Computer Vision

Lu Yang

Email: yanglu@uestc.edu.cn

School of Automation Engineering, UESTC

Reference Material/1

- Text Book

  -Computer Vision-A Modern Approach

  David Forsyth & Jean Ponce

  http://luthuli.cs.uiuc.edu/~daf/book/book.html

  -Computer Vision: Algorithms and Applications

  Richard Szeliski

  http://szeliski.org/Book/

Reference Material/2

- Others

  Introduction Techniques for 3D Computer Vision, E. Trucco & A. Verri

  Vision Science, R. Palmer

  Multiple View Geometry, Hartley & Zisserman

  The Geometry of Multiple Images, Faugeras

- Computer Vision Home Page

  http://www.cs.cmu.edu/~cil/vision.html

- Compendium of vision:

  http://homepages.inf.ed.ac.uk/rbf/CVonline/

- IEEE/Elsevier/Springer

  CVPR/ICCV/ECCV/TPAMI/IJCV/PR/PRL/CVIU/MVA

- Homeworks

  - 2 week period

  - Matlab implementation + problem

- Final Project

  - Literature search + implementation

- Matlab

  - Matlab primer

- Math

  -Basic math tools: Linear Algebra, Matrix, Basic optimization/calculus, probability(basic probability, density/estimation, Gaussian distributions).

Course Schedule/1

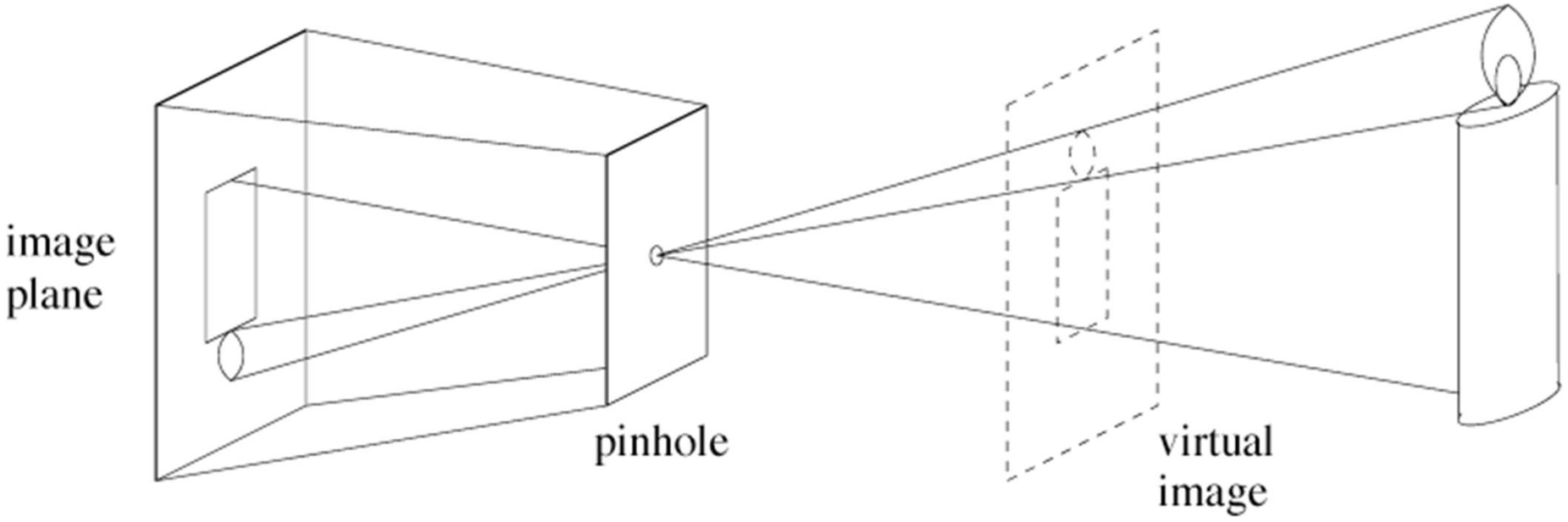
- Introduction (1)
- Image Formulation and Image Models(4)
  - Camera Geometric Model
  - Camera Calibration
- Basic Image Processing and Representation (6)
  - Linear Filtering
  - Edge extraction
  - Texture
  - Interesting Points

Course Schedule/2

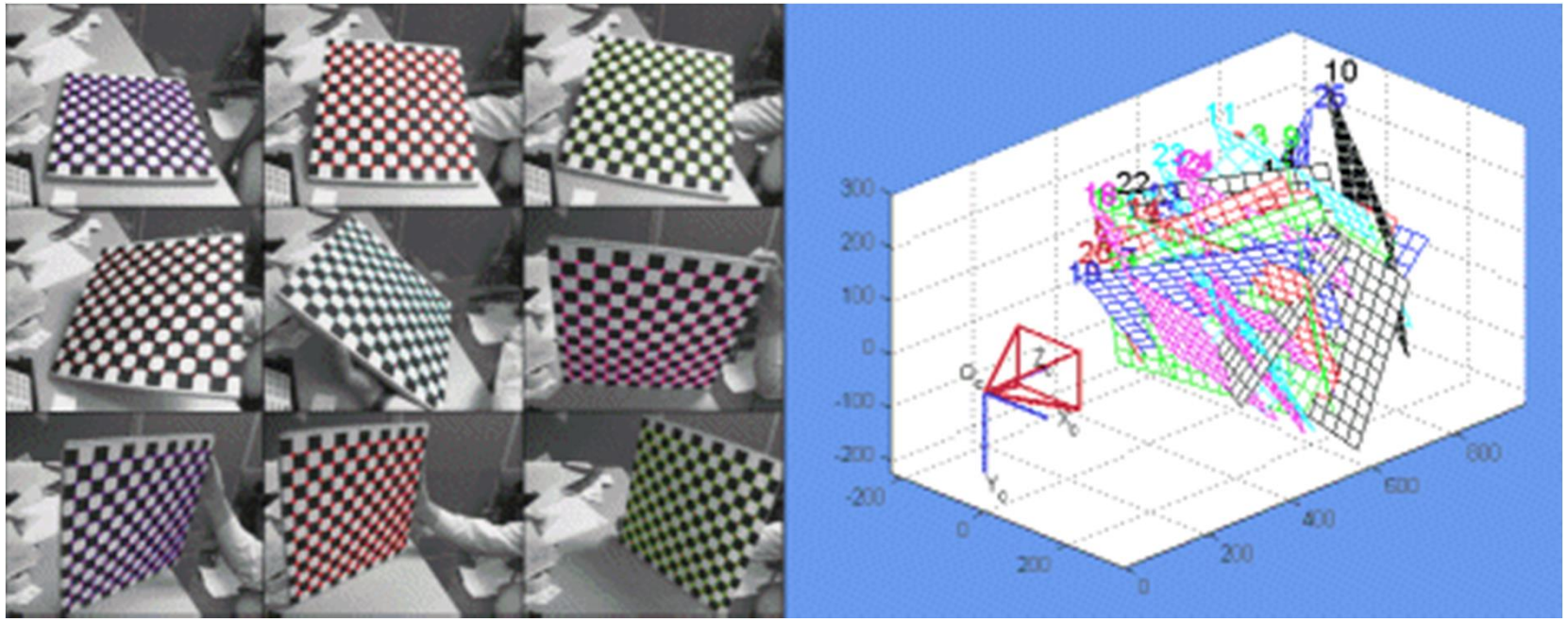
- Image sequence processing(15)
  - Multi-view Geometry; Stereo vision
  - Structure from motion; Optical flow
  - Object tracking
- Middle Level Computer Vision(6)
  - Clustering based image segmentation/fitting
  - Model based segmentation/fitting
  - Randomized algorithm based segmentation/fitting
- High level computer vision(8)
  - Template based object recognition
  - Bag-of-Feature based object recognition
  - Part based object recognition

Pinhole cameras

- Abstract camera model - box with a small hole in it
- Pinhole cameras work in practice

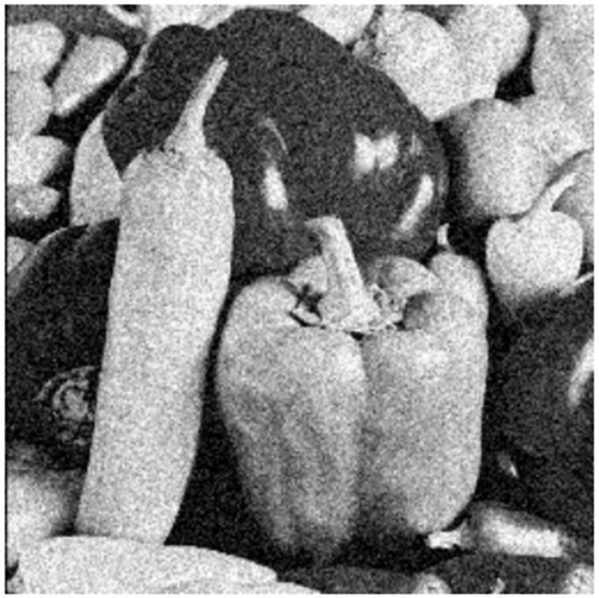


Camera Calibration



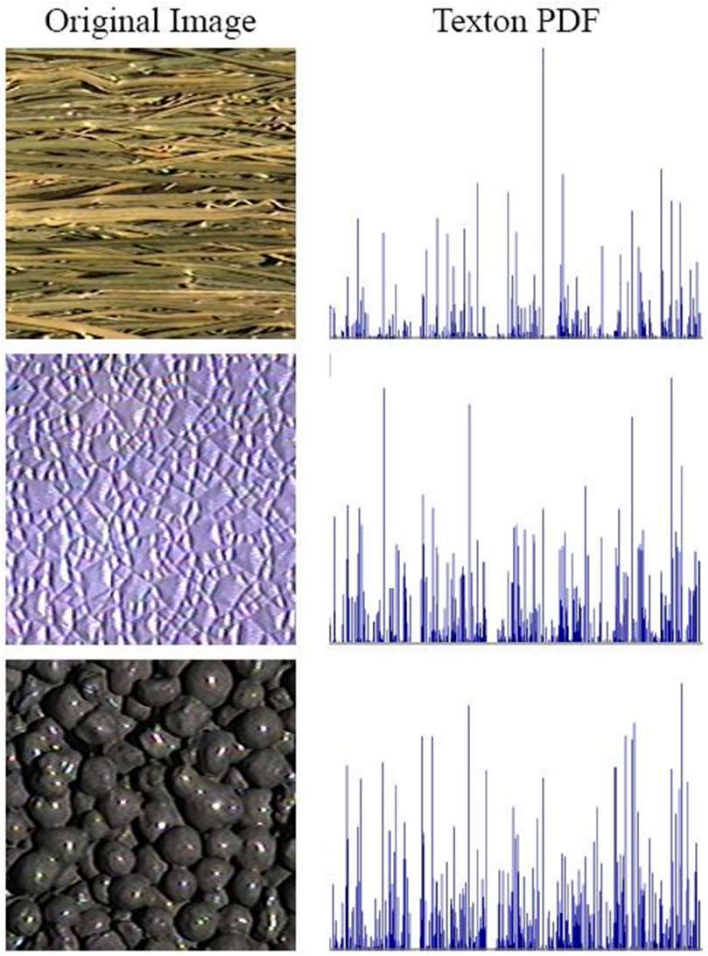
Matlab Camera Calibration Toolbox:
http://www.vision.caltech.edu/bouguetj/calib_doc/#system

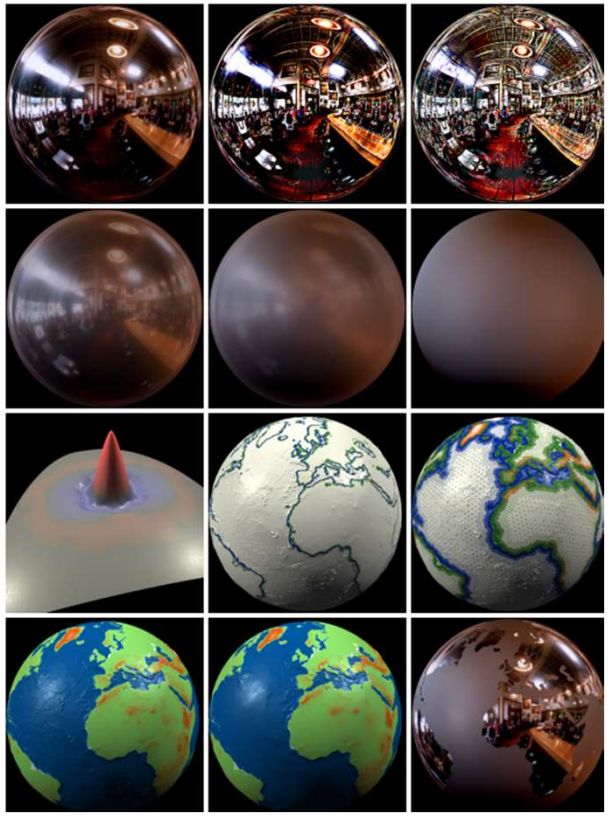
Image Filtering



Edge Detection

Texture



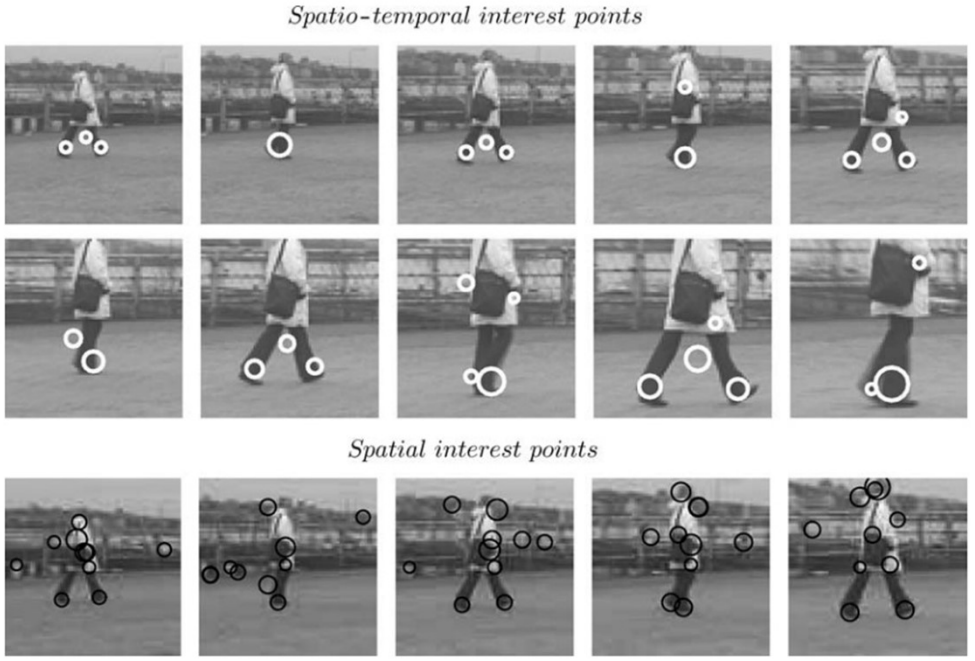


Spherical Texture

http://homepages.inf.ed.ac.uk/s0346435/projects/mrf/mrf_texture_project.htm

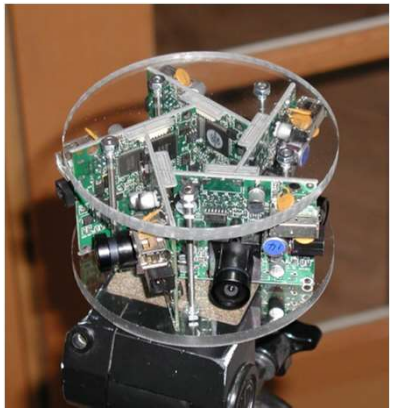
Interesting Points

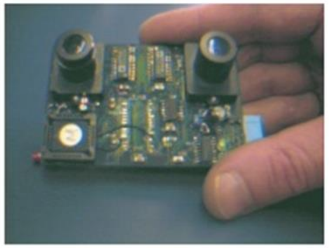
Harris Interest Points

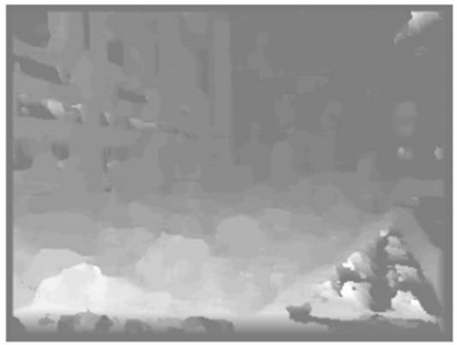


STIP, I. Laptev, IJCV 2005

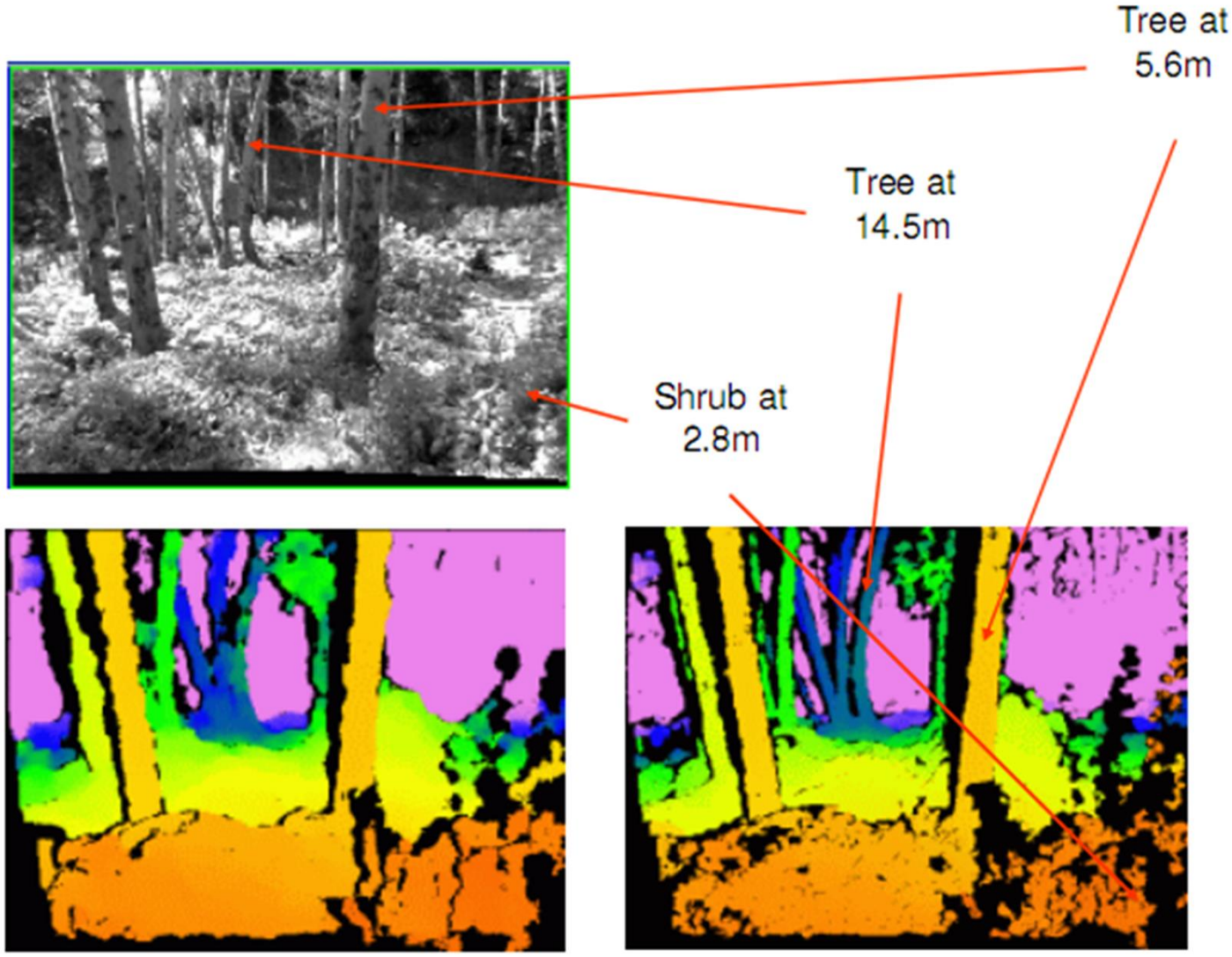
Multi-Camera Geometry/1







Multi-Camera Geometry/2



Multi-Camera Geometry/3

Background replacement using z-keying with a bi-layer segmentation algorithm(Kolmogorov et al. 2006)

Multi-Camera Geometry/4

Multi-Camera Geometry/5

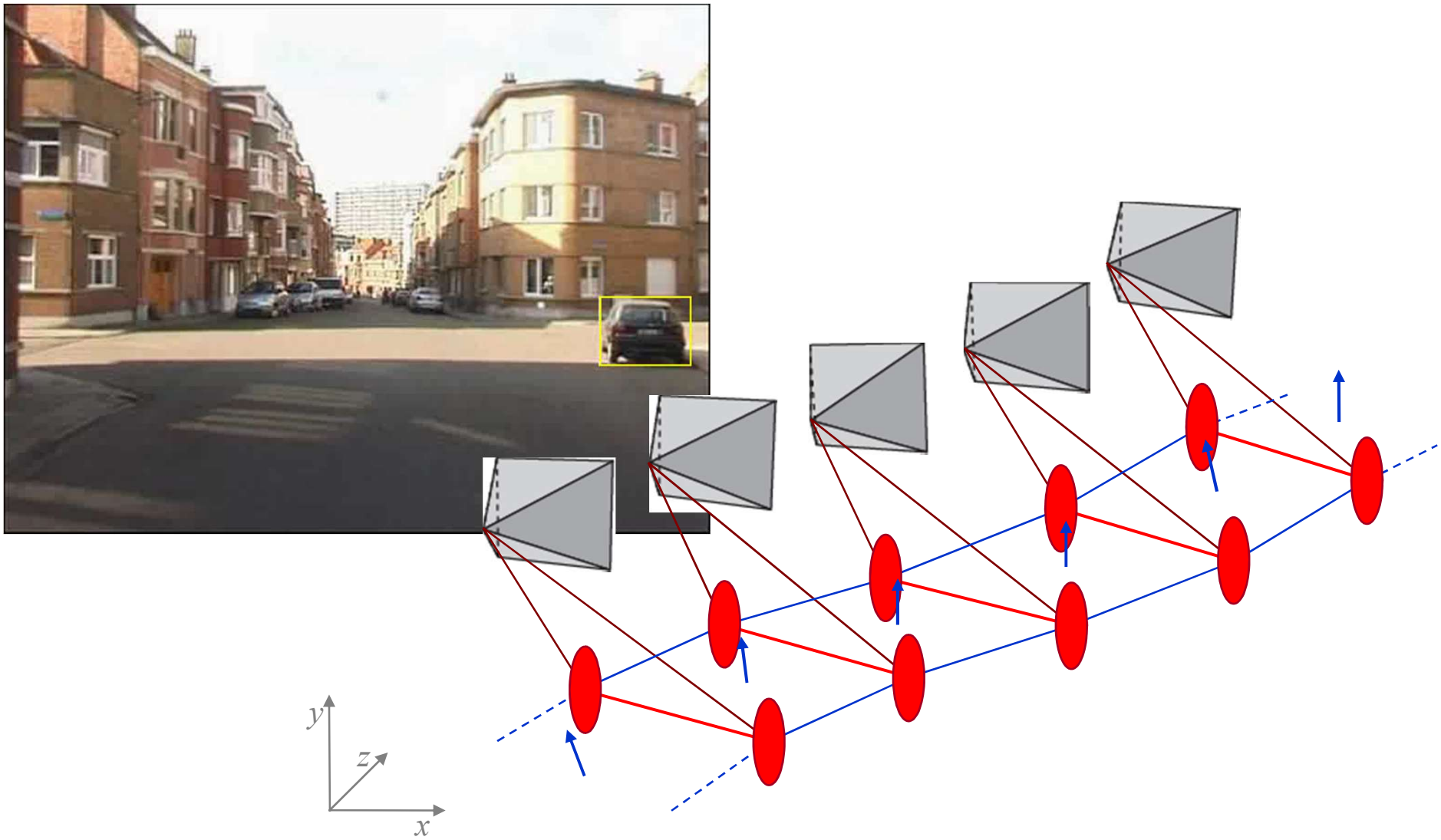
- Real-time human pose reocnition in parts from single depth images, CVPR 2011 best paper.

KINECT for XBOX 360.

Figure **Synthetic and real data**. Pairs of depth image and ground truth body parts. Note wide variety in pose, shape, clothing, and crop.

Figure **Example inferences.** Synthetic (top row); real (middle); failure modes (bottom). Left column: ground truth for a neutral pose as a reference. In each example we see the depth image, the inferred most likely body part labels, and the joint proposals show as front, right, and top views (overlaid on a depth point cloud). Only the most confident proposal for each joint above a fixed, shared threshold is shown.

Structure from Motion



http://www.vision.ee.ethz.ch/~bleibe/cvpr07/index.html

Tracking

Optical Flow

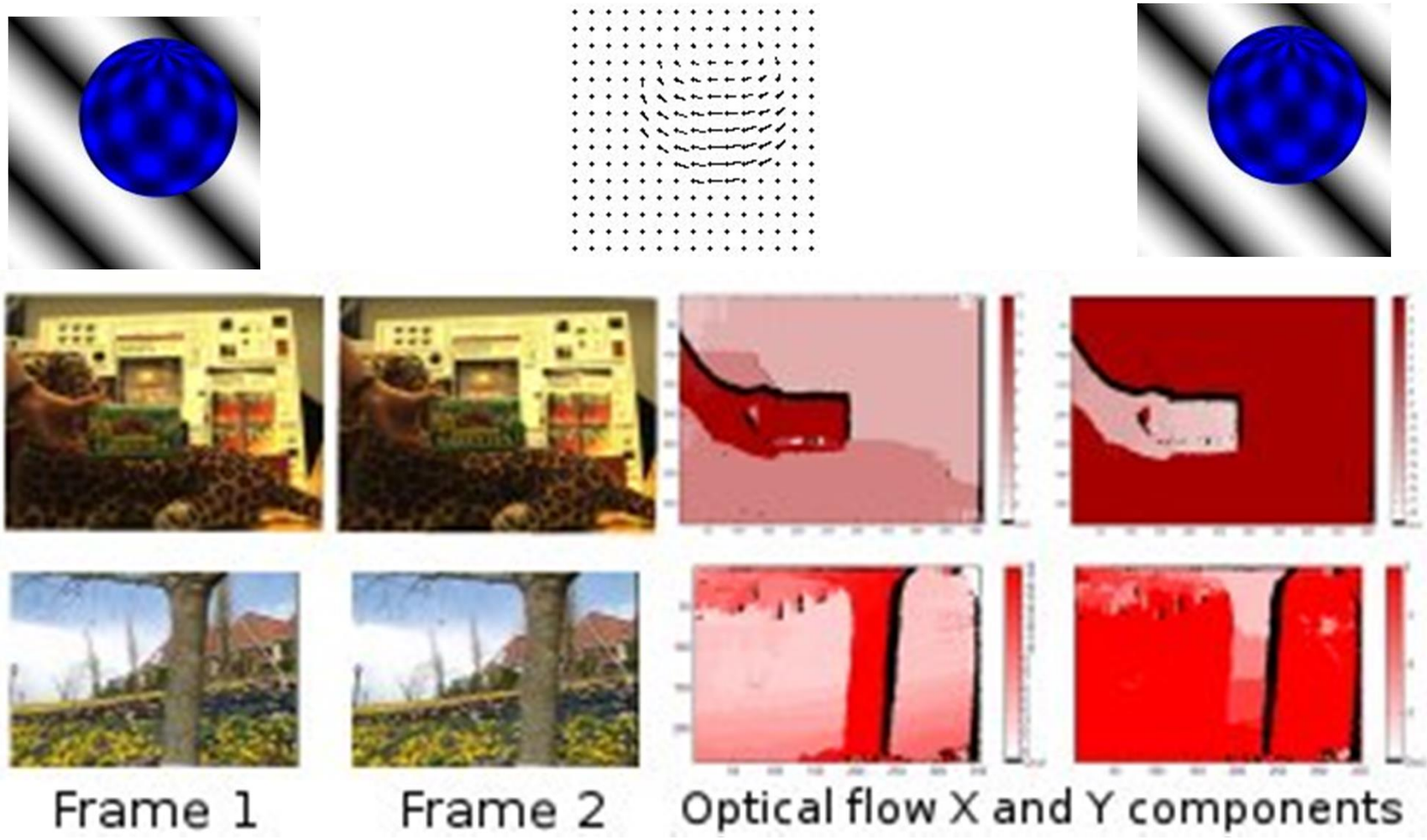
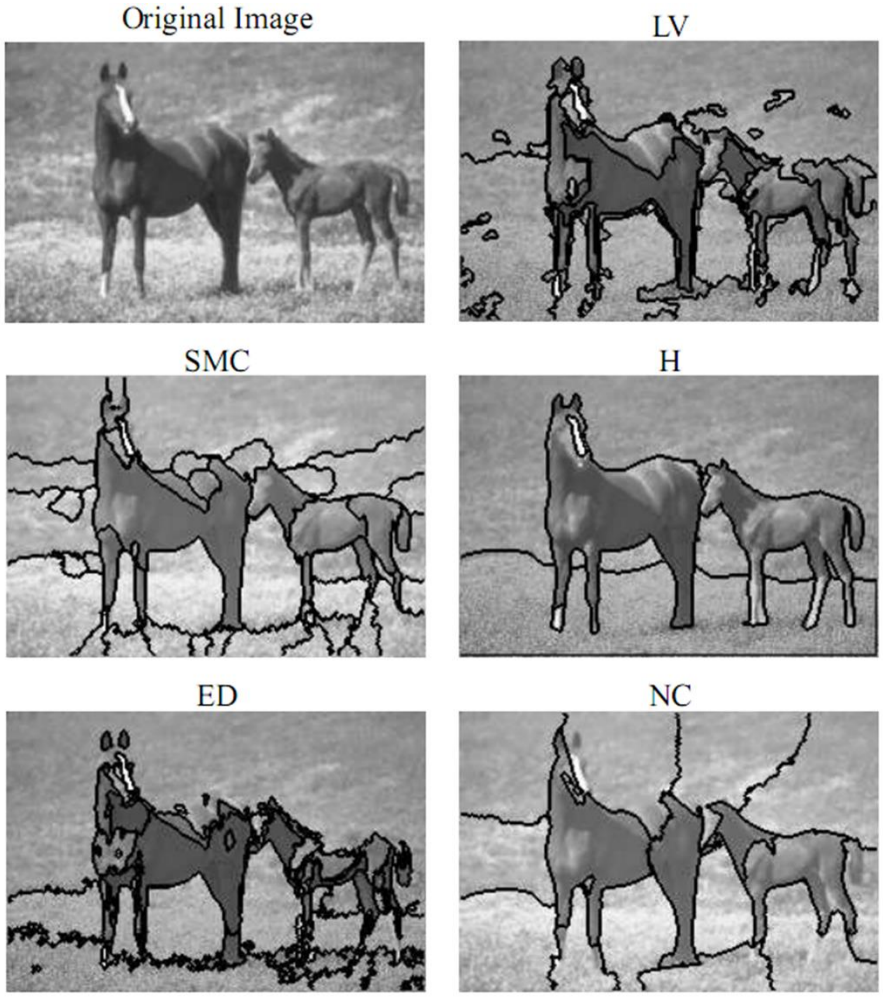
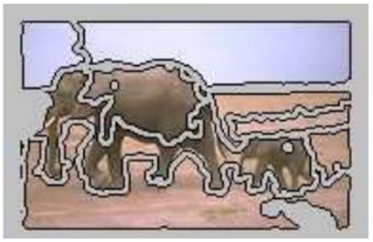
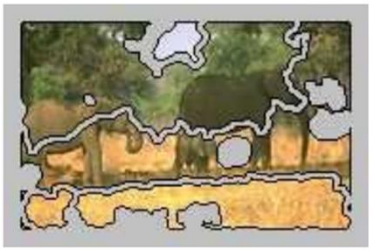


Image Segmentation/Fitting







LV: Local variation

SMC: Spectral min-cut

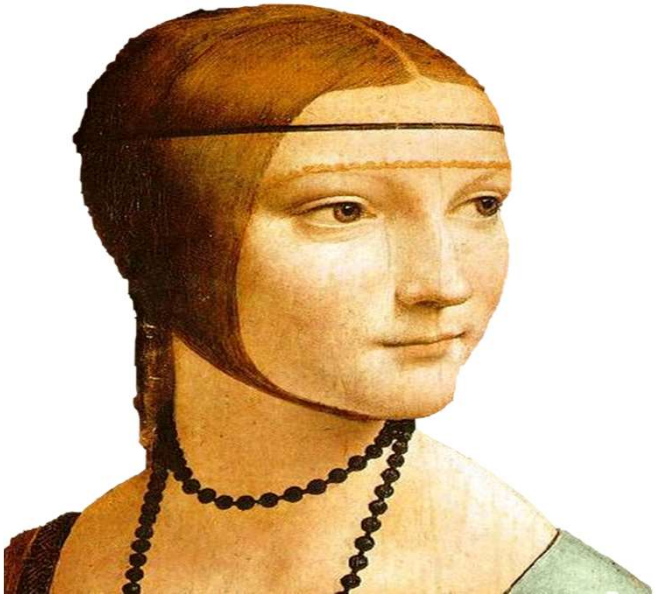
H: Human

ED: Edge augmented mean shift

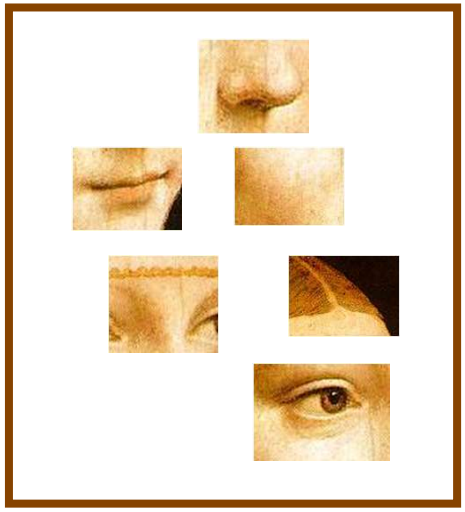
NC: Normalized cut

Object Recognition/1

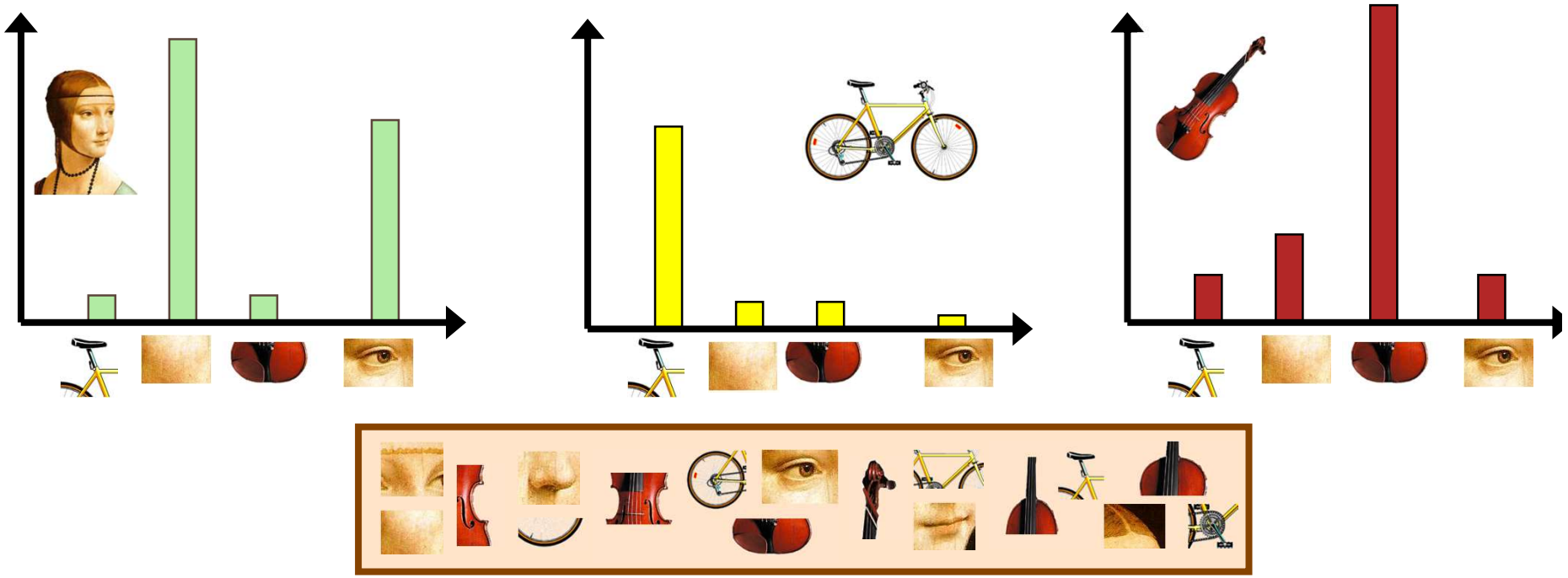
Object → **Bag of 'words'**

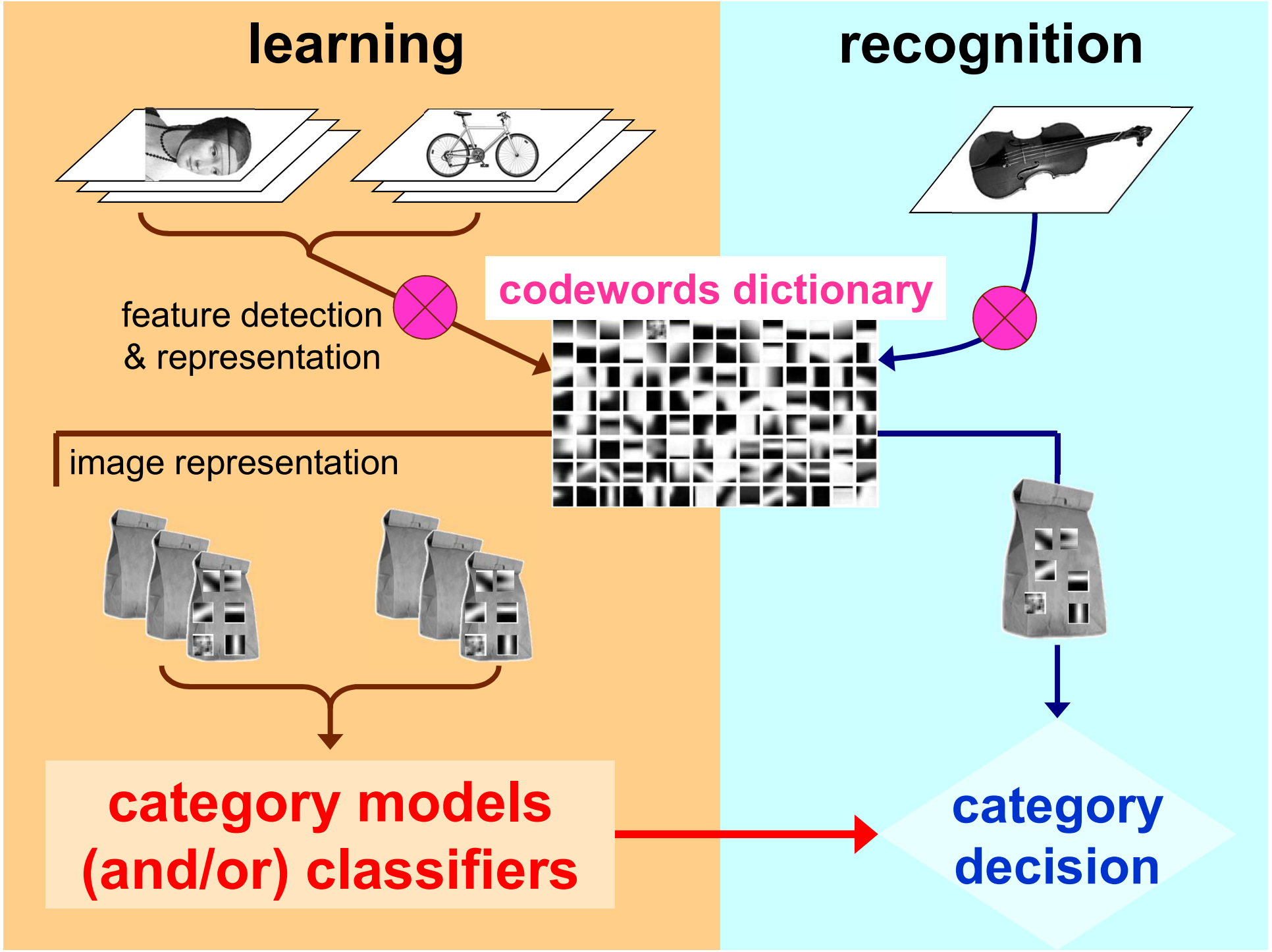


Object Recognition/2



Object Recognition/3




learning
recognition
feature detection
& representation
codewords dictionary
image representation
category models
(and/or) classifiers
category
decision

Vehicle Plate Recognition

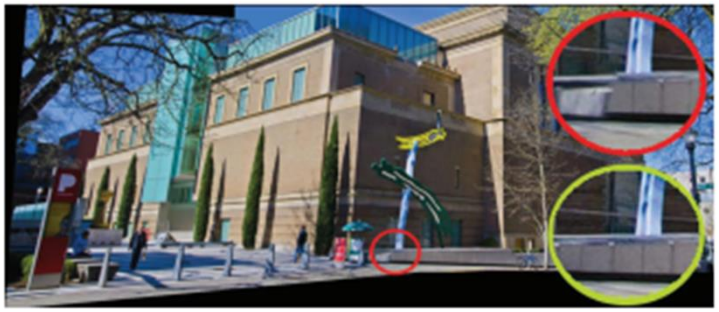
Detecting Texts from Videos/Images

Vision based HCI

Hand Trajectory Dataset: UESTC-HTD

Image stitching

(a) The original images from [3]



(b) The result using SEAM [4]

(c) The result using APAP [9]

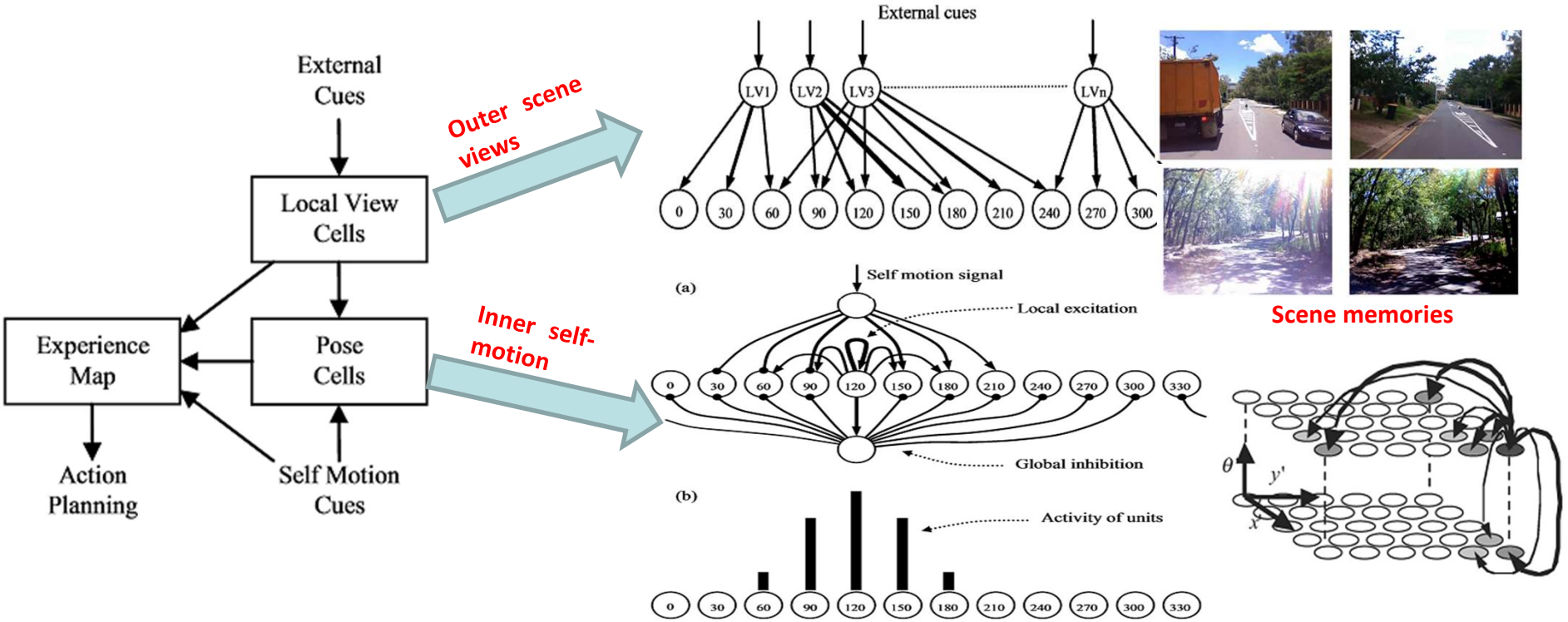
(d) The result using proposed approach

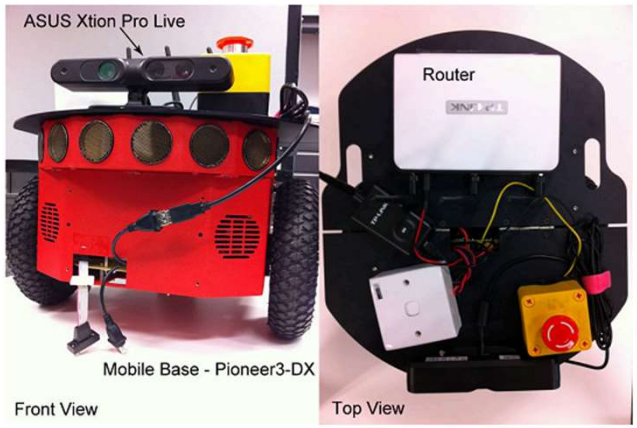
Deep Learning & Computer Vision

- **Direct Perception for Deep Driving - Cambridge**

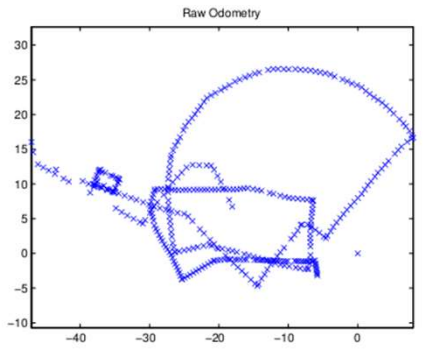
V. Badrinarayanan *et al.*, SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation, *arXiv:1511.00561v2*, 2015.

SLAM & Computer Vision





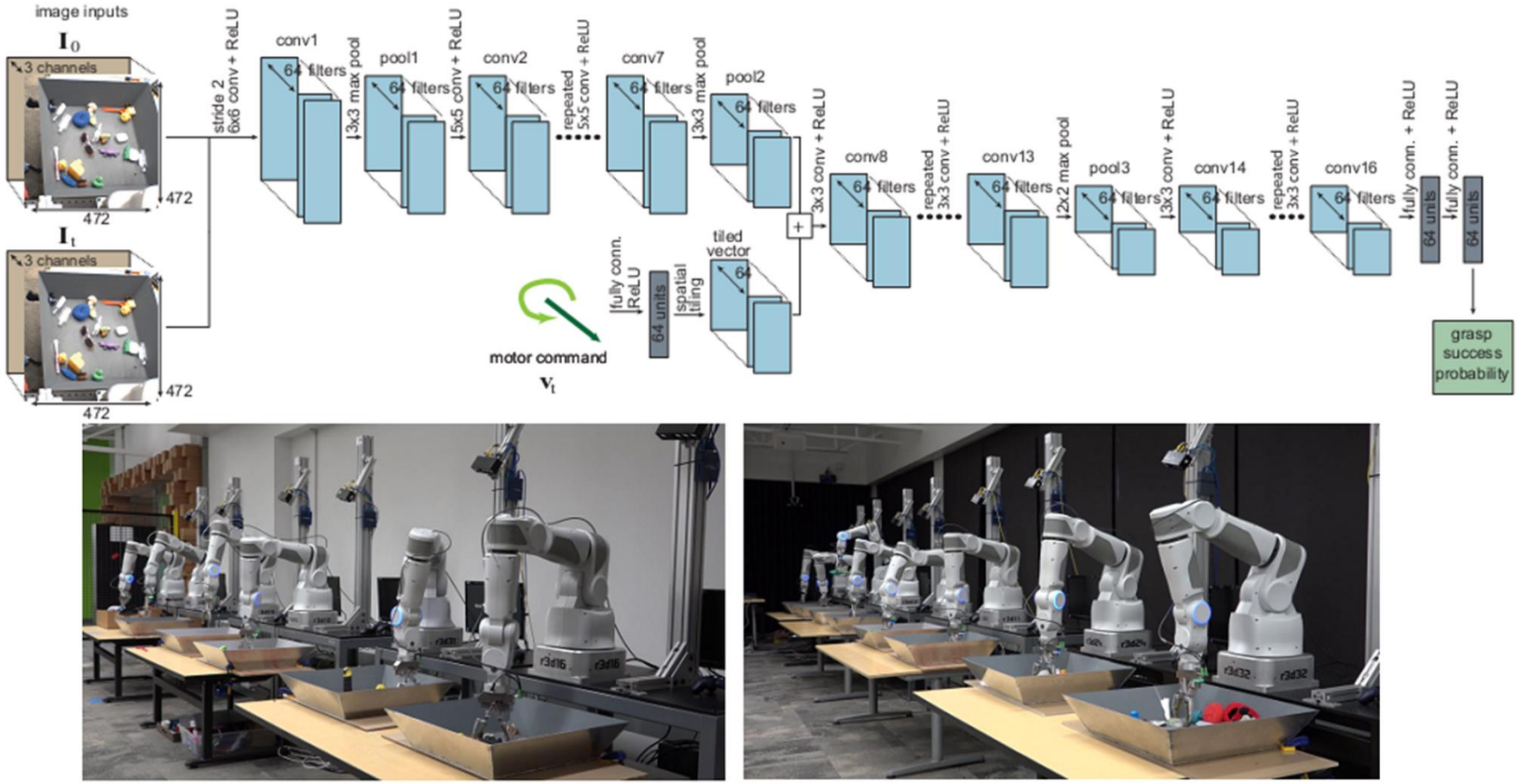
3D position and pose



M. Milford , G. Wyeth, Hippocampal Models for Simultaneous Localisation and Mapping on an Autonomous Robot, ACRA, 2003.
B. Tian, V. A. Shim, M. Yuan, C. Srinivasan, H. Tang, H. Li, RGB-D based Cognitive Map Building and Navigation, IEEE/RSJ IROS, 2013.

Deep Learning & Robotics

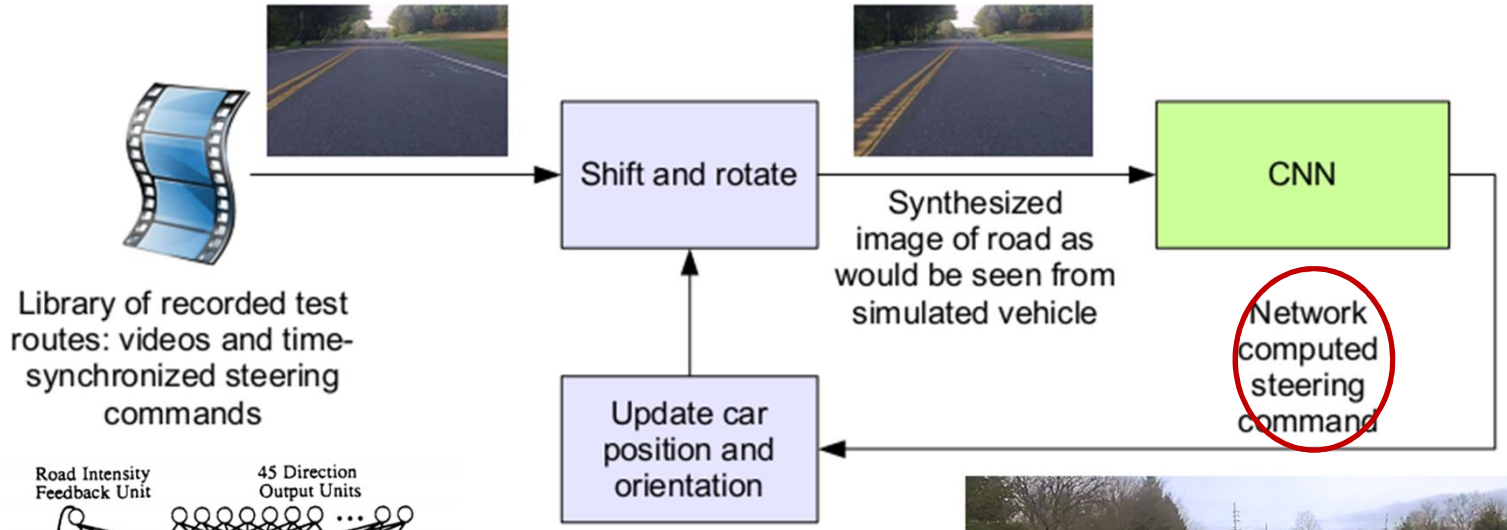
- **Deep Grasping - Google**

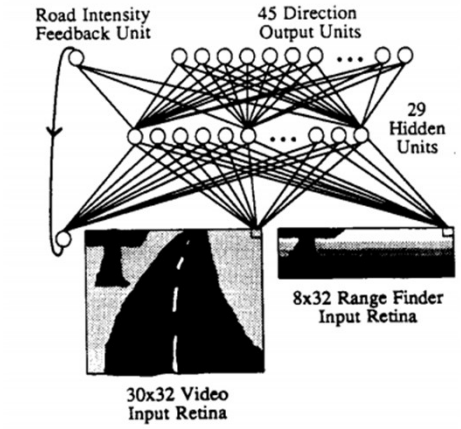


S. Levine *et al.*, Learning Hand-Eye Coordination for Robotic Grasping with Deep Learning and Large-Scale Data Collection, *arXiv:1603.02199v3*, 2016.

Deep Learning & Driving

- **Behavior Reflex - NVIDIA**





M. Bojarski *et al.*, End to End Learning for Self-Driving Cars, arXiv:1604.07316v1,2016.