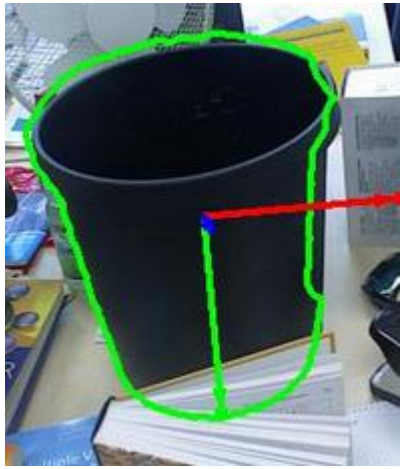
Poselets

Lubomir Bourdev and Jitendra Malik
EECS, U.C. Berkeley and Adobe Systems

PRESENTED BY : SUGEERTH MURUGESAN

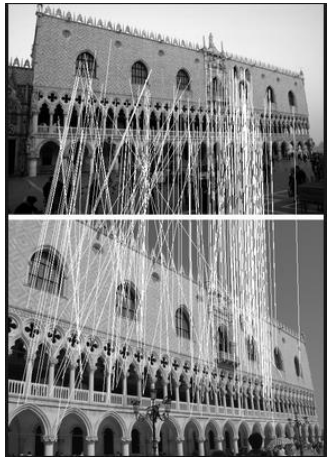
Motivation

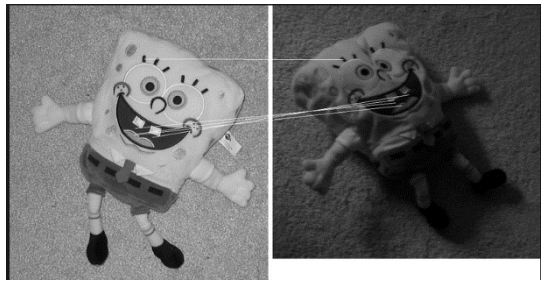
- Pose has:
  - Appearance Space – pixel values
  - Configuration Space – 3D coordinates parts



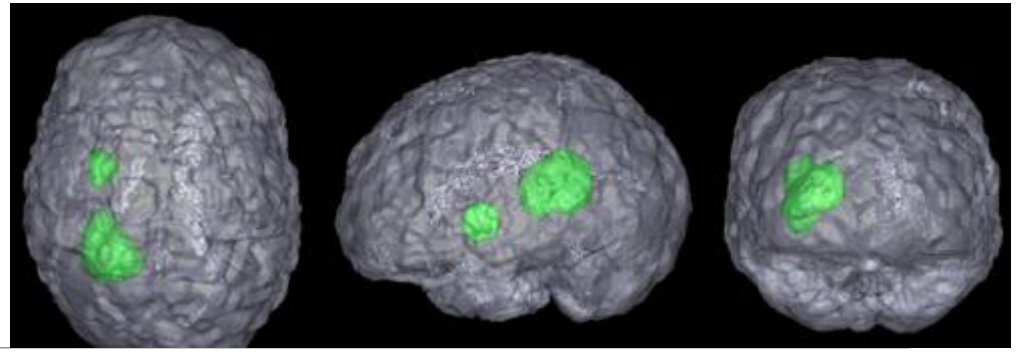
Appearance space is function of:

- Clothing
- Viewpoint
- Occlusion
- Illumination



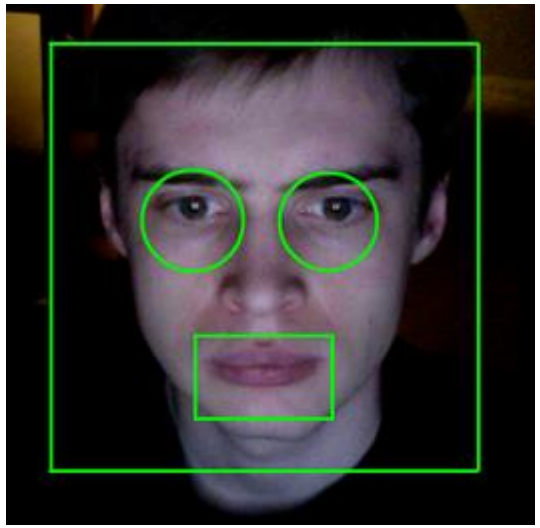


Motivation



Challenges in computer vision:

- Detection
- Joint Localization
- Segmentation

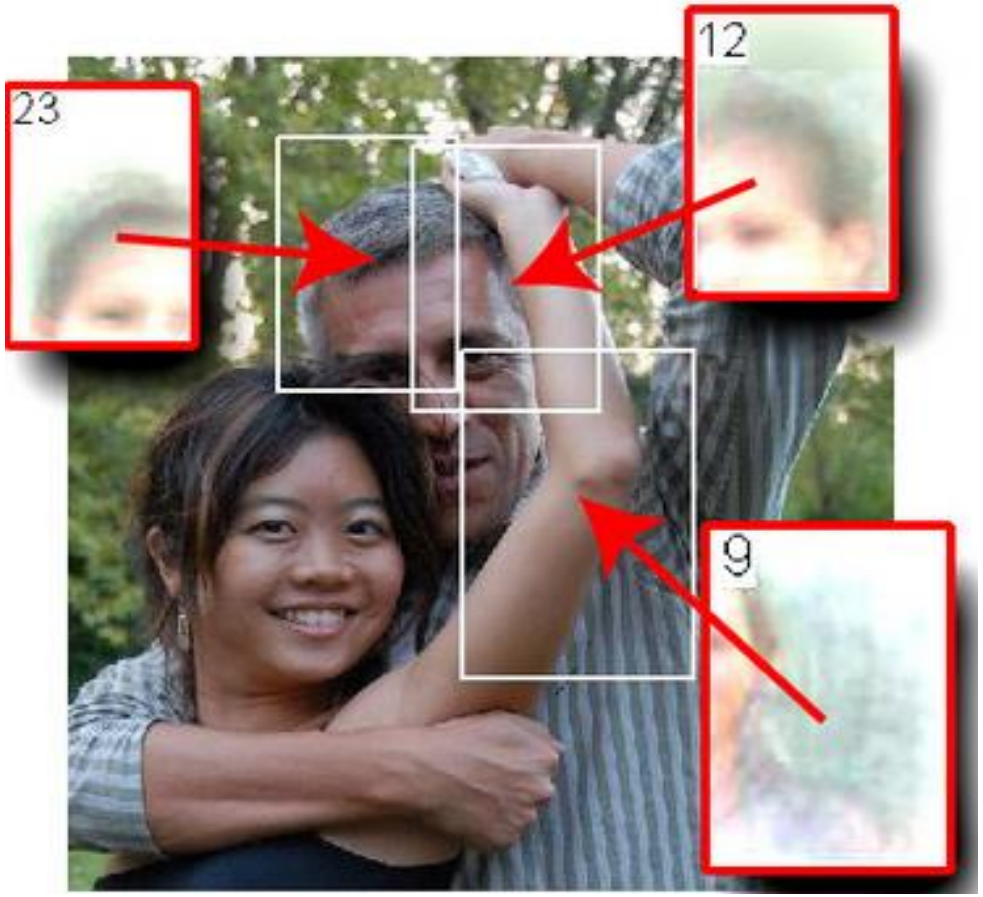


Motivation

What if ?

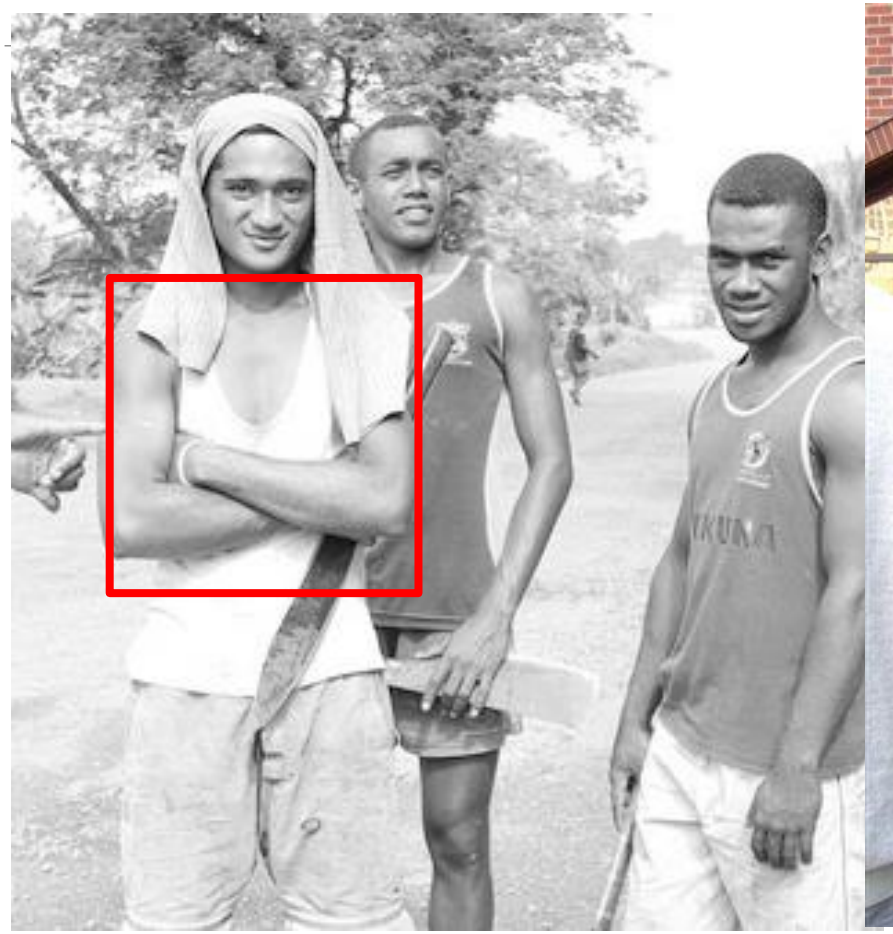
Use Divide and Conquer

- to detect “parts” to factor complexity
- Need for:
  - Introducing Poselets
  - Algorithm to select Poselets



Bourdev, Lubomir, and Jitendra Malik. "Poselets: Body part detectors trained using 3d human pose annotations."

Where is the bounding box?



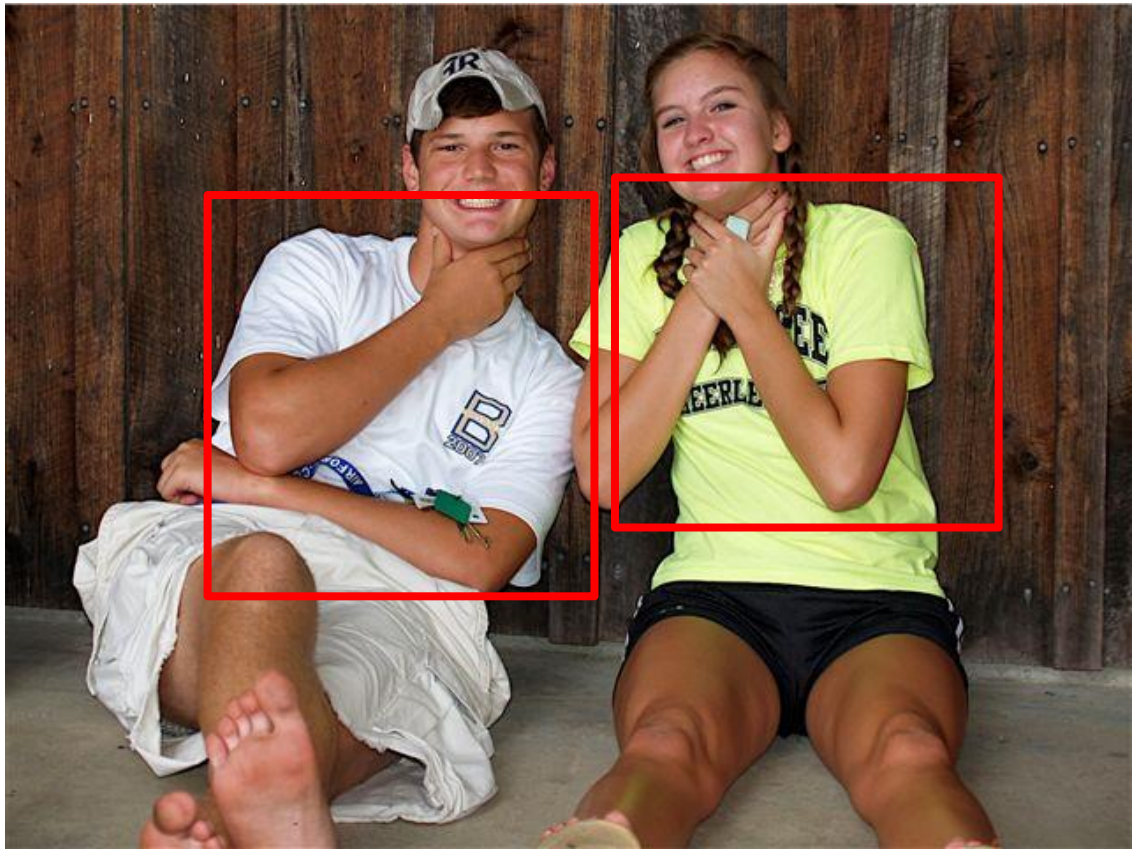
Poselets

Poselets capture part of the pose from a given viewpoint

Poselets

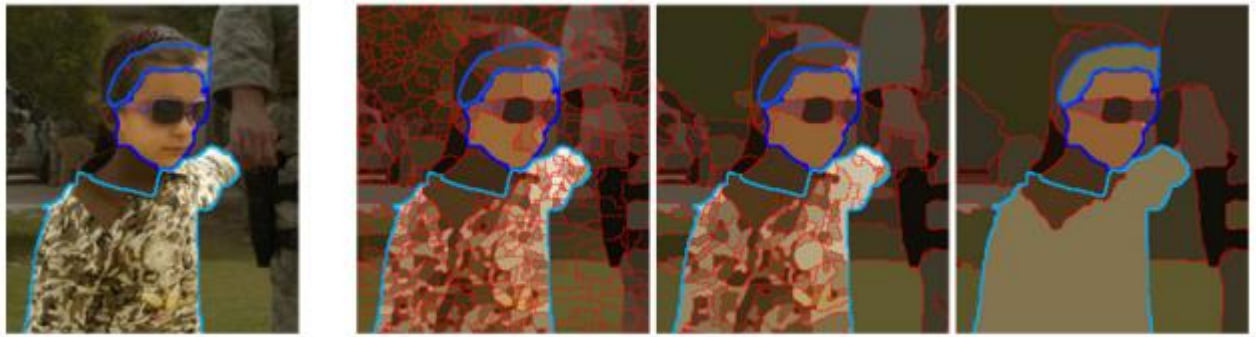
Examples may differ visually but have common semantics

How do we train a poselet for a given pose configuration?

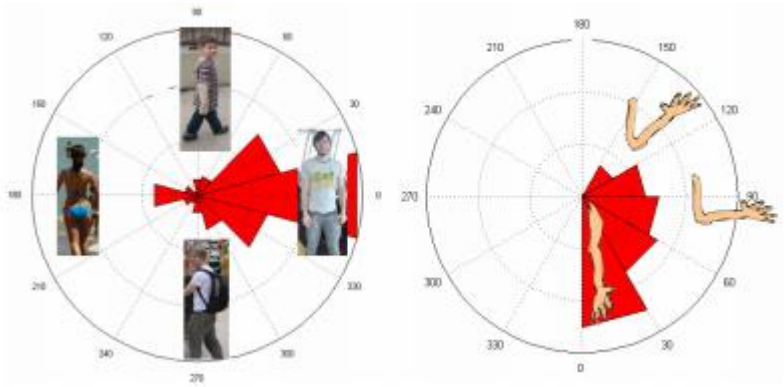


H3D dataset

3D annotation of keypoints



Pixel level labels

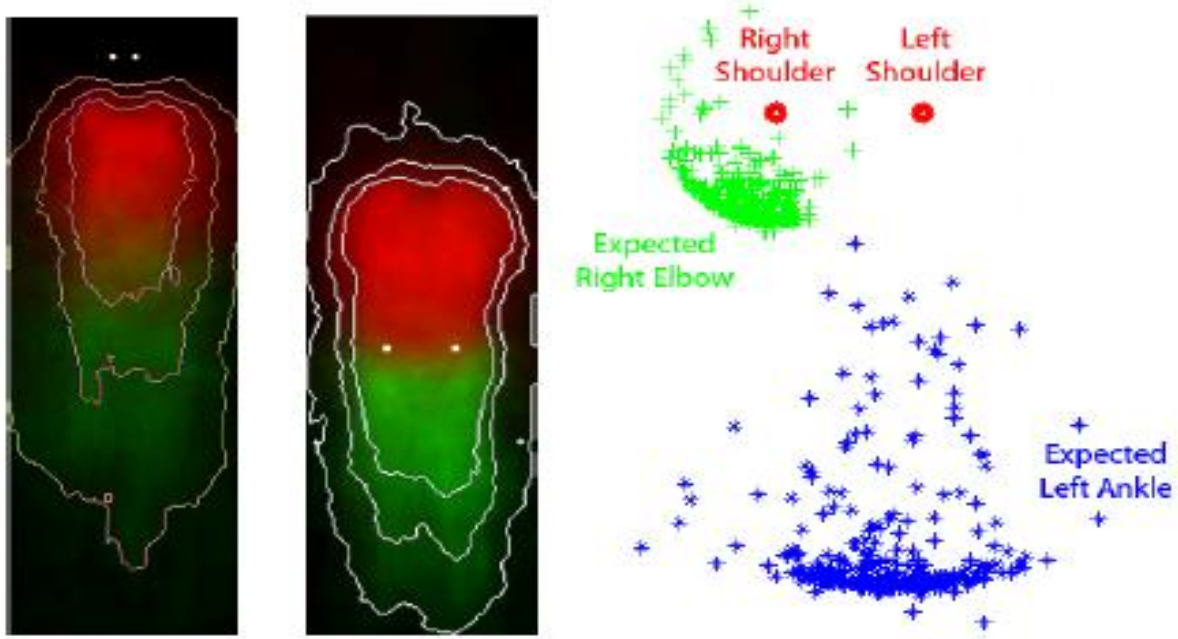


Azimuth angles of frontal profiles and expected arm bending

[1] Maire, Michael, et al. "Using contours to detect and localize junctions in natural images."

[2] Taylor "Reconstruction of articulated objects from point correspondences in a single uncalibrated image."

H3D dataset



Conditional probability masks

Appearance queries

Bourdev, Lubomir, and Jitendra Malik. "Poselets: Body part detectors trained using 3d human pose annotations."

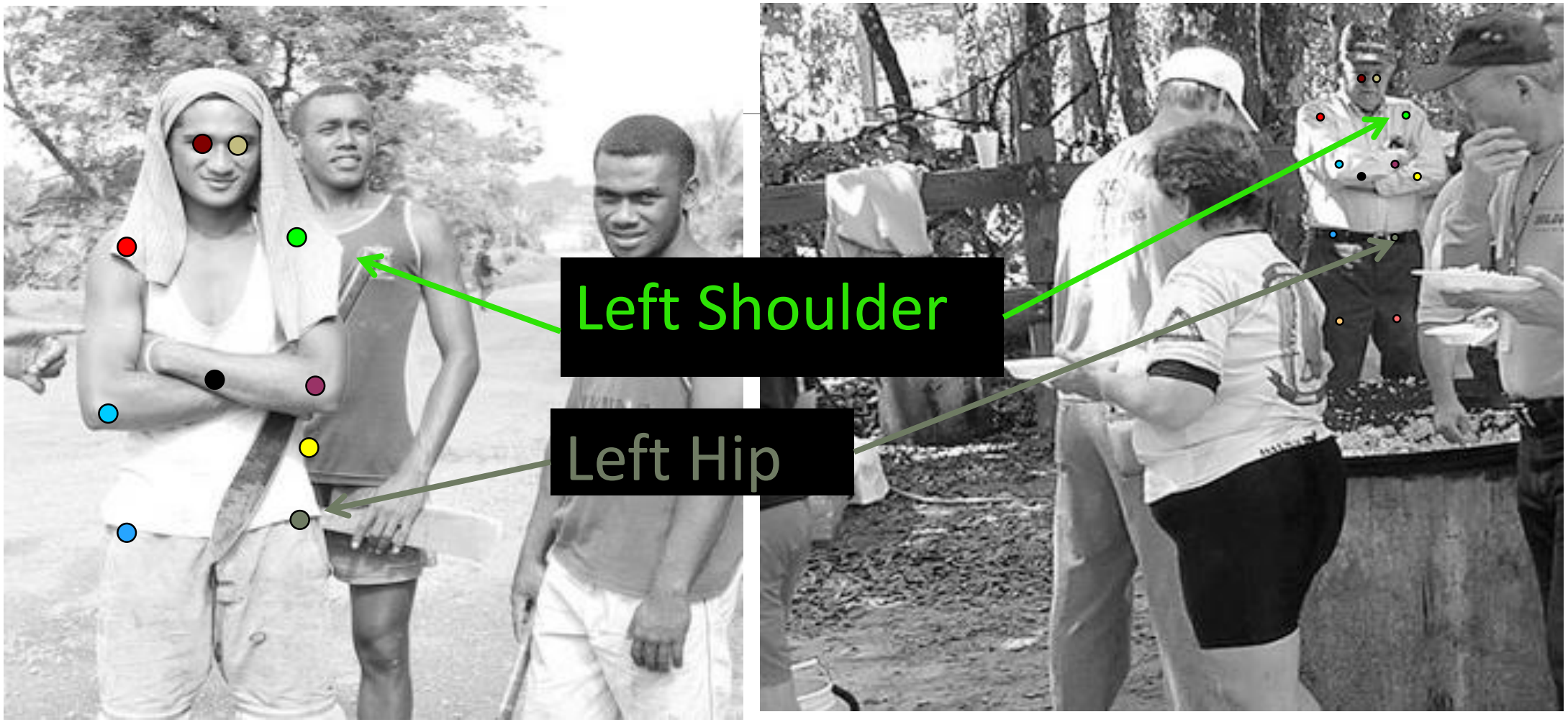
Finding correspondences at training time

Given part of a human pose

How do we find a similar pose configuration in the training set?

Bourdev, Lubomir, and Jitendra Malik. "Poselets: Body part detectors trained using 3d human pose annotations."

Finding correspondences at training time



We use keypoints to annotate the joints, eyes, nose, etc. of people

Bourdev, Lubomir, and Jitendra Malik. "Poselets: Body part detectors trained using 3d human pose annotations."

Method (Finding Candidates)

Distance from example *s* to example *r* as:

$$d_s(r) = \sum_i w_s(i) ||\mathbf{x_s}(i) - \mathbf{x_r}(i)||_2^2 (1 + h_{s,r}(i))$$

Where $X_s(i)$ = [x, y, z] are normalized 3D coordinates of the *i*-th keypoint of the example *s*. The weight is $w_s(i)$ a Gaussian with mean at the center of the patch.

The term $h_{s,r}(i)$ is a penalty based on the visibility mismatch of keypoint *i* in the two examples.

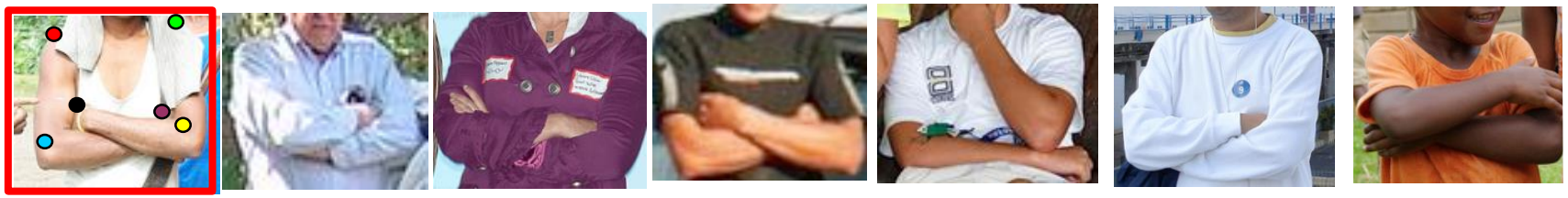
Finding correspondences at training time

Residual Error

We use keypoints to annotate the joints, eyes, nose, etc. of people

Bourdev, Lubomir, and Jitendra Malik. "Poselets: Body part detectors trained using 3d human pose annotations."

Training poselet classifiers

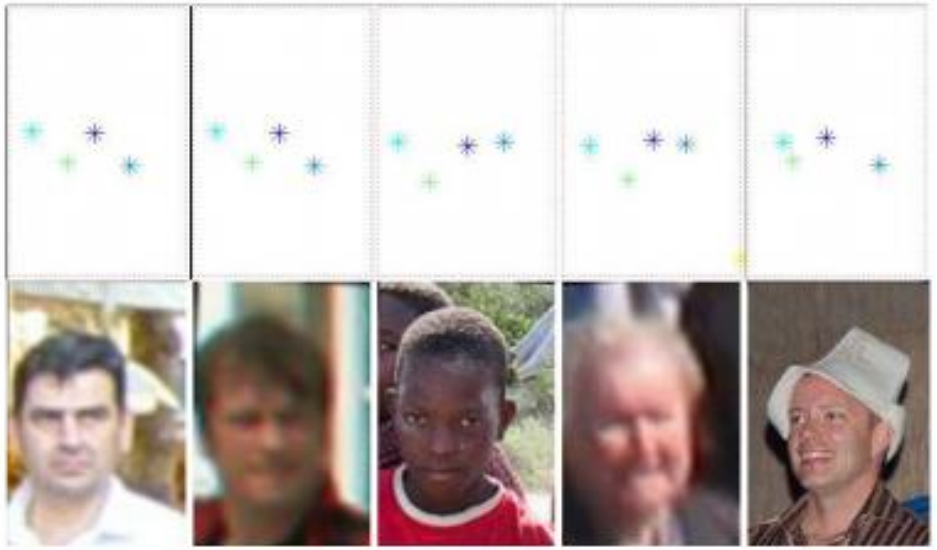


Residual Error:	0.15	0.20	0.10	0.85	0.15	0.35

1. Given a seed patch
2. Find the closest patch for every other person
3. Sort them by residual error
4. Threshold them
5. Use them as positive training examples to train a linear SVM with HOG features

Bourdev, Lubomir, and Jitendra Malik. "Poselets: Body part detectors trained using 3d human pose annotations."

Examples



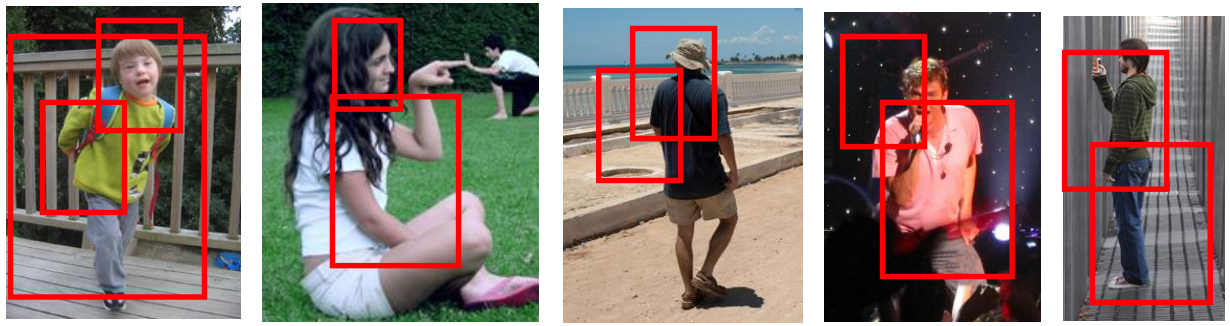
Top: 3D configuration space,
Bottom: Appearance space

Example query regions (left column) and the corresponding closest matches in configuration space generated by H3D.

Bourdev, Lubomir, and Jitendra Malik. "Poselets: Body part detectors trained using 3d human pose annotations."

Which poselets should we train?

Choose thousands of random windows, generate poselet candidates, train linear SVMs

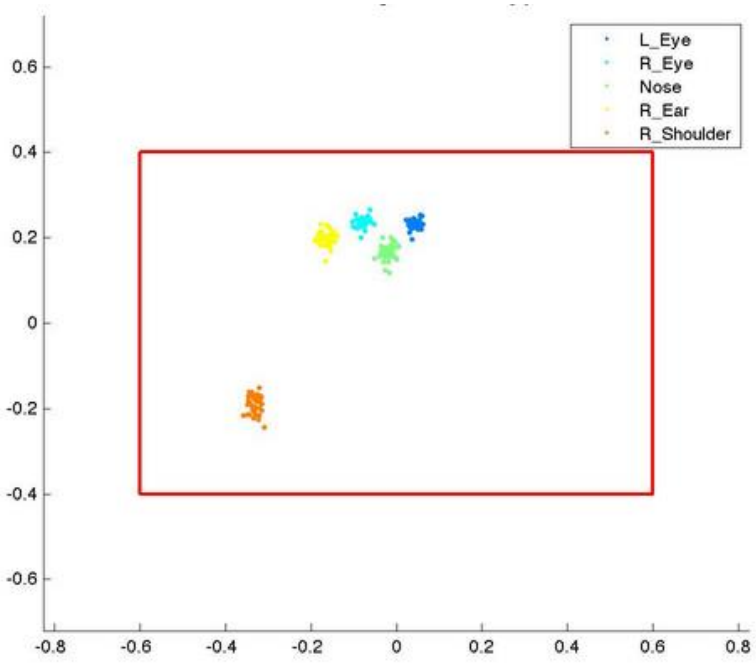


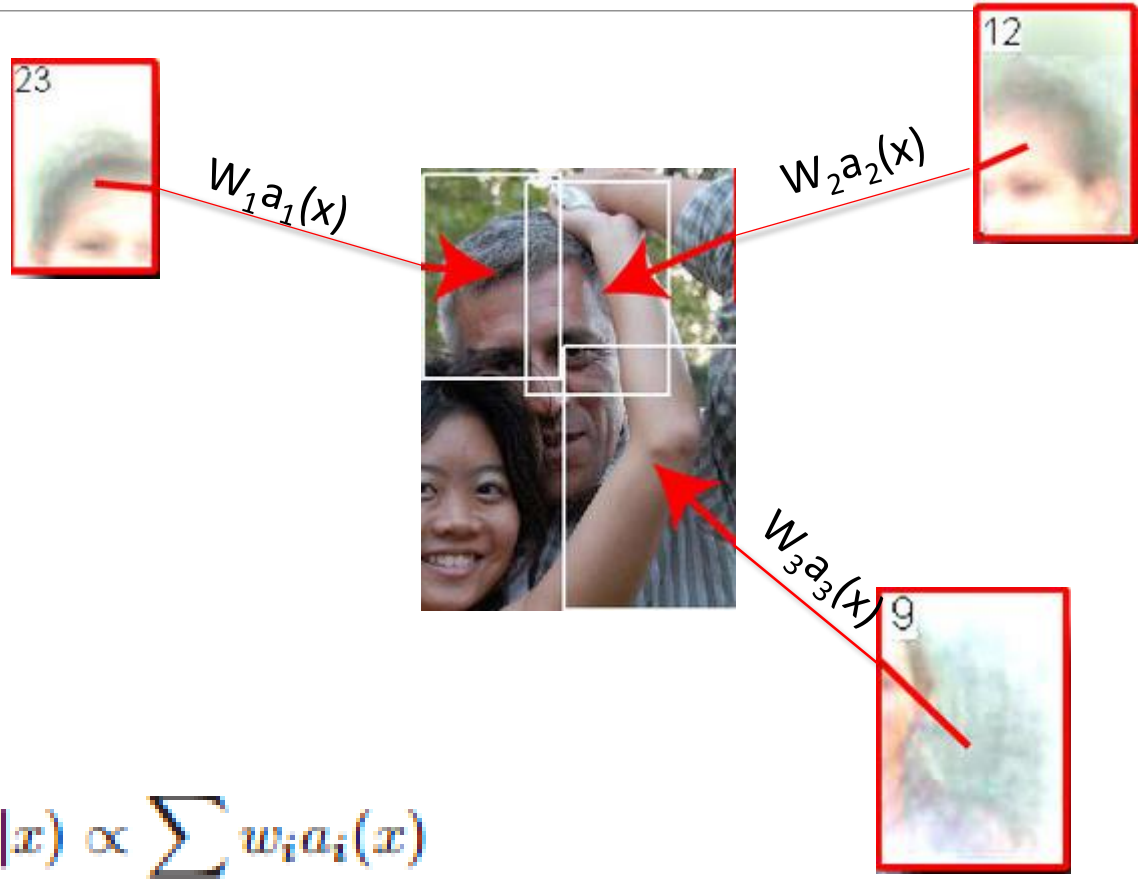
Select a small set of poselets that are:

- Individually effective
- Complementary

Bourdev, Lubomir, and Jitendra Malik. "Poselets: Body part detectors trained using 3d human pose annotations."

Object localization

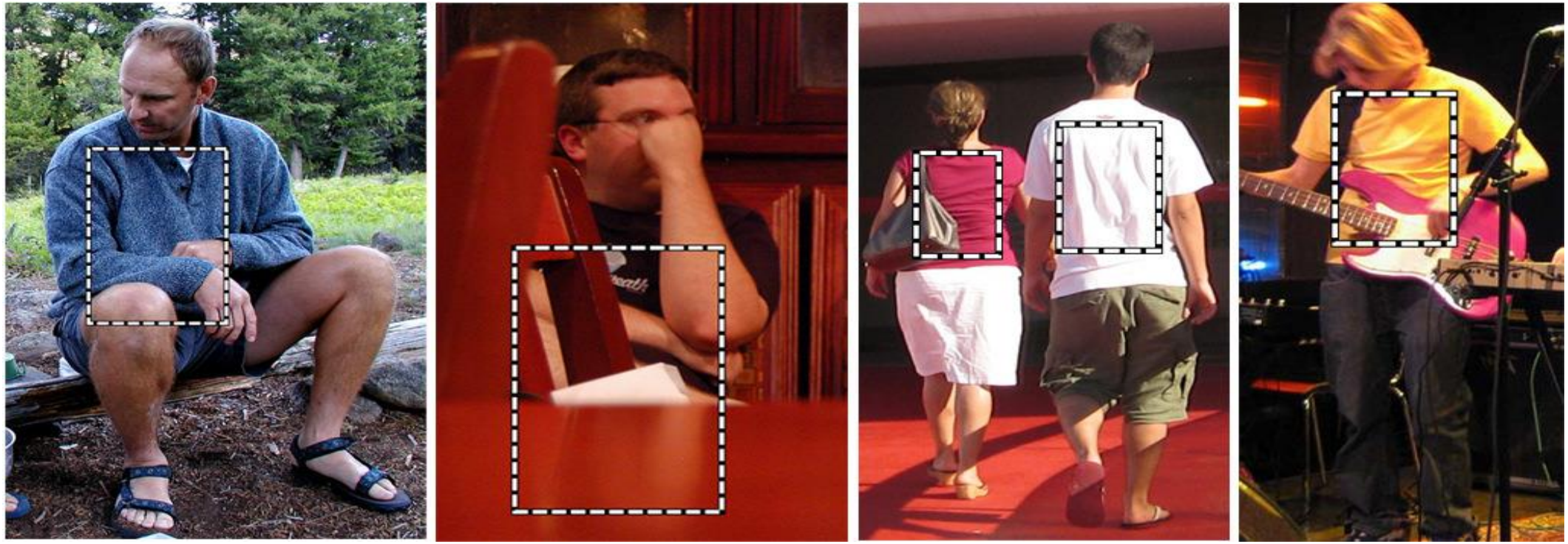




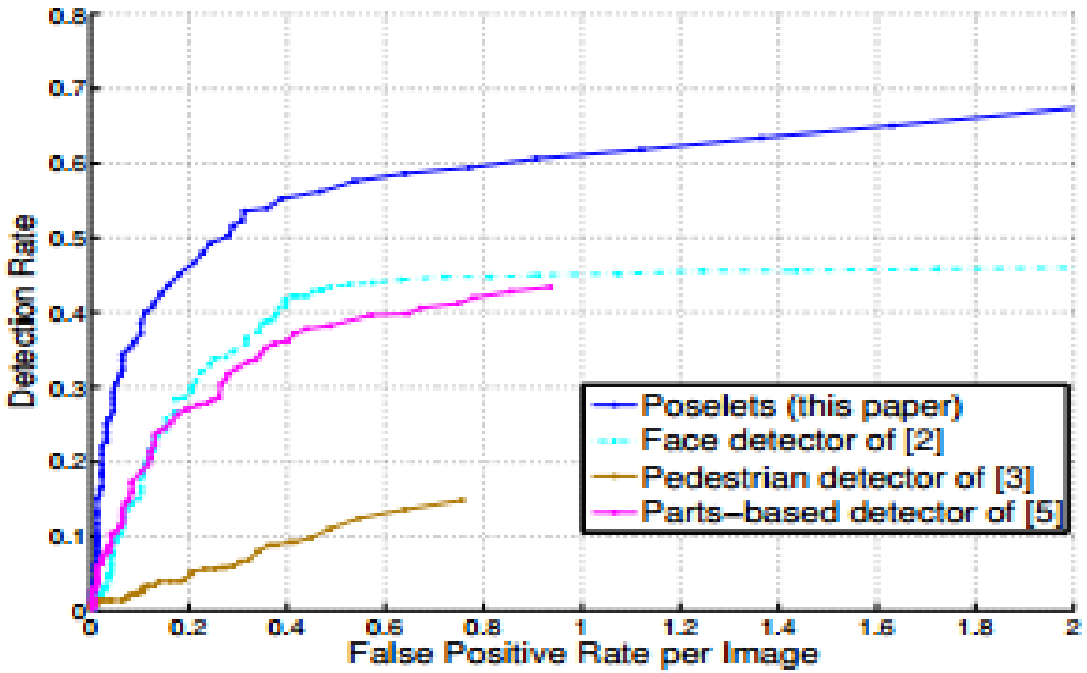
$$P(O|x) \propto \sum_i w_i a_i(x)$$

Bourdev, Lubomir, and Jitendra Malik. "Poselets: Body part detectors trained using 3d human pose annotations."

Experiments:Torso Detection



State of the art in Torso Detection



Frontal detector of Bourdev and Brandt [2]

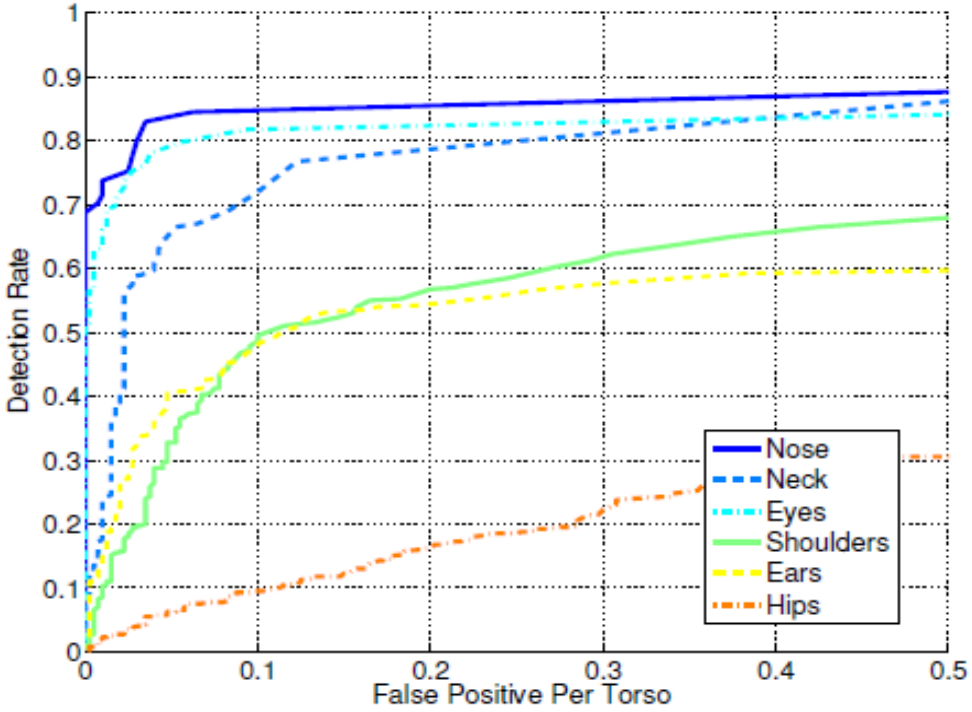
Pedestrian detector of Dalal and Triggs [3]

The parts-based deformable detector of Felzenszwalb, Mcallester [5]

Bourdev, Lubomir, and Jitendra Malik. "Poselets: Body part detectors trained using 3d human pose annotations."

Experiments

Detecting Keypoints



Detection rate of some keypoints conditioned on true positive detection.

Bourdev, Lubomir, and Jitendra Malik. "Poselets: Body part detectors trained using 3d human pose annotations."

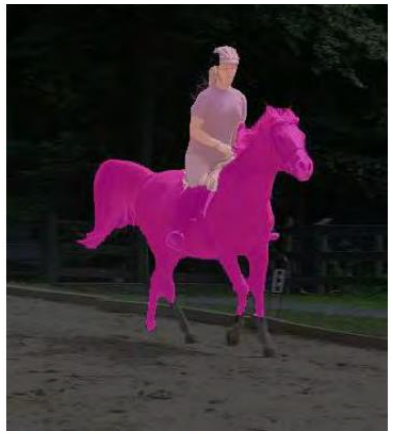
Other Uses of Poselets

Object Segmentation (CVPR 2011)

- Predict area using “soft masks”
- Deformation to match image edges
- Extends poselets to other objects, Birdlets

Activity Recognition (CVPR 2011)

- Recognition from a single image
- Use which poselets fired and to what extent to predict the activity



Bourdev, Lubomir, and Jitendra Malik. "Poselets: Body part detectors trained using 3d human pose annotations."

Some Poselets

Conclusion & Future Work

Conclusion

- The authors propose a two-layer classification/ regression model for detecting people and localizing body components.
- 3D annotation guides the search for good parts.

Future work

- Use H3D more widely
- Using poselets for recognizing complex tasks

Poselets website

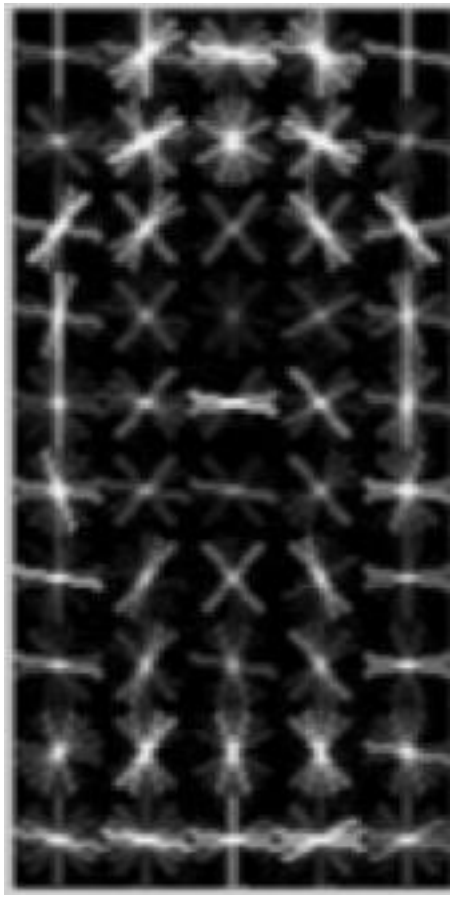
http://eecs.berkeley.edu/~lbourdev/poselets

- The set of published poselet papers
- H3D data set + Matlab tools
- Java3D annotation tool + video tutorial
- Matlab code to detect people using poselets
- Our latest trained poselets

Prior work

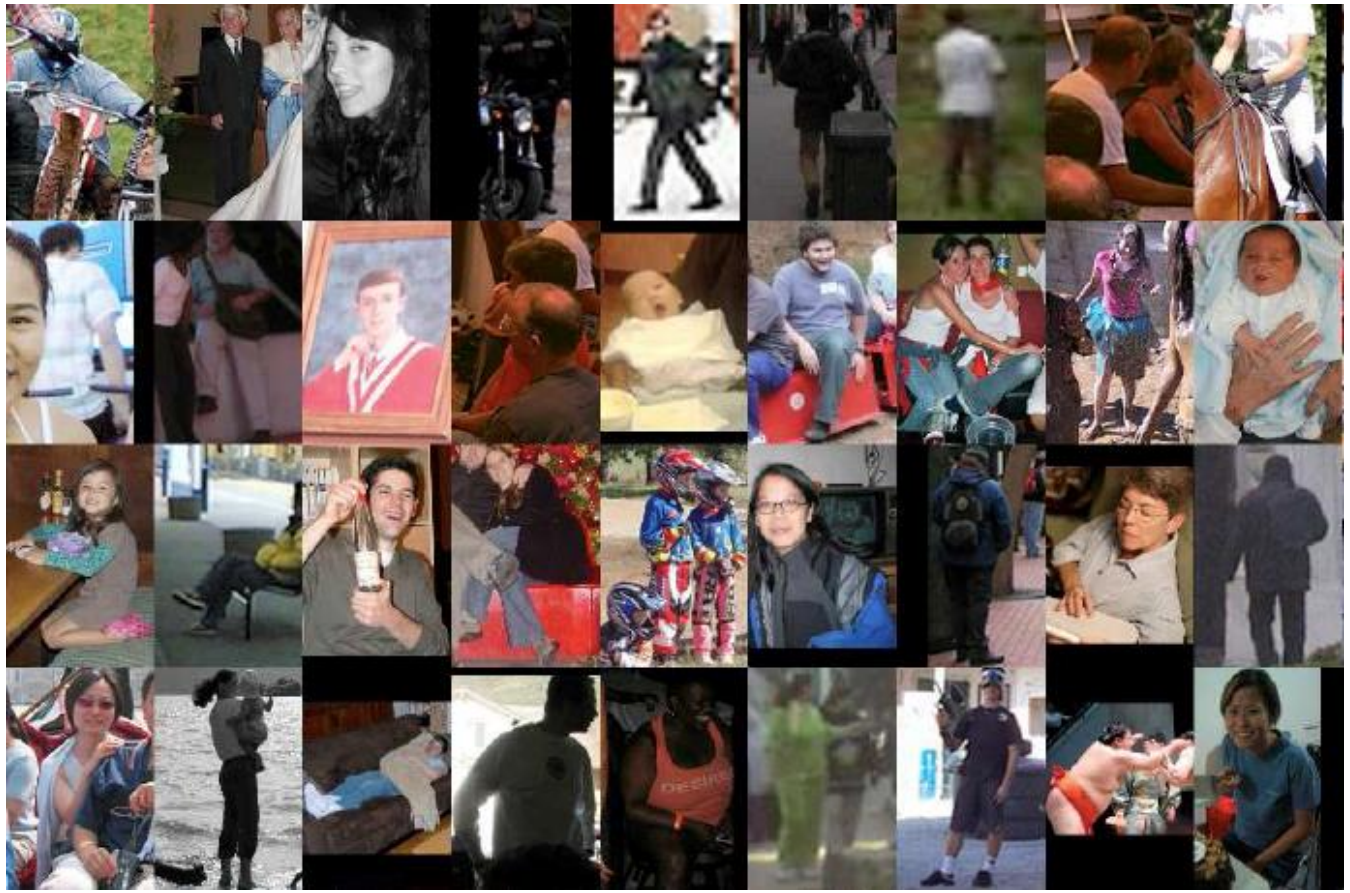
Dalal and Triggs ‘05

- Learn to classify pedestrians vs. background
- HOG + linear SVM
- Doesn’t account for variations in body pose and viewpoint



Bourdev, Lubomir, and Jitendra Malik. "Poselets: Body part detectors trained using 3d human pose annotations."

Which image has the poselet?



Bourdev, Lubomir, and Jitendra Malik. "Poselets: Body part detectors trained using 3d human pose annotations."

Which poselets are discriminative for gender?

Preferred by human subjects

Preferred by our system

[Bourdev et al., ICCV11]

Selecting a small set of complementary poselets