



CS231A – Computer vision

Project proposals

- Types of Projects
- Where to Get Projects
- Project Proposal
- Project Ideas
- Where to find data

Types of Projects

- Replicate an interesting paper
- Comparing different methods to a test bed
- A new approach to an existing problem
- Original research

Where to Get Projects

- Project suggestions posted on Piazza
- Find an interesting vision paper and replicate it
 - CVPR, ECCV, ICCV
- Find a computer vision research project to join, email the professor or the grad students
 - Silvio Savarese, Fei-Fei Li, Sebastian Thrun, Daphne Koller
- Come up with your own ideas
- Questions?

Grading

- Course project: 40%
 - mid term progress report 5%
 - final report 30%
 - presentation 10%

Project Proposal

- Project Proposal Format
- - max 4 pages;
- - 3 sections:
 - * title and authors
 - * sec 1. intro: problem you want to solve and why
 - * sec 2. technical part: how do you propose to solve it?
 - * sec 3. milestones (dates and sub-goals)
 - * references
- - final format: pdf

To-Do

- Form a team of 1-3 people
- Decide on a project
- Write up project proposal (due next Thursday, 1/31)
- Questions?

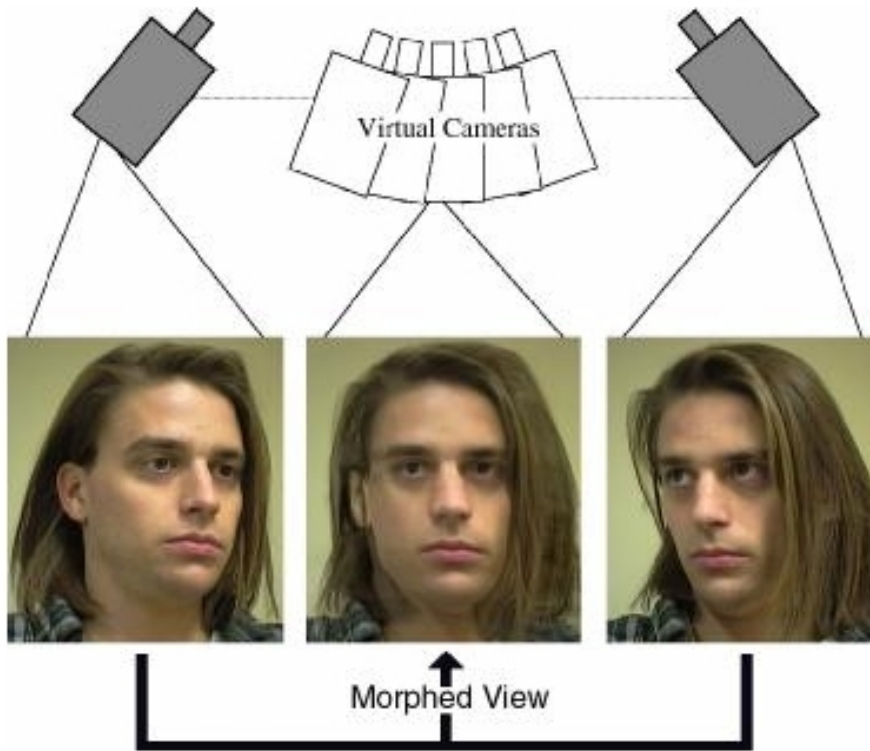


CS231A – Computer vision

Project proposals

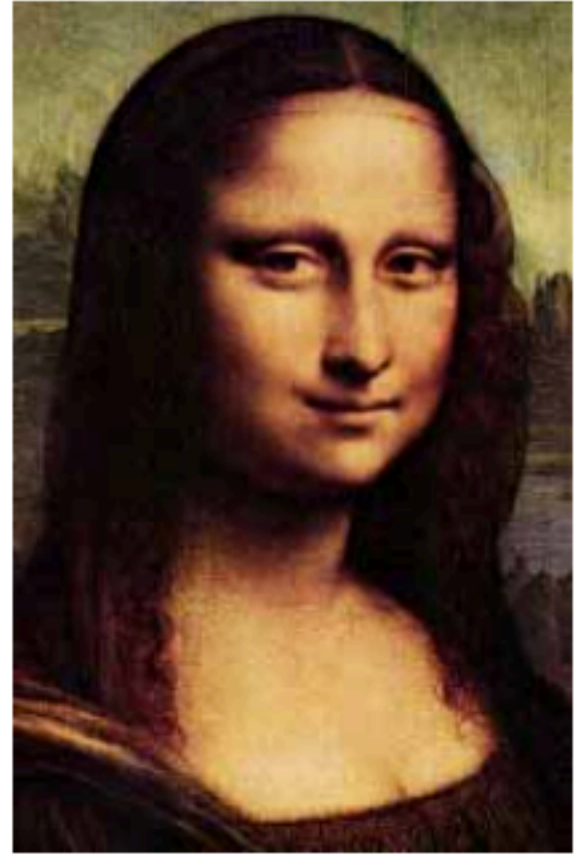
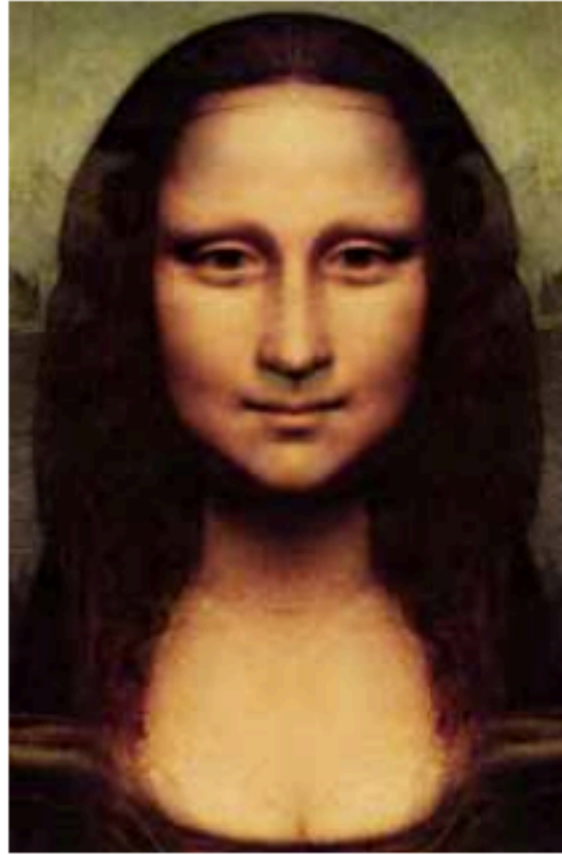
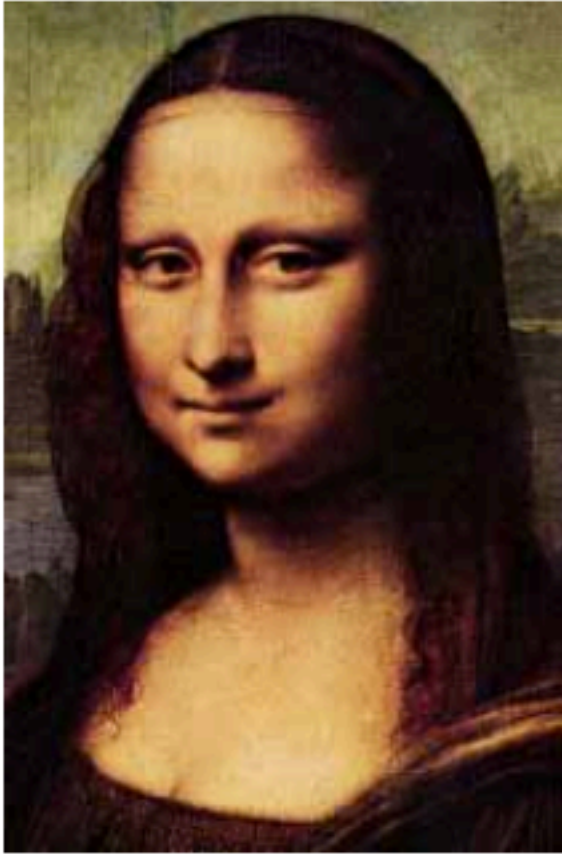
- Geometry
- Matching & modeling
- Mid level processing
- Recognition
- Databases
- Hardware/system

Geometry

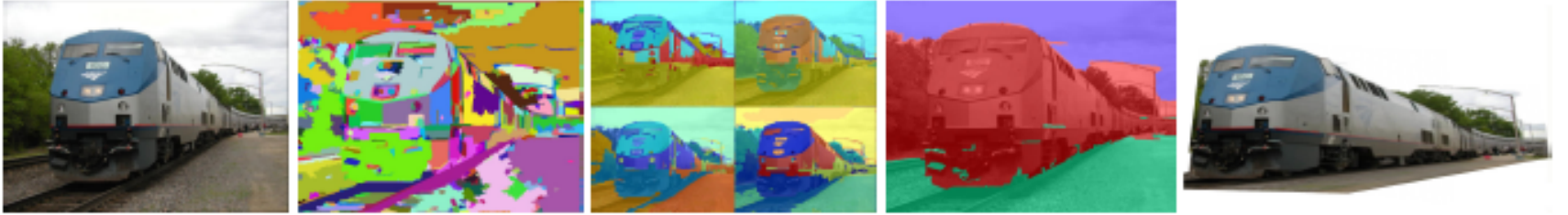


S. M. Seitz and C. R. Dyer,
Proc. SIGGRAPH 96, 1996,
21-30



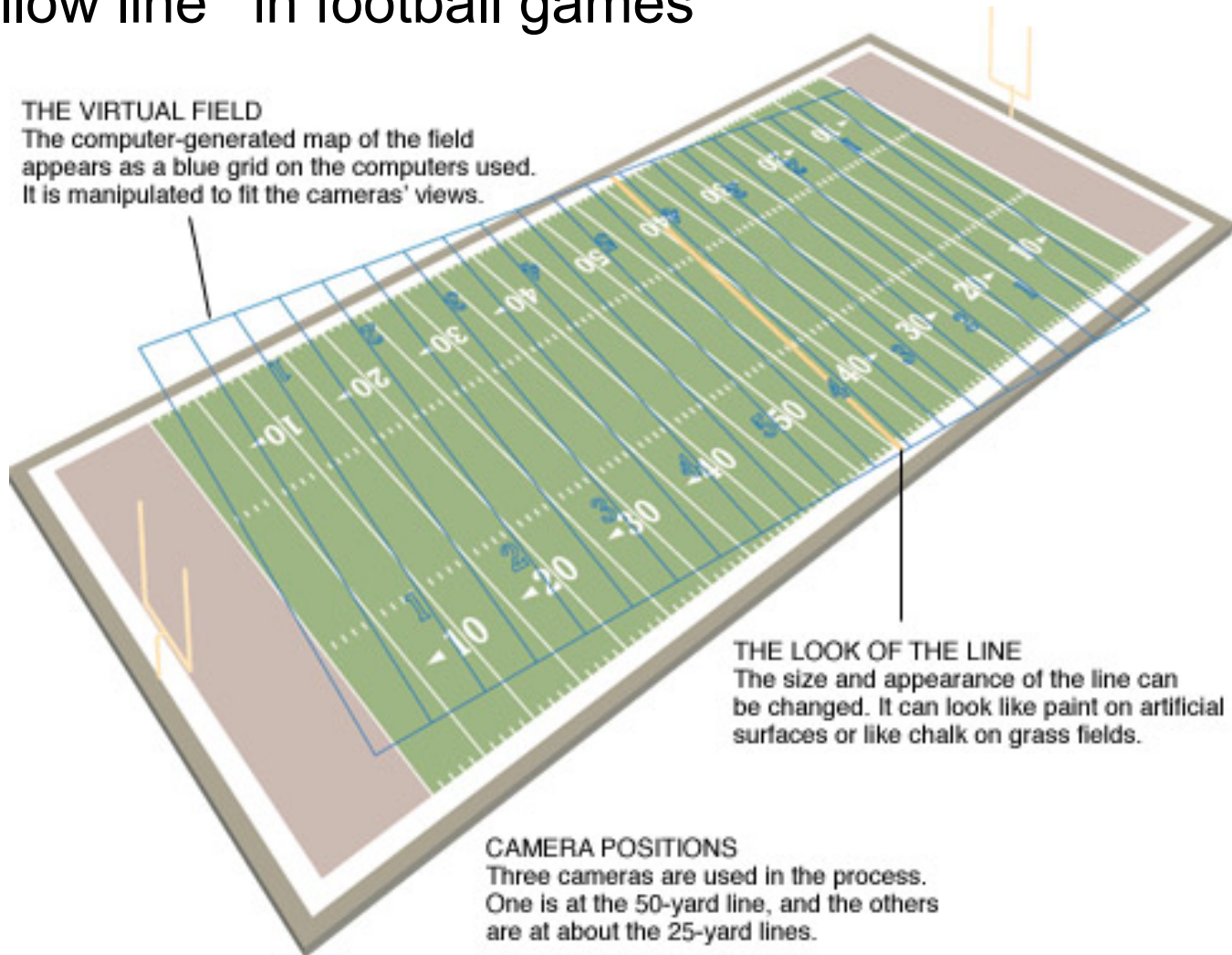


Geometry



Geometry

The 'yellow line' in football games



Geometry

- Augmented reality in sports events
- Extracting basic geometrical attributes (planar surfaces, occlusion boundaries, etc..) from an image/video

Geometry

- View point synthesis: view-morphing

Ref: S. M. Seitz and C. R. Dyer, *Proc. SIGGRAPH 96*, 1996, 21-30

- Scene pop-out

Ref: D. Hoiem, A.A. Efros, and M. Hebert, "Automatic Photo Pop-up",
ACM SIGGRAPH 2005

- Adding “yellow lines” or augmentation reality in sport events

MORE:

- Extracting basic geometrical attributes (planar surfaces, occlusion boundaries, etc..) from an image/video

Matching

- Matching images: Panorama Stitching.

Ref: M. Brown and D. G. Lowe, Recognizing Panoramas, ICCV 2003.

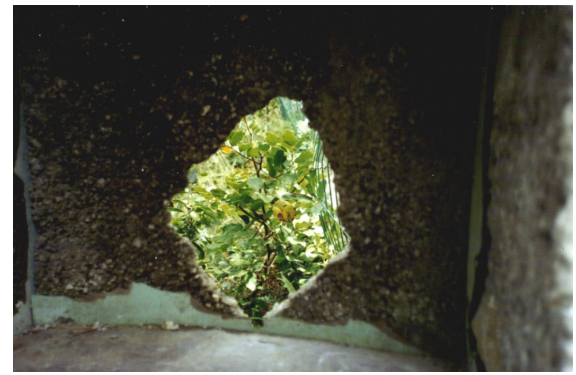
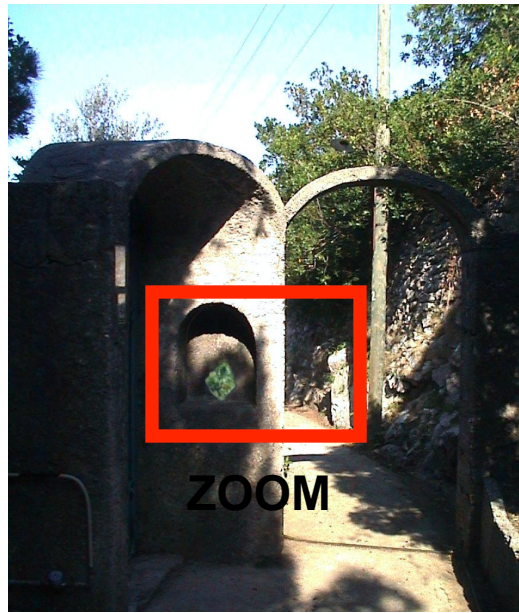
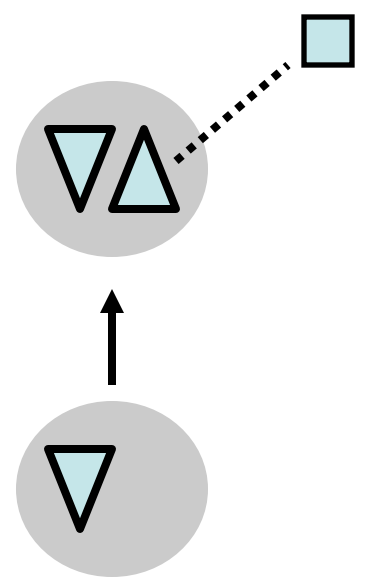
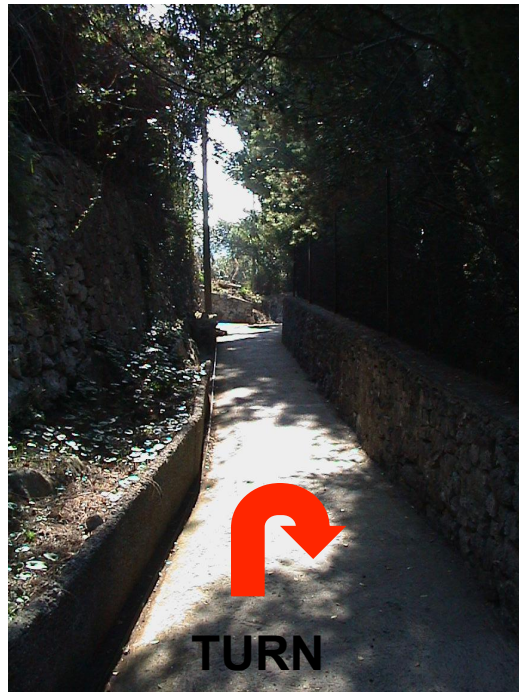


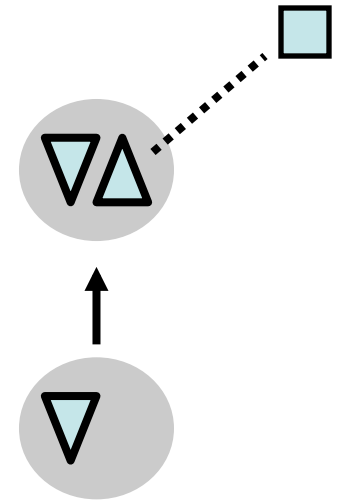
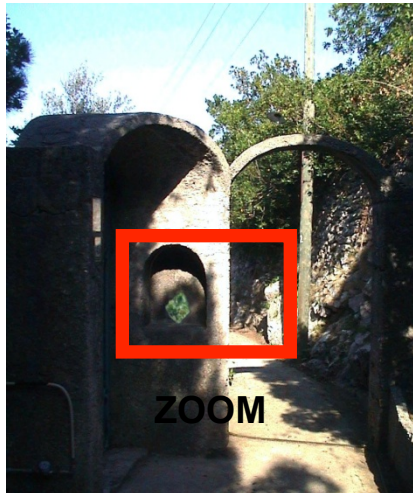
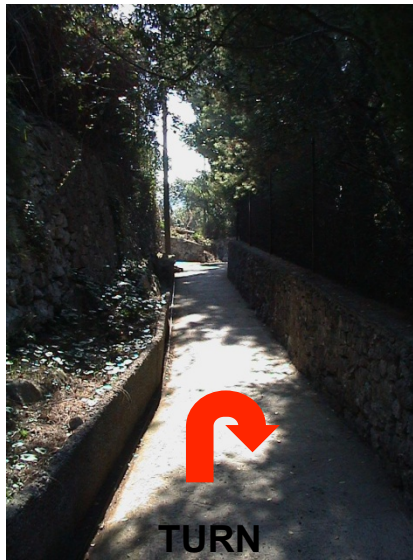
Matching

- Where I am ? Integrating imagery with GPS information
- Building 2d maps from collection of images of one environment



Unsorted
set of
images





**GOAL: estimate this topological map
by estimating partial 3d reconstruction**

Mid level processing

- **Image segmentation: low level segmentation**

- Ref: J. Shi & J. Malik, Normalized Cuts and Image Segmentation, PAMI 2000.

- Ref: Pedro F. Felzenszwalb and Daniel P. Huttenlocher, [Efficient Graph-Based Image Segmentation](#), International Journal of Computer Vision, Volume 59, Number 2, September 2004

- Ref: D. Comaniciu, P. Meer, [Mean shift: A robust approach toward feature space analysis](#), IEEE Trans. Pattern Anal. Machine Intell., 24, 603-619, May 2002.



Mid level processing

- Image completion



REF: “James Hays, Alexei A. Efros. Scene Completion Using Millions of Photographs. ACM Transactions on Graphics (SIGGRAPH 2007). August 2007, vol. 26, No. 3. ”

- Image completion with geometrical constraints

Discovering object symmetries



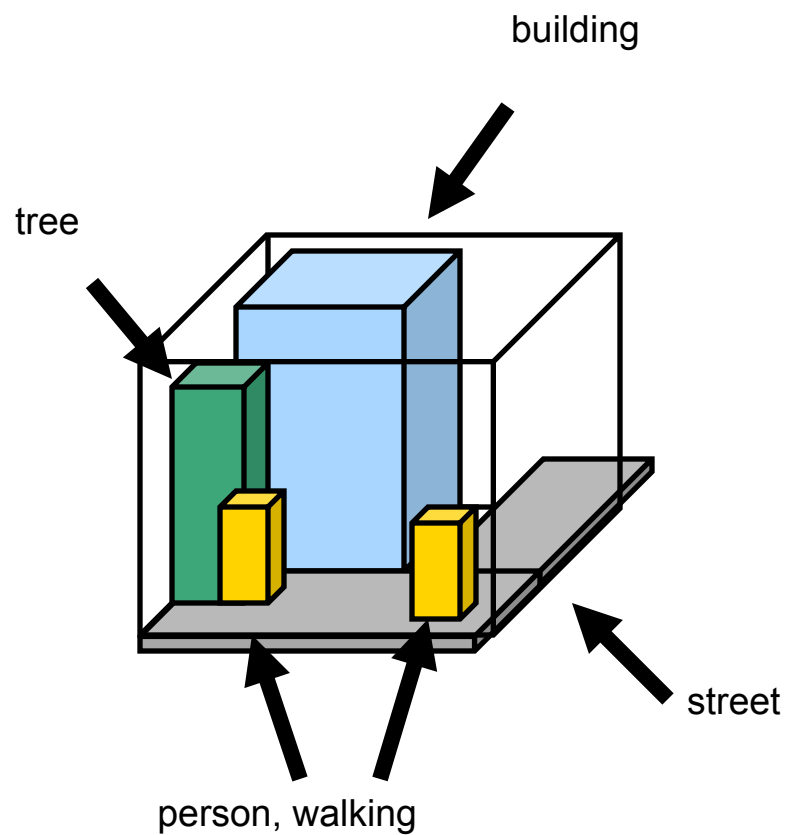
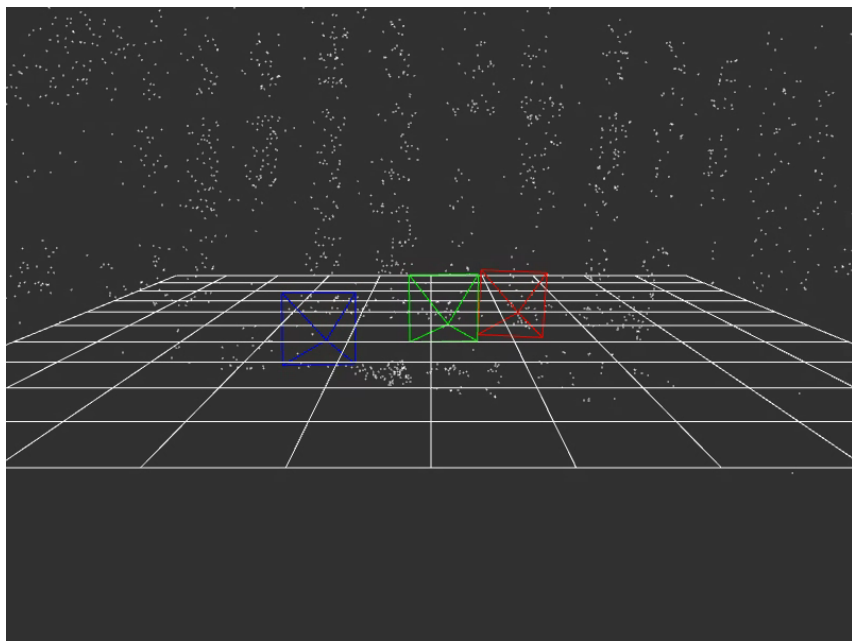
- Combine appearance and geometry to extract axis of symmetries
- Use these for image completion

Multi-view Scene Understanding



- Extract optical flows from 3d views of objects
- Foreground/background segmentation

Scene interpretation from video shots

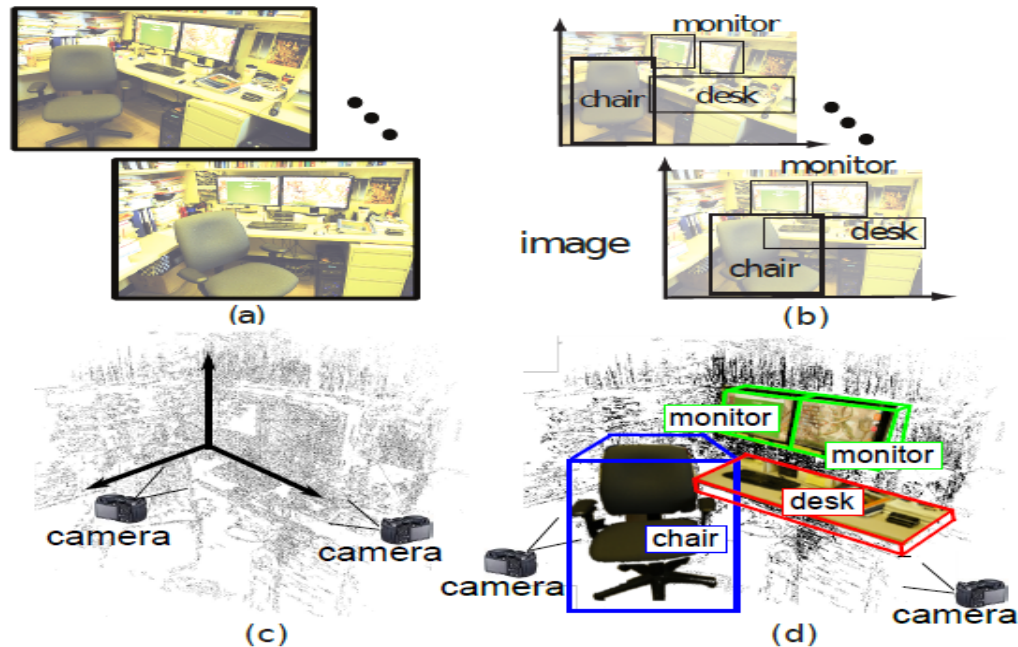


Scene category: *campus*

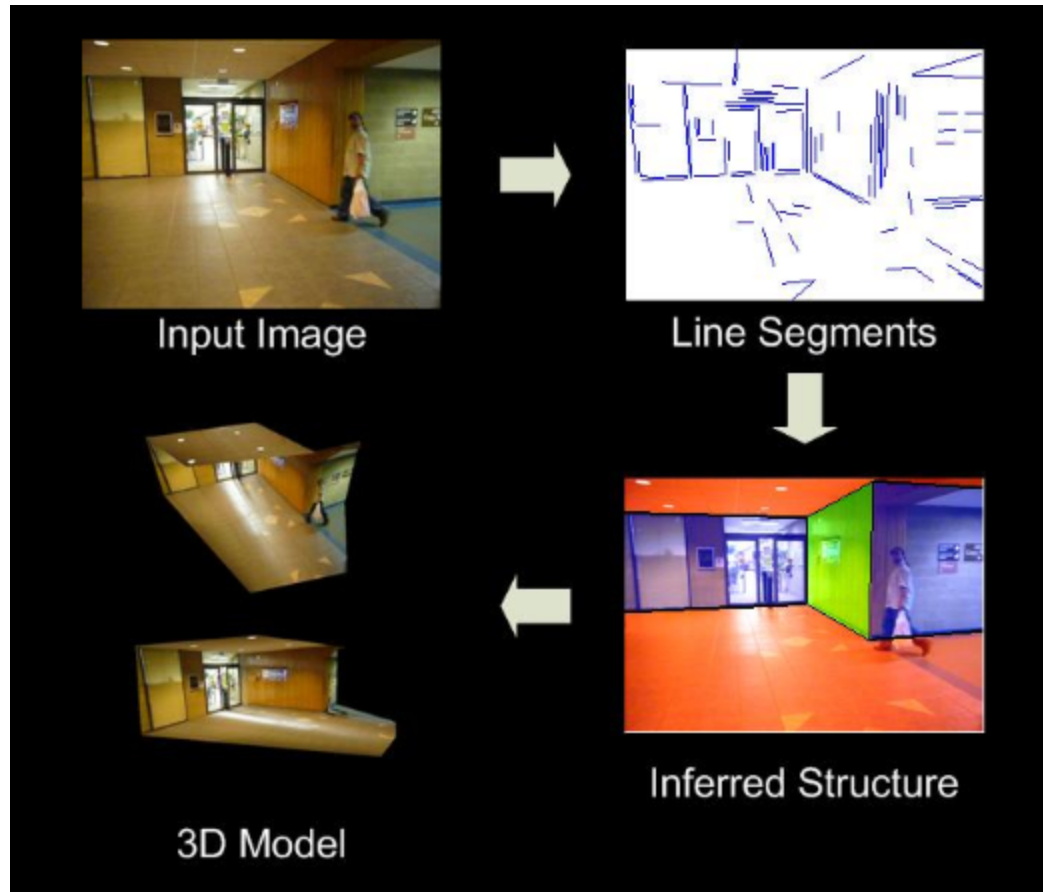
• Layout Estimation

REF : Sid Ying-Ze Bao, Min Sun, Silvio Savarese, [Toward Coherent Object Detection And Scene Layout Understanding](#), CVPR 2010.

REF : Sid Ying-Ze Bao, Silvio Savarese, [Semantic Structure from Motion](#), CVPR 2011.



Indoor layout estimation



Recognition [objects]

- Detecting faces or humans from one image

Ref: P. Viola and M. J. Jones, Robust Real-Time Object Detection, IJCV 2004.

Ref: Navneet Dalal, Bill Triggs,

[Histograms of Oriented Gradients for Human Detection](#), CVPR 2005

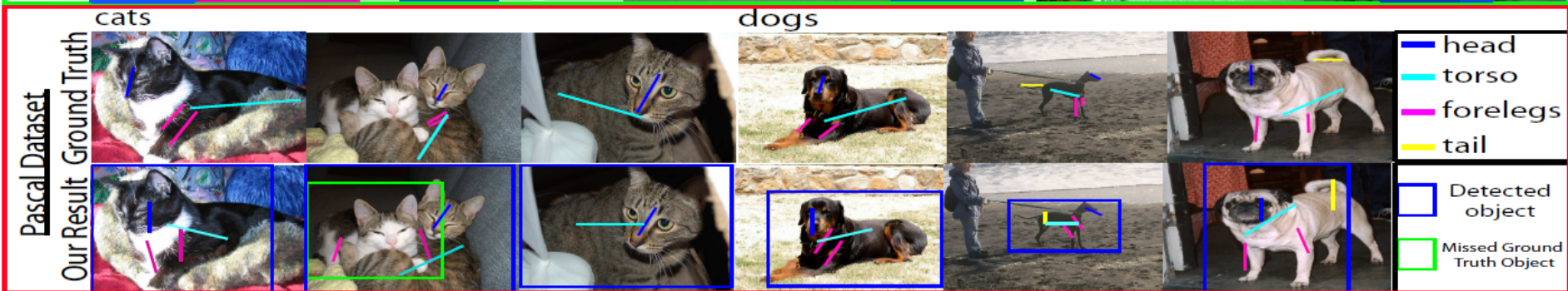
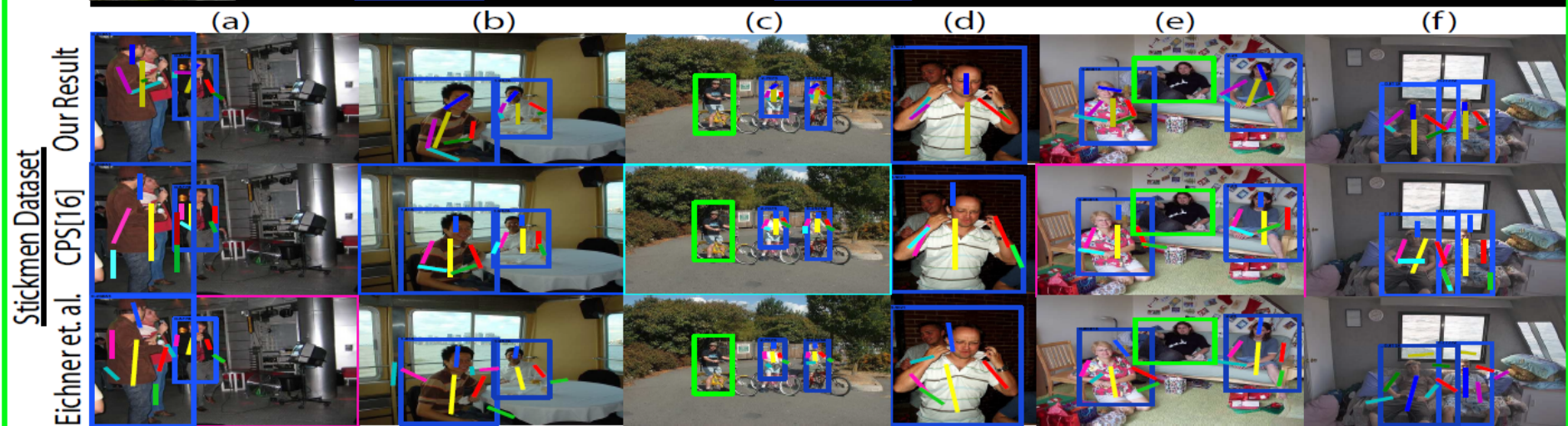
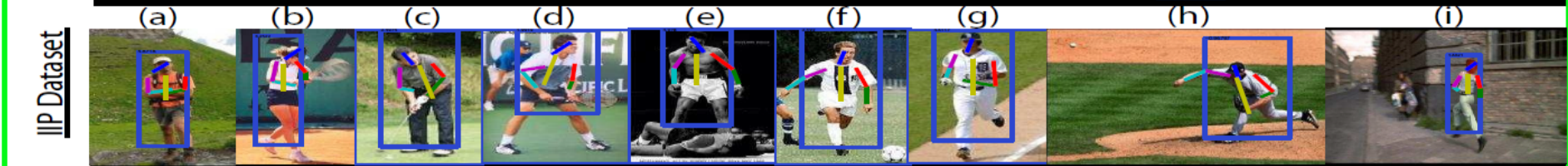


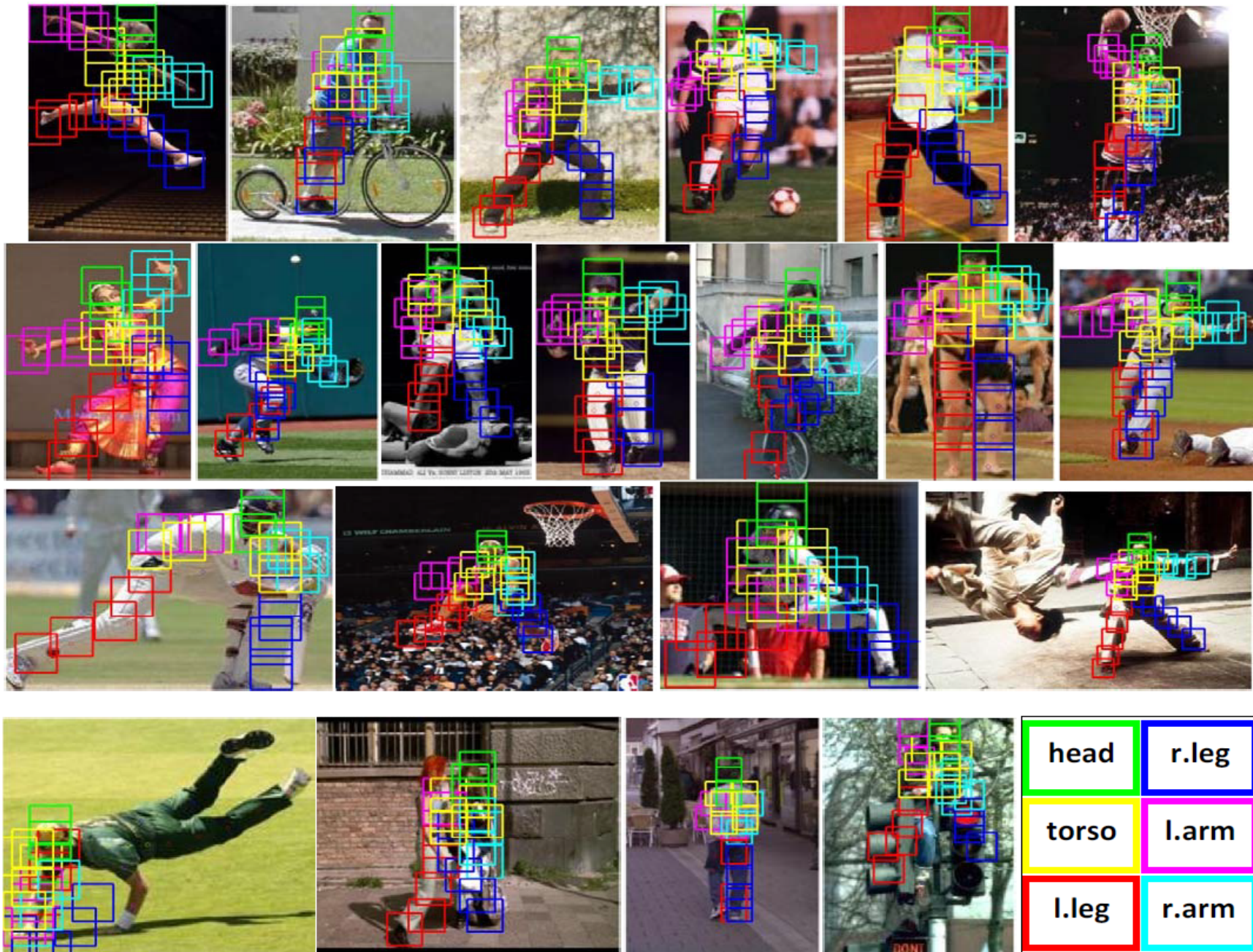
- Object recognition and Pose estimation

Ref: Min Sun and Silvio Savarese, [Articulated Part-based Model for Joint Object Detection](#) and Pose Estimation, ICCV 2011.

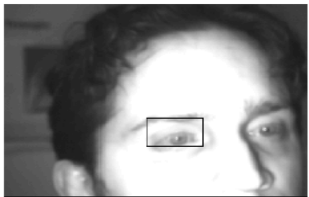
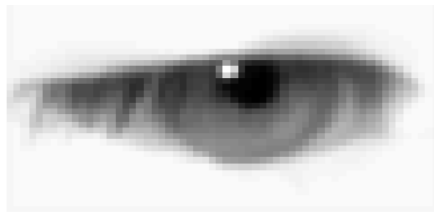
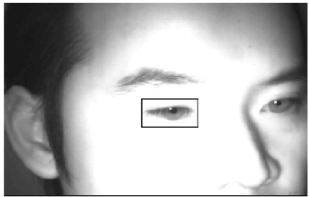
Ref: Benjamin Sapp, Alexander Toshev and Ben Taskar, [Cascaded Models for Articulated Pose Estimation](#), ECCV 2010.

Ref: Y. Yang, D. Ramanan. "[Articulated Pose Estimation using Flexible Mixtures of Parts](#)" Computer Vision and Pattern Recognition (CVPR) Colorado Springs, Colorado, June 2011.





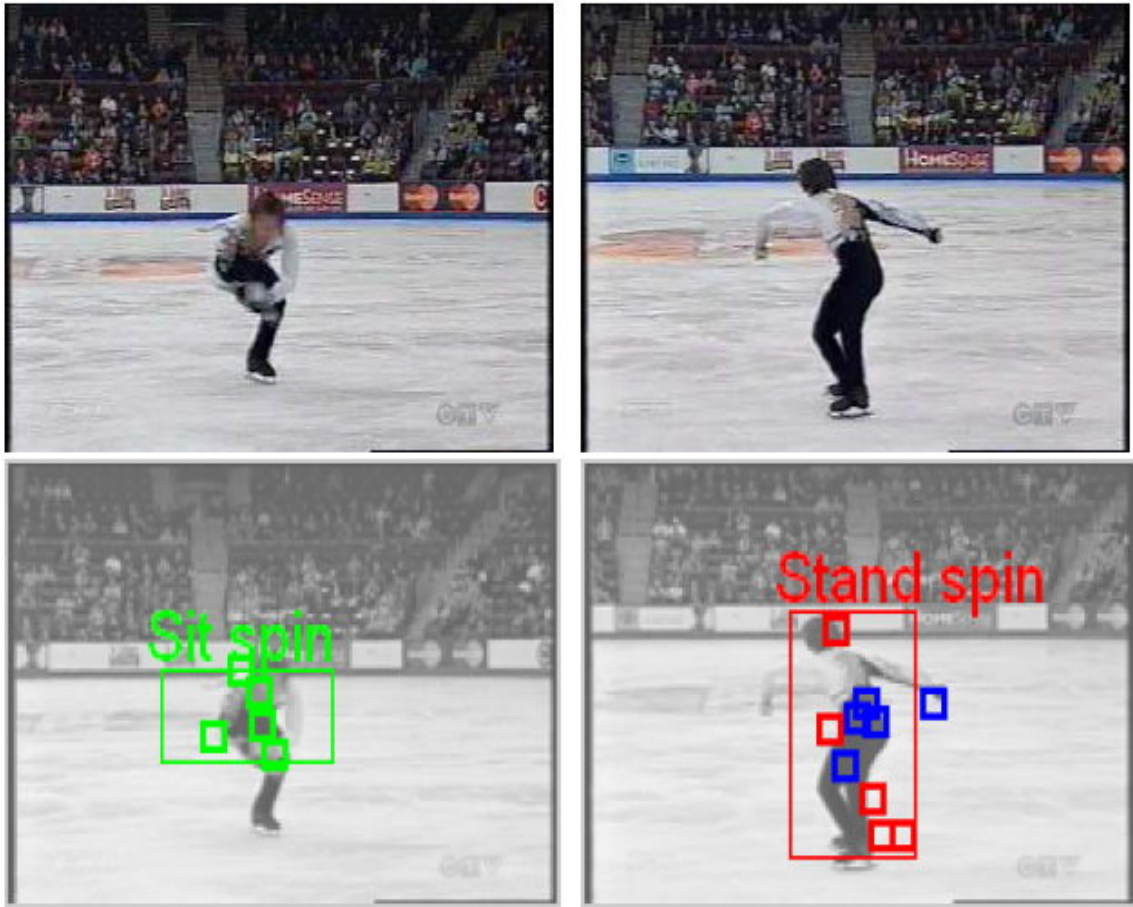
Human – gaze recognition



Recognition [objects]

- Object categorization: recognizing cats vs dogs
 - Constellation models
 - Bag of World models
 - ISM models

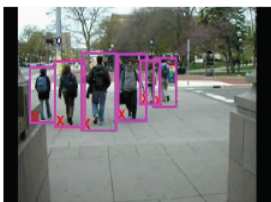
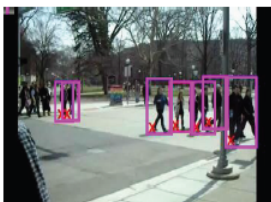
Recognition [human actions]



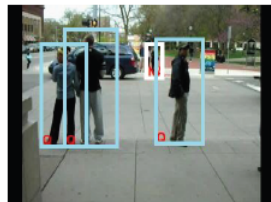
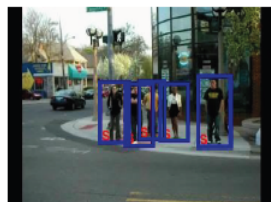
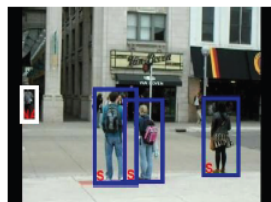
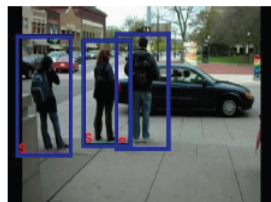
Ref: J. Niebles, H. Wang and L. Fei-Fei, **Unsupervised Learning of Human Action Categories**, in Video Proceedings, *IJCV 08*

Many others!

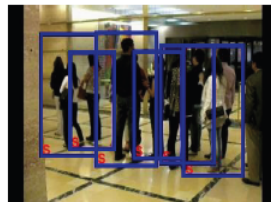
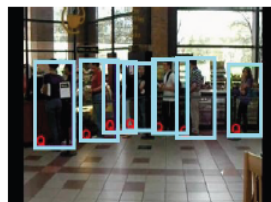
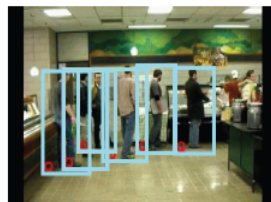
Crossing



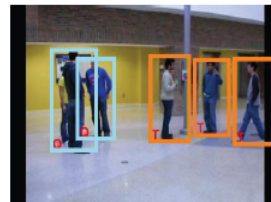
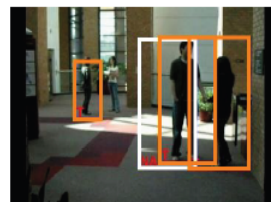
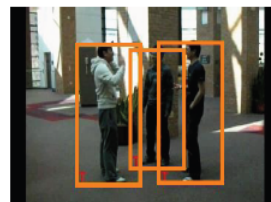
Waiting



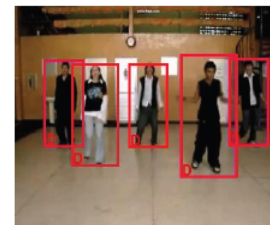
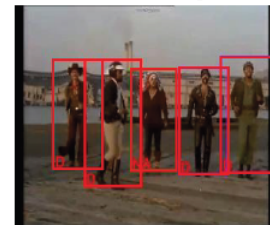
Queuing



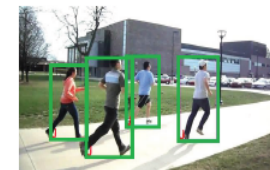
Talking



Dancing



Jogging



- **ACTION RECOGNITION**

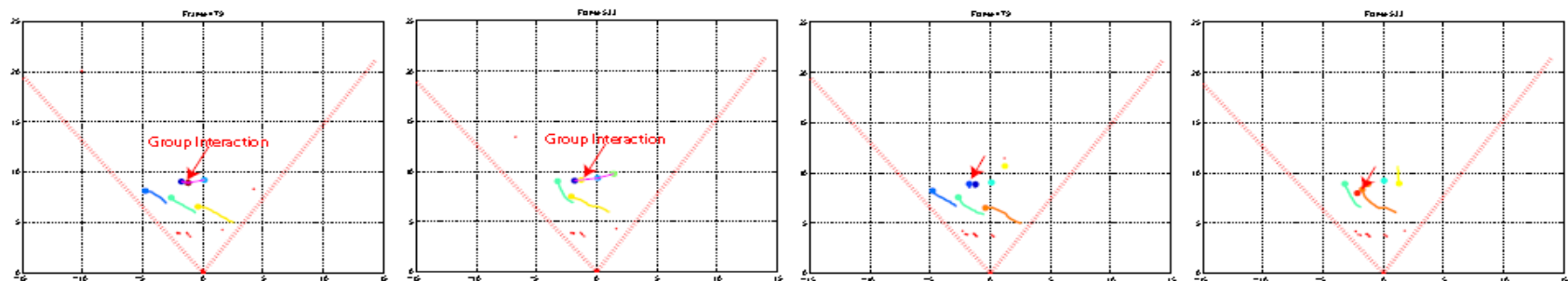
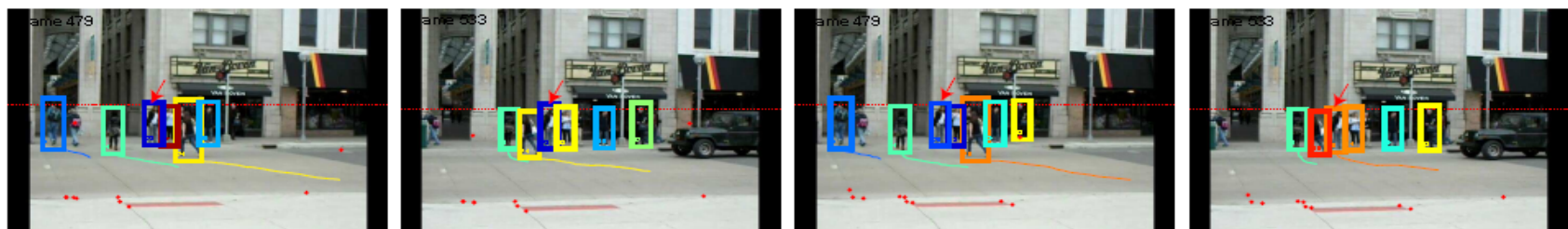
REF : Juan Carlos Niebles, Hongcheng Wang and Li Fei-Fei. [Unsupervised Learning of Human Action Categories Using Spatial-Temporal Words](#). In British Machine Vision Conference (BMVC), volume 3, pages 12449-1258, Edinburgh, UK, September 2006.

REF : J. Liu, B. Kuipers, S. Savarese, [Recognizing Human Actions by Attributes](#), Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, 2011 (To Appear). oral presentation.

REF : W. Choi, K. Shahid, and S. Savarese, [Learning Context for Collective Activity Recognition](#), Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, 2011.

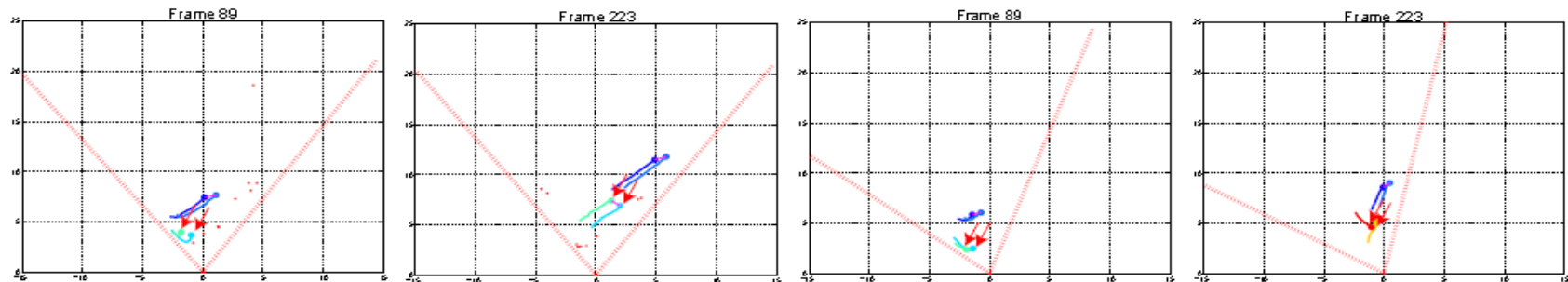
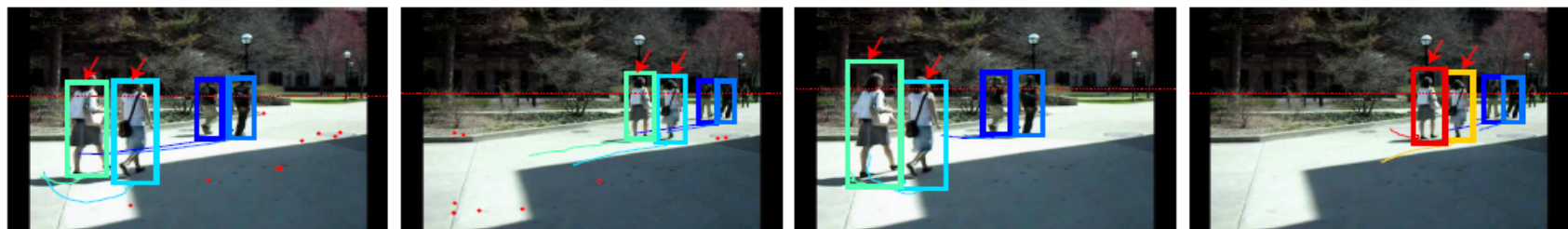
With Interaction

Without Interaction



With Ground Features

Without Ground Features



- **Multi target Tracking papers**

REF : W. Choi and S. Savarese, [Multiple Target Tracking in World Coordinate with Single, Minimally Calibrated Camera](#), Proc. of European Conference of Computer Vision, 553-567, 2010.

REF : W. Choi, C. Pantofaru, S. Savarese, [Detecting and Tracking People using an RGB-D Camera via Multiple Detector Fusion](#), Workshop on Challenges and Opportunities in Robot Perception (in conjunction with ICCV-11).

REF : A. Ess, B. Leibe, K. Schindler, and L. van Gool. [Robust Multi-Person Tracking from a Mobile Platform](#), in Transactions PAMI 2009.

REF : H. Pirsiavash, D. Ramanan, C. Fowlkes. "[Globally-Optimal Greedy Algorithms for Tracking a Variable Number of Objects](#)" Computer Vision and Pattern Recognition (CVPR) Colorado Springs, Colorado, June 2011.

Recognition [scenes]

- Scene classification: indoor VS outdoor

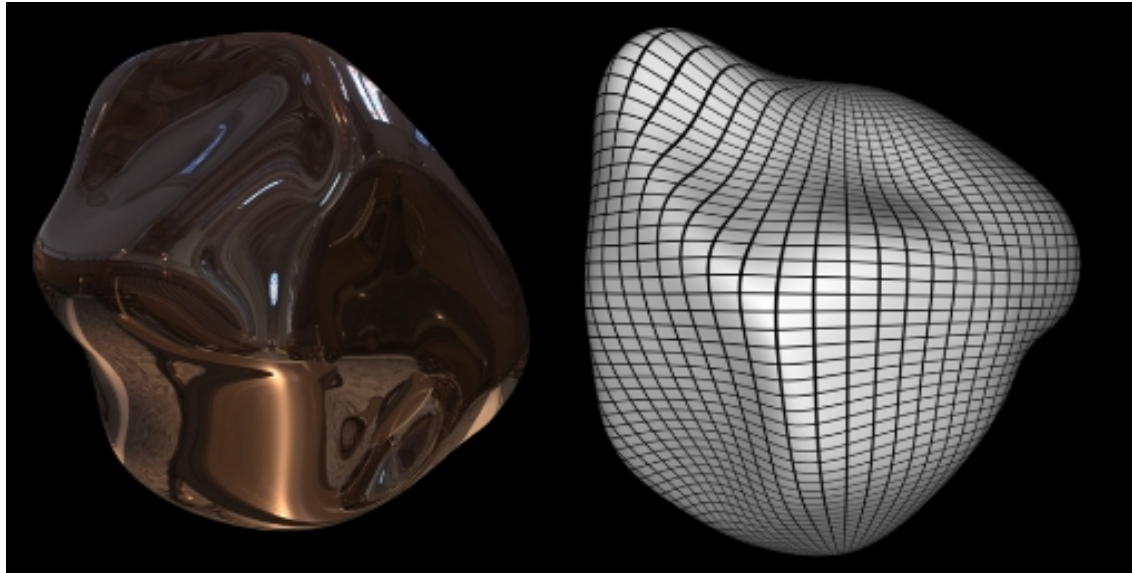
Ref: L. Fei-Fei and P. Perona. A Bayesian Hierarchical Model for Learning Natural Scene Categories, CVPR 2005.

Ref: Beyond Bags of Features: Spatial Pyramid Matching for Recognizing Natural Scene Categories, S. Lazebnik, C. Schmid, and J. Ponce, CVPR 2006



The image cannot be displayed. Your computer may not have enough memory to open the image, or the image may have been corrupted. Restart your computer, and then open the file again. If the red x still appears, you may have to delete the image and then insert it again.

Recognition [materials]

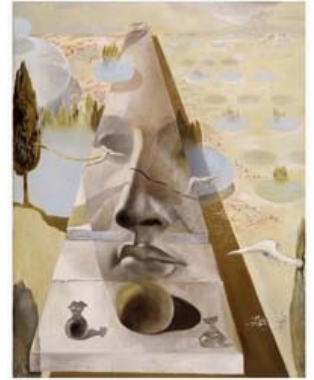


- What is reflective, what is not?

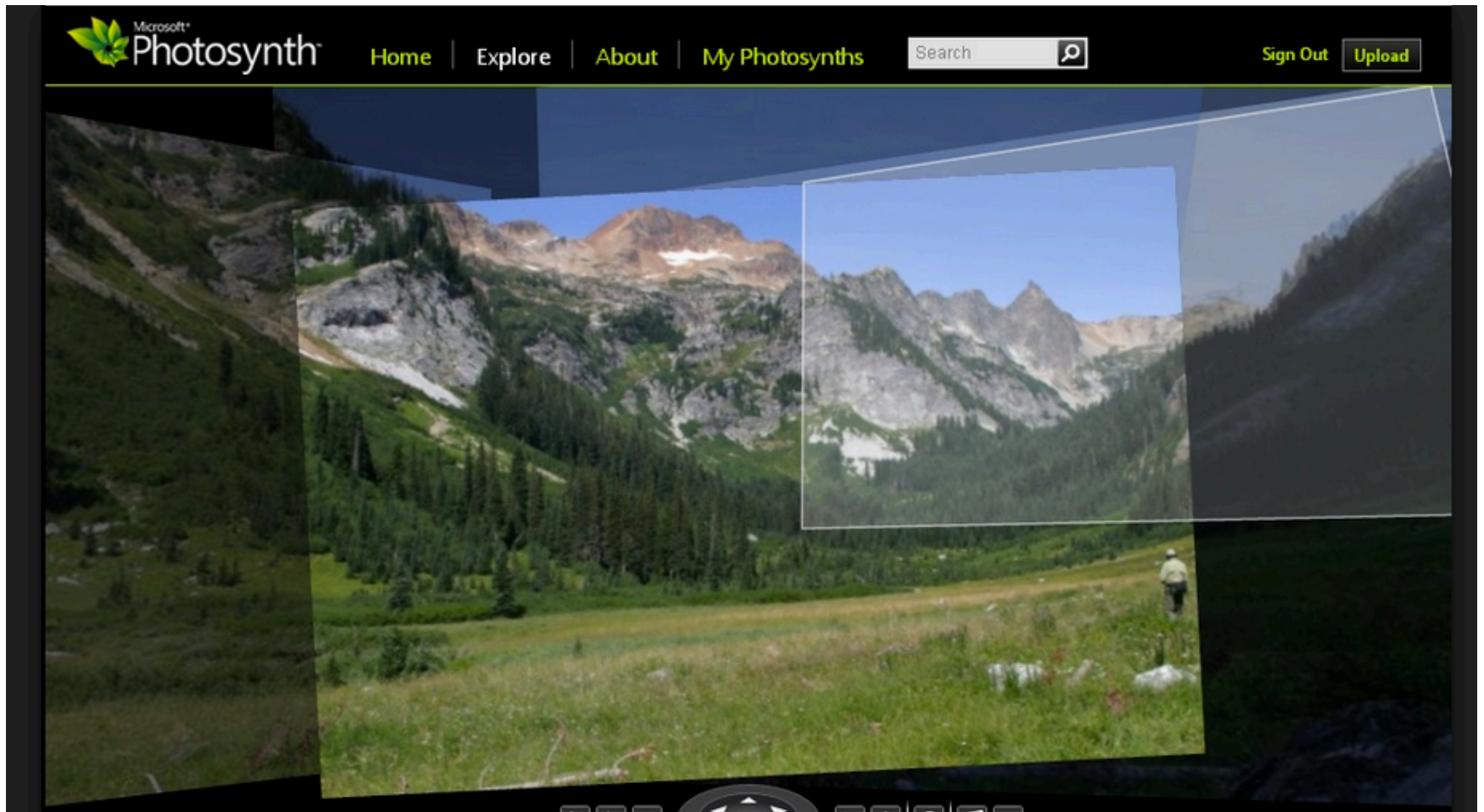
•Ref: A. DelPozo, and **S. Savarese**, "Detecting Specular Surfaces on Natural Images", *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, Minnesota, USA, June 2007

Web based

- Design system to crawl the web and extract/classify visual **static imagery** from webpages; query by image
- Design system to crawl the web and extract and classify **videos** from webpages



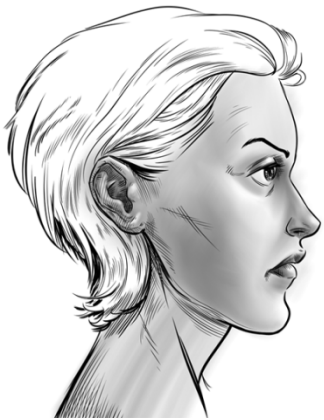
Photosynth



Object recognition



HCI: articulating the visual query



Can I find this shoe at a better price somewhere else?
Type of leather?

WEB



Search by visual query

Hardware & systems

- Setting up and calibrate a system of N cameras for people or object tracking
- Setting up a calibrated system for 3d shape acquisition
 - Mount cameras
 - Setup hardware
 - Build software to control hardware
 - Run experiments to validate that the system works

Datasets

- 360 deg. videos of objects for 3d object categories
- Videos of environments for scene categorization
- Video of human activities for action categorization

- Collect images
- Annotate images
- Build software for indexing images/data

3D Objects Search Engines



3D Search
3D Chat
3D Avatars
3D Social Networks
3D Virtual Worlds
ExitReality Plaza



Google 3D warehouse



Princeton Shape Retrieval and Analysis Group
3D Model Search Engine

LabelMe



now in 3D!