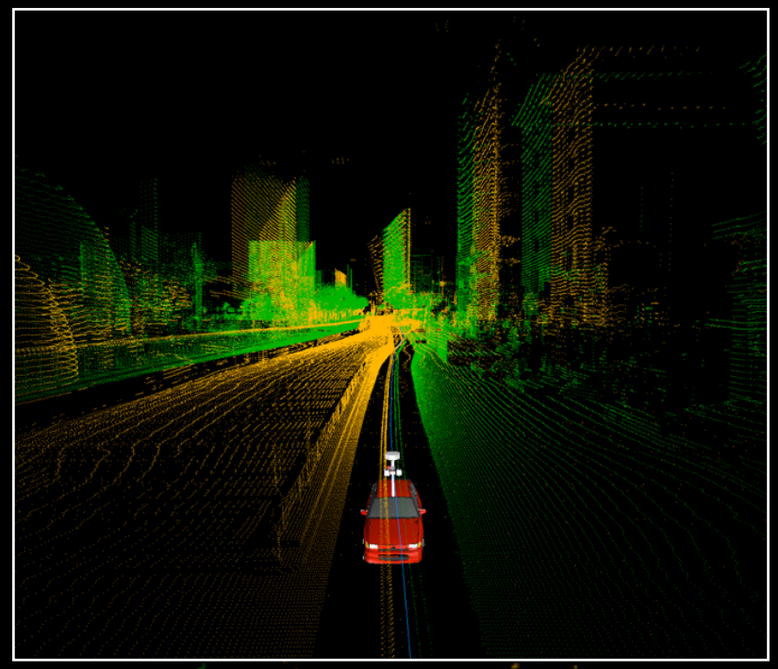
San Francisco (Financial District)

Example: one street in Hong Kong

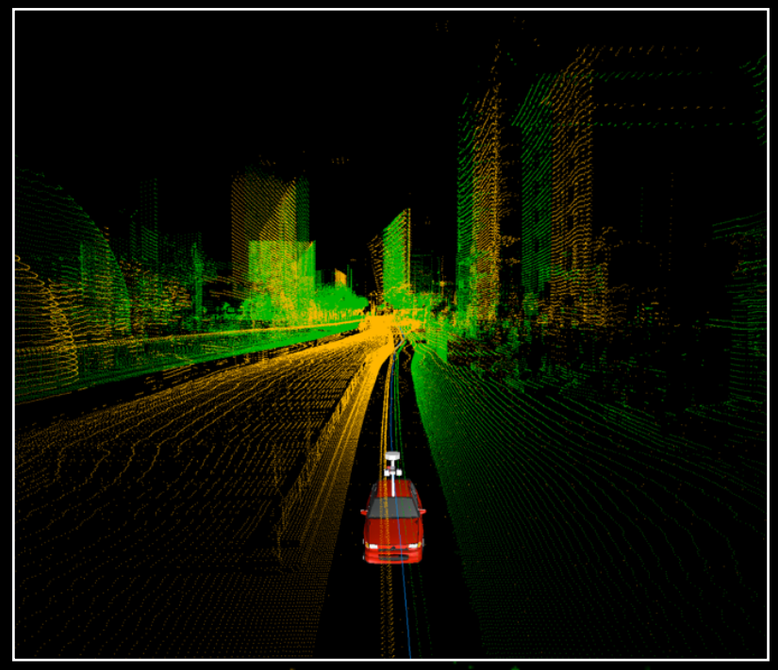



©2012 Google - Panoramio - Imagery Date: September 2009





Misalignments
between passes



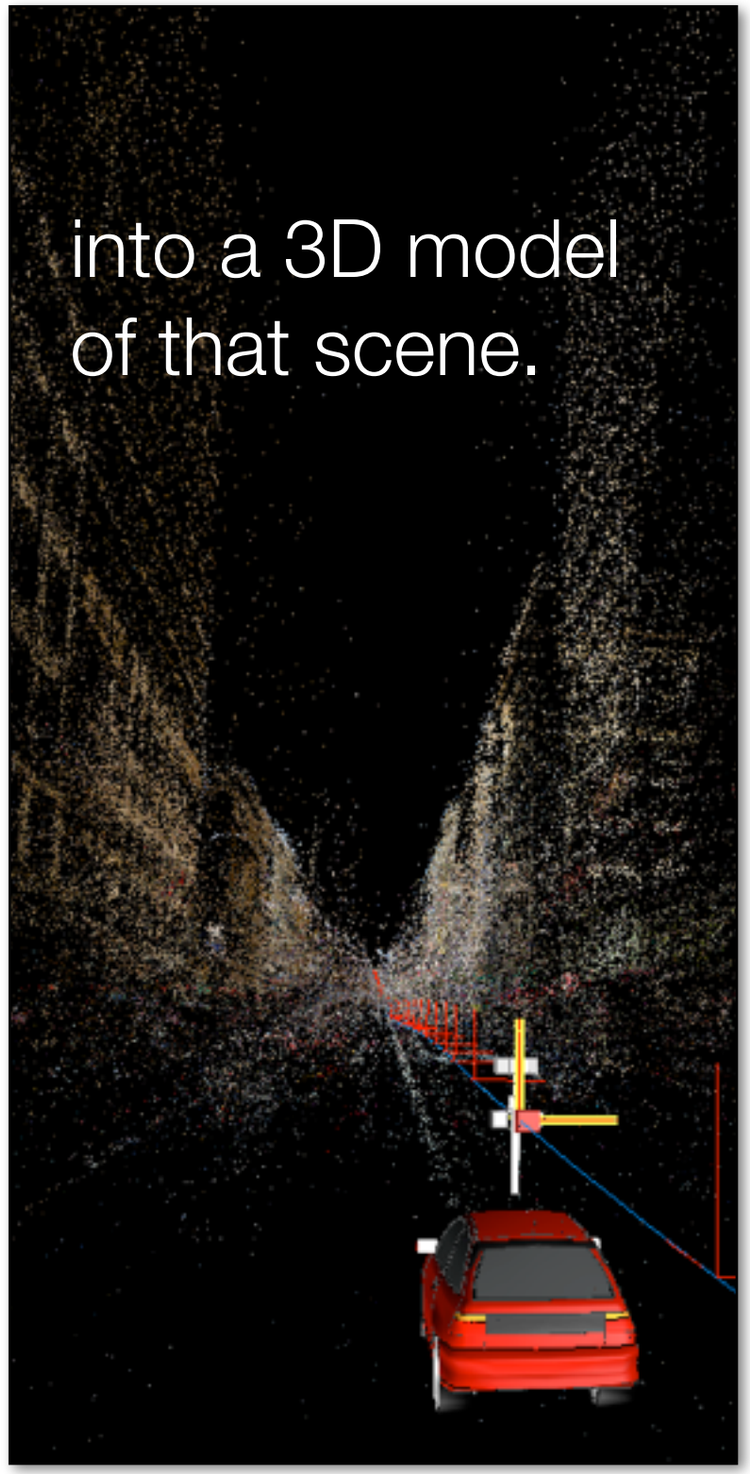
Improved alignment between passes

The tool: **Structure from Motion**

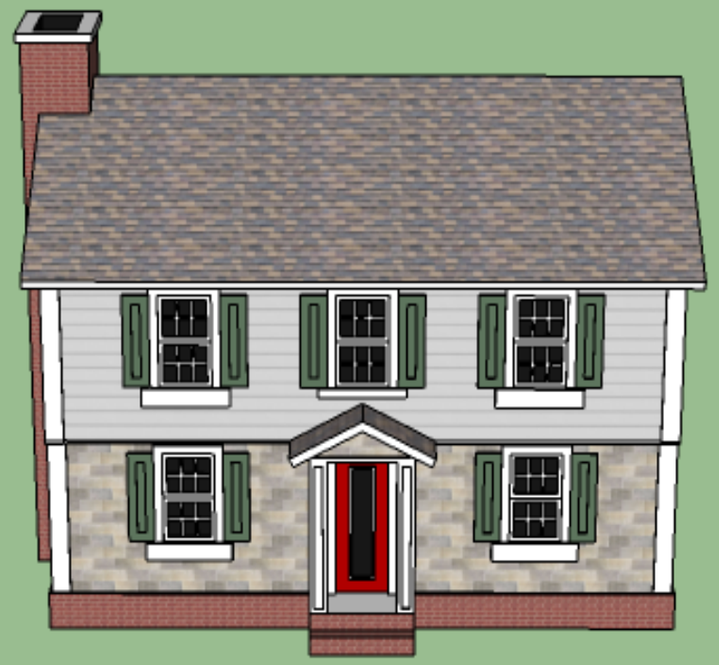
“Structure from motion”

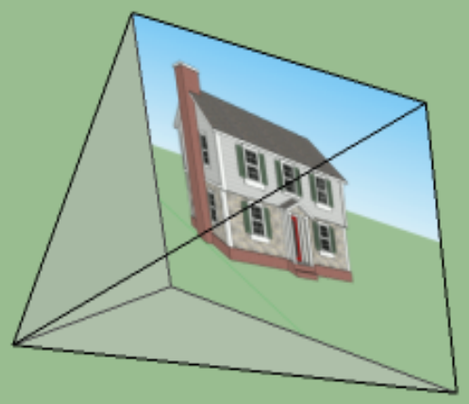
Transforms many images of the same scene

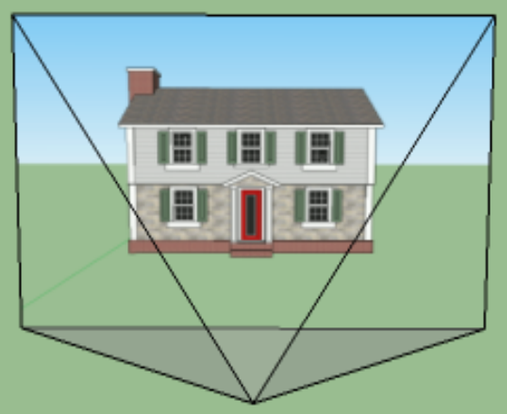
into a 3D model of that scene.

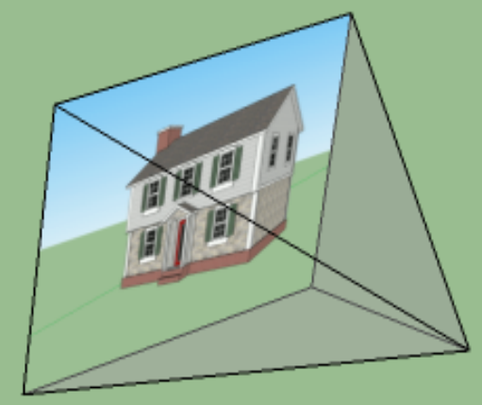


Structure from motion

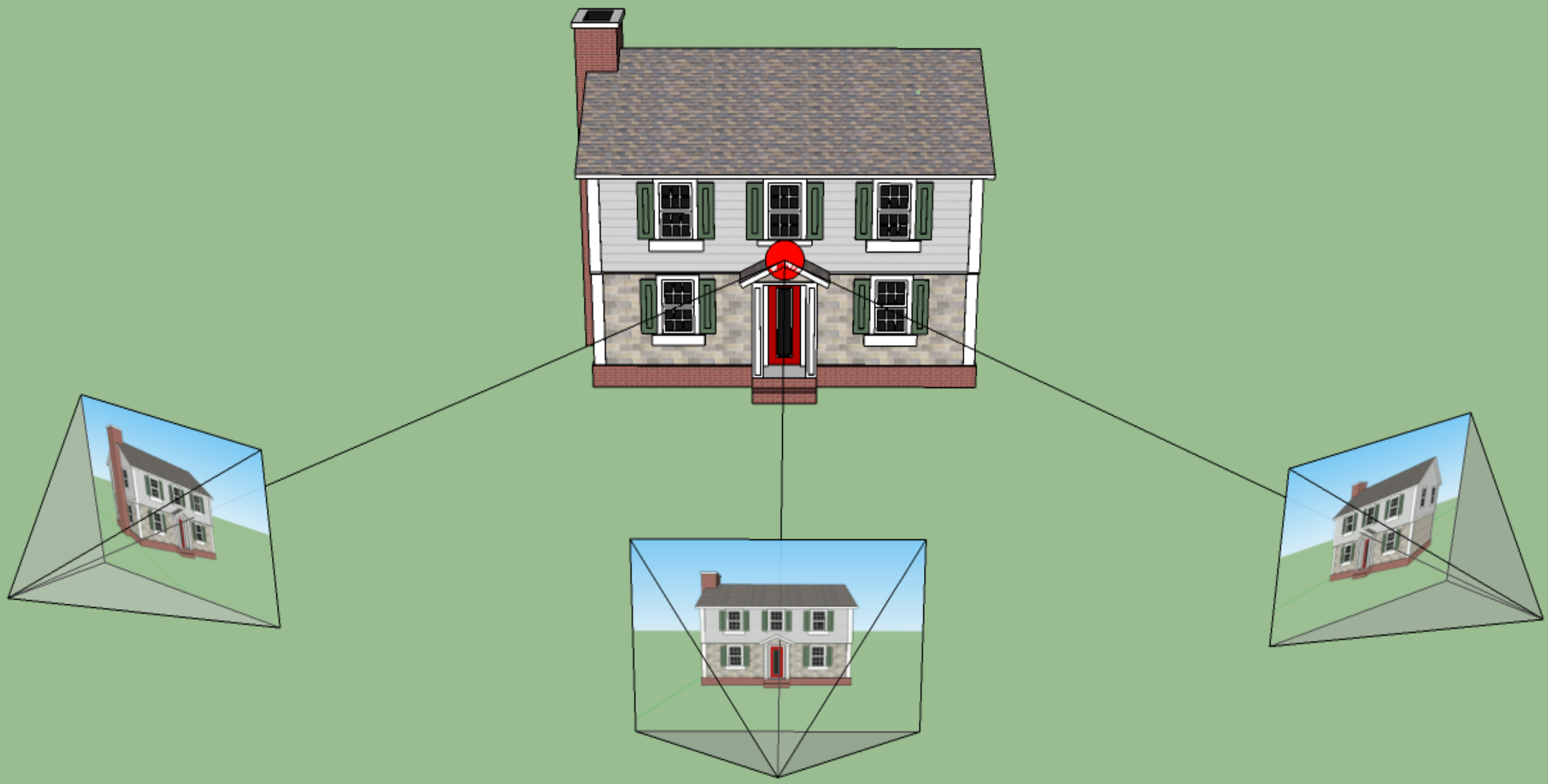




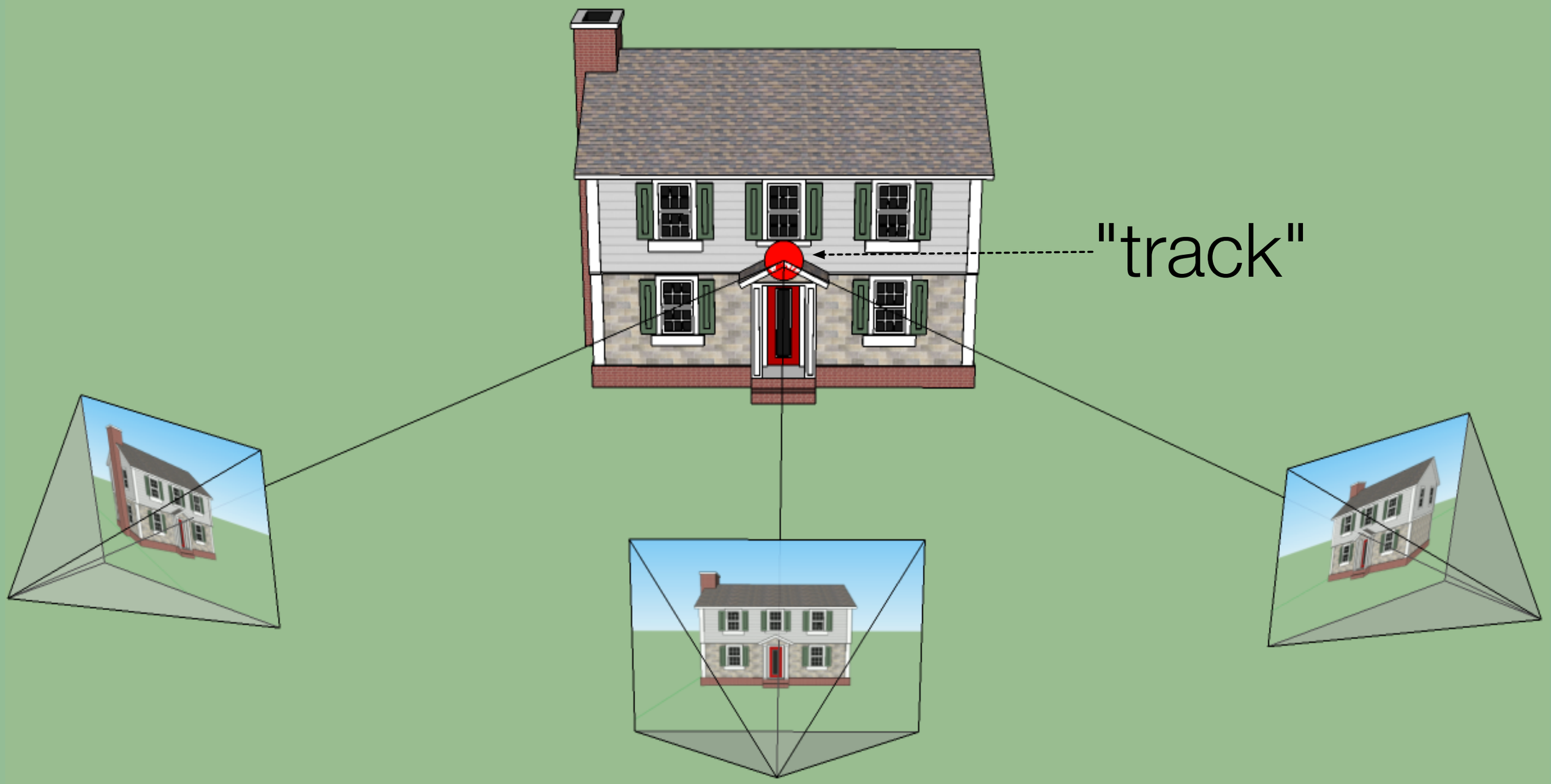




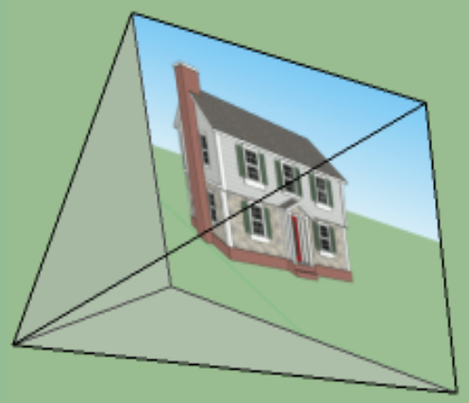
Structure from motion

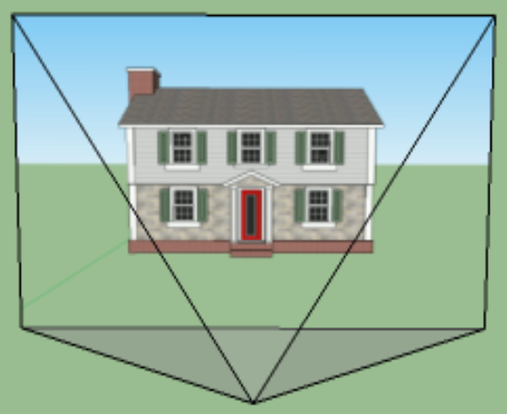


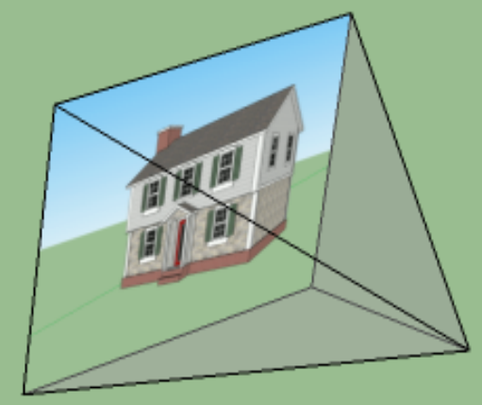
Structure from motion



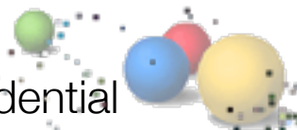
Structure from motion

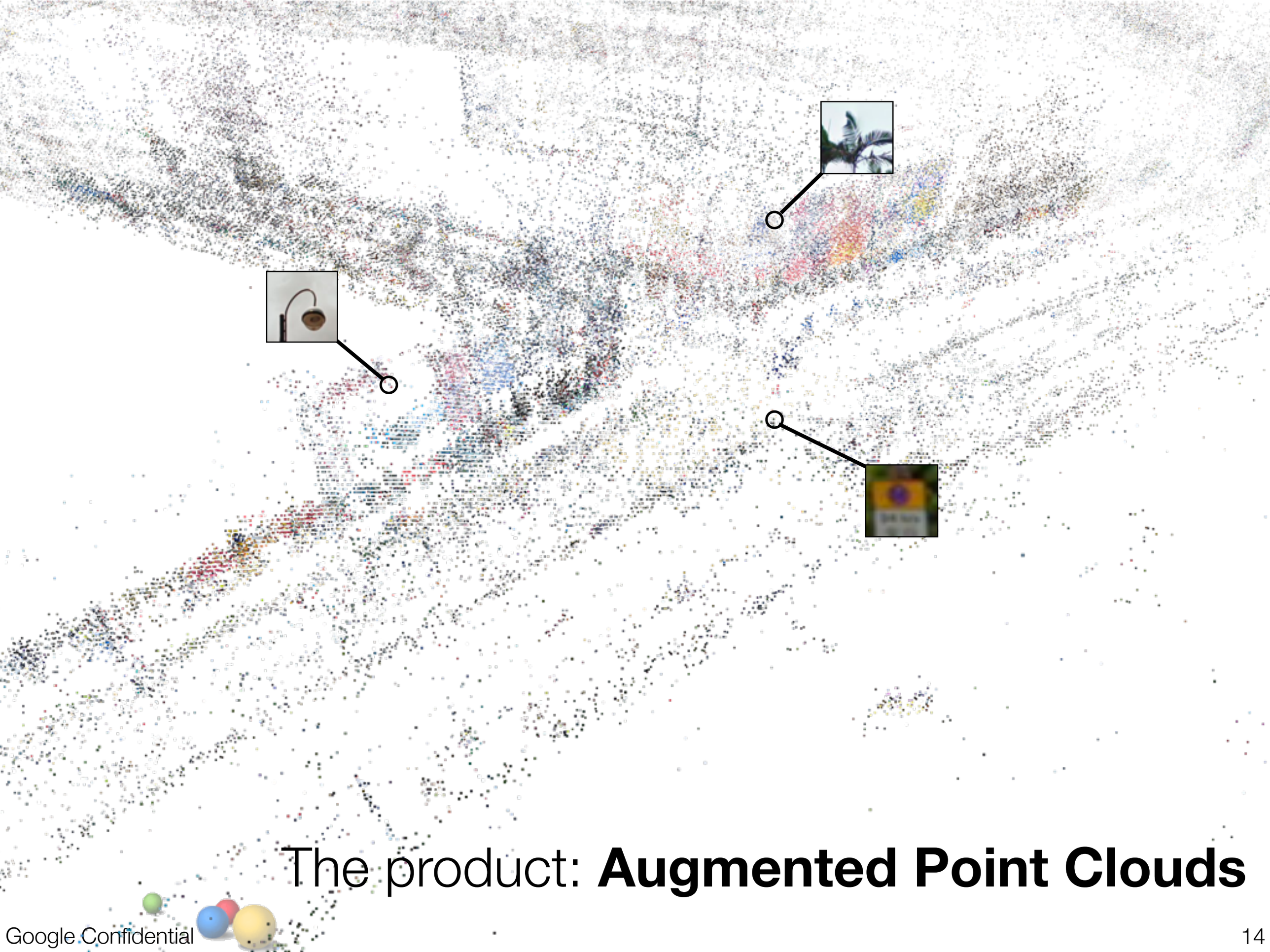






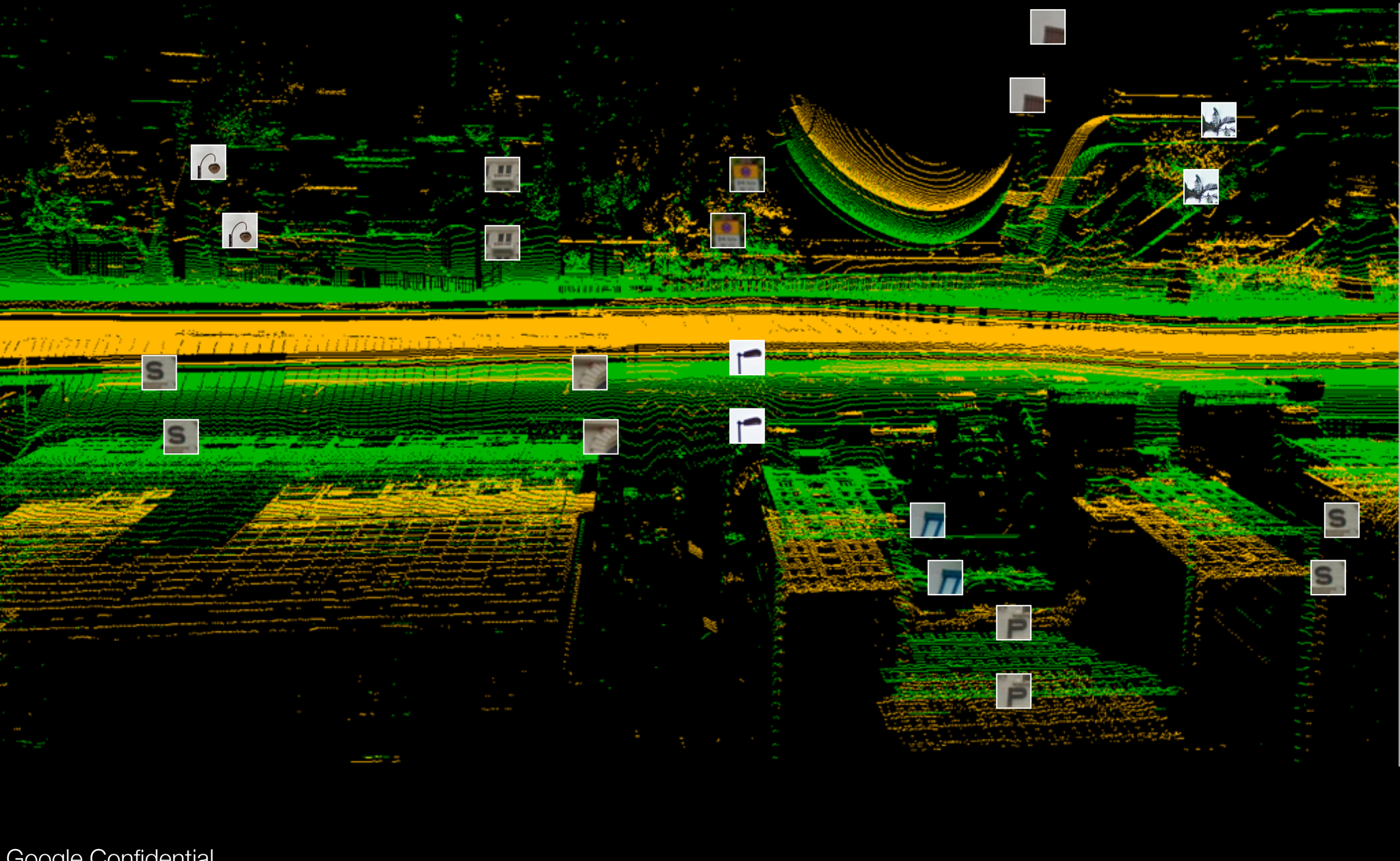
The product: **Augmented Point Clouds**



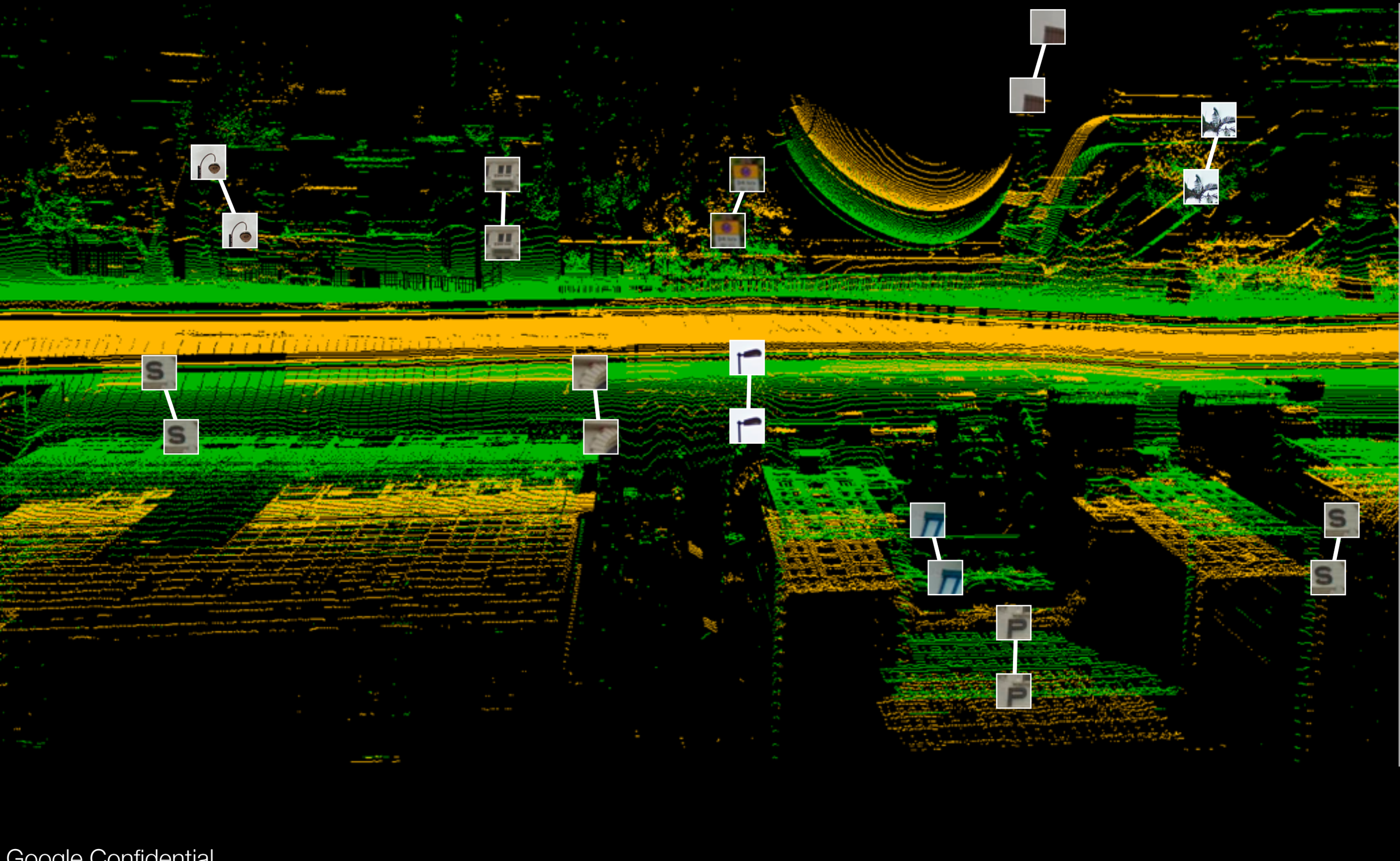


The product: **Augmented Point Clouds**

SfM “**augmented points**” with appearance and location



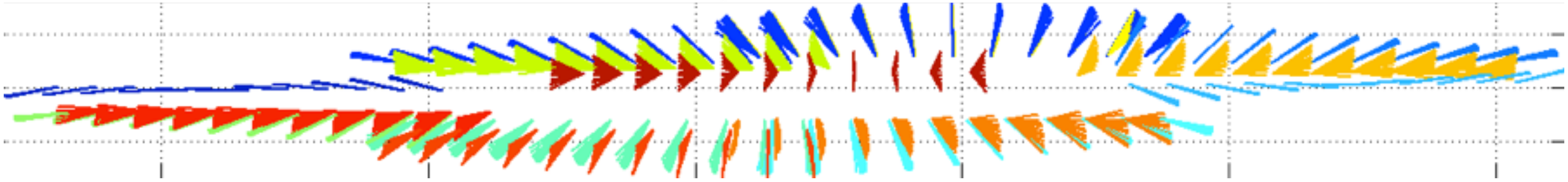
Find **correspondences** with **appearance**



Align with **location**

Two key challenges:

Rolling shutters.

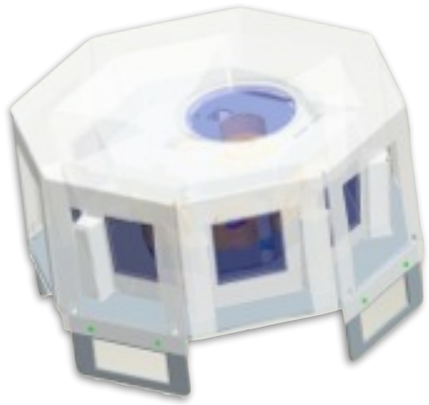


Google scale. **10B** panos, **1e17** pixels

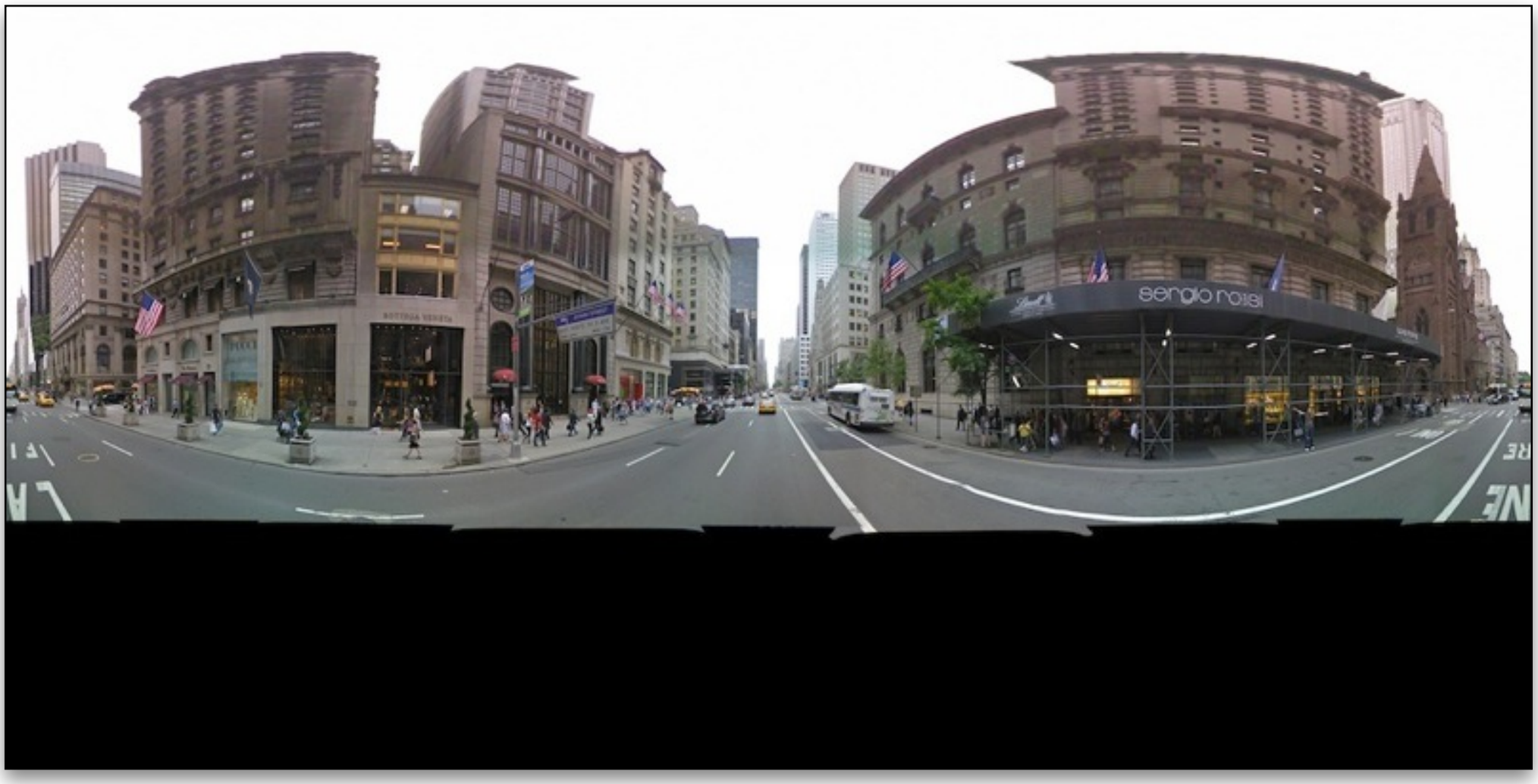
Raw materials: Imagery

Ladybug2
6 Cameras
4.7 Megapixels

Raw materials: Imagery

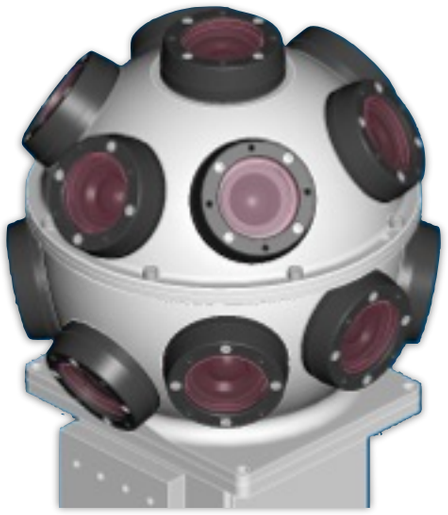


R5
9 Cameras
43 Megapixels



Raw materials: Imagery

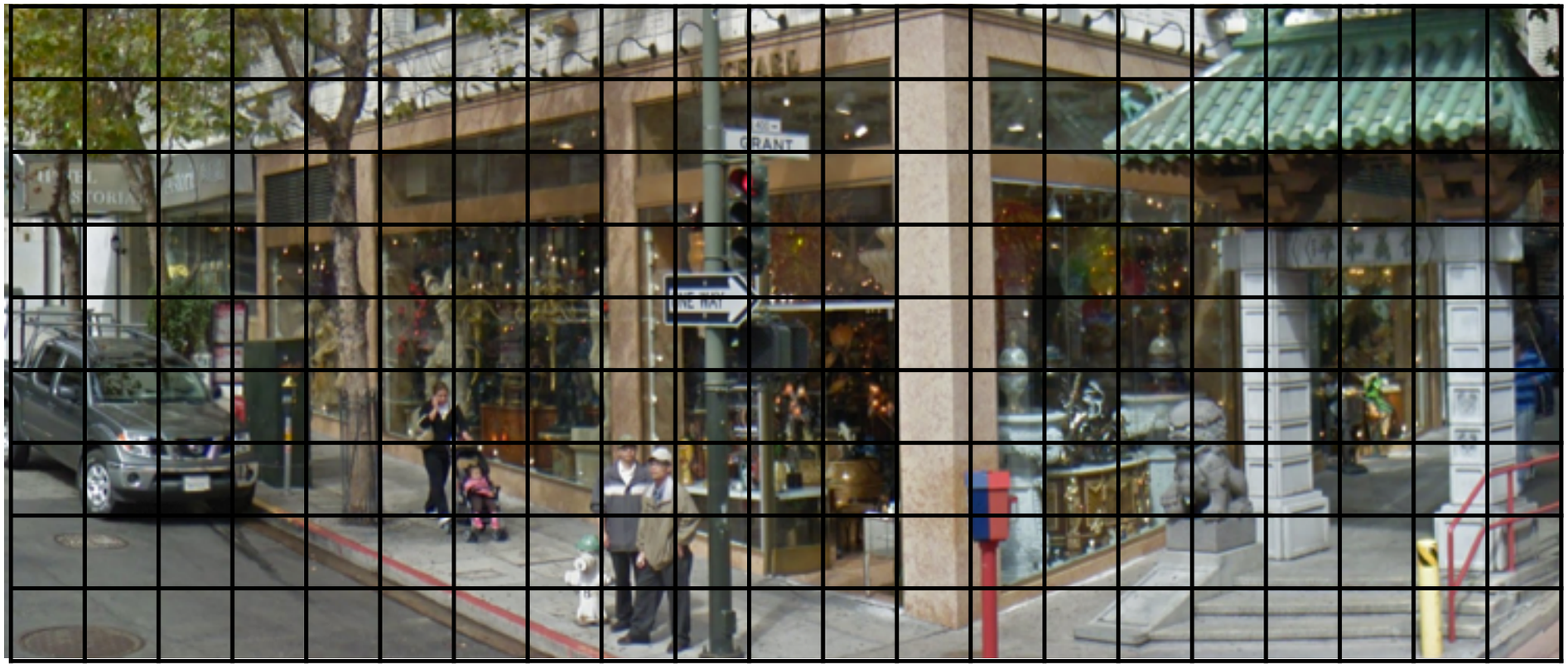
R7
15 Cameras
75 Megapixels



R5 and R7 cameras use **rolling shutters**

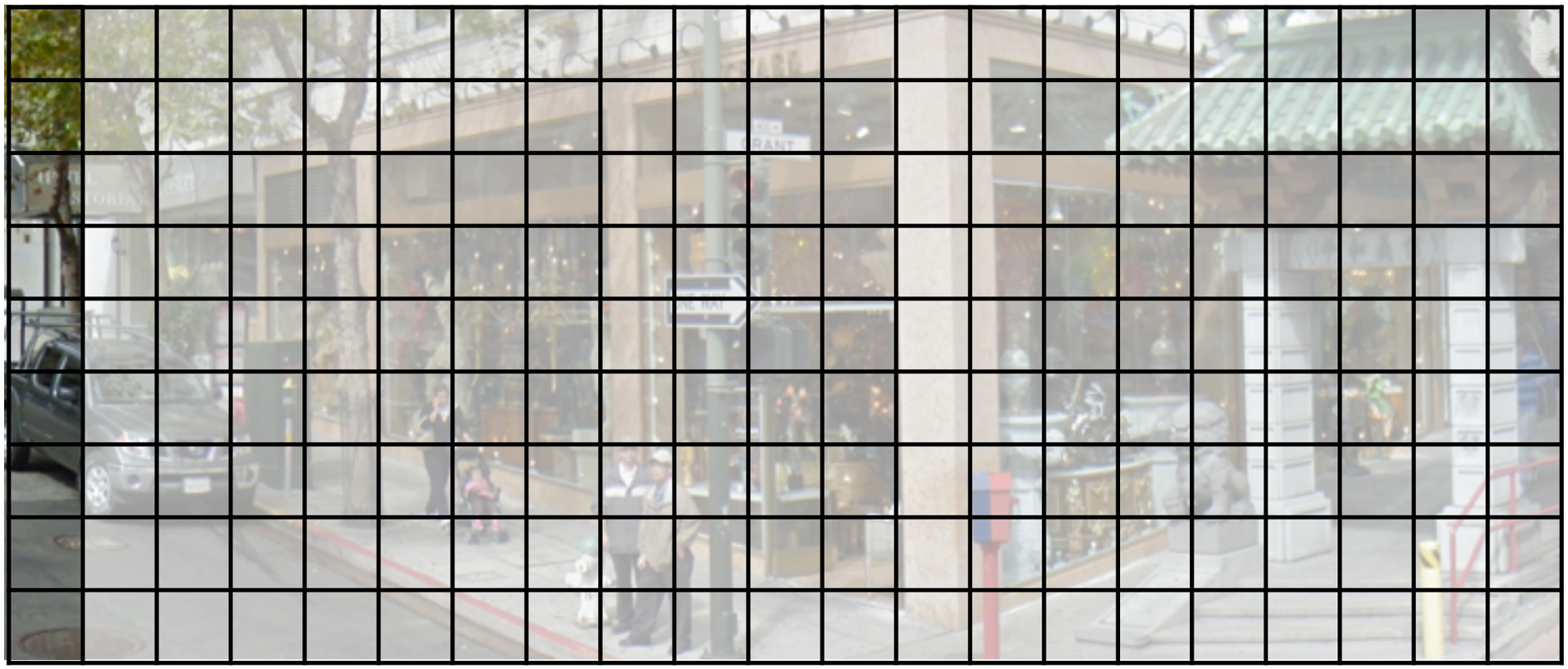
courtesy: Wikipedia, *Rolling Shutter.*

Instant shutters



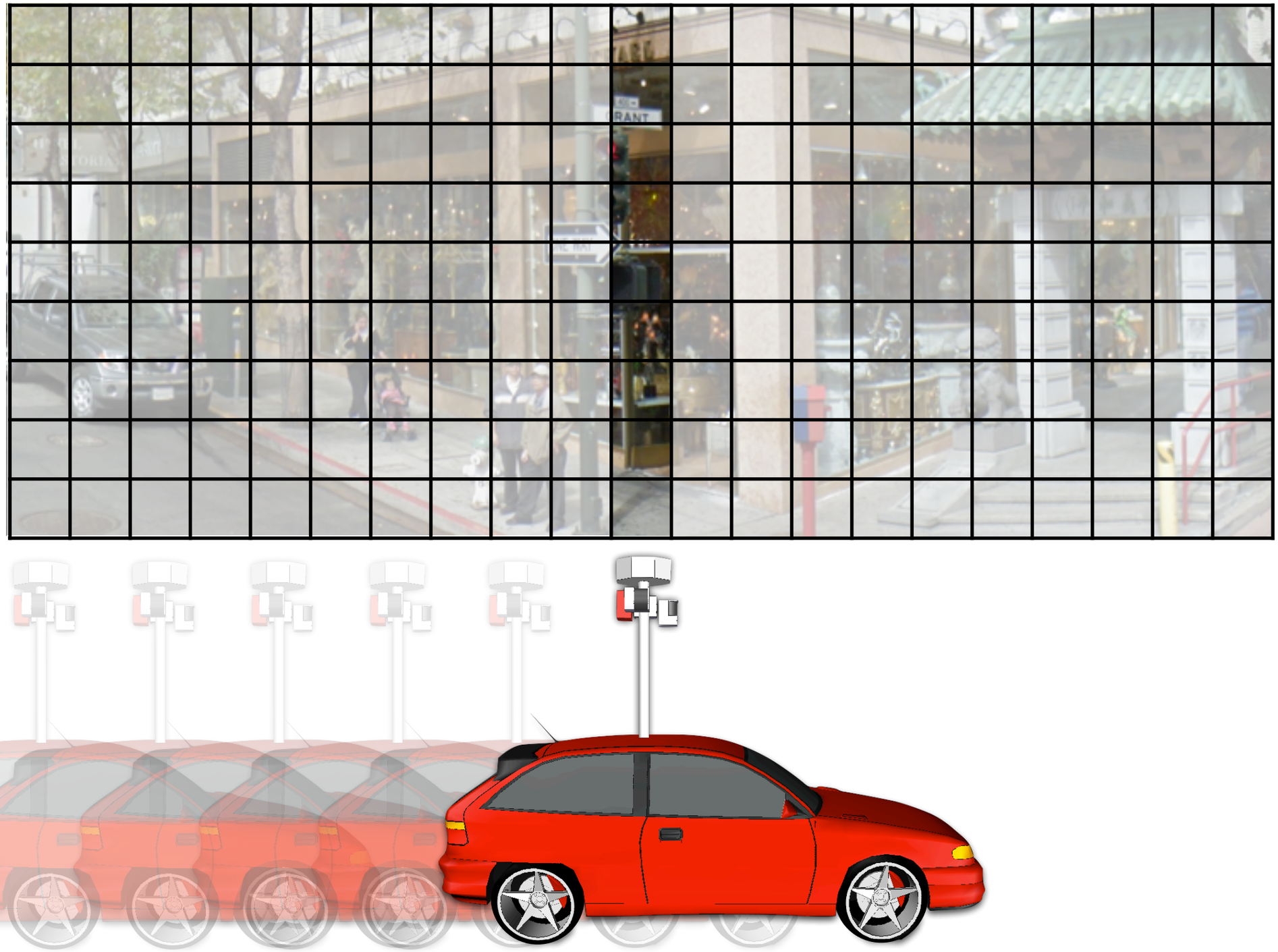
all pixels captured simultaneously

Rolling shutters



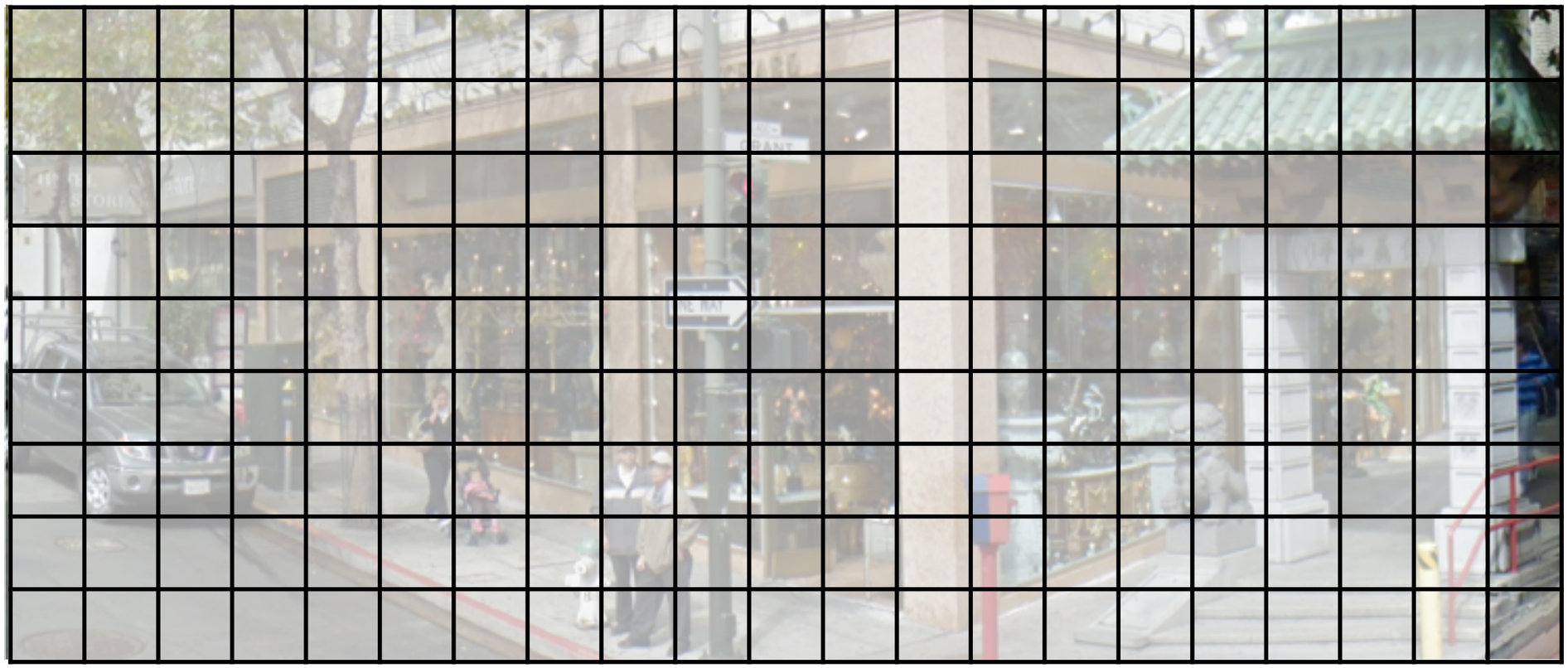
exposure starts

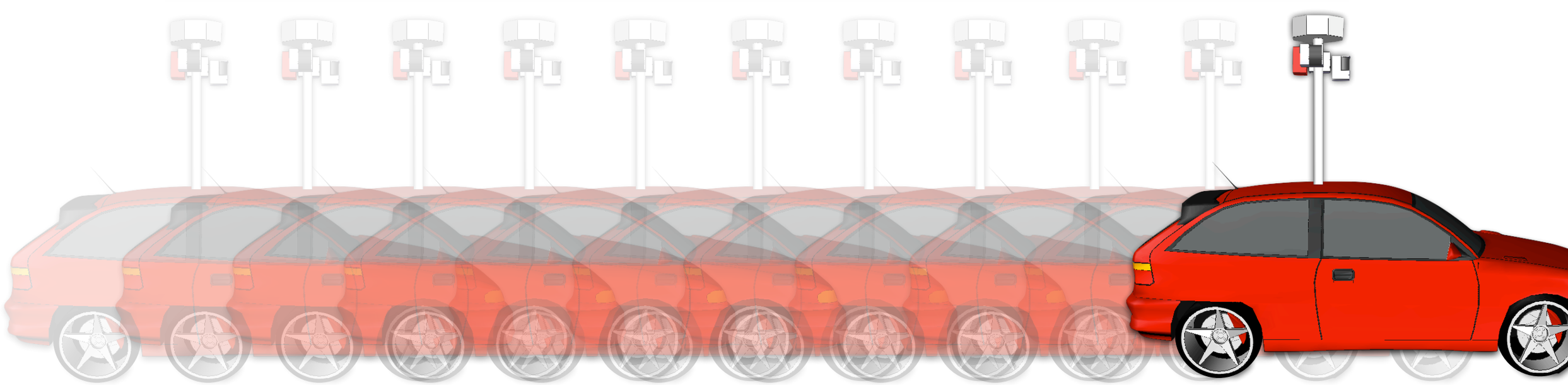
Rolling shutters



each column of pixels captured at a different time

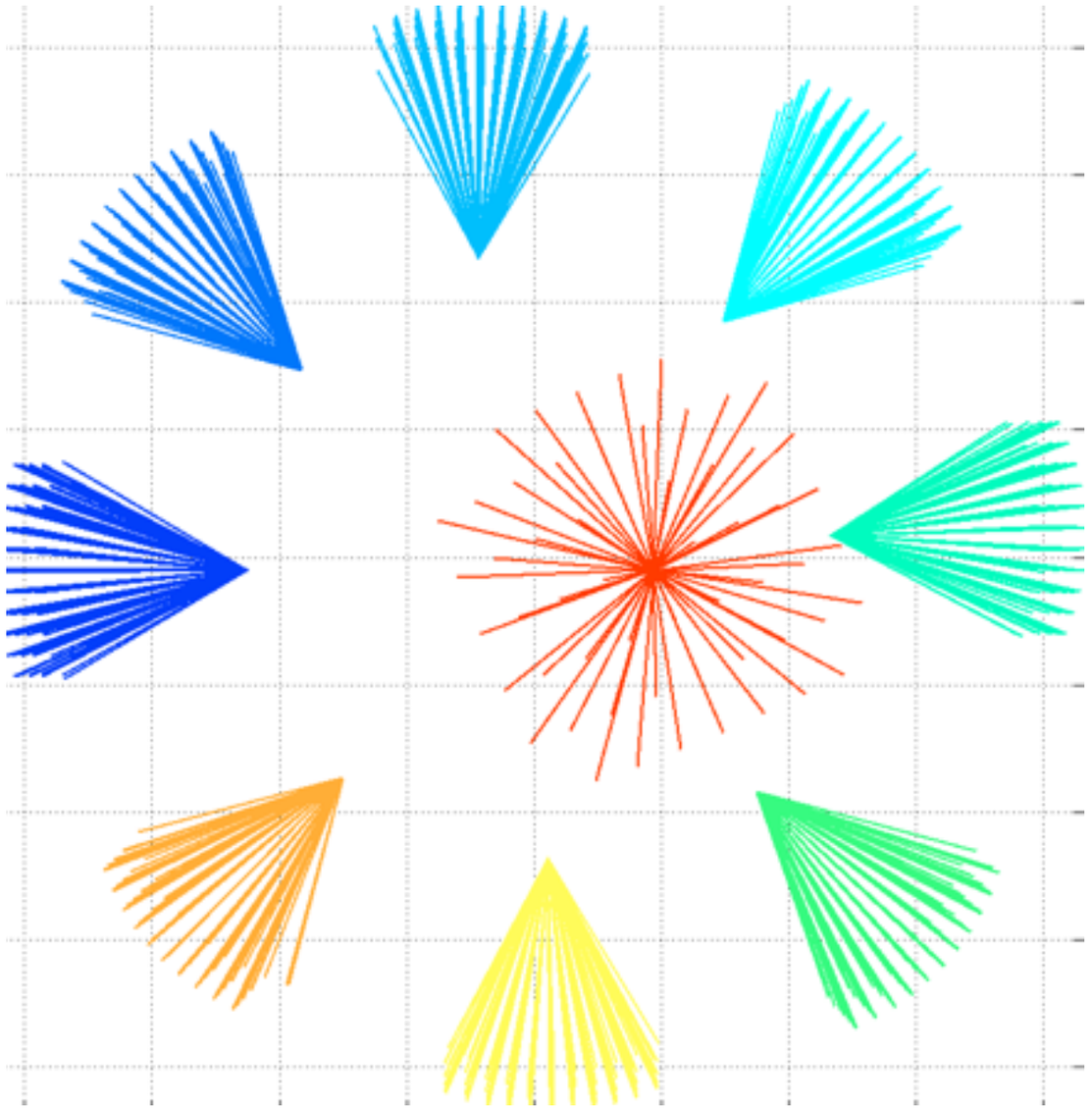
Rolling shutters





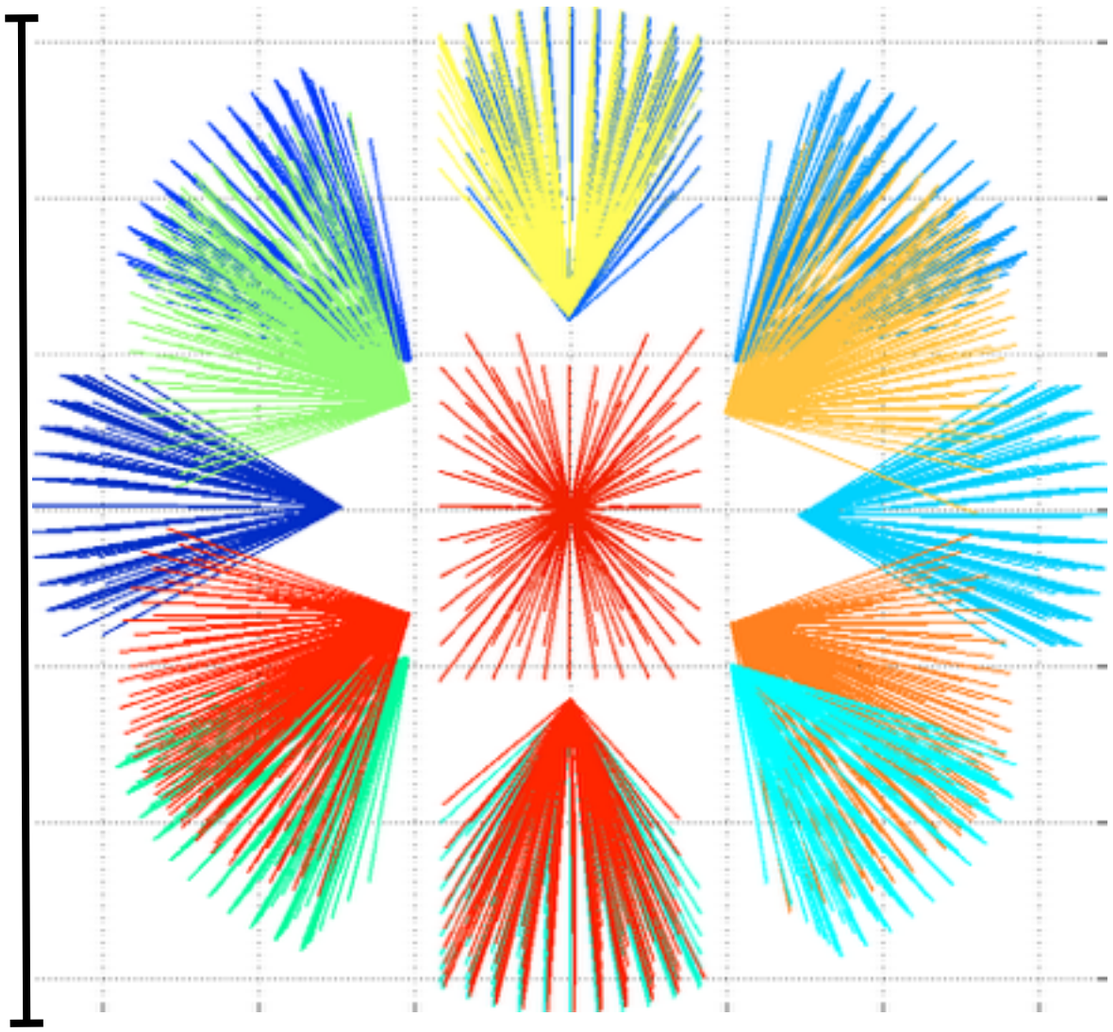
at the end of the exposure, ~300ms have elapsed
and the car may have traveled several meters.

Camera rays if we had instant shutters

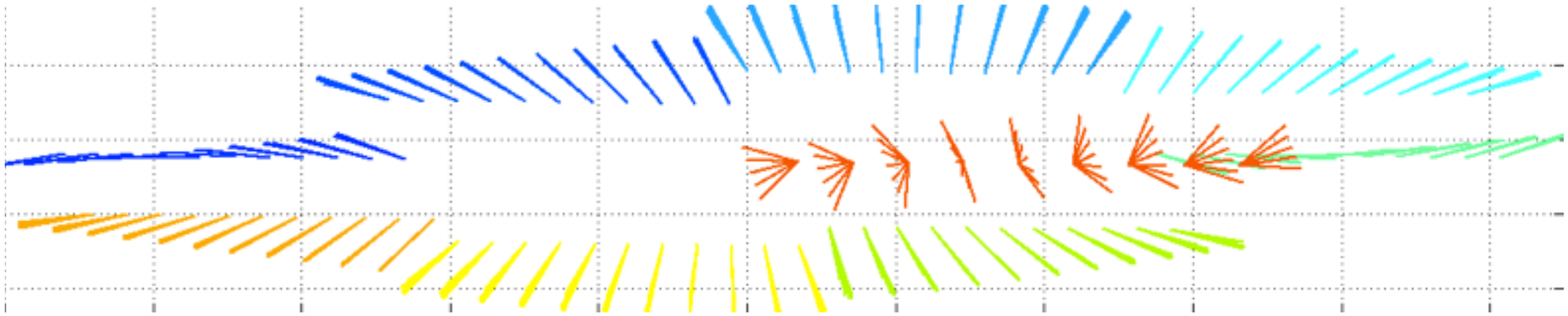


R5 (9 cameras)
45cm

R7 (15 cameras)
30cm

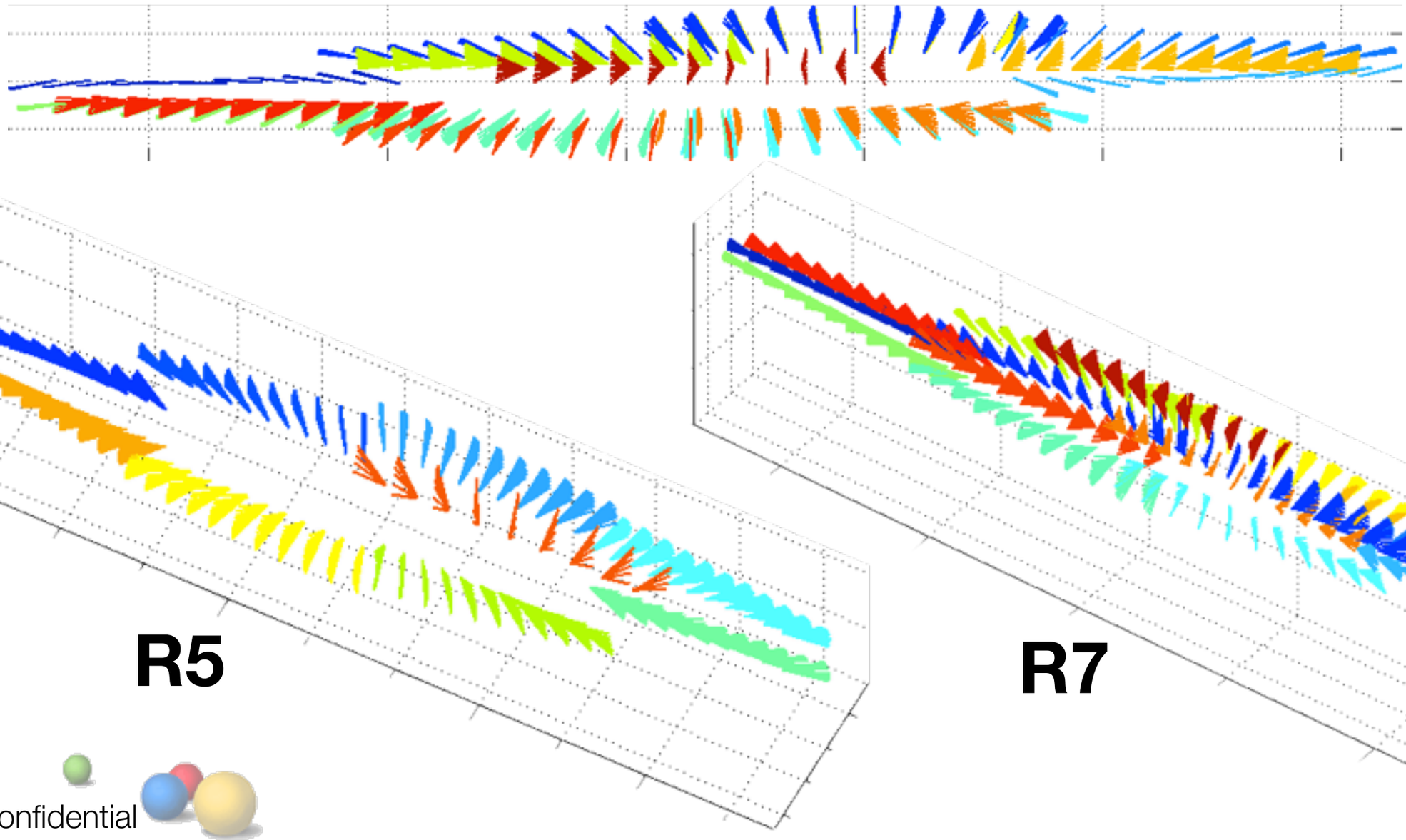


Camera rays from rolling shutters



R5

> 3m

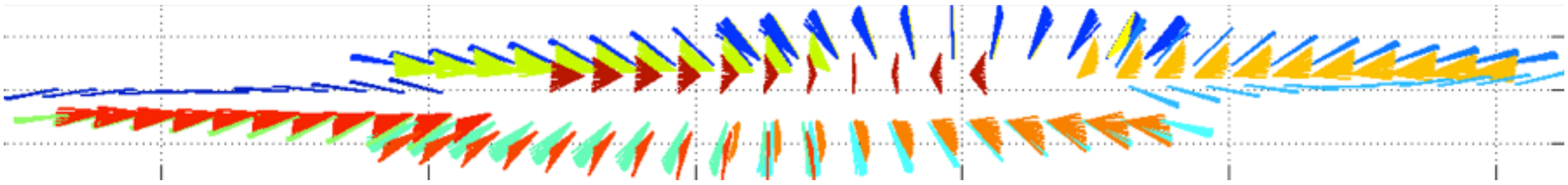


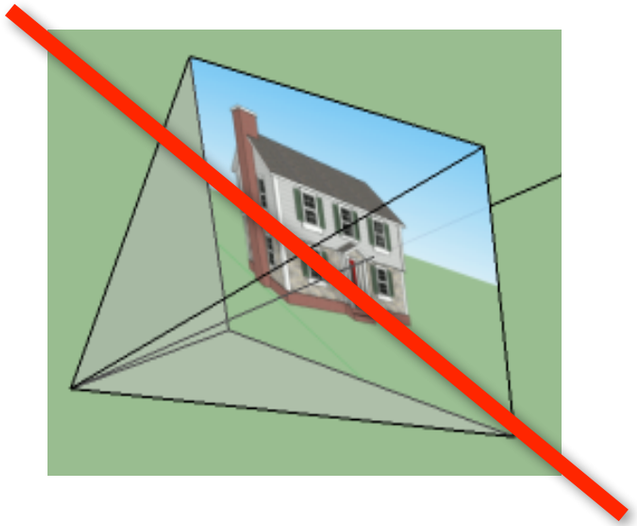
R7

R5

R7

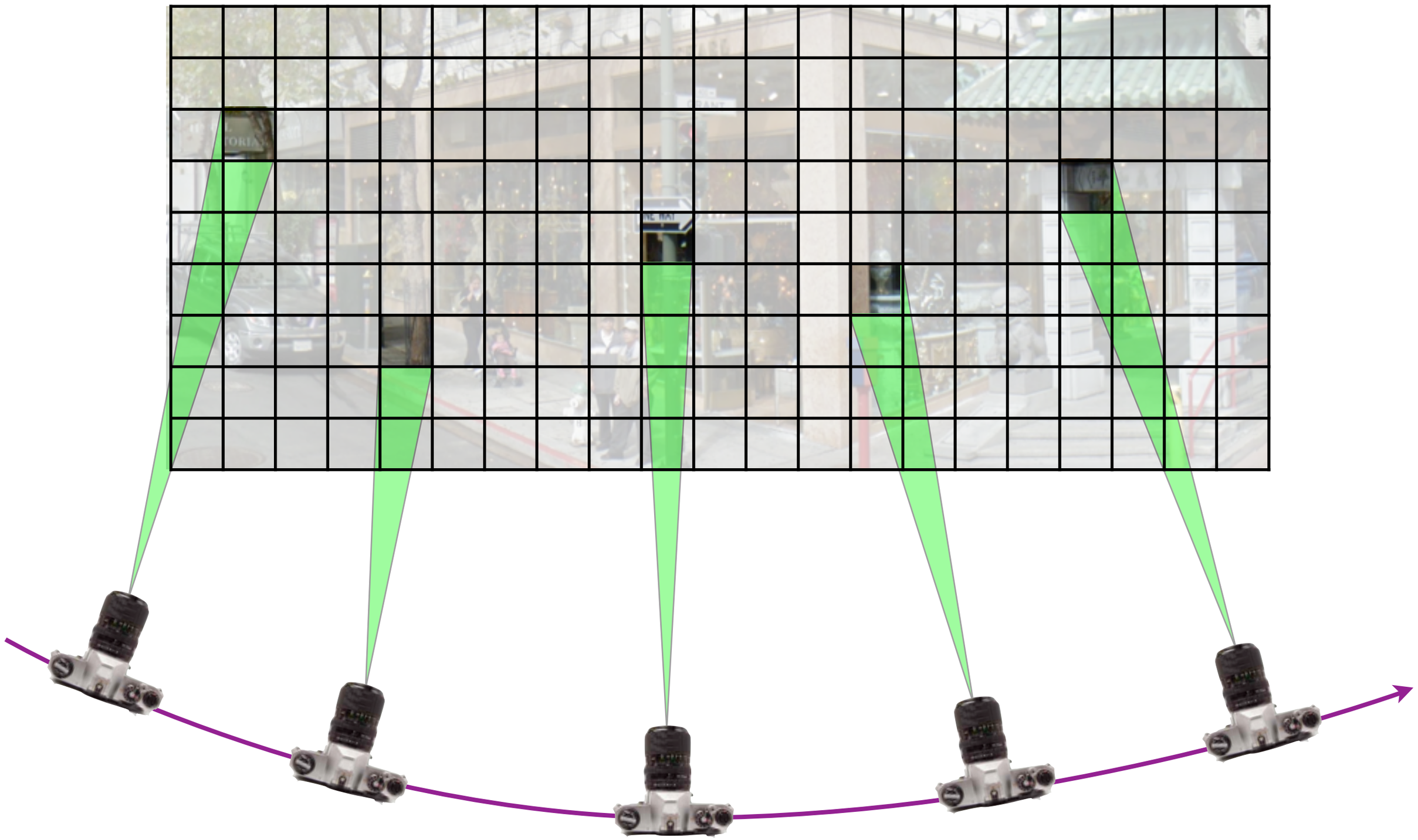
Rolling shutters break normal SFM!





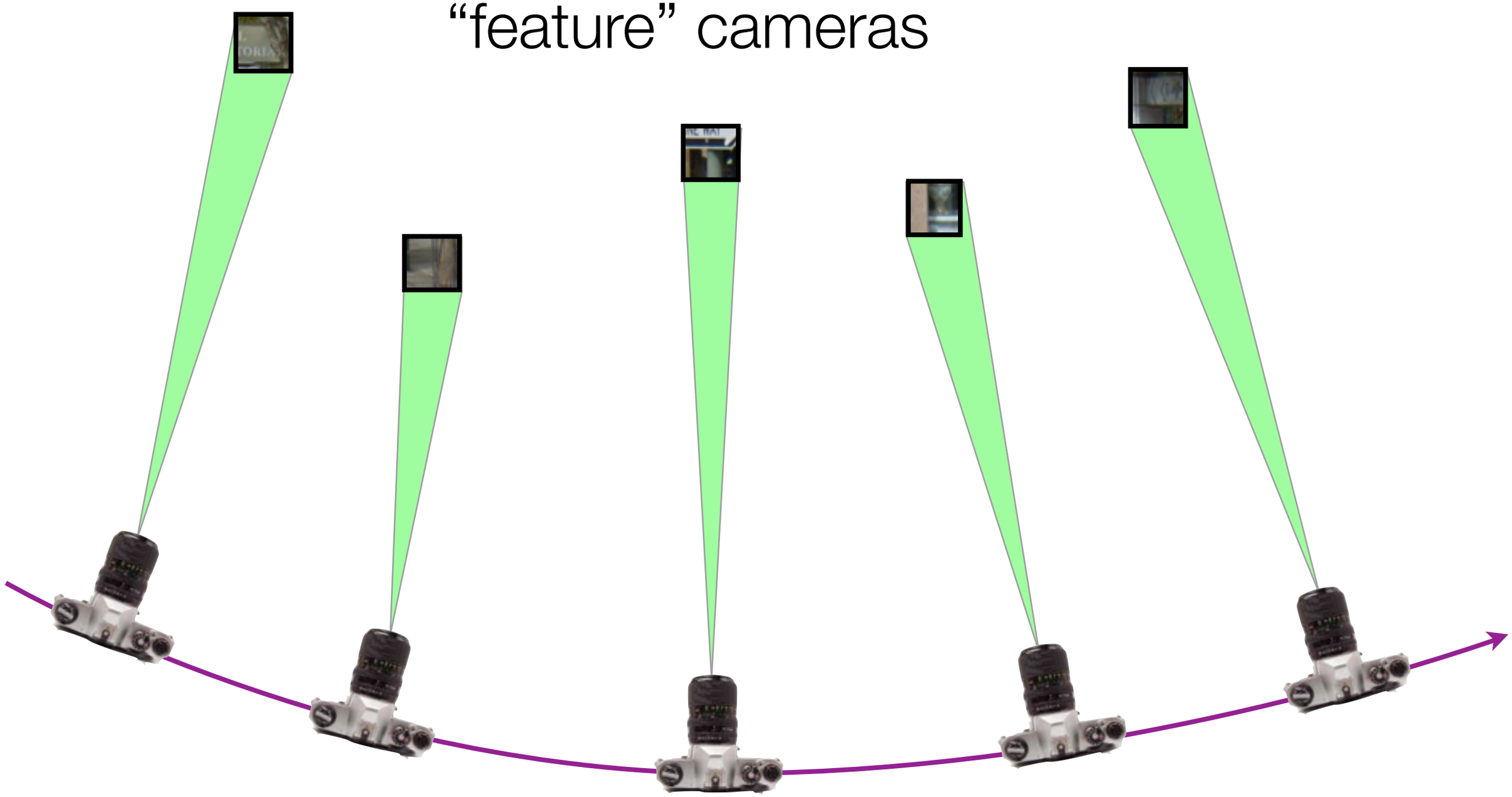
No single camera center.
No linear camera model.
Hard to triangulate points.

Our simple solution: every feature defines a camera



trajectory of rig over single exposure

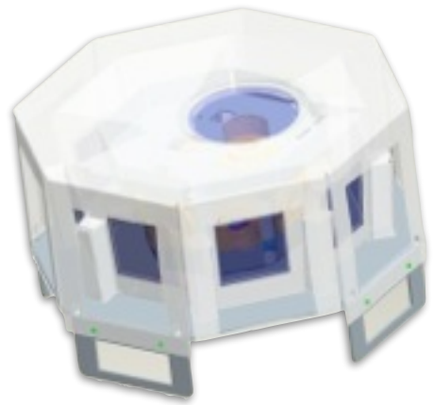
Our simple solution: every feature defines a camera



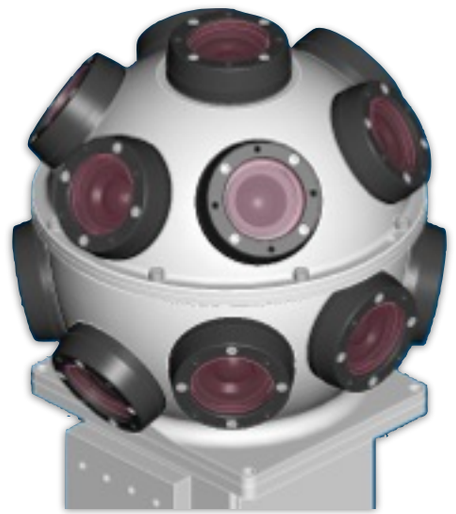
locally instantaneous and linear - SFM works again!

feature cameras work for any model, instant or rolling.
one implementation for all rigs.

and

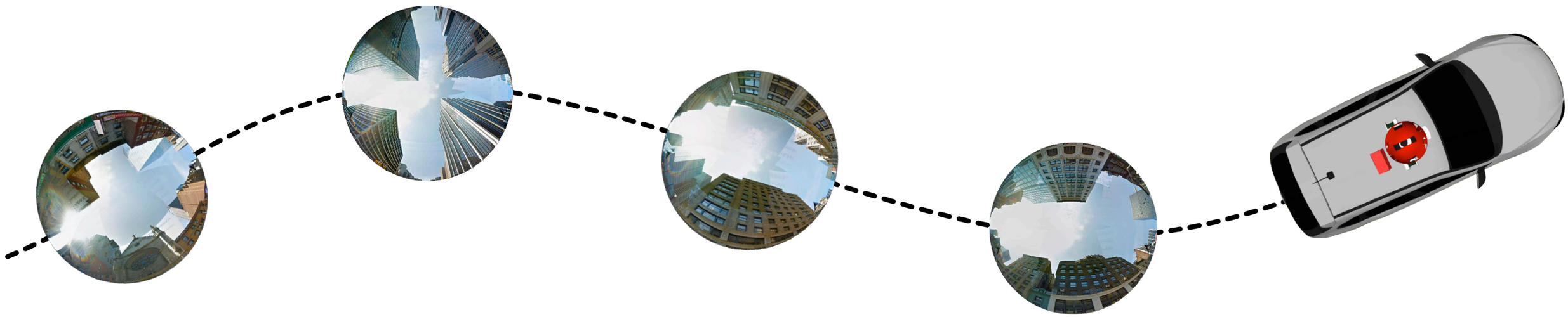


and



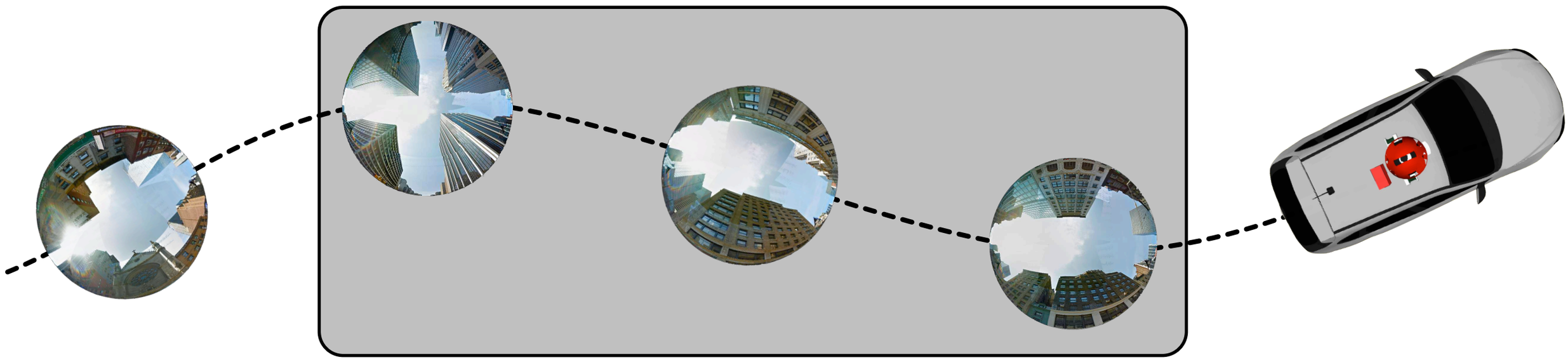
Managing Scale

We start with vehicle pose that's:
locally accurate (cm), **globally** okay (10m)



GPS - rough global estimate of location
Inertial measurement unit - tracks vehicle acceleration
Wheel encoder - how fast is the vehicle moving?

Managing Scale



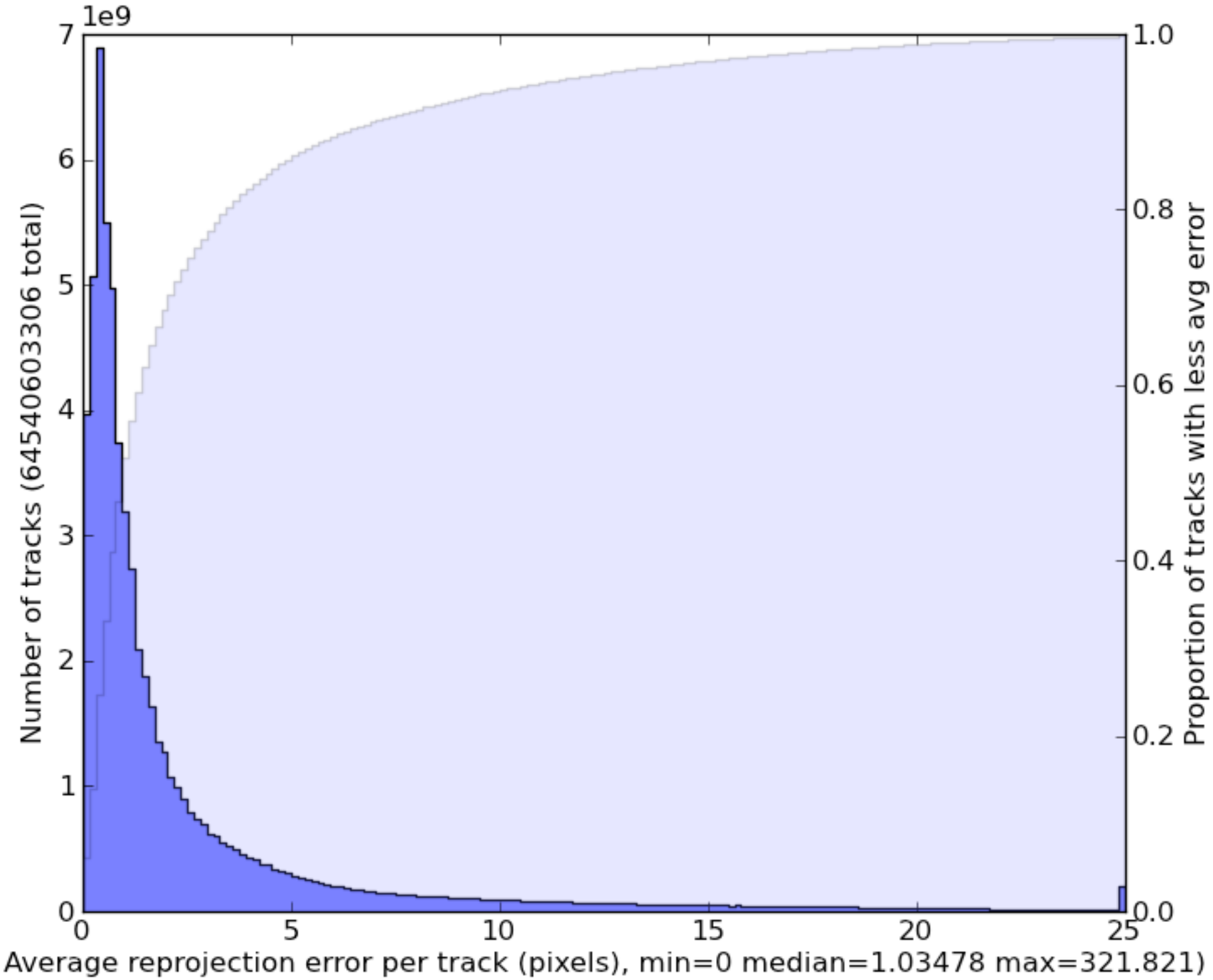
Traditional SfM	Our Approach
Match each image to all others: $O(n^2)$ effort	Fixed matching window along vehicle path: O(n) effort
Visual odometry / RANSAC to estimate relative pose	Good relative pose prior from inertial sensors

Progress so far:

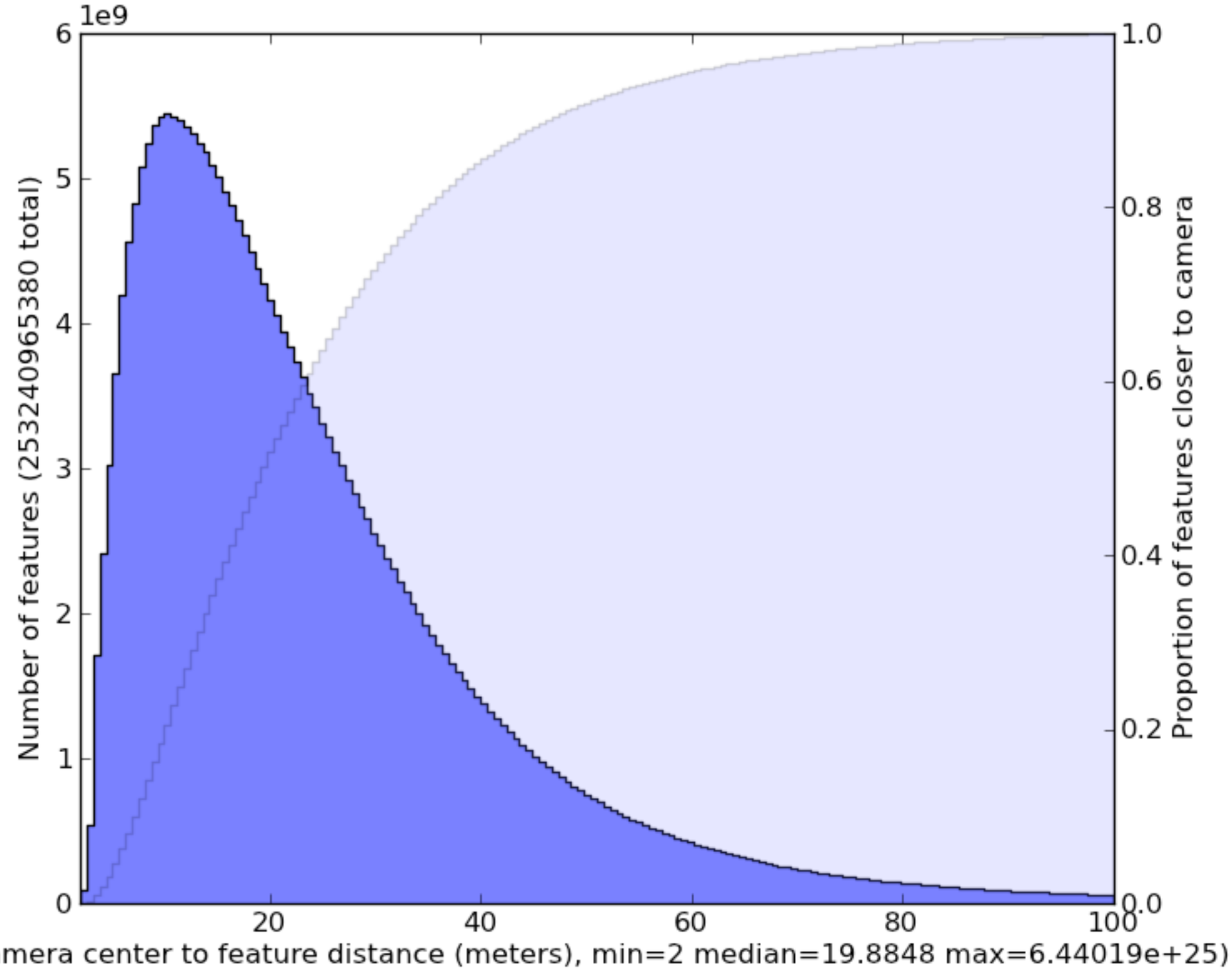
Panoramas	**9B**
Points	**800B**
Views	**1.5T**
Compute time	**1920 core-years**

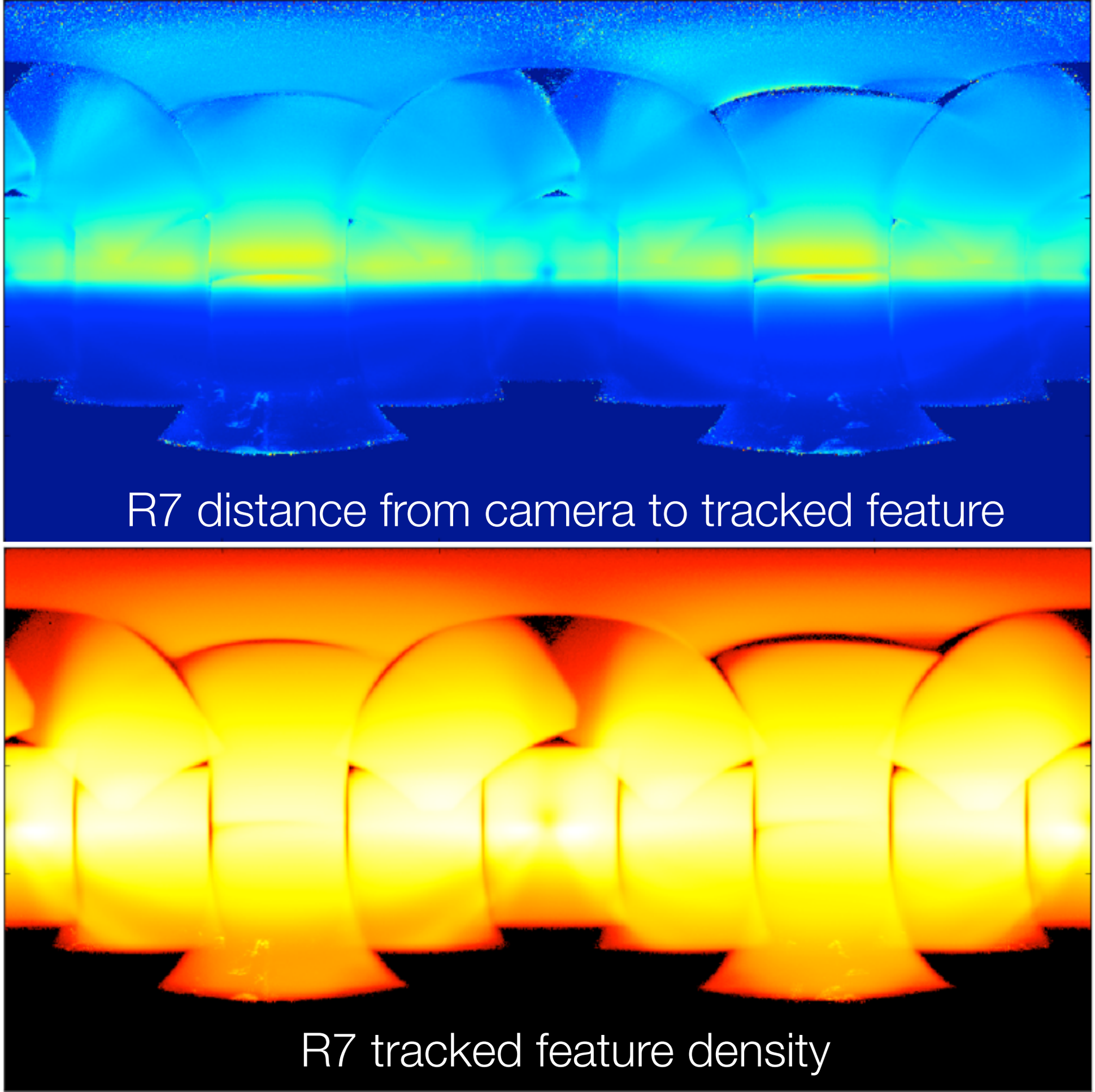
Point Cloud Demo

Average reprojection error per tracked feature



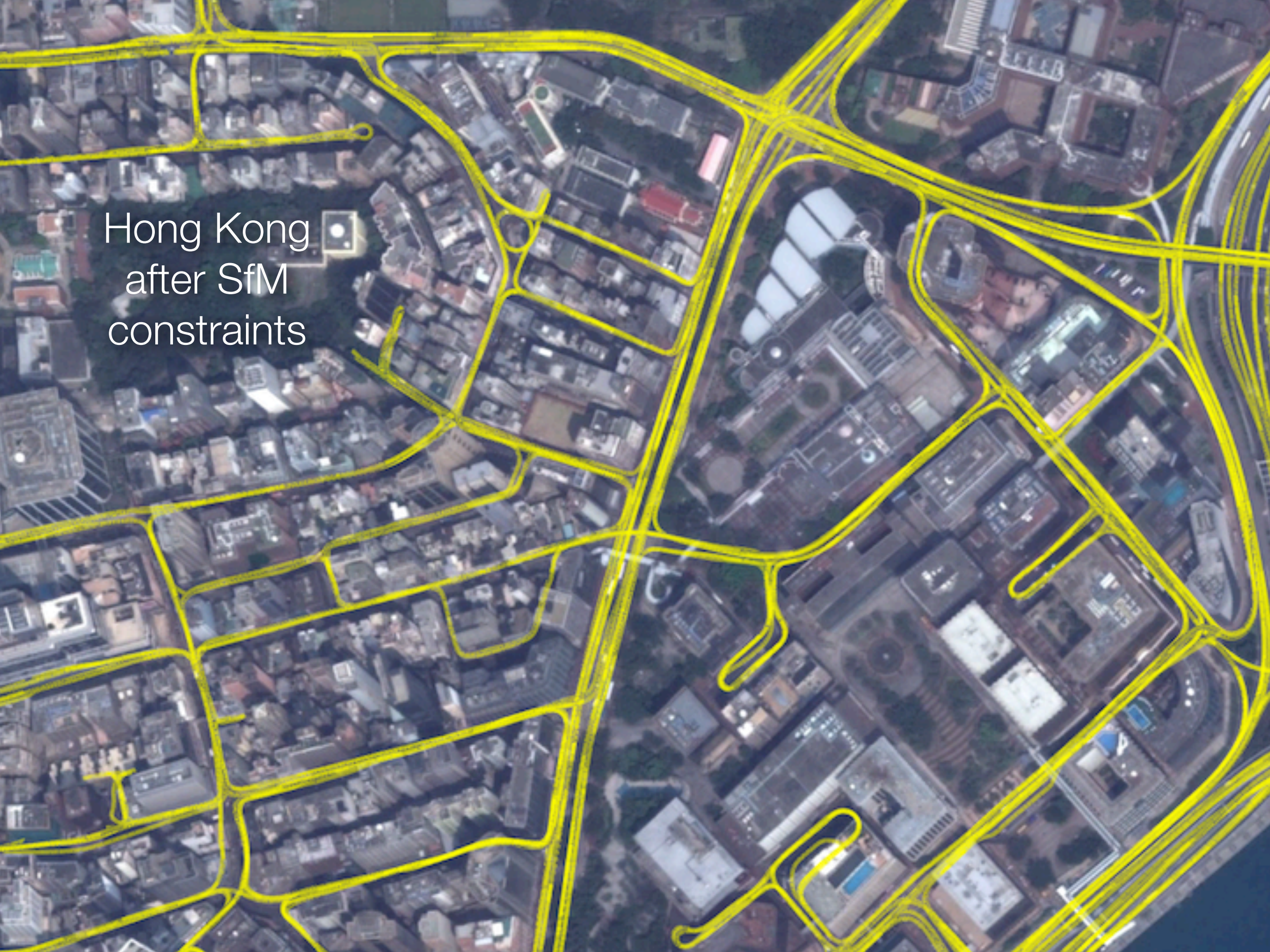
Distance from camera to tracked feature




R7 distance from camera to tracked feature
100
90
80
70
60
50
40
30
20
10
R7 tracked feature density
1e+06
1e+05
1e4
1000
100
10
1

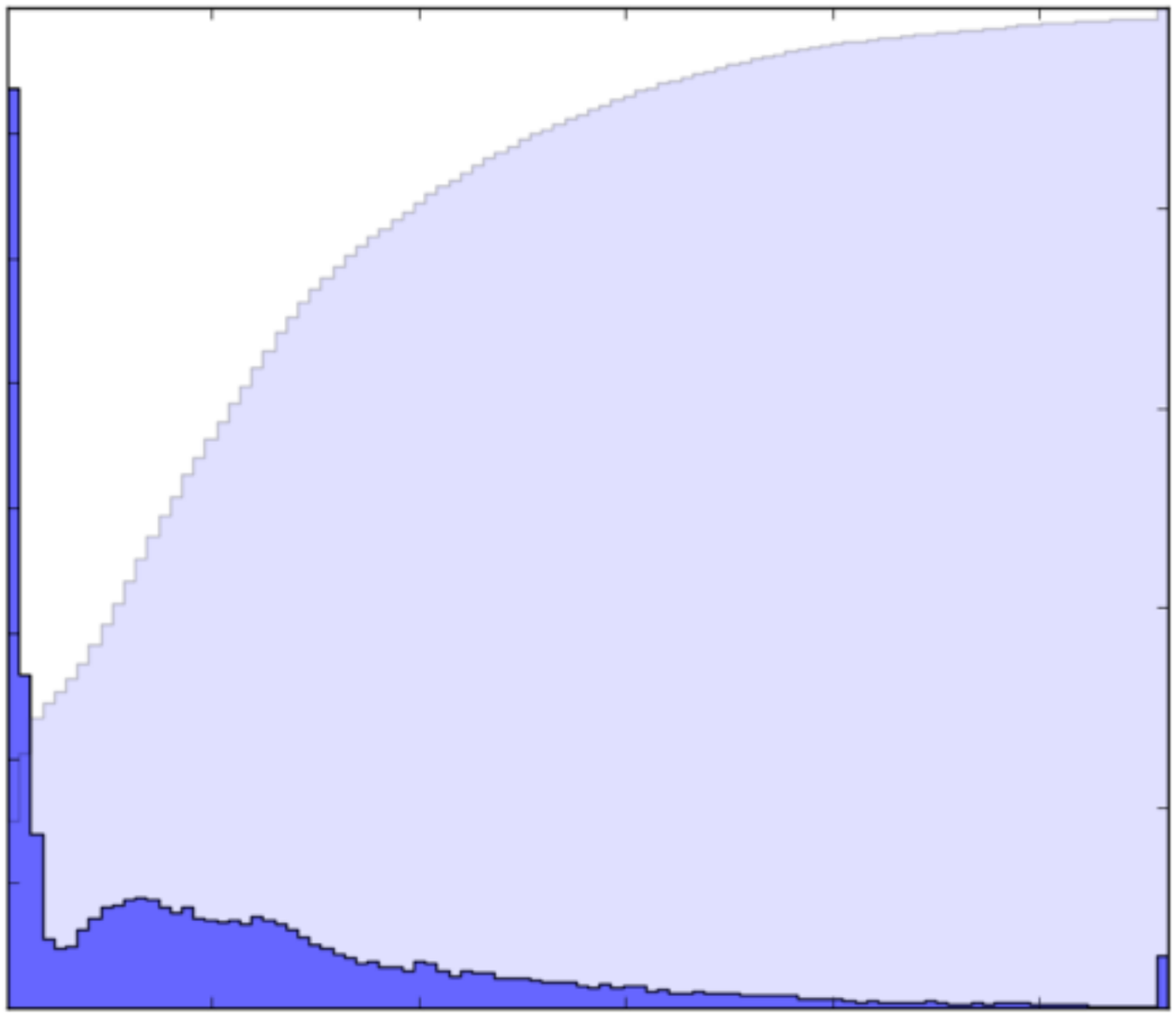
Goog

Hong Kong
before SfM
constraints


Hong Kong
after SfM
constraints

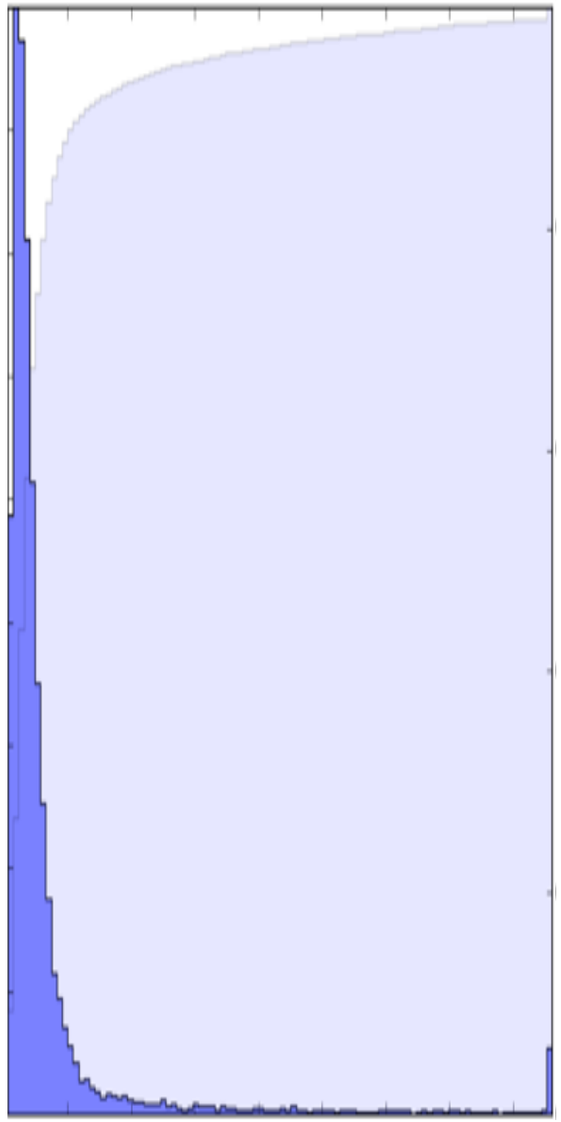
SfM constraint error

before



90% < **5m**, **1.4º**

after



90% < **0.3m**, **0.4º**

Project 2: Stereo

- **I am traveling next week!**
- My office hours are **today**, 5pm
- Work with your groups next week in class