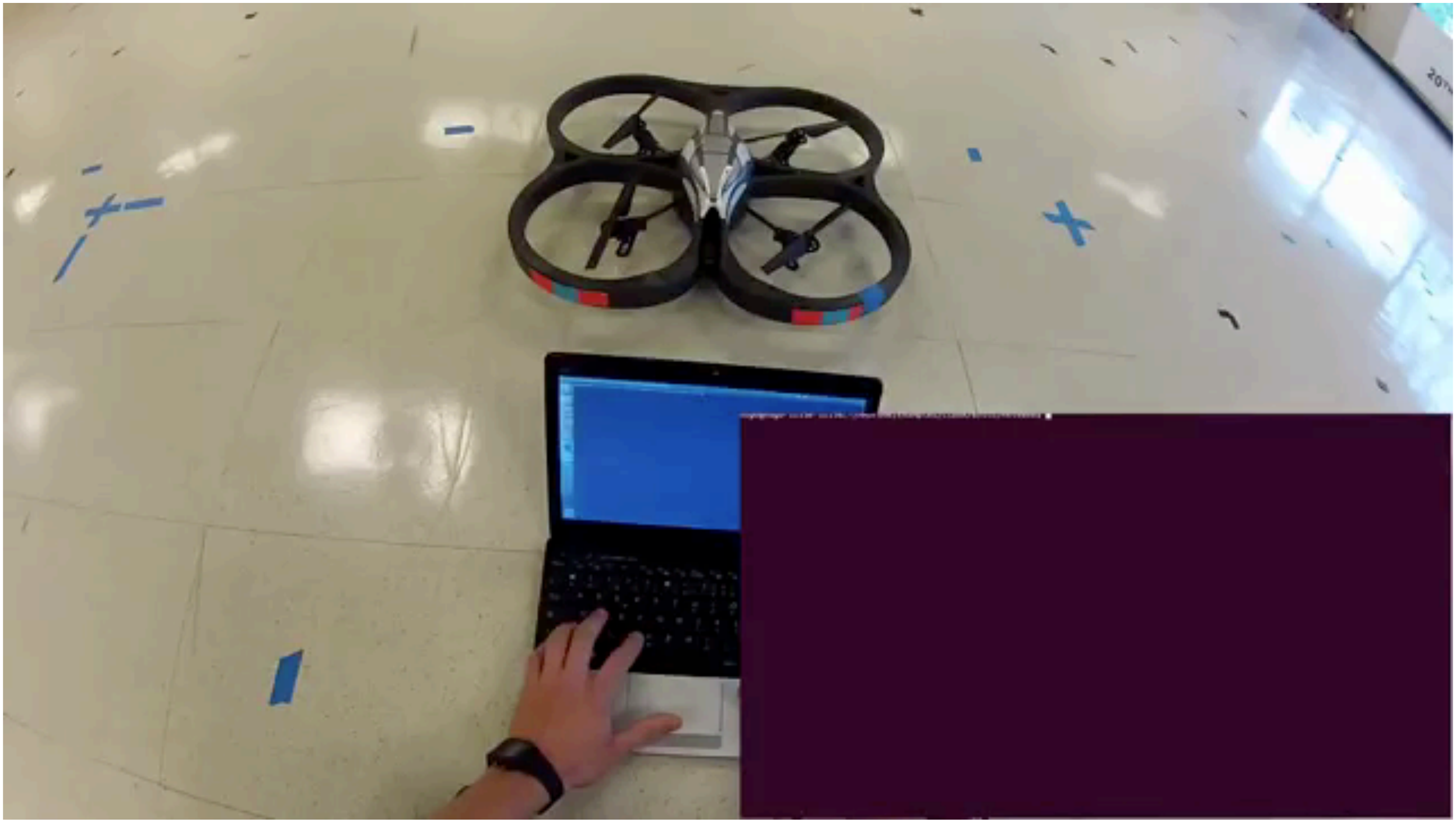
Video and Motion Analysis

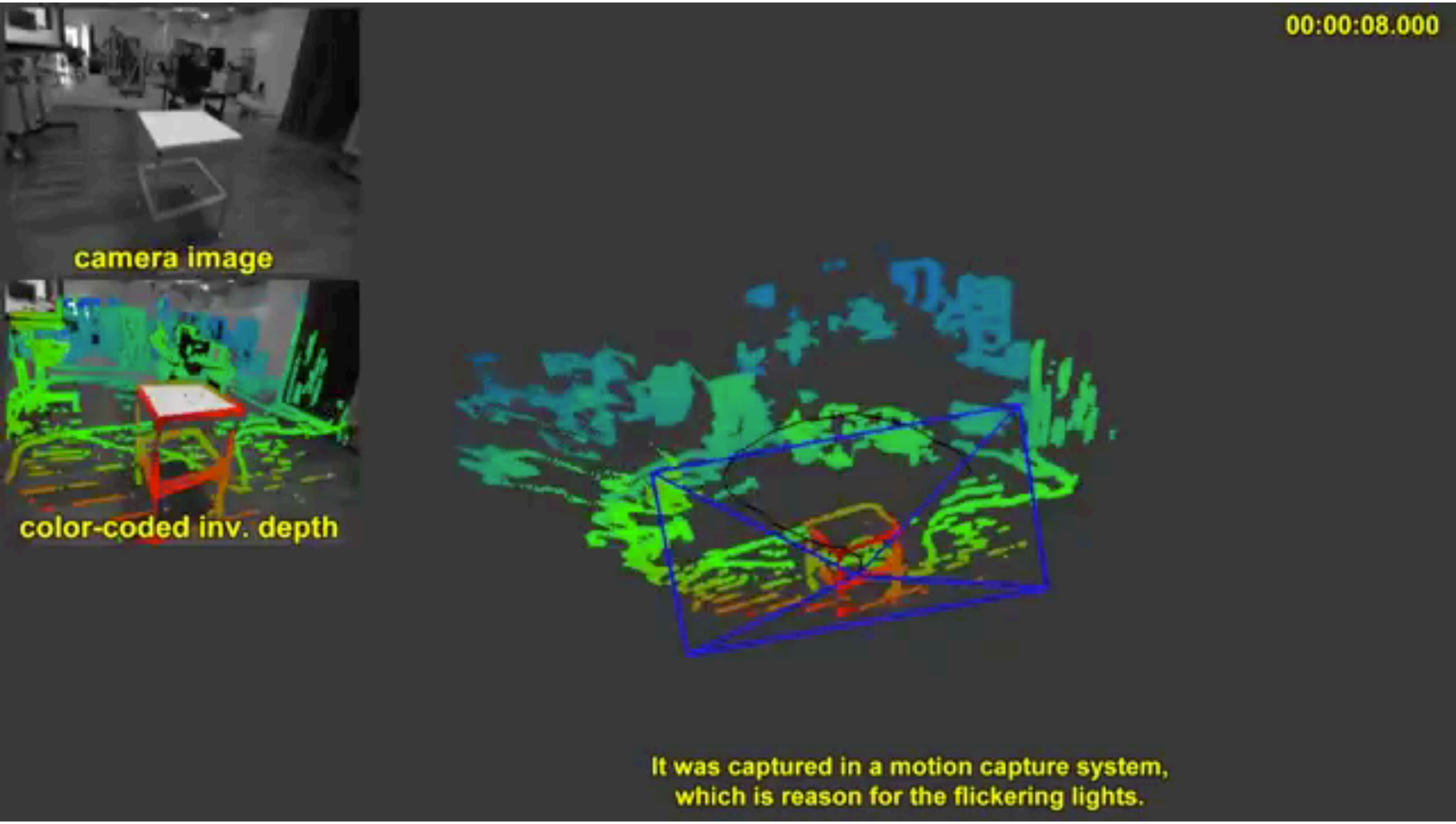
16-385 Computer Vision (Kris Kitani)

Carnegie Mellon University

Optical flow used for feature tracking on a drone



optical flow used for motion estimation in visual odometry



Roadmap

(Where we have been and where we are going)

Image filtering

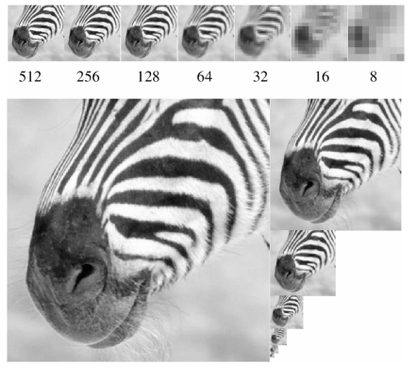


image pyramids

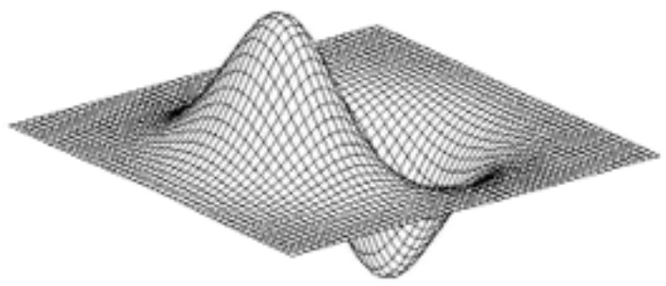
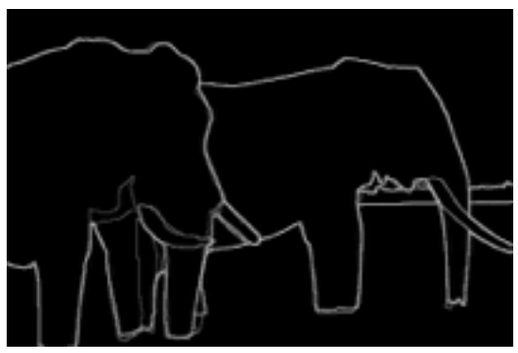
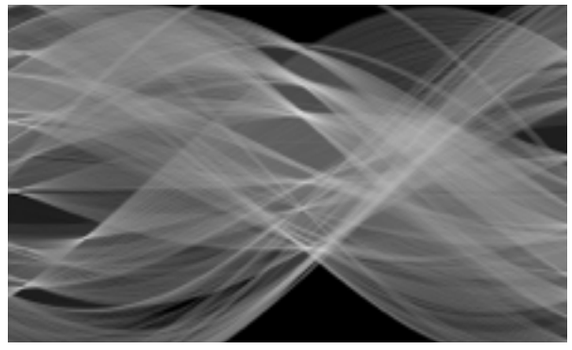


Image gradients



Boundaries



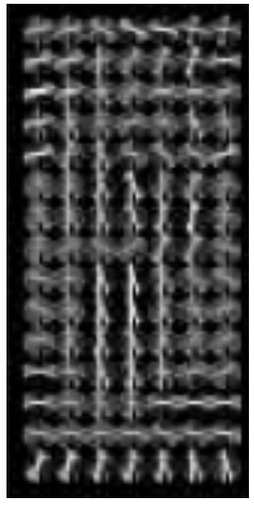
Hough Transform

Image Manipulation (January)

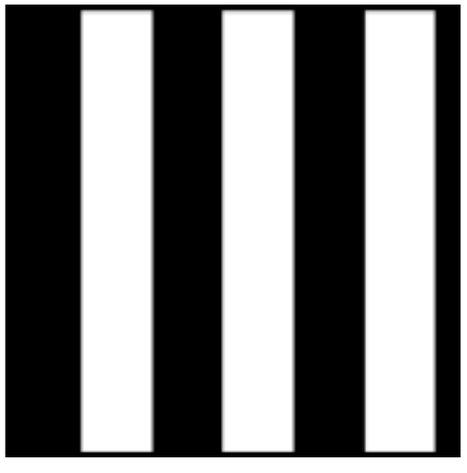
Corner detection

Multi-scale detection

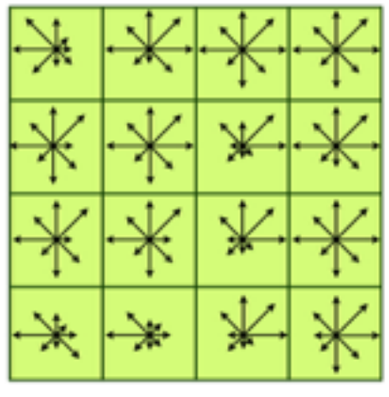
Haar-like



HOG



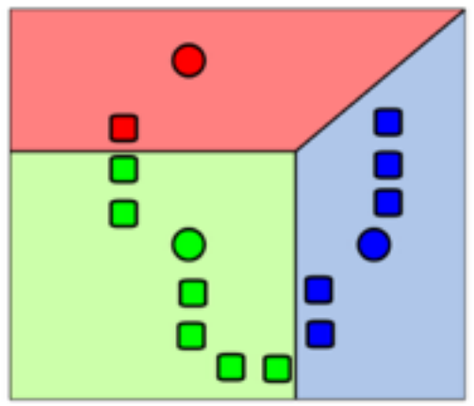
SURF



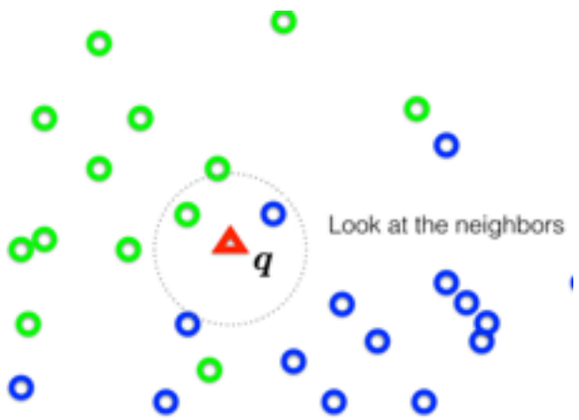
SIFT

Image Features (February)

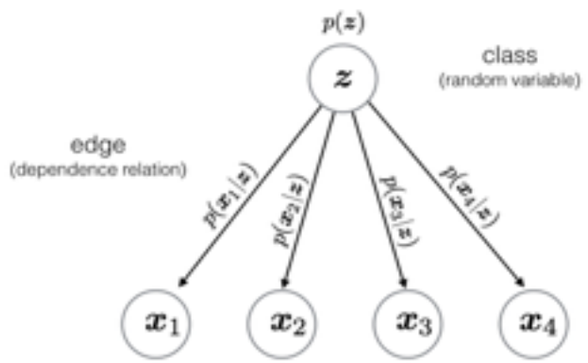
Bag-of-words



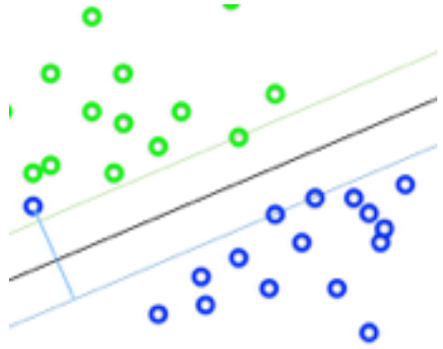
K-means



Nearest Neighbor

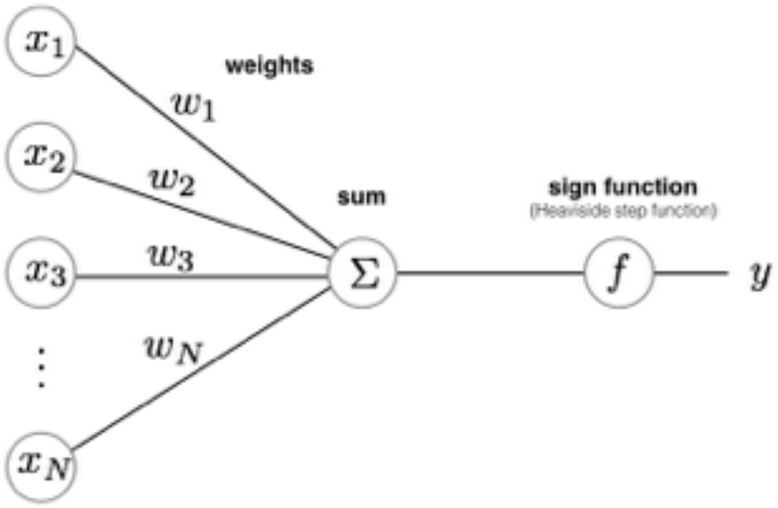


Naive Bayes

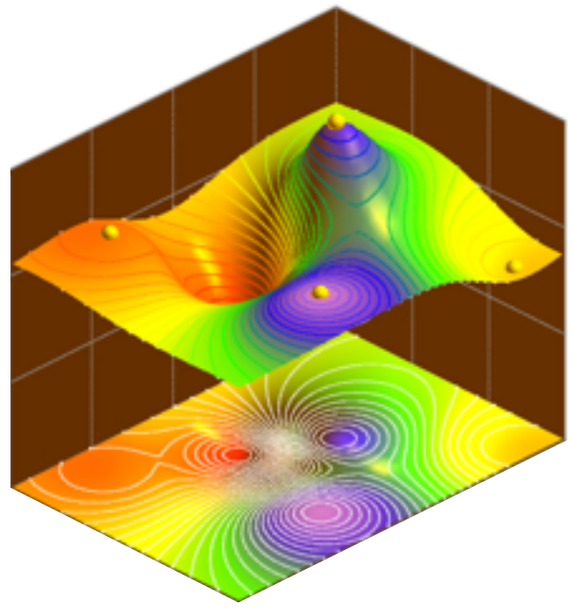


SVM

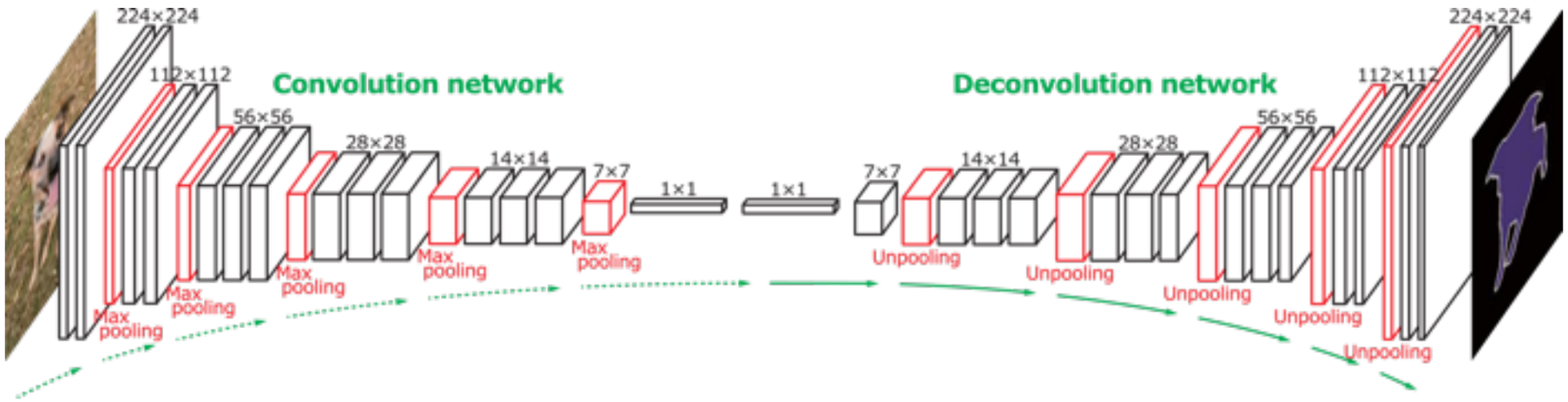
Object Recognition (February)



Perceptron



Gradient Decent



Convolutional Neural Networks (February)

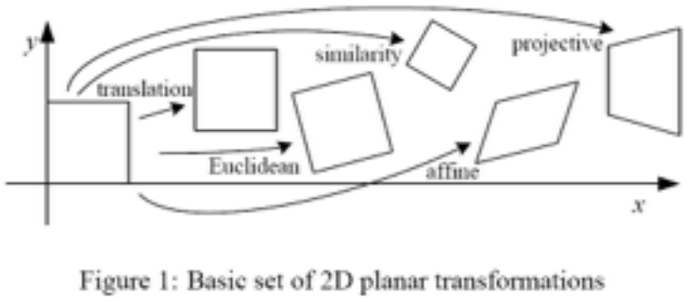
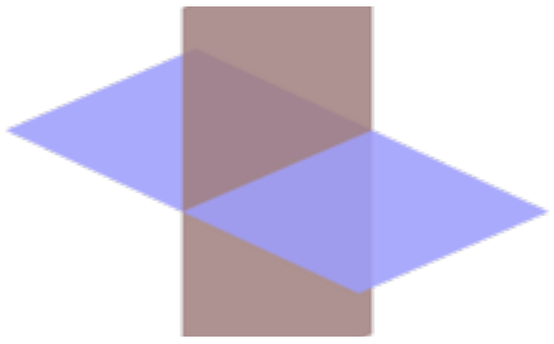


Figure 1: Basic set of 2D planar transformations

2D Transforms



DLT

RANSAC

$$H$$

Homography

2D Alignment (March)

$$x = \mathbf{P}\mathbf{X}$$

camera matrix

$$\mathbf{P}$$

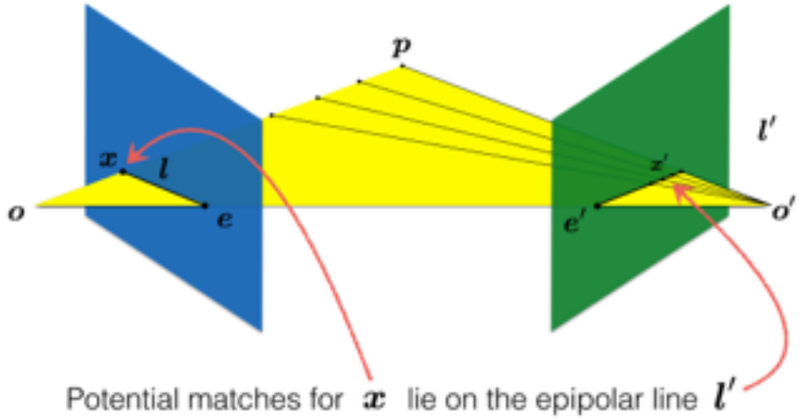
pose estimation

$$\mathbf{X}$$

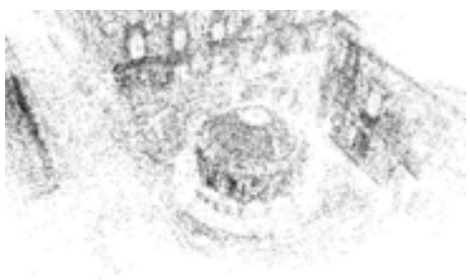
triangulation

$$\mathbf{F}$$

fundamental matrix

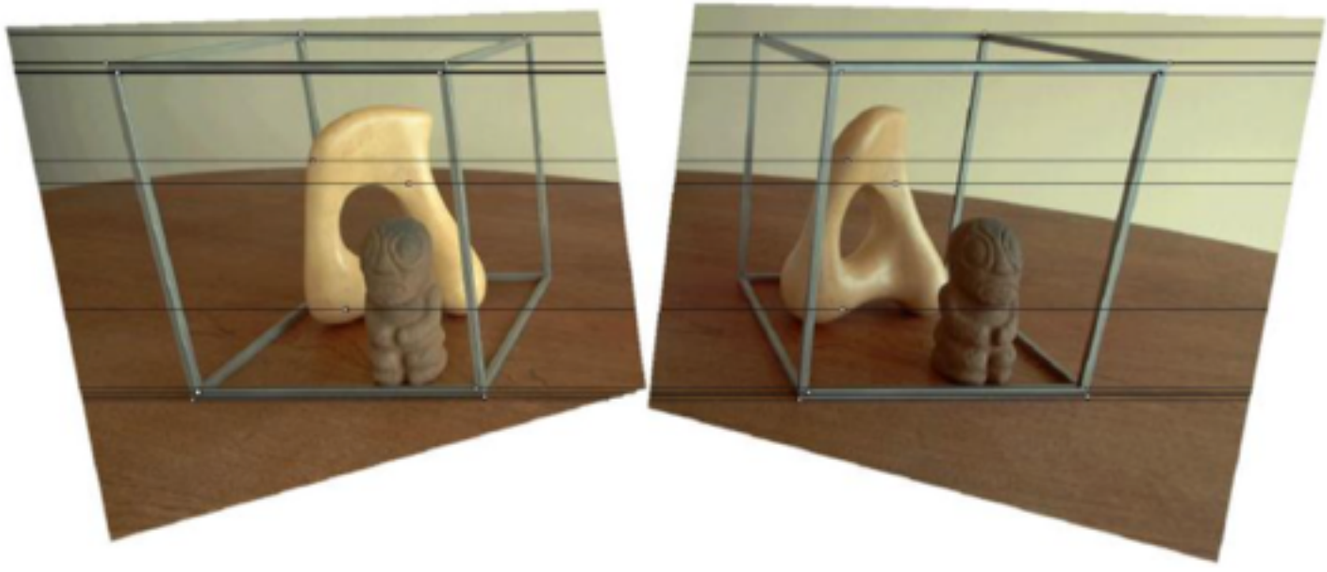


epipolar geometry

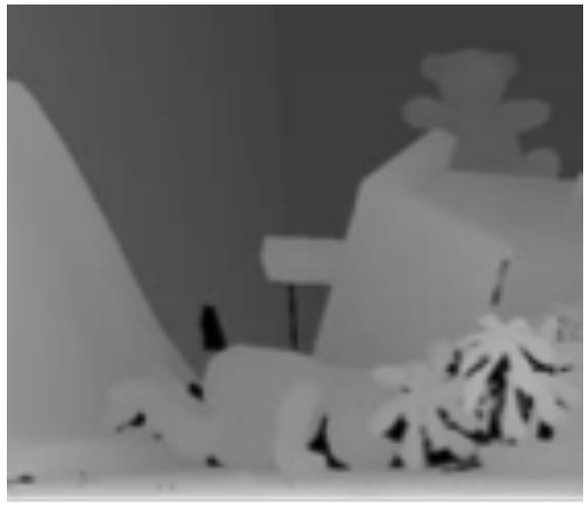


Reconstruction

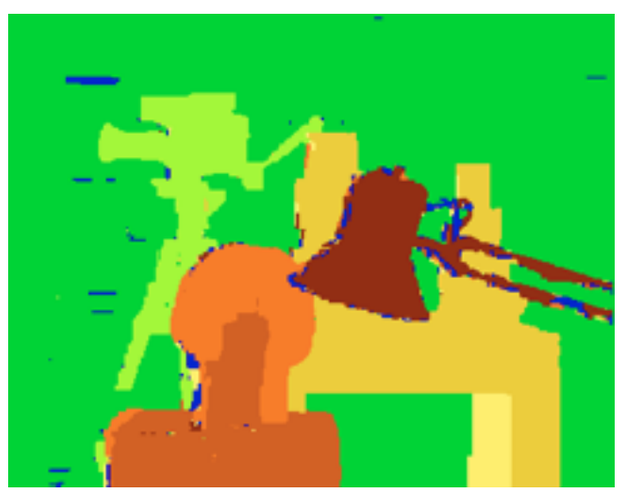
2 view geometry (March)



Stereo Rectification



Block matching



Energy minimization

Stereo (April)

What you can do now

- Detect lines (circles, shapes) in an image
- Recognize objects using a bag-of-words model
- Recognize using Deep Convolutional Neural Networks
- Automatic image warping (homography) and basic AR
- Reconstruct 3D scene structure from two images

What you will learn next

Computer Vision for Video

(a.k.a., working with sequential images)

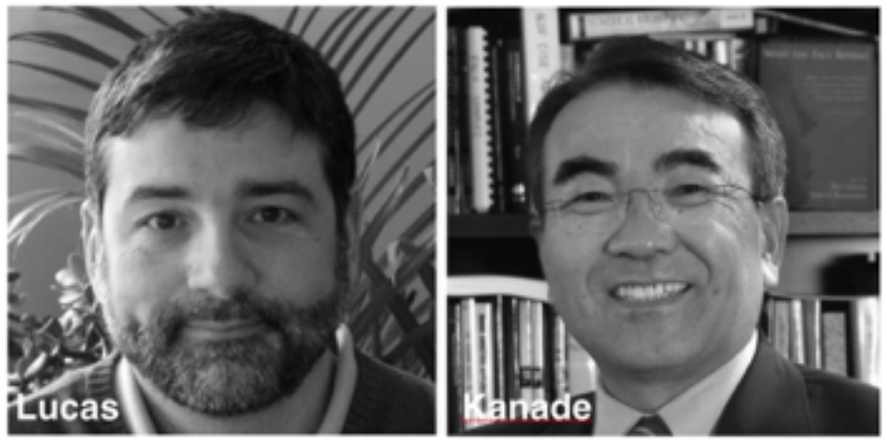
$$\begin{bmatrix} I_x(\boldsymbol{p}_1) & I_y(\boldsymbol{p}_1) \\ I_x(\boldsymbol{p}_2) & I_y(\boldsymbol{p}_2) \\ \vdots & \vdots \\ I_x(\boldsymbol{p}_{25}) & I_y(\boldsymbol{p}_{25}) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} I_t(\boldsymbol{p}_1) \\ I_t(\boldsymbol{p}_2) \\ \vdots \\ I_t(\boldsymbol{p}_{25}) \end{bmatrix}$$

$$\min_{\boldsymbol{u},\boldsymbol{v}} \sum_{ij} \left\{ E_d(i,j) + \lambda E_s(i,j) \right\}$$

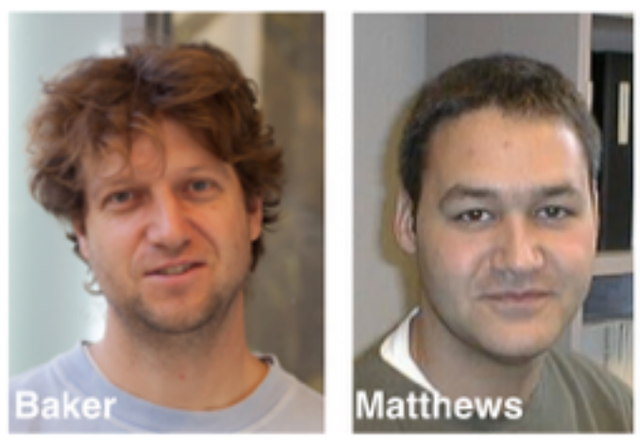
Constant Flow

Horn-Schunck

Optical Flow (April)



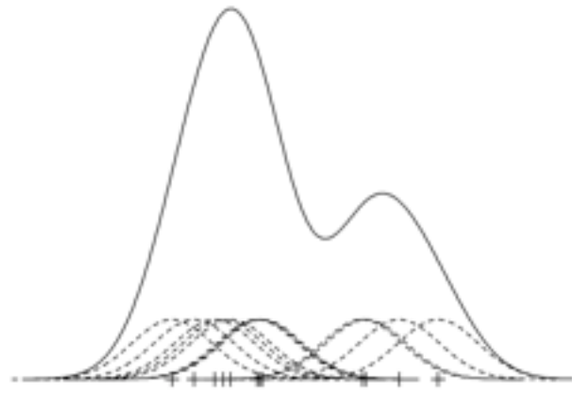
Lucas-Kanade
(Forward additive)



Baker-Matthews
(Inverse Compositional)

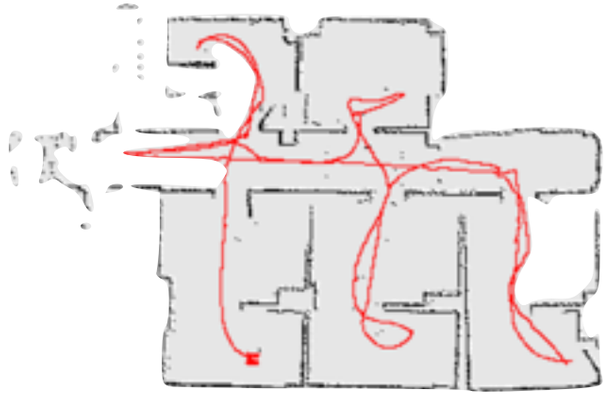
Image Alignment (April)

KLT



Mean shift

Kalman Filtering



SLAM

Tracking in Video (April)