RoboJackets

FIRST - IGVC - BattleBots - RoboCup

2009 TE Sessions Supported by

CATERPILLAR®

NATIONAL INSTRUMENTS™

The Arthur M. Blank Family Foundation

Rockwell Automation

The George W. Woodruff School of Mechanical Engineering

Georgia Tech College of Computing

Introduction to Computer Vision

Andy Bardagjy

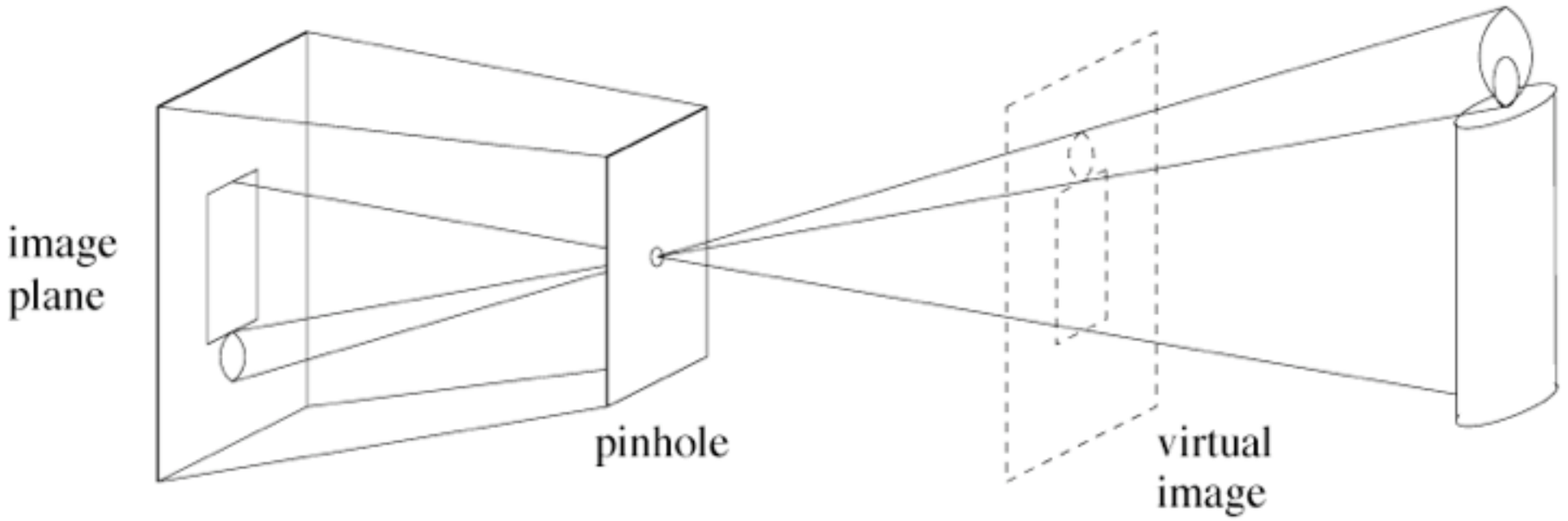
October 5, 2009

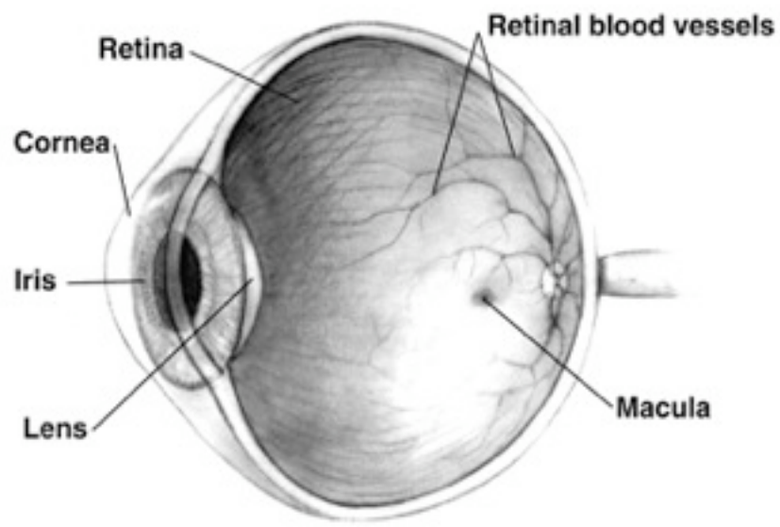
www.robojackets.org

Why Use Vision

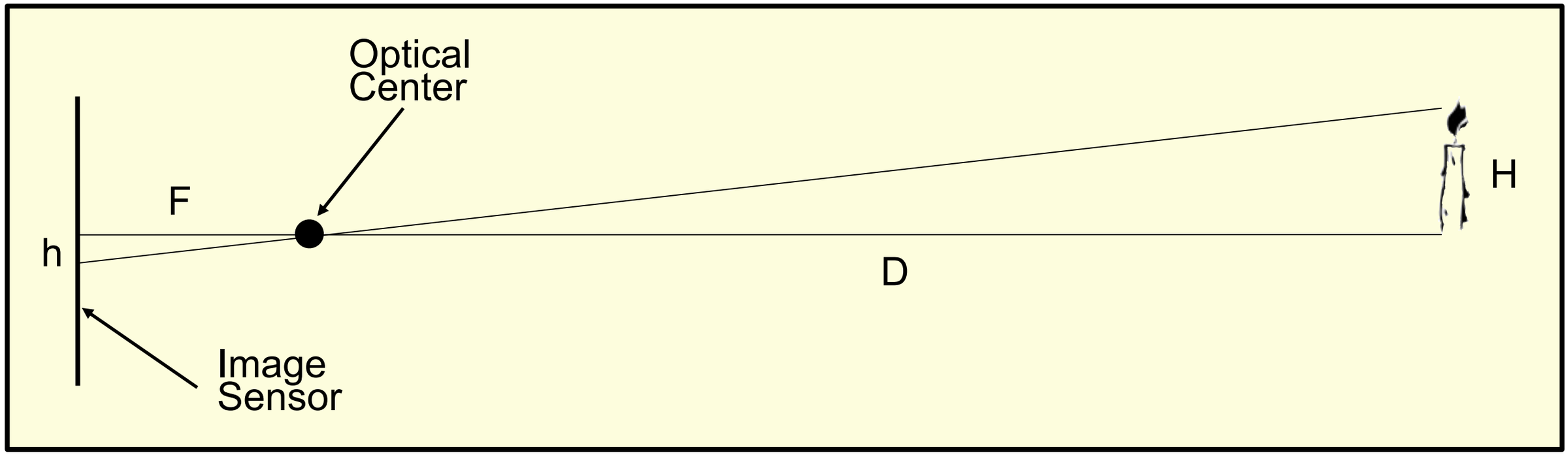
- Tremendous amount of information
  - –Spatial
  - –Temporal
  - –Radiometric
- Cost
- Passive
- Size
- Our “primary” sensor

Image Formation





Geometry

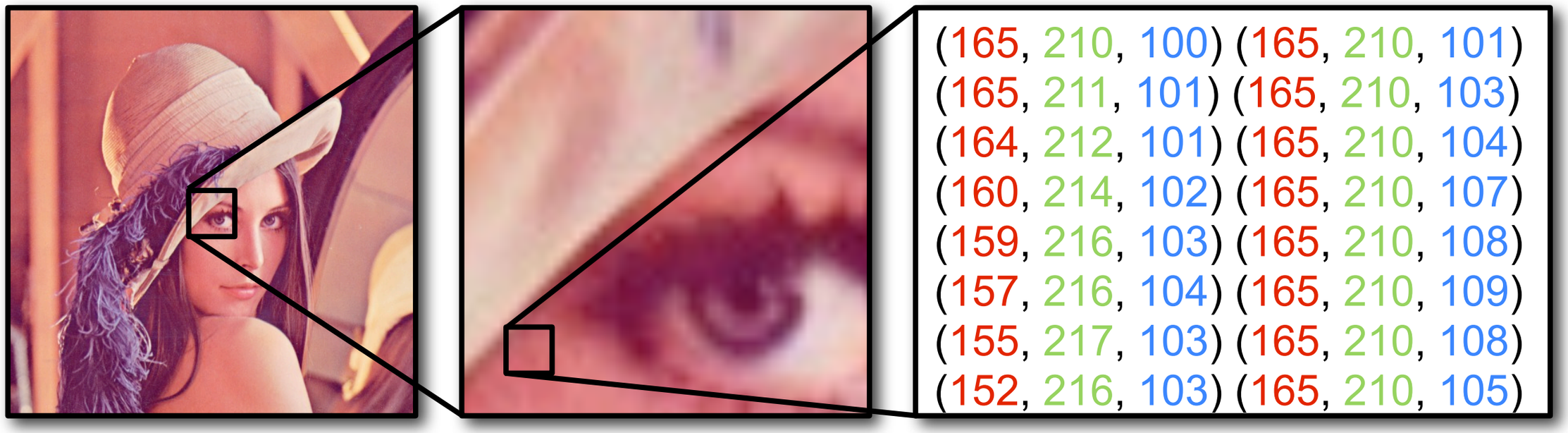


F = focal length
D = distance to object
h = displacement on sensor
H = height of object

By Similar Triangles

$$\frac{h}{f} = \frac{H}{D}$$

Image Representation

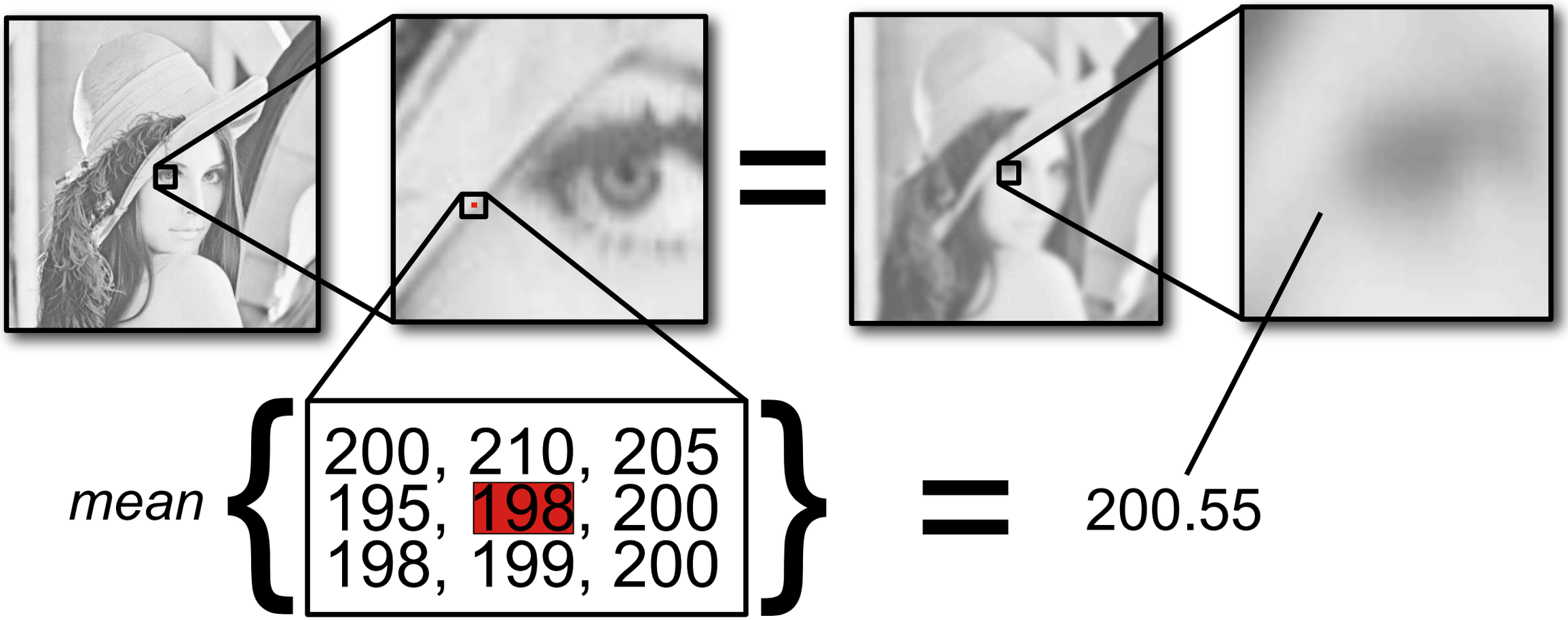


- Matrix representation of image data
- Data “cube”
- Origin at upper left

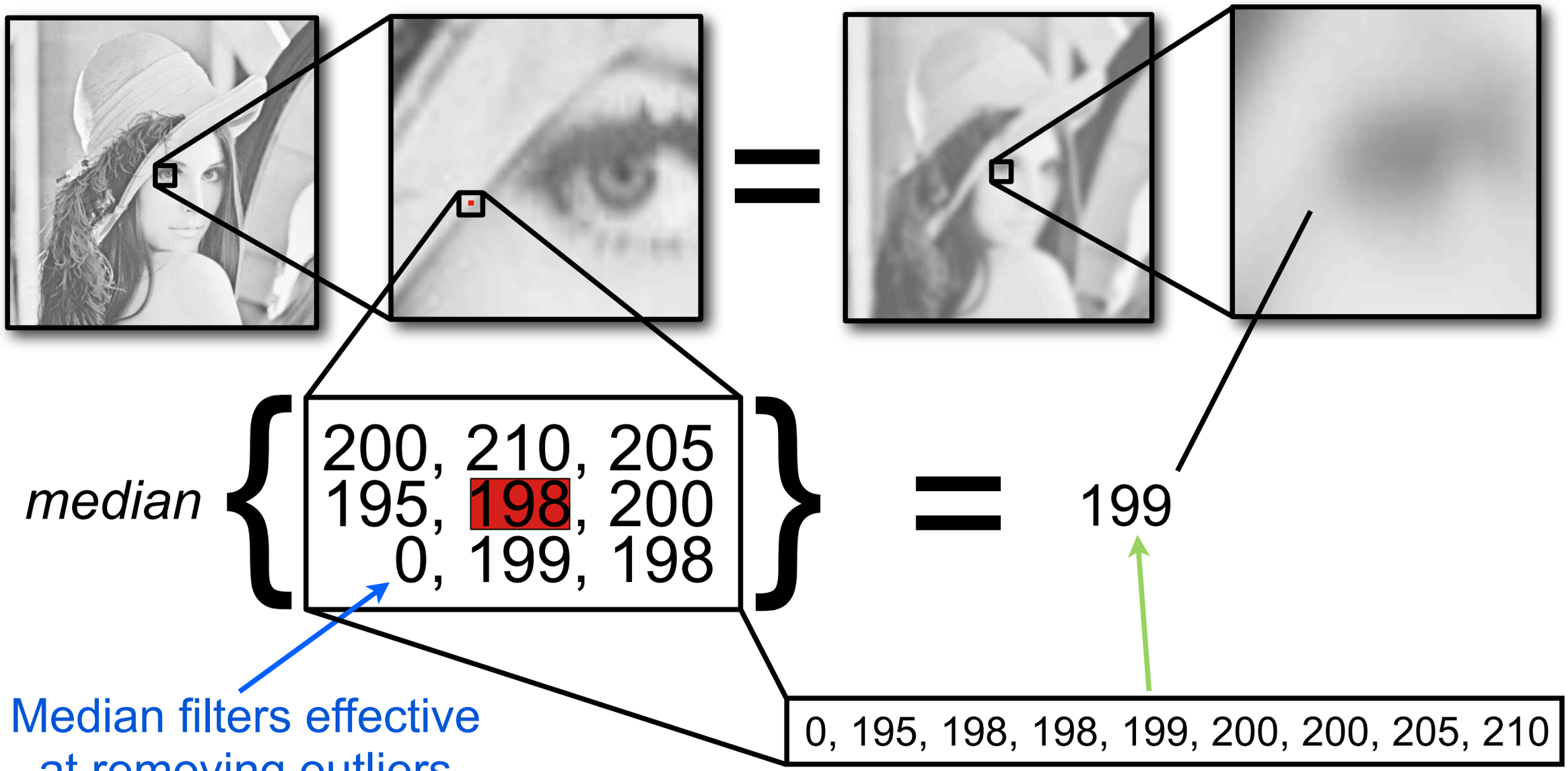
Filtering

- Median filter
- Averaging filter
- Convolution
- Thresholding

Averaging Filter



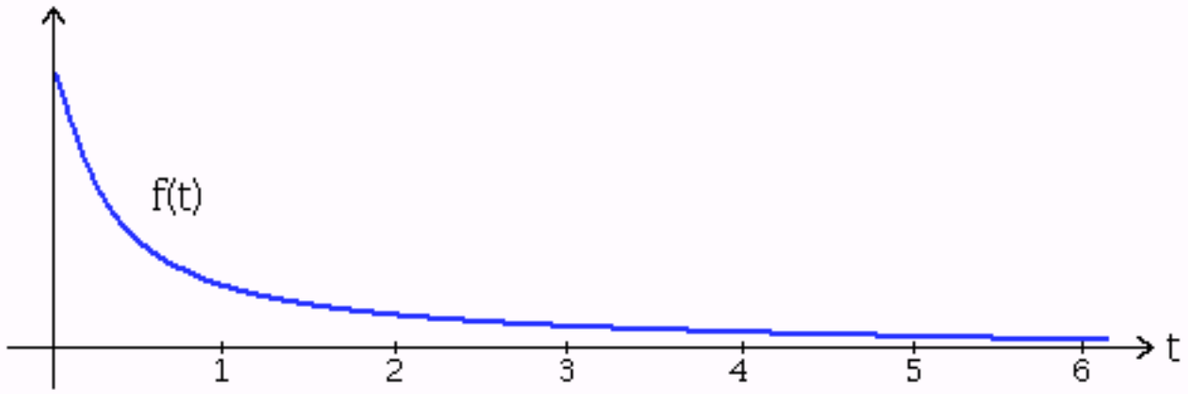
Median Filter

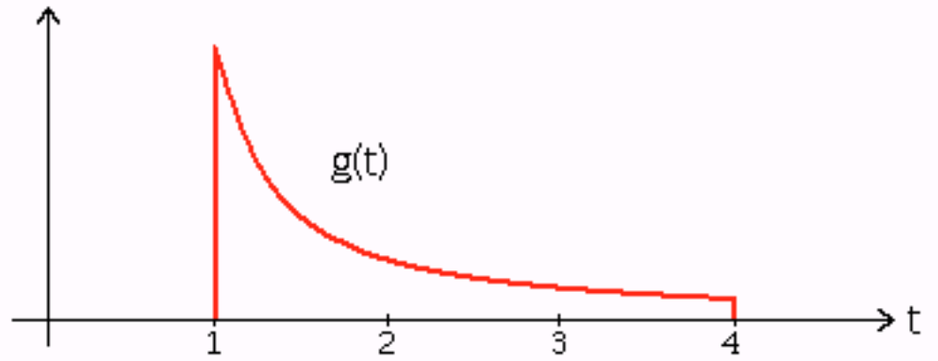


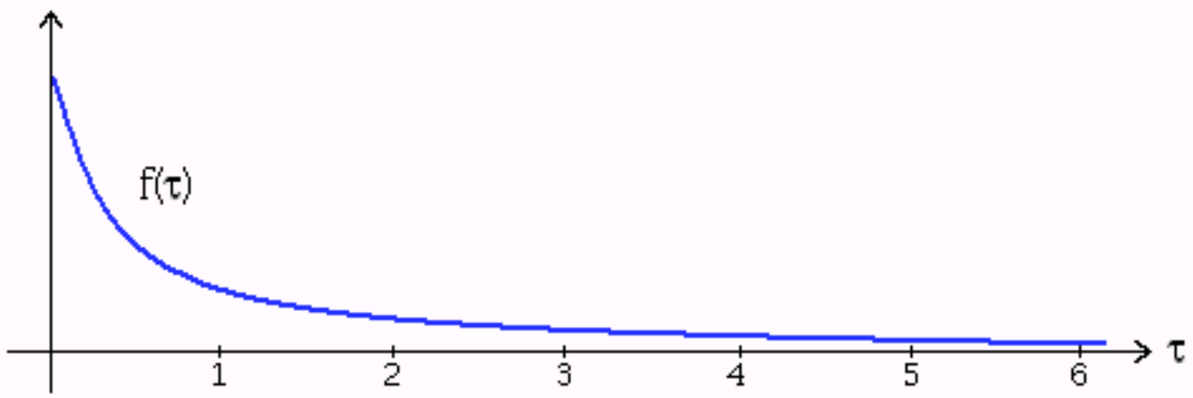
mean = 178.3

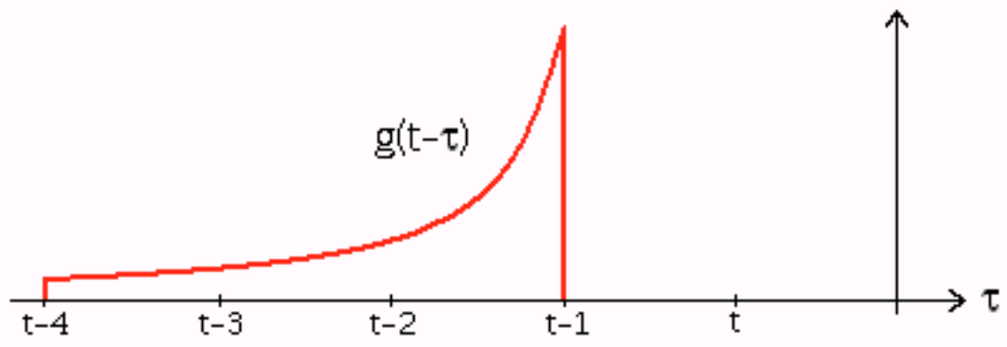
RoboJackets

Convolution

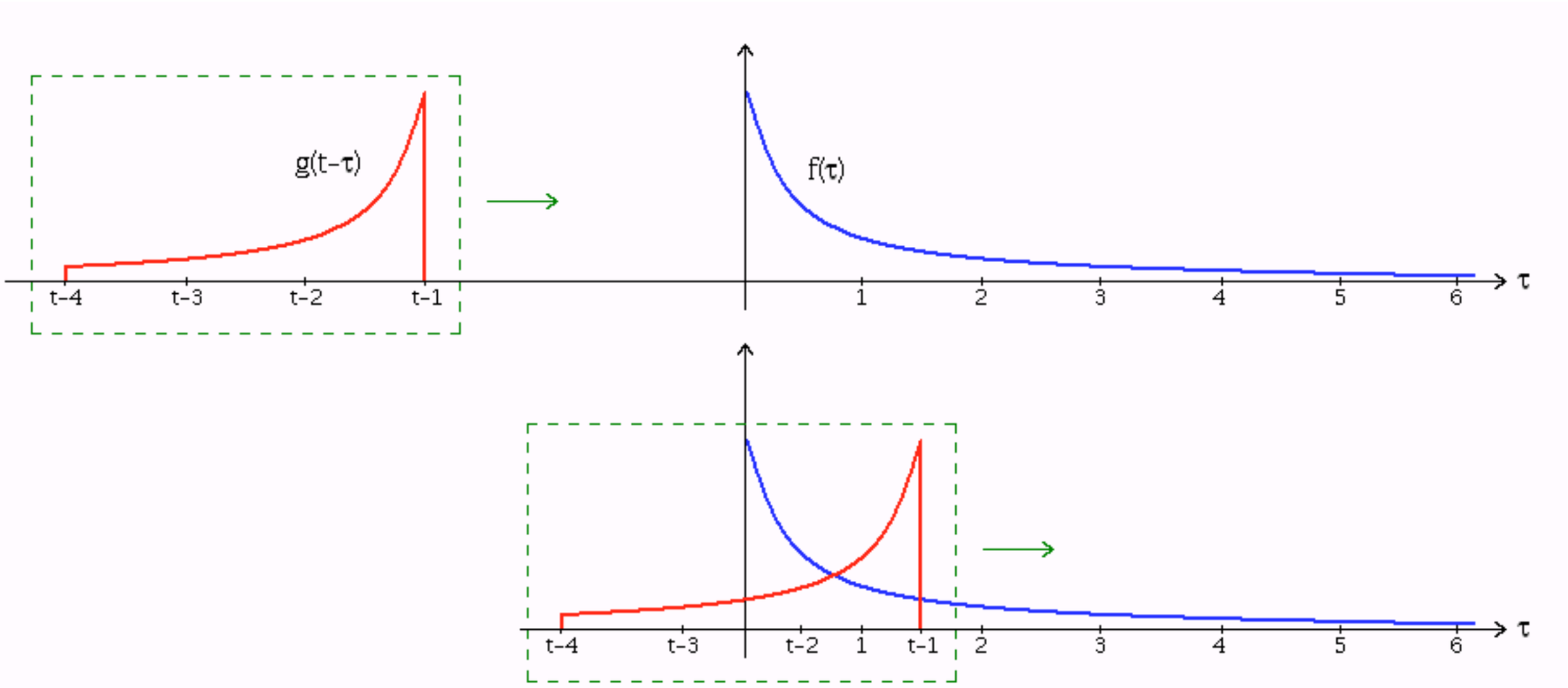




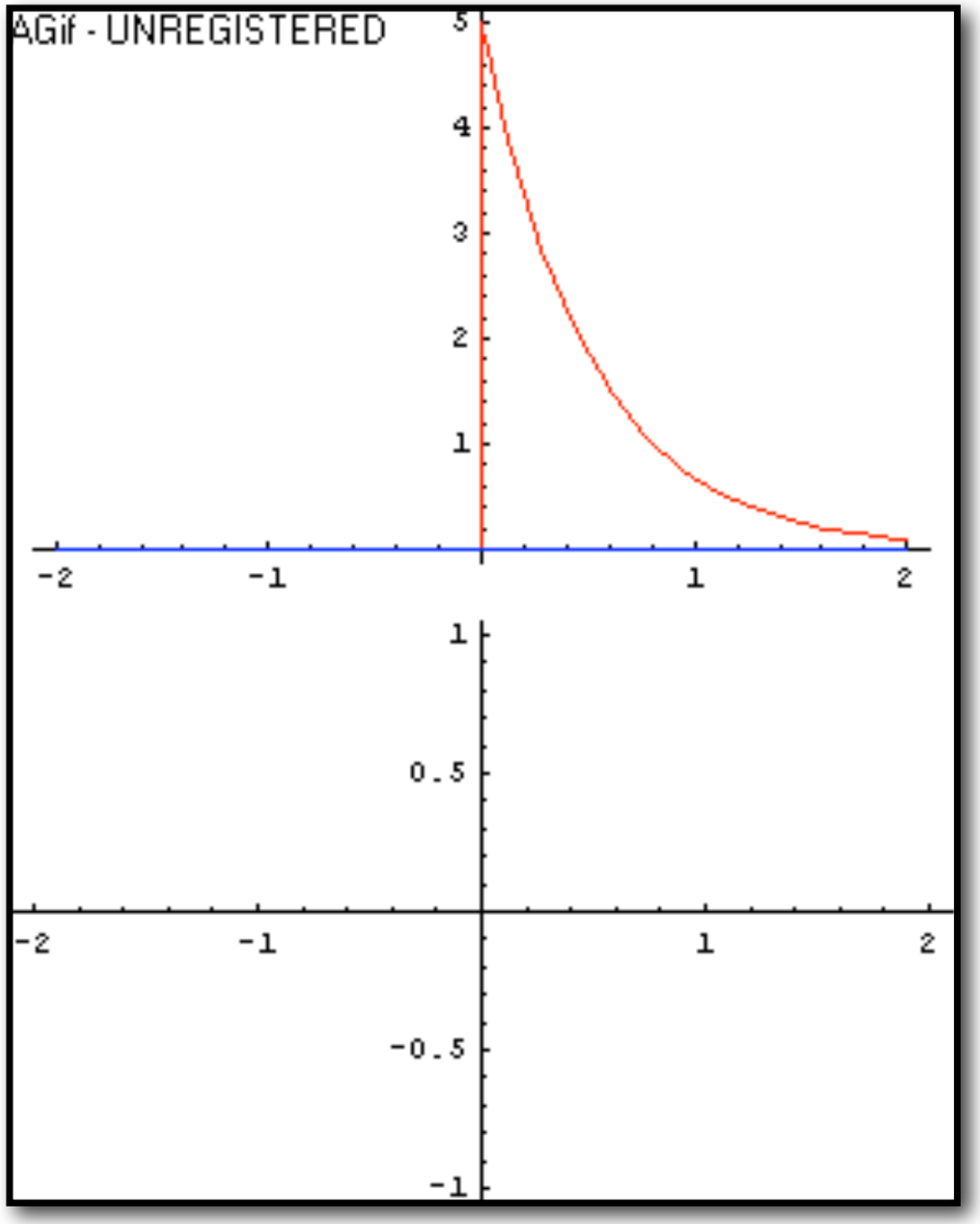




Convolution

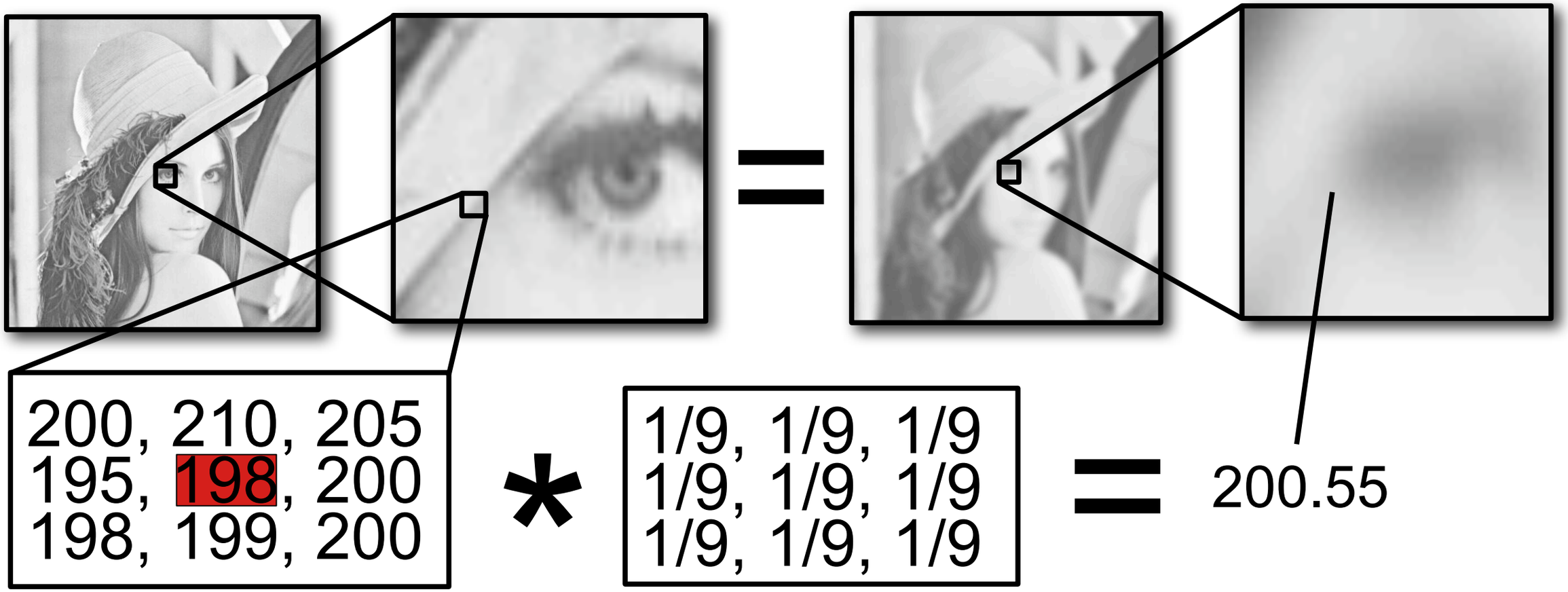


Convolution



wikipedia

Convolution



200, 210, 205
195, 198, 200
198, 199, 200

*

1/9, 1/9, 1/9
1/9, 1/9, 1/9
1/9, 1/9, 1/9

= 200.55

* This is an example of an averaging ("mean") filter.

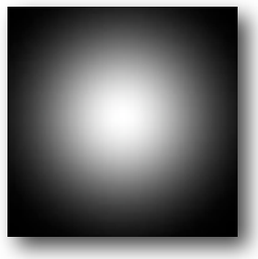
$$(f \star g)(t) = \int_{-\infty}^{\infty} f(\tau) \cdot g(t - \tau)\, d\tau$$

$$(f \star g)[n] = \sum_{m=-\infty}^{\infty} f[m] \cdot g[n - m]$$

RoboJackets

Other Masks...

Gaussian Blur



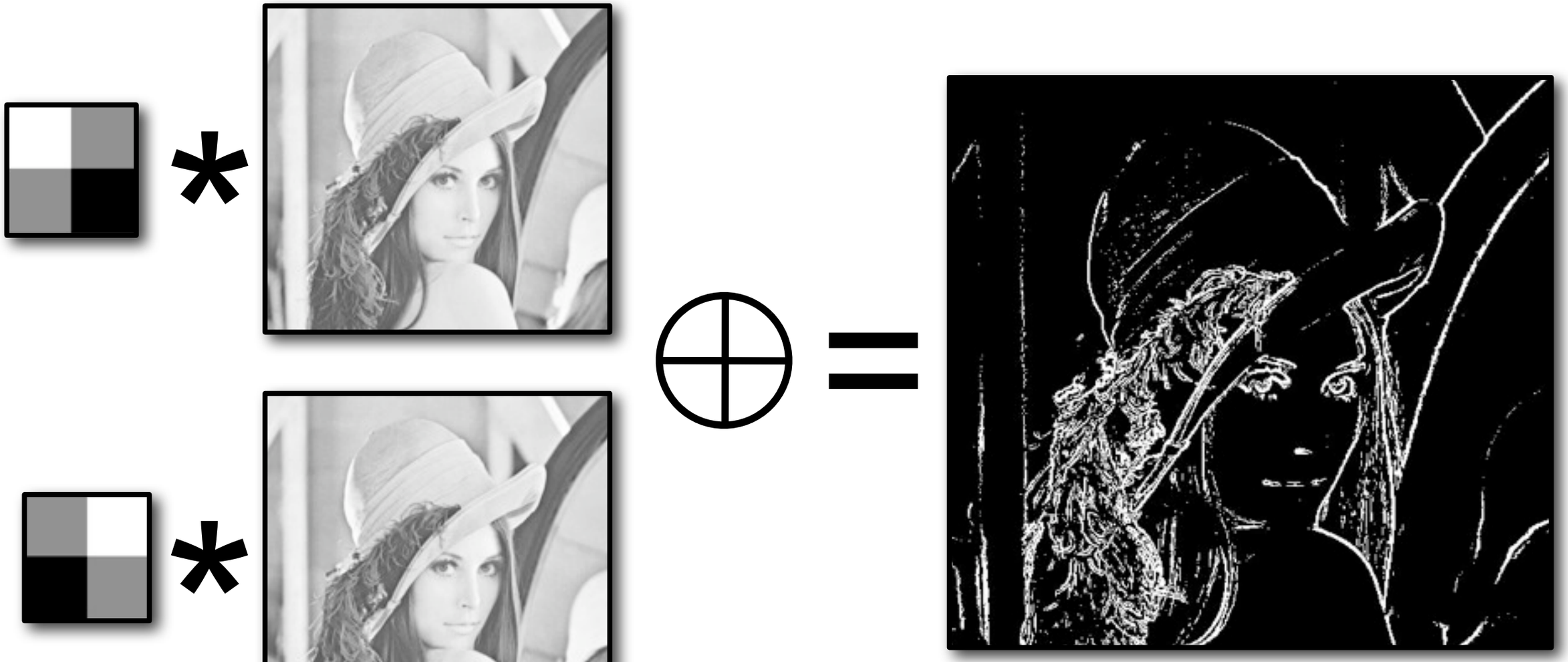
*

=

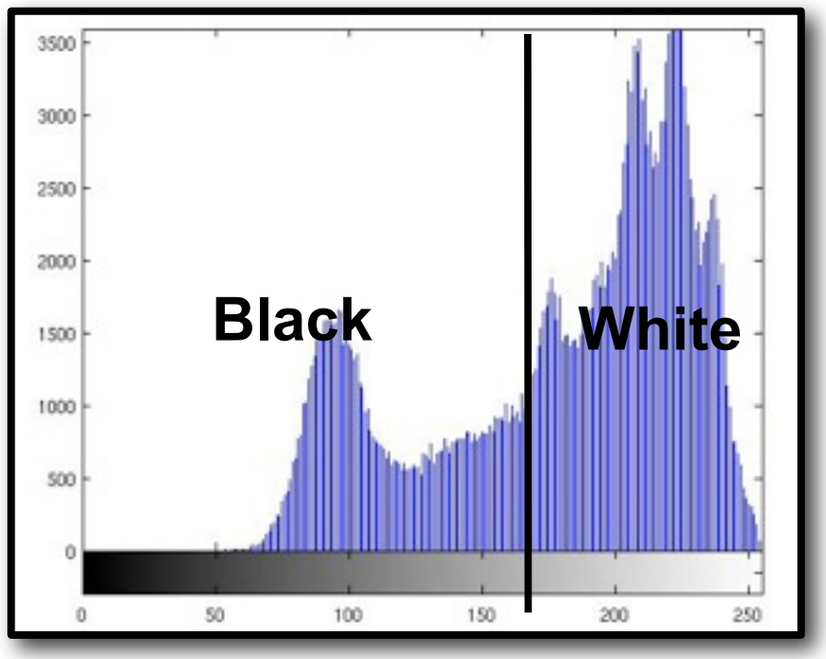
- More weight given to center pixel
- Approximates image resizing, real world blur.
- Resistant to outliers
- Enemy of noise

Other Masks...

Robert's Edge Detector



Thresholding



Algorithm

```
if (pixel > threshold)
        white
else
        black
```

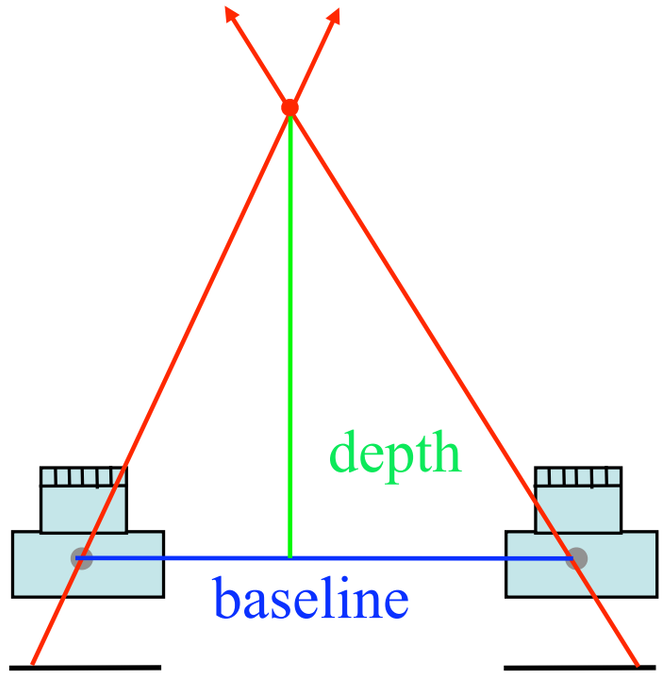
Useful for

- Color recognition
- Crude image compression

Stereo

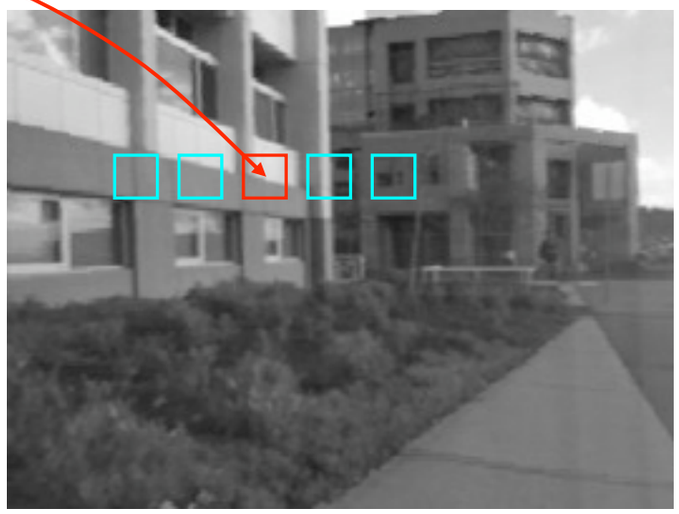
Stereo

Triangulate to find depth from the same feature in two (or more) images.



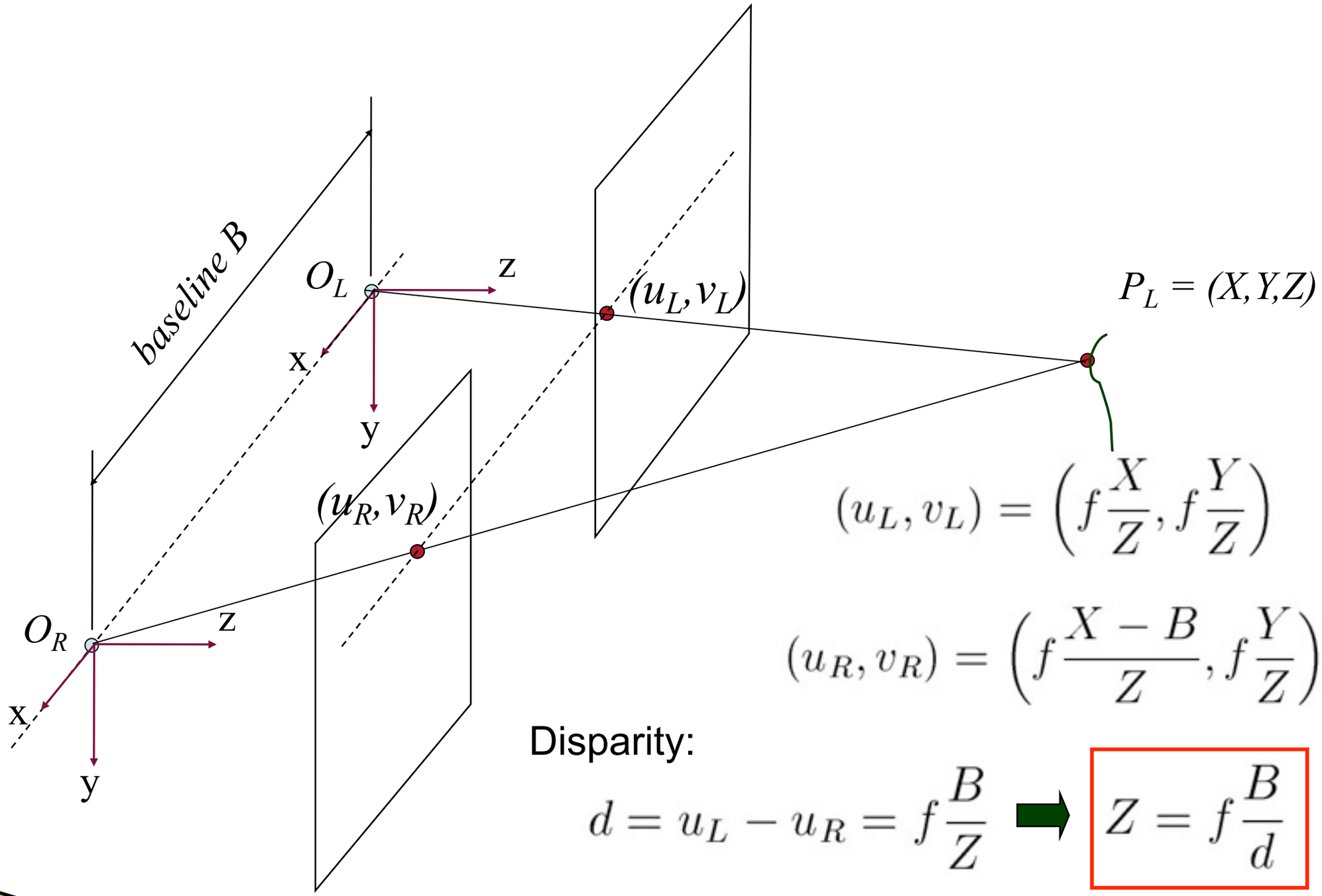
Left

Right



Requires
- Feature detection and matching across views (correspondence points)
- Calibrated cameras*

Stereo



$$(u_L, v_L) = \left(f\frac{X}{Z}, f\frac{Y}{Z}\right)$$

$$(u_R, v_R) = \left(f\frac{X-B}{Z}, f\frac{Y}{Z}\right)$$

Disparity:

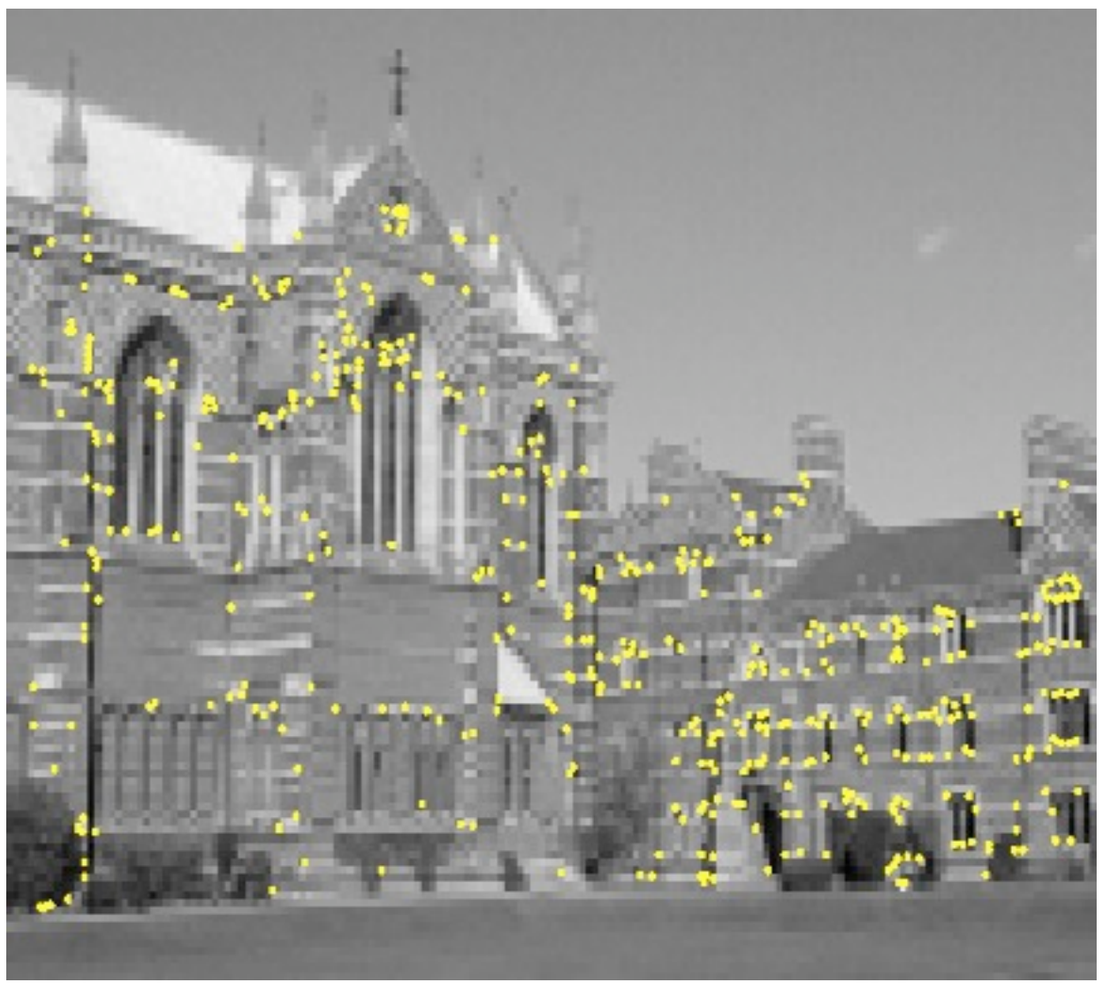
$$d = u_L - u_R = f\frac{B}{Z} \Rightarrow \boxed{Z = f\frac{B}{d}}$$

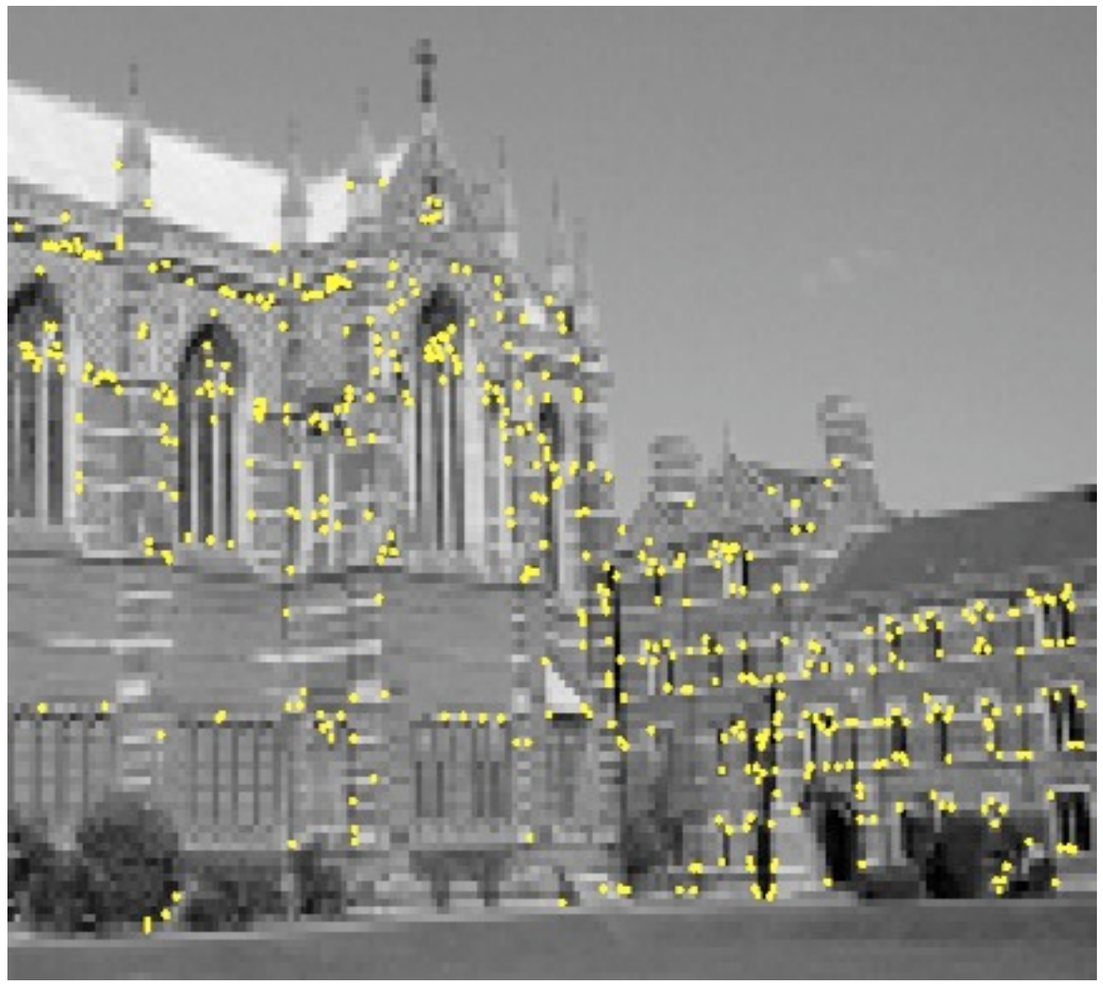
Stereo

Algorithm

1. Feature detection
2. Feature matching
3. Model fitting

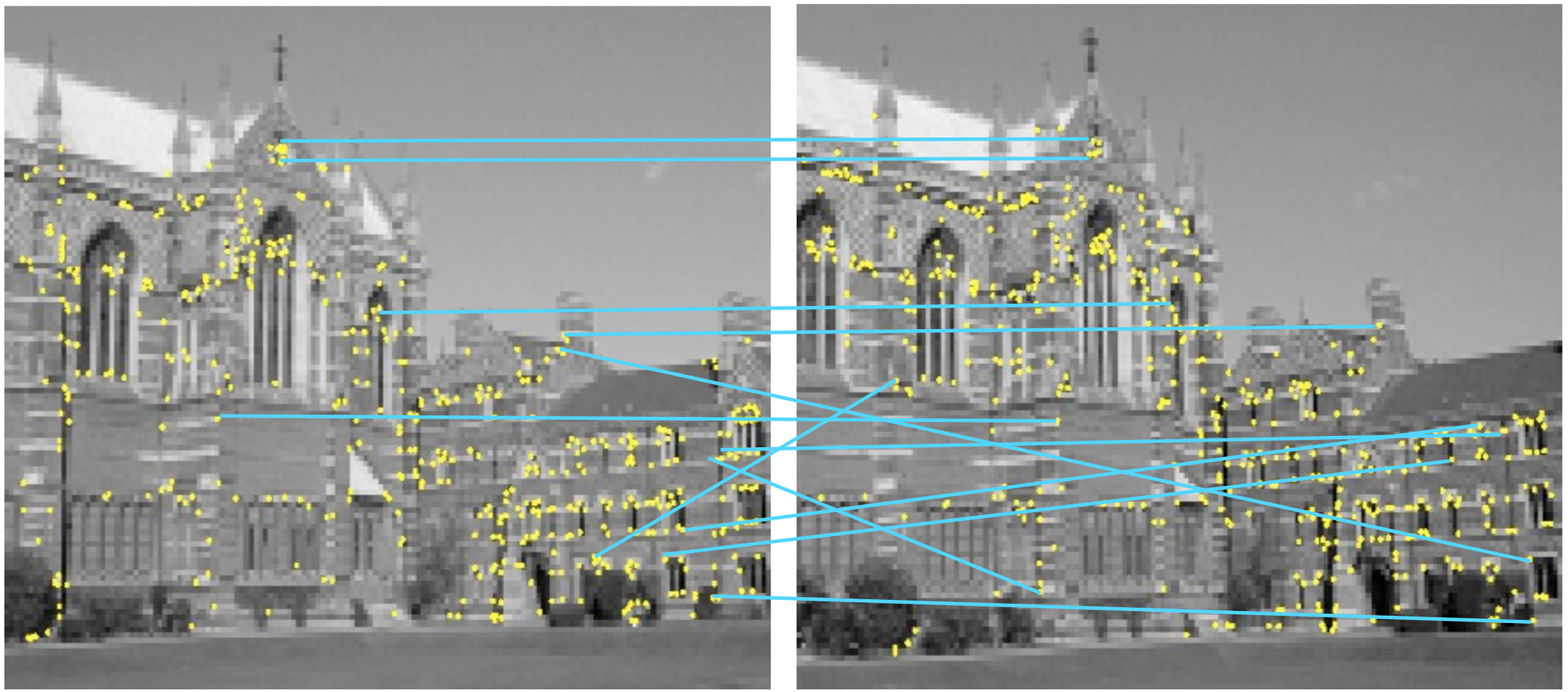
Feature Detection





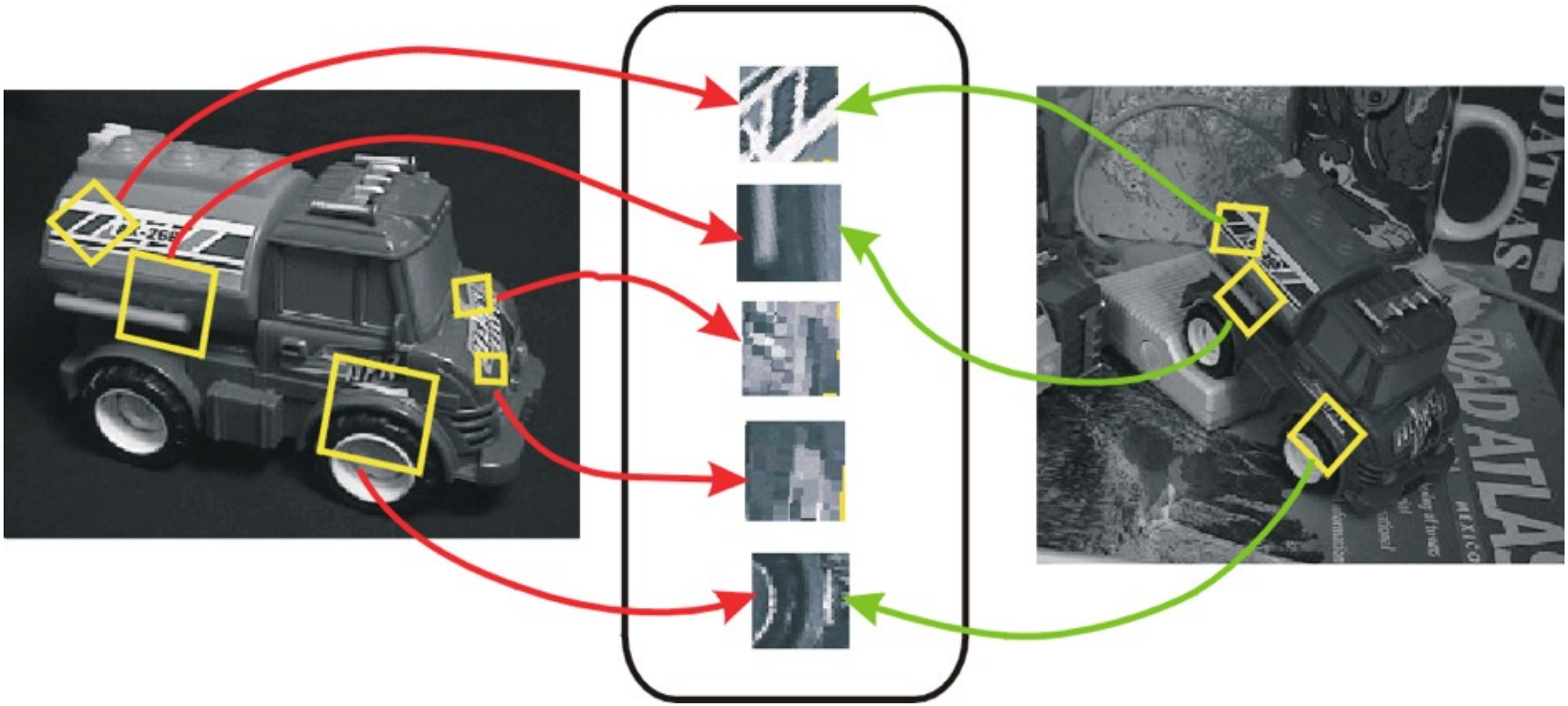
- Harris Corner Detector
- SIFT Feature Detector
- SURF
- By hand...
- Actively researched...

Feature Matching

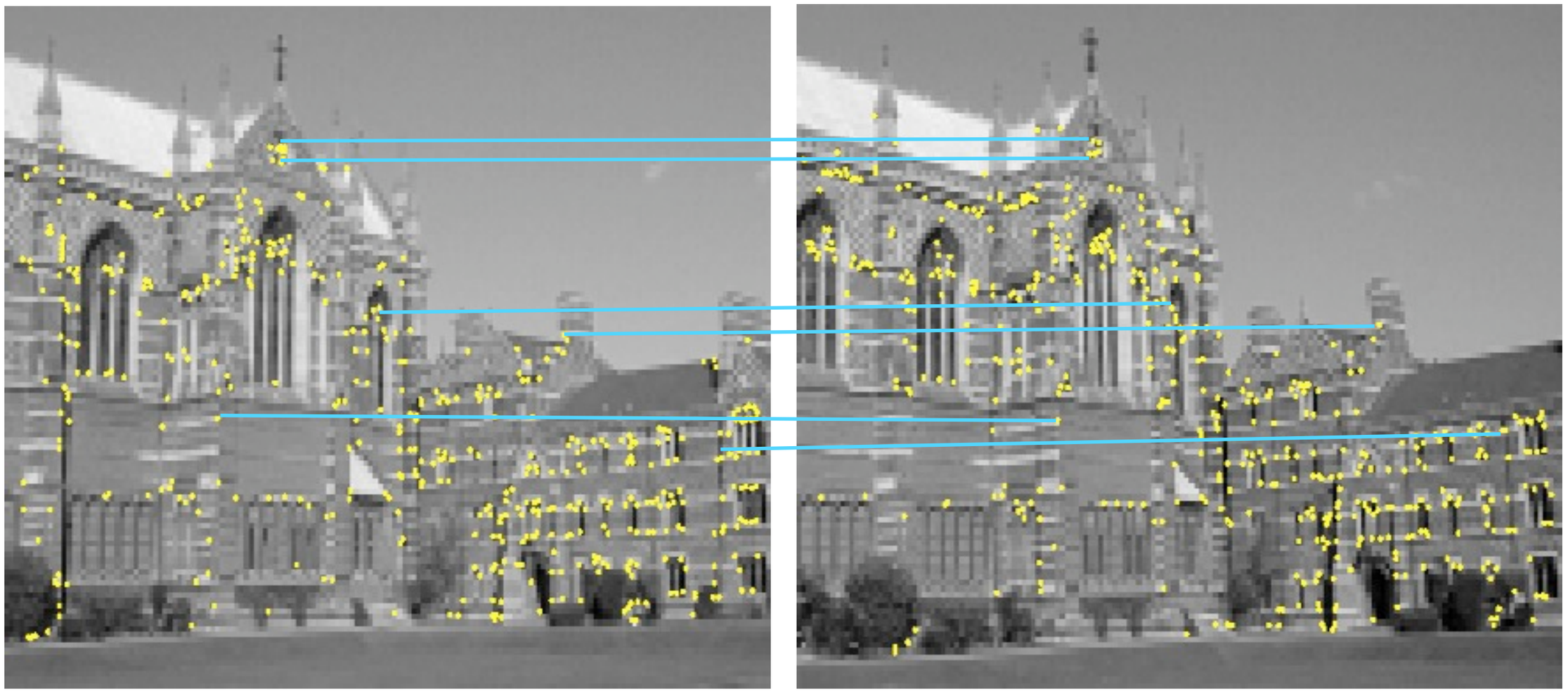


- Match features
- Depends on feature type
- ie color, texture, brightness

Feature Matching



Model Fitting



- RANSAC
- Particle Filtering
- Bundle Adjustment
- Expectation Maximization

RANSAC

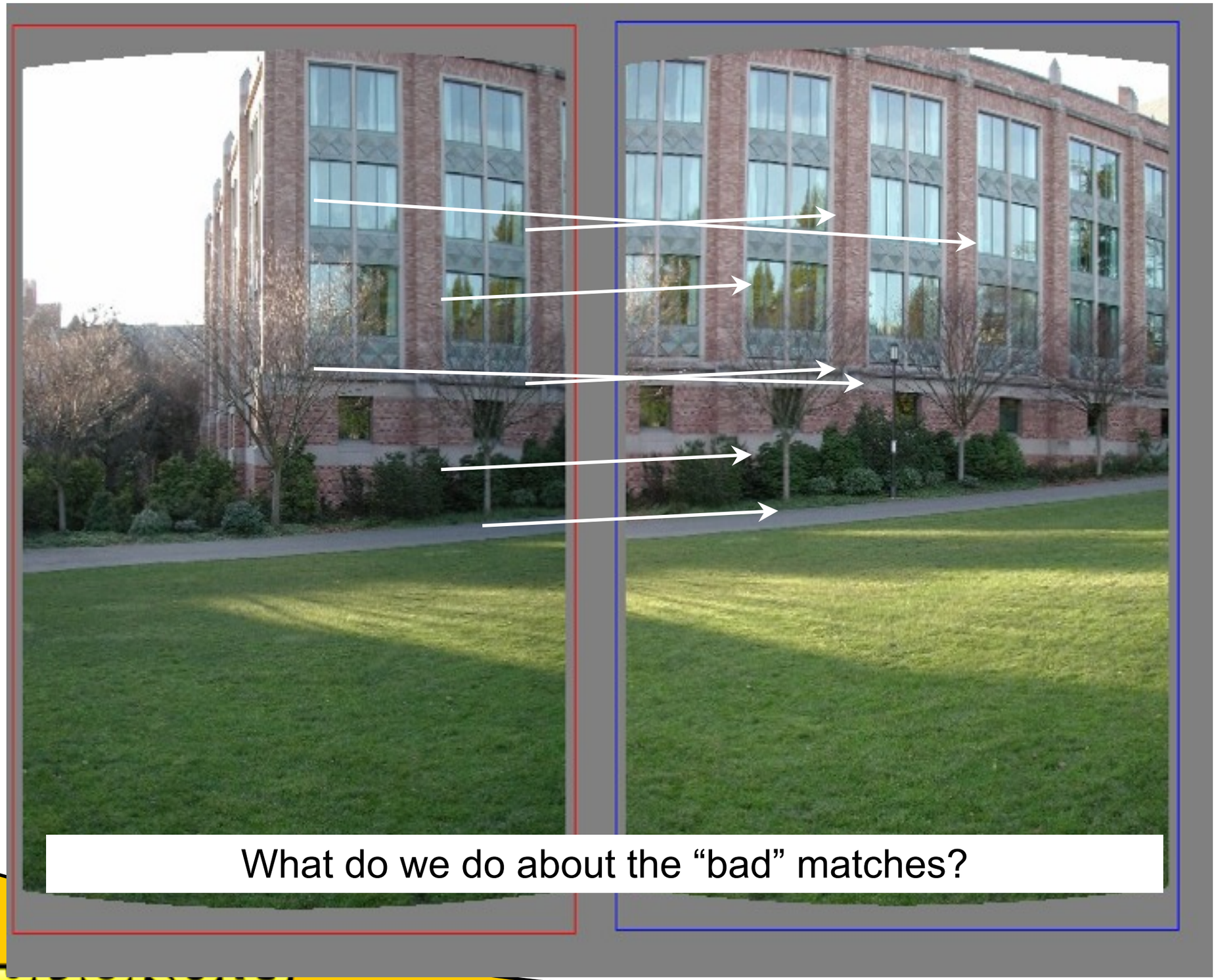
–Objective:

- Robust fit of a model to data S

–Algorithm

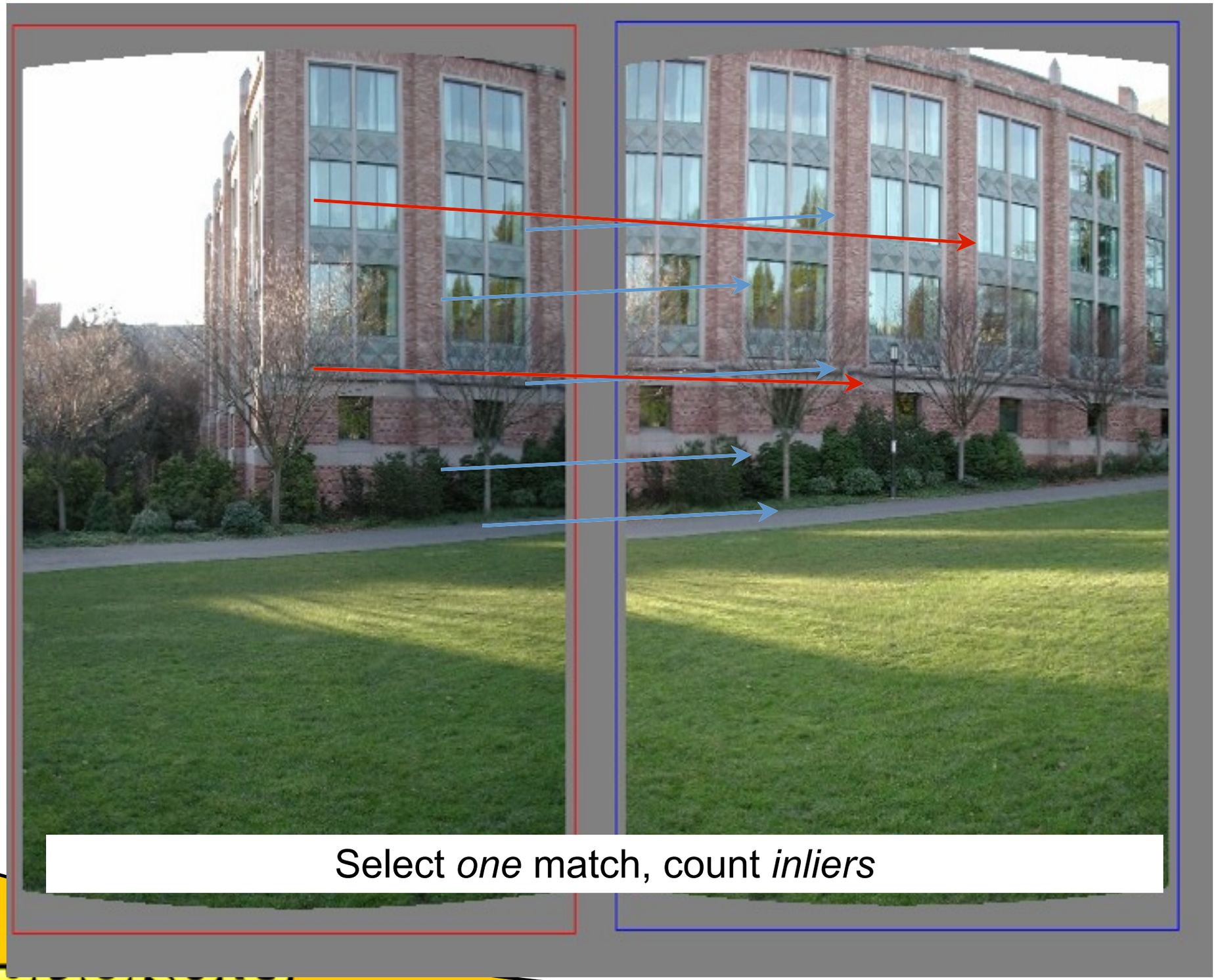
- Randomly select s data points
- Make a model with those points
- Get consensus set S (which are well described by model)
- If |S|>T, terminate and return model
- Repeat for N trials, return model with max |S|
- Optional: refine returned model

Matching features



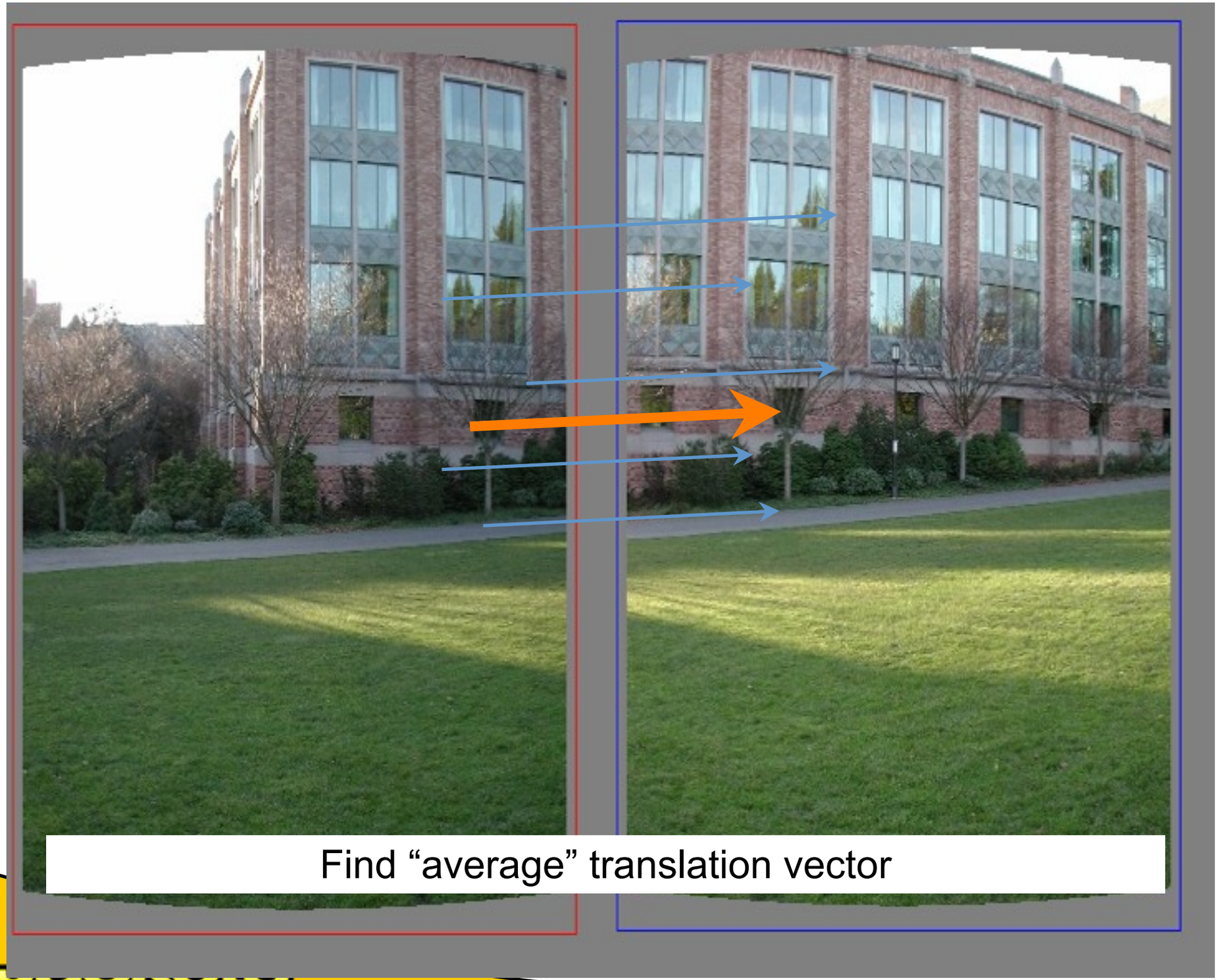
What do we do about the “bad” matches?

RANSAC



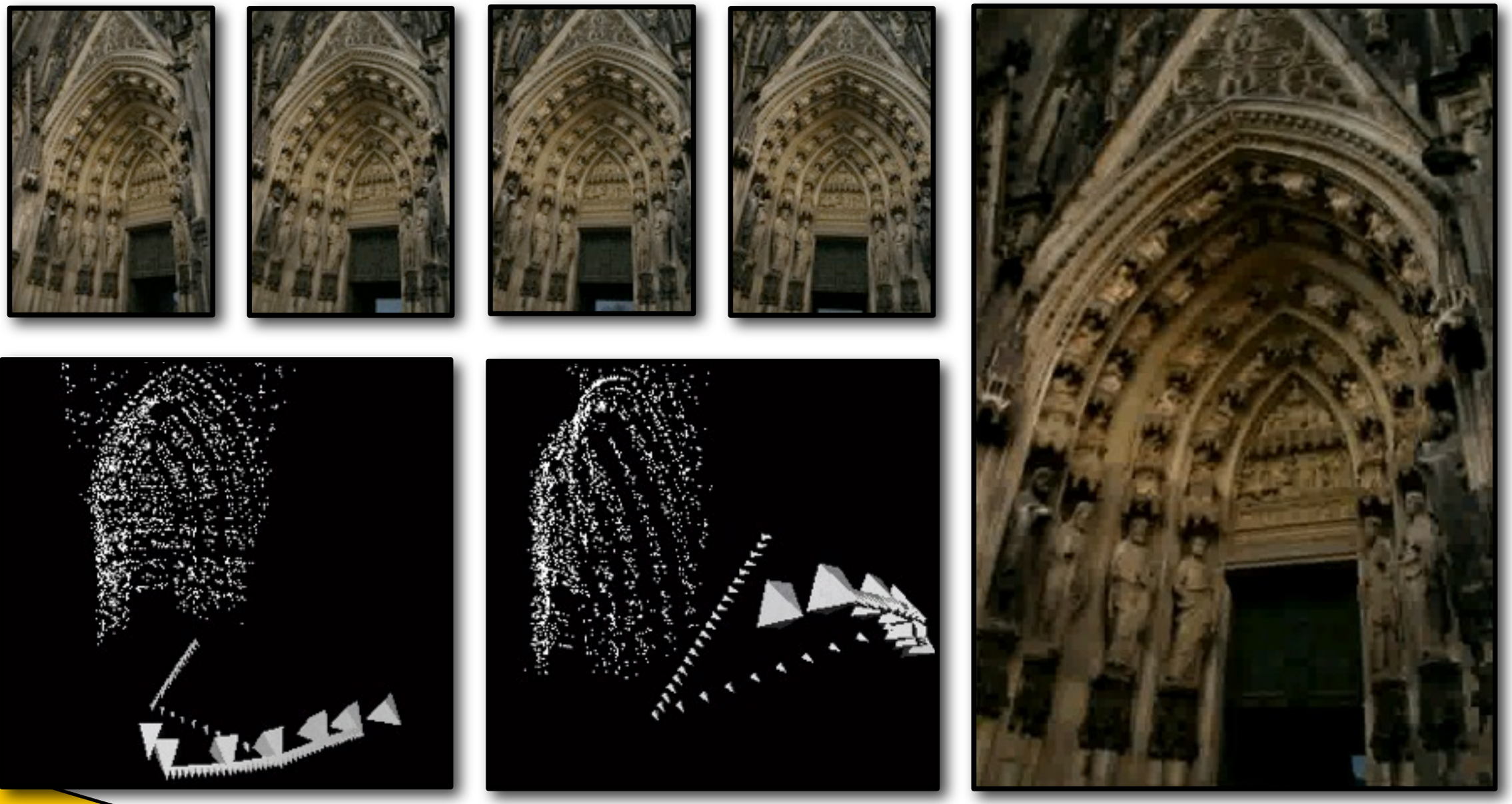
RANSAC

Refine Model Using Inliers



Find “average” translation vector

Wide Baseline Stereo

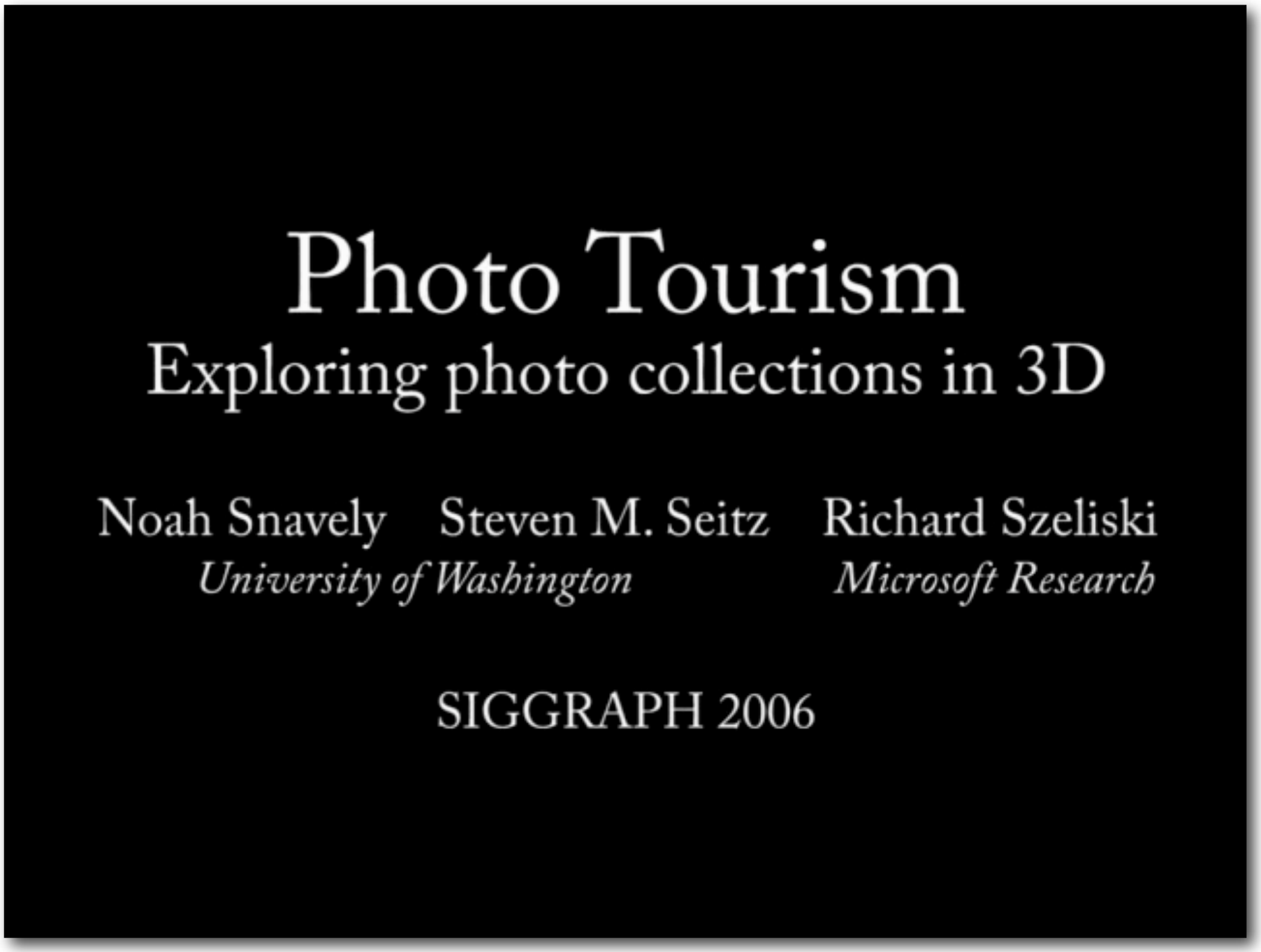


RoboJackets

Photosynth

http://photosynth.net/

Photosynth

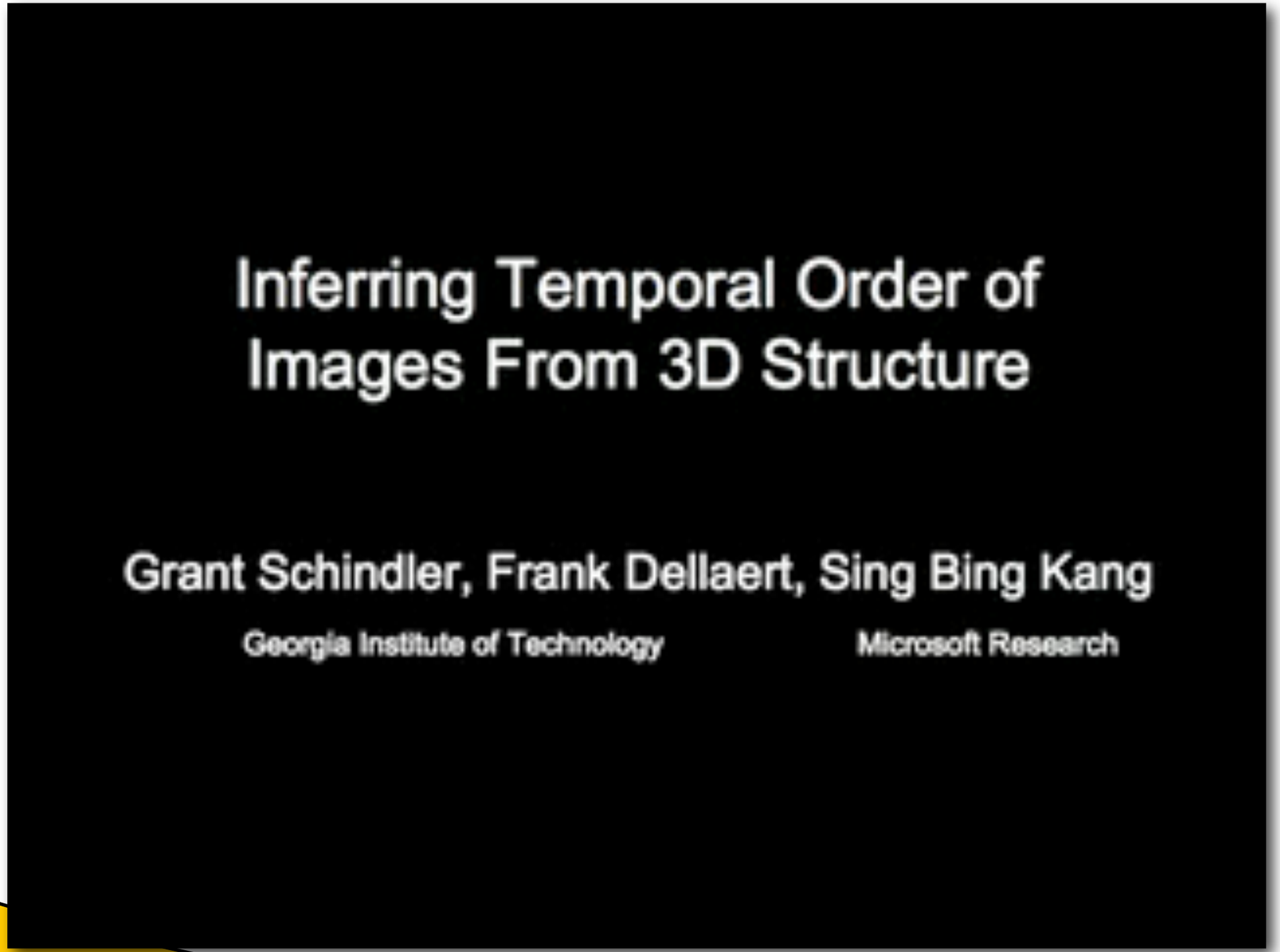


http://photosynth.net/

Building Rome in a Day

(from the Photosynth Guys)

4D Cities


Inferring Temporal Order of
Images From 3D Structure
Grant Schindler, Frank Dellaert, Sing Bing Kang
Georgia Institute of Technology
Microsoft Research

Mosaicking

Techniques Used

- Feature Detection and Matching
- Model Fitting
  - Expectation Maximization
  - Graph Cuts (optional)

RoboJackets

Images Courtesy Frank Dellaert

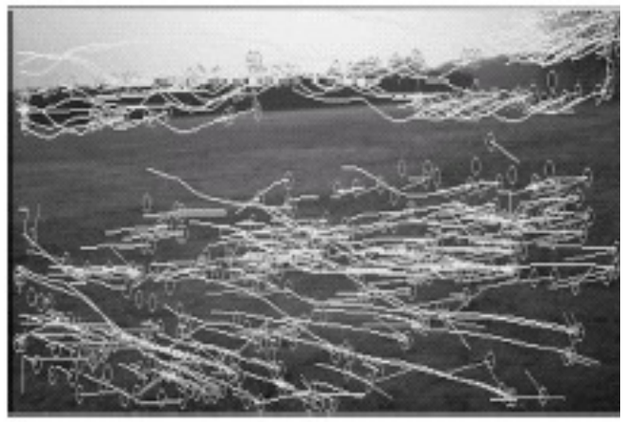
SLAM

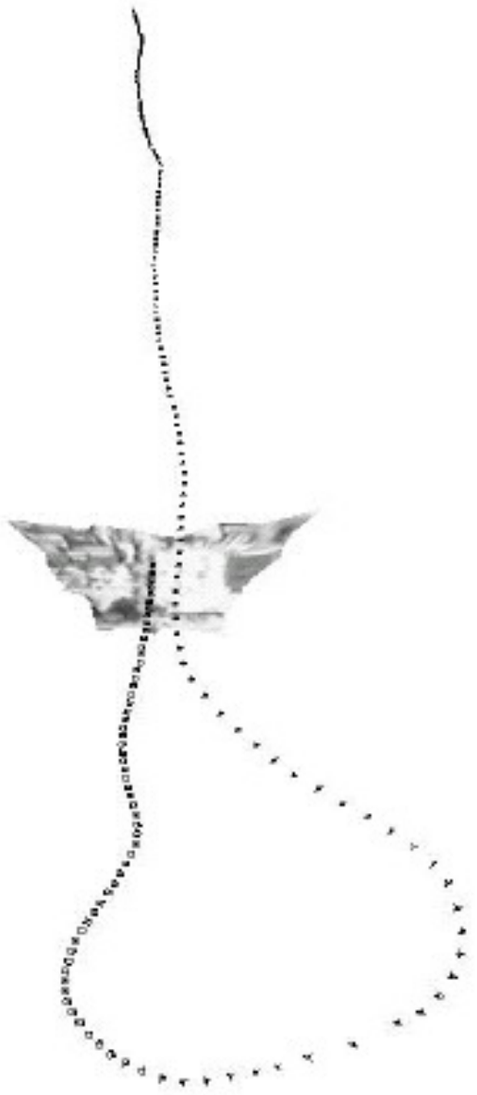
Self Localization and Mapping

Uses features and stereo equations to compute its location and map it's environment

Equations solved, features still need work.

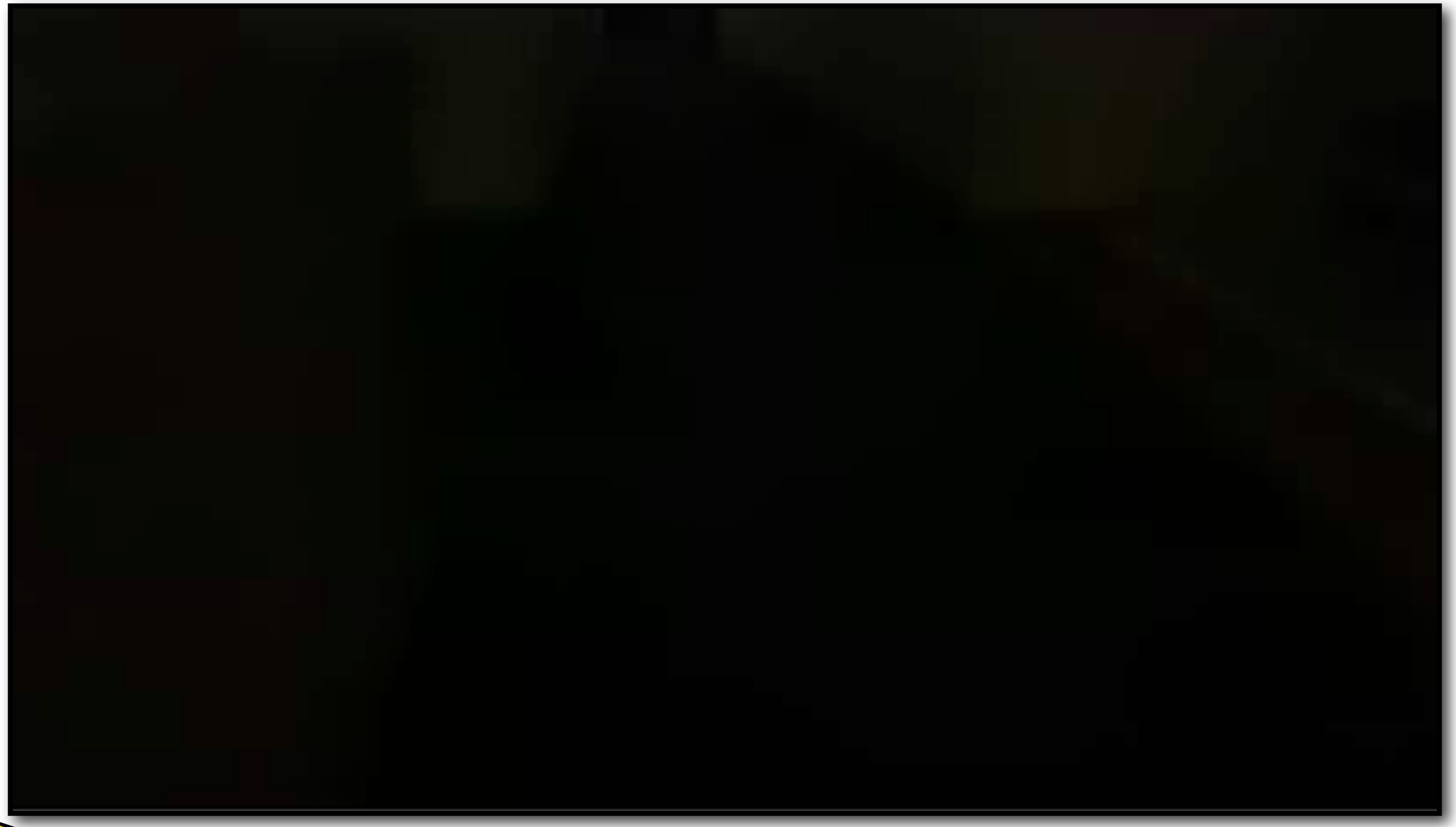
Without other sensors, no sense of scale.



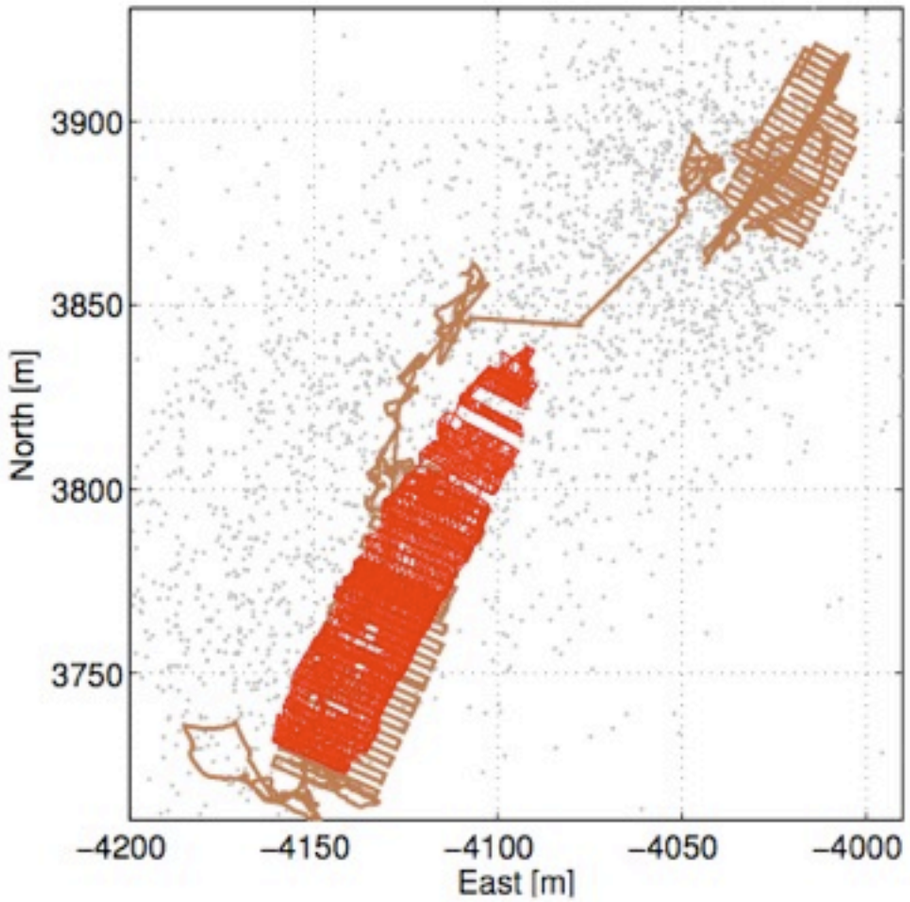


RoboJackets

Helicopter SLAM



Visually Navigating the RMS Titanic with SLAM Information Filters



Bullet Time!

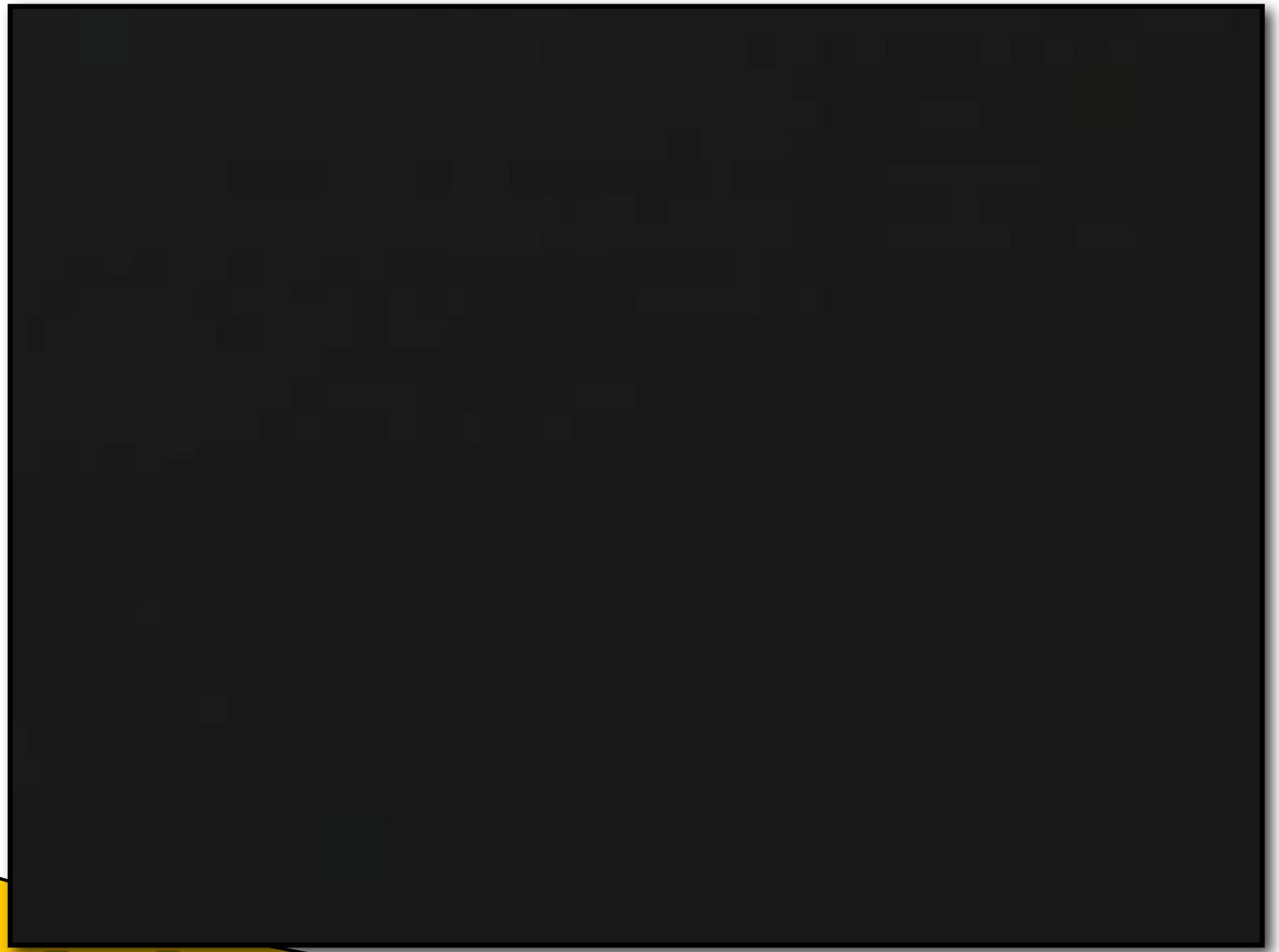
Video: The Matrix - Bullet Time

- Stereo
- View Morphing

RoboJackets

Image Courtesy Kang

Bullet Time

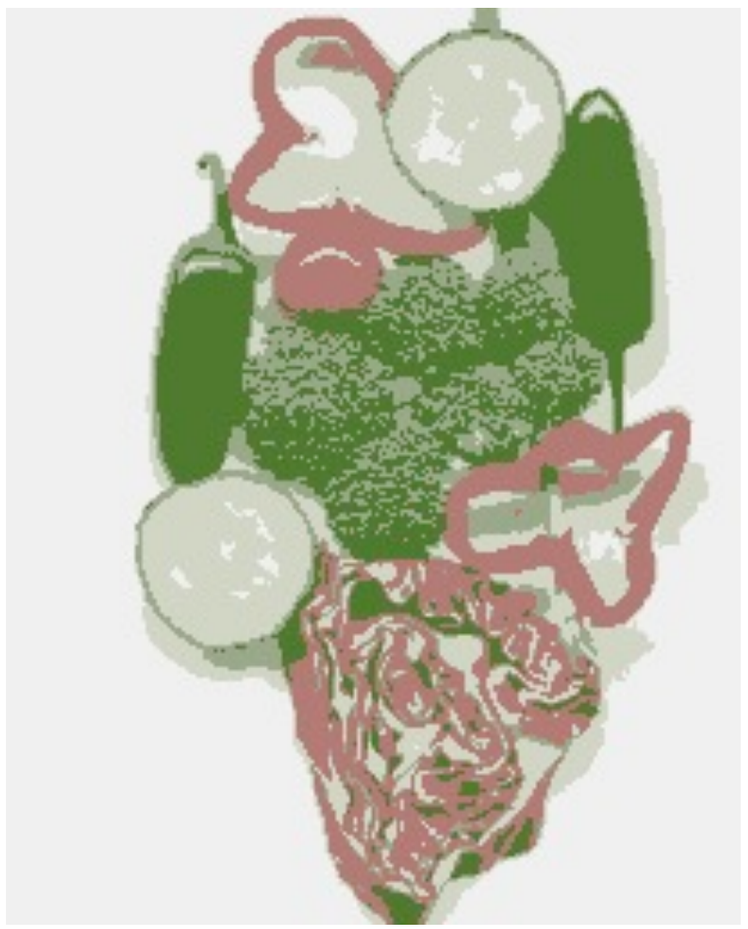


Clustering

Image

Clusters on intensity

Clusters on color



K-means clustering on intensity and color

K-Means Clustering

2006 Warner Independent Pictures

Algorithm

- Fix cluster centers and allocate points to closest center.
- Find centroid of clusters and recompute.
- Stop when no points change allegiances.

A Scanner Darkly

Rotoscoping

A Scanner Darkly

2006 Warner Independent Pictures

Done semiautomaticlly in the movie, techniques exist to do it automatically.

Techniques Used
- Linear Filtering
- Clustering
- Edge Detection

Projections

Catadioptric Camera
(latin for mirror + lens)

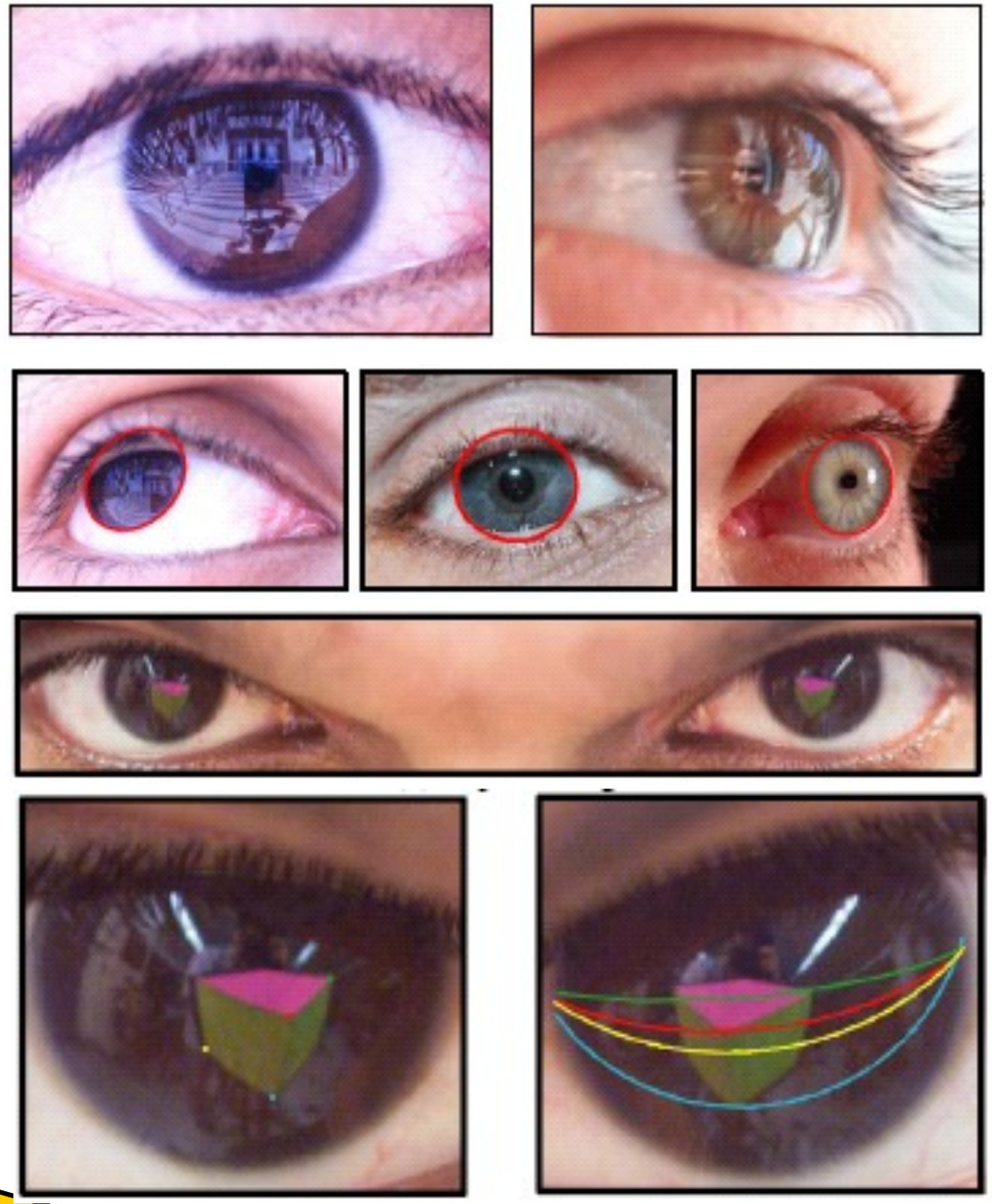
Catadioptric Imaging

Multiview Radial Catadioptric Imaging for Scene Capture

Sujit Kuthirummal
Shree K. Nayar

Columbia University

The World in an Eye



(a) Eye image (cropped)

(b) Spherical Panorama (cropped)

(c) Foveated retinal image (45° FOV)

(I) Interacting with a Computer

Plenoptic Imaging

Driving

Competitions
- DARPA Grand Challenge
- DARPA Urban Challenge
- IGVC
- LAGR

Applictions
- Structured and unstructured road following
- Lane detection
- Pedestrian detection avoidance
- Signal detection
- Cruise control
- Merge assistance
- Driver impairment detection

Video Google

Retrieved key-frames from three different shots