Event prediction

CS 590v

Applications

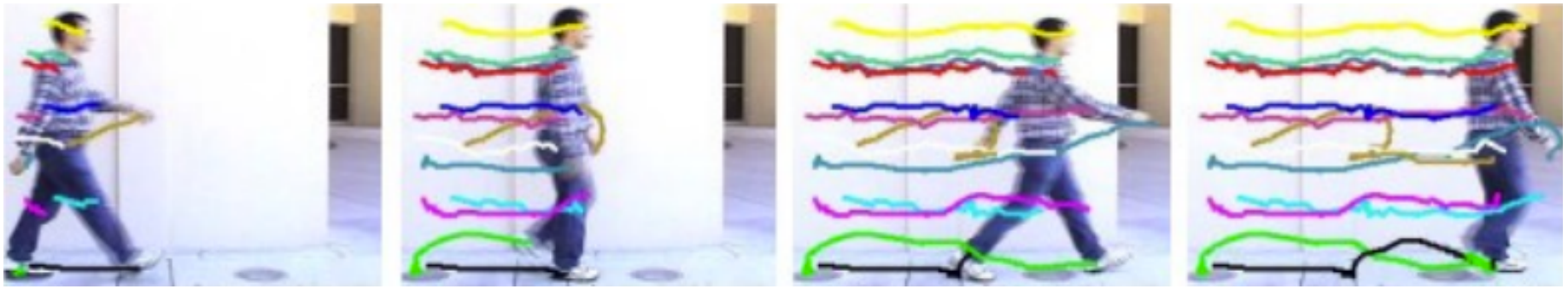
- Video search
- Surveillance
  - Detecting suspicious activities
  - Illegally parked cars
  - Abandoned bags
- Intelligent environments
  - Healthcare: fall detector
  - Healthcare: hand-washing prompter

Common approaches

- Compare keypoint (joint) trajectory shapes
- Various spatio-temporal features with classifier

Trajectories of keypoints

- Yilmaz and Shah, 2005 – UCF
- Joint trajectories in XYT space
- Compare trajectory shapes to classify actions



Trajectories of keypoints

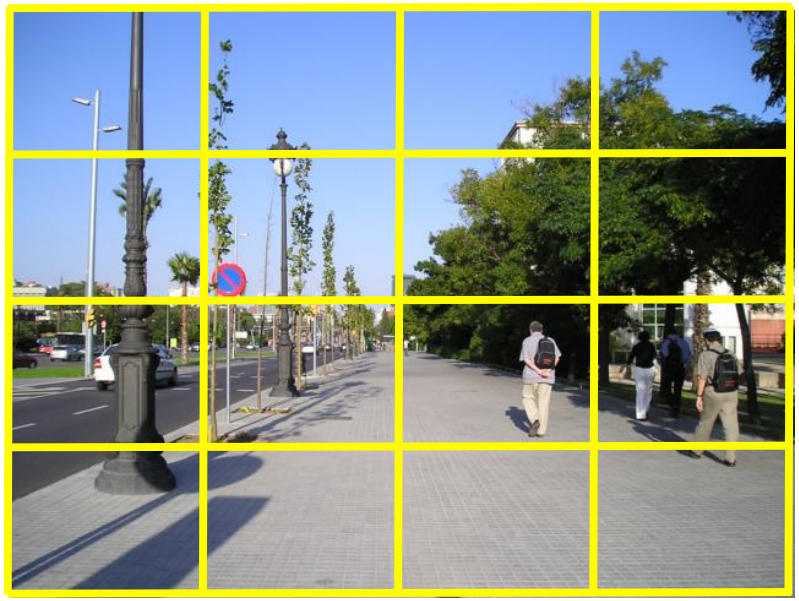
- Another approach: modeling trajectories likelihood
  - Flag low likelihood events

Scene context

- Idea: transfer event information only from similar images
- Use context to determine expected motion
  - E.g. climber on a rock wall vs. climber on a building

Features for matching images: Gist

Oliva and Torralba, 2001



- Apply oriented Gabor filters over different scales
- Average filter energy in each bin

8	orientations
4	scales
x 16	bins
512	dimensions

- Used for scene recognition
- Similar to SIFT (Lowe 1999)

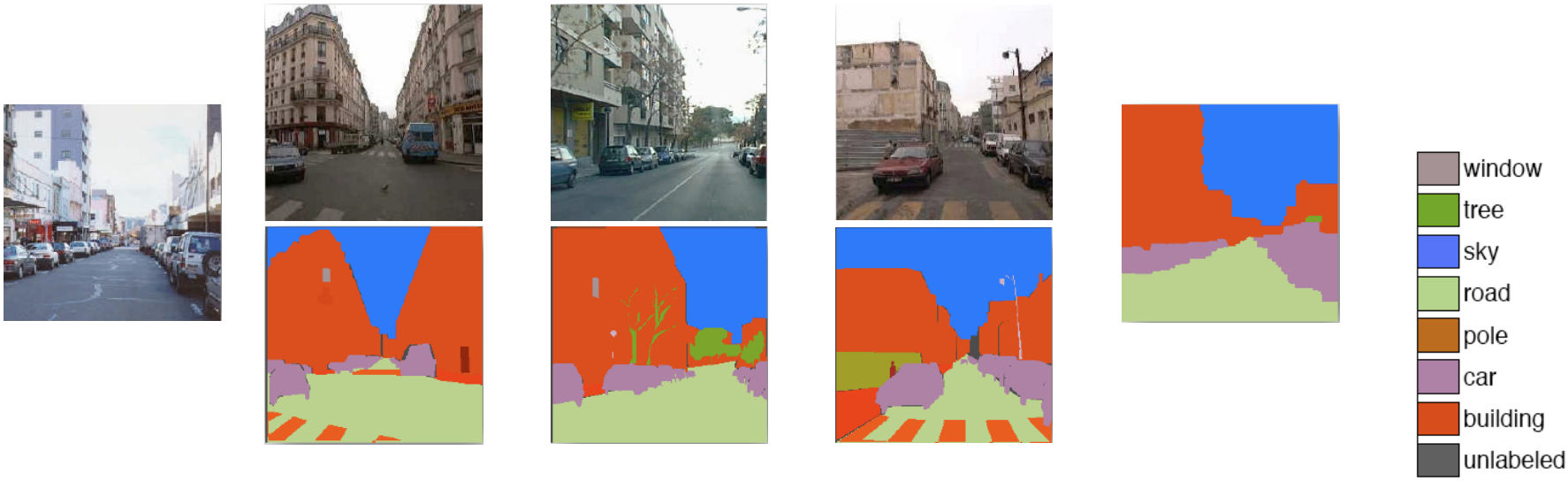
Retrieving similar images from large image databases

Image completion using Flickr images

Hays and Efros, 2007

Original Image | Input | Scene Matches | Output

Transfer of knowledge: SIFT-flow



C. Liu, J. Yuen, A. Torralba, 2009

- Can "copy and paste" segmentation labels from similar labeled example to unknown

Task

- Predict motion from static images
- Predict semantic event from static images

Approach: high-level

- Model video as trajectories of keypoints
  - Cluster of trajectories for each object

- Global similarity measure

Tracking key points

- KLT feature tracker
  - Solve for the displacement *d* that optimizes a dissimilarity metric, per pair of consecutive frames
- Cluster trajectories into objects
  - By average distance between them

Comparing track clusters

- Create spatial histogram of cluster velocities
- Sum of histogram intersections is the similarity

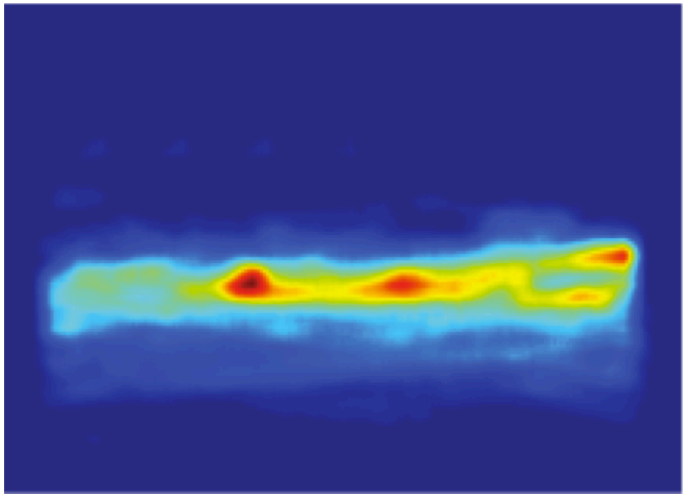
$$\mathbf{S}_{clust}(\mathbf{C}_1, \mathbf{C}_2) \equiv \mathbf{I}(H_1, H_2) = \sum_{i \in \mathbf{G}} \sum_{b=1}^{8} min\Big(H_1(i,b), H_2(i,b)\Big)$$

- 1-level spatial pyramid matching

Predicting local motion

- Average motion over *N* nearest neighbors

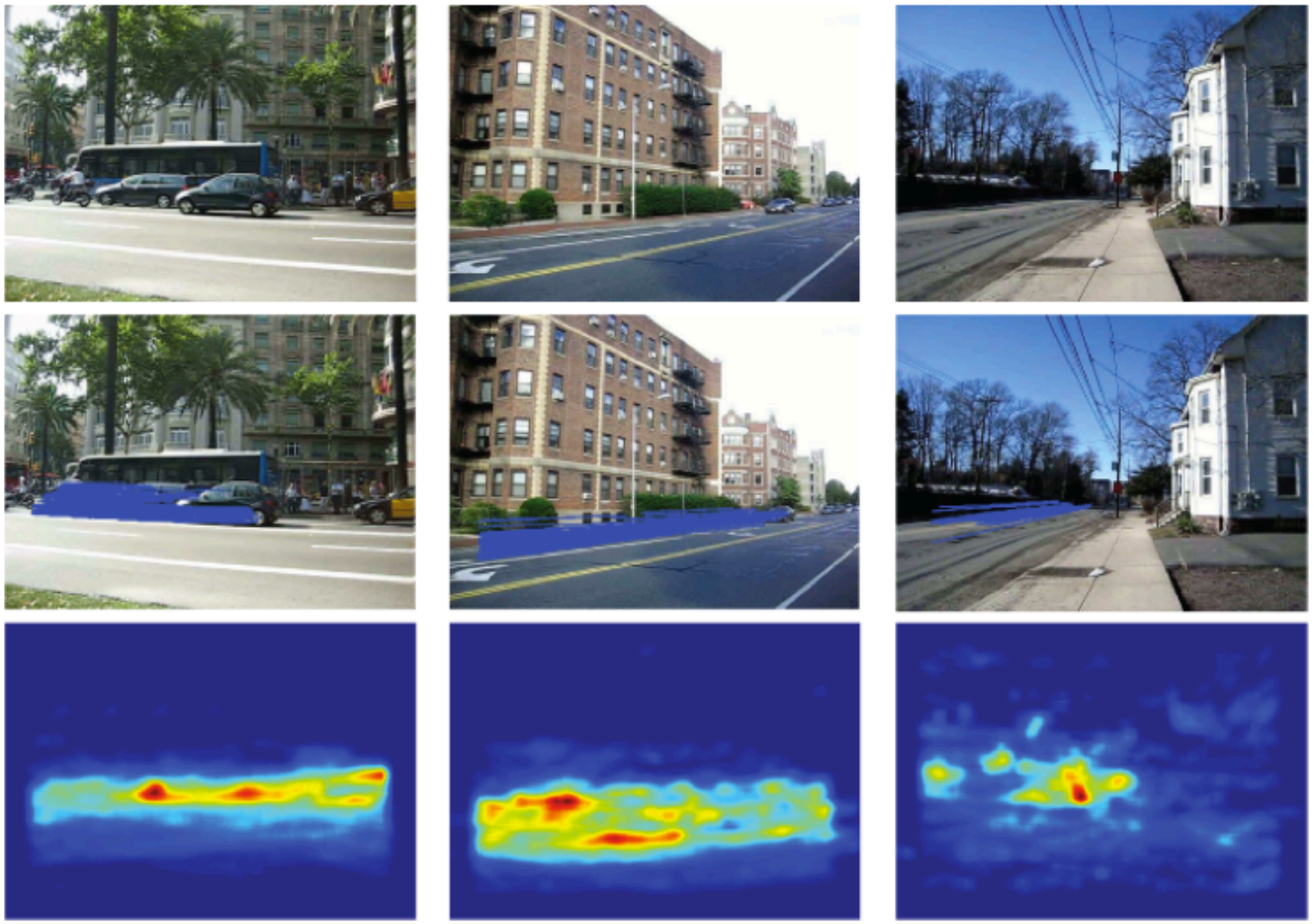
$$p(motion|x, y, \text{scene}) = \frac{1}{N}\sum_{i}^{N}\frac{1}{M_i}\sum_{j}^{M_i}\sum_{t\in D} K(x - x_{i,j}(t), y - y_{i,j}(t); \sigma)$$



Event prediction

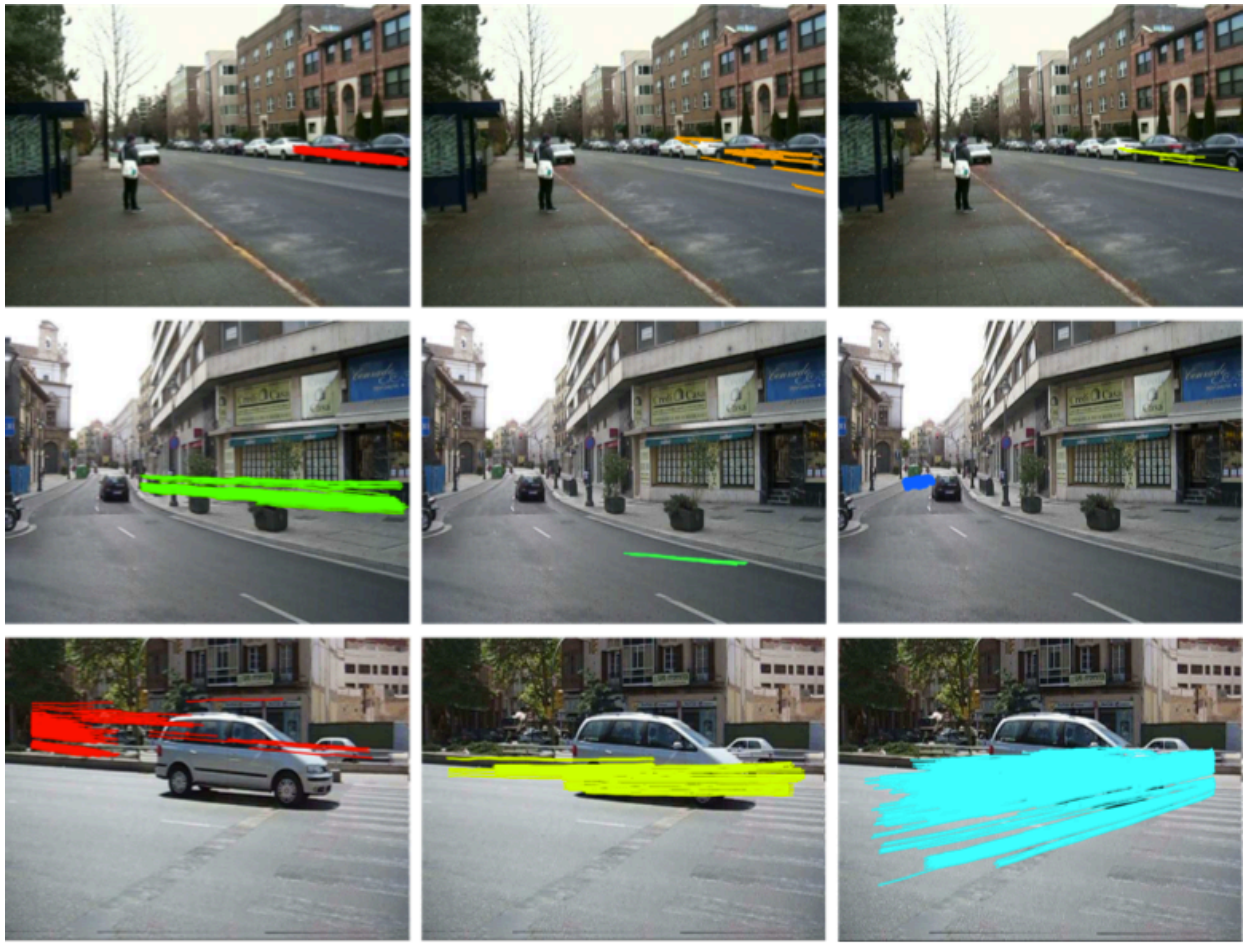
- Match query with similar track clusters from database
- Cluster retrieved tracks to reduce redundant results

Results: local motion prediction

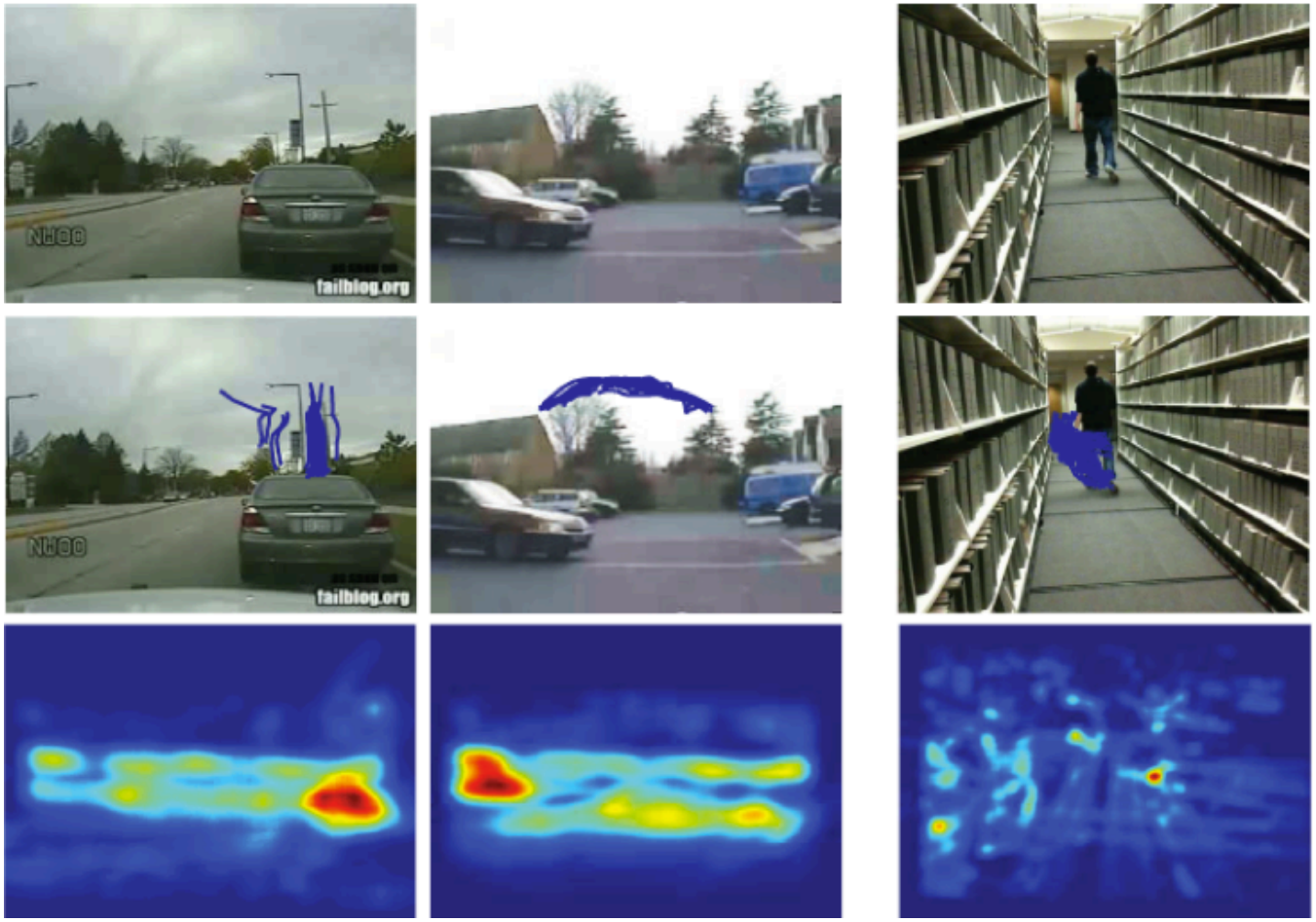


Results: event prediction

Results: event prediction



Unusual events



Overall discussion

- Still need training data from scenes similar to target
- Possible corruption by unusual events in the training set?
- Applications
  - Unusual event detector
  - Knowledge transference between environments

The end