Detecting Categories in News Video Using Image Features

Slav Petrov, Arlo Faria, Pascal Michaillat, Alex Berg,
Andreas Stolcke, Dan Klein, Jitendra Malik

System Overview

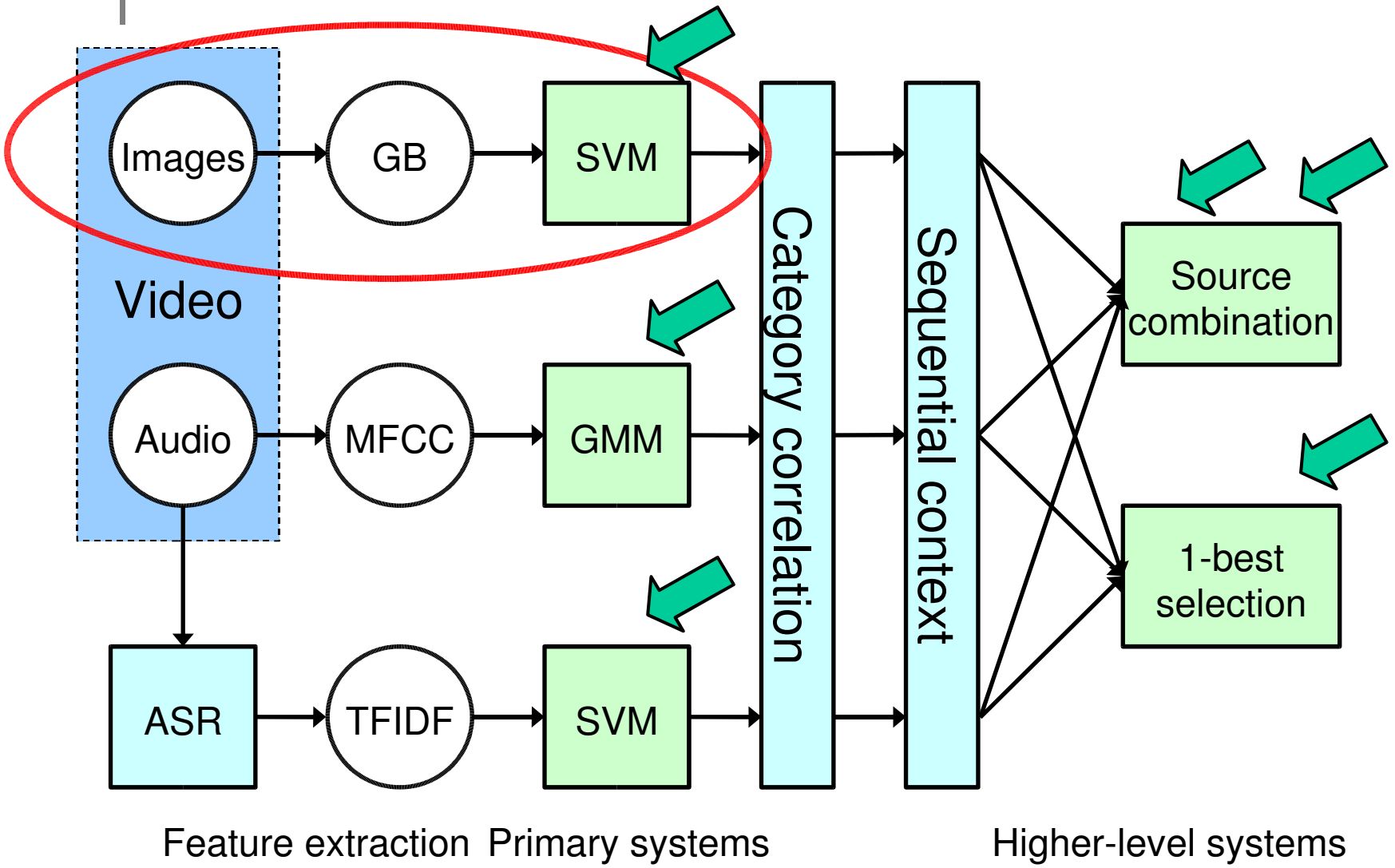


Image Features in TrecVid '05

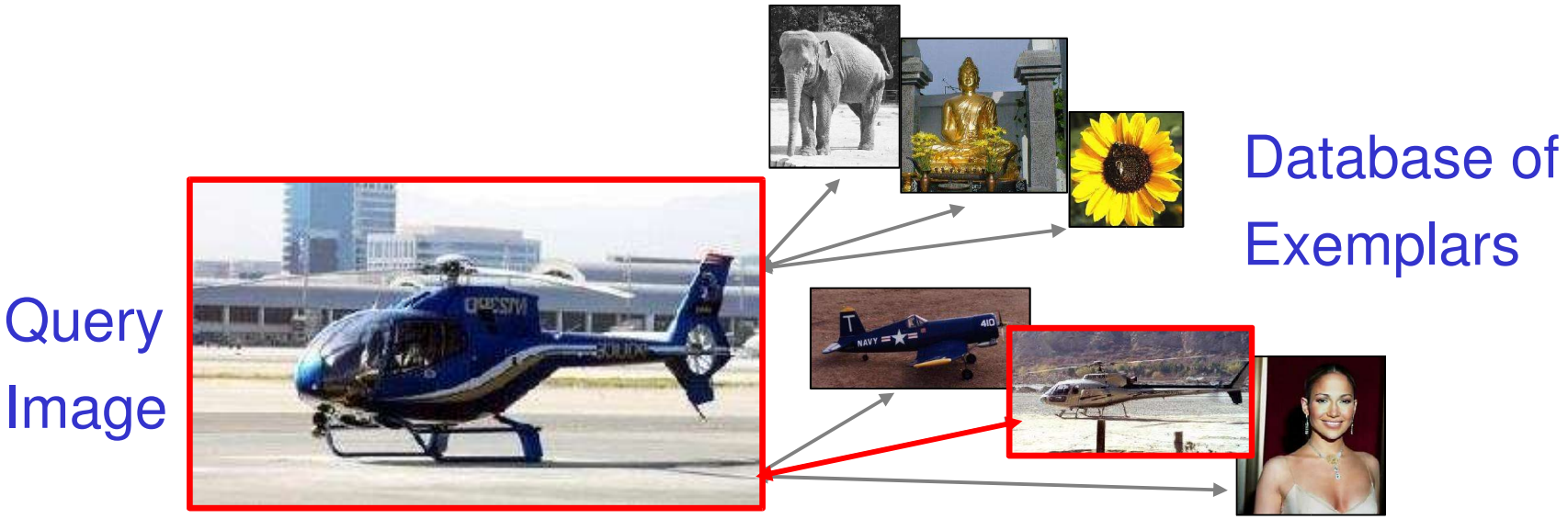
- IBM:
  - Color Histogram
  - Color Correlogram
  - Color Moments
  - Co-occurence Texture
  - Wavelet Texture Grid
  - Edge Histogram Layout
- CMU (local):
  - Color Histograms (in different color spaces)
  - Texture Histograms
  - Edge Histograms
- Columbia (part based model):
  - Color
  - Texture
  - Size
  - Spatial Relation
- Tsinghua (local and global):
  - Color Auto-Correlograms
  - Color Coherence Vectors
  - Color Histograms
  - Color Moments
  - Edge Histograms
  - Wavelet Texture

Image Features in TrecVid '05

		IBM	CMU	Columbia	Tsinghua
Color	Histograms	✓	✓	✓	✓
	Moments	✓			✓
	Correlograms	✓			✓
Texture	Histograms	✓	✓	✓	
	Wavelets	✓			
Edge Histograms		✓	✓		✓

Exemplars for Recognition

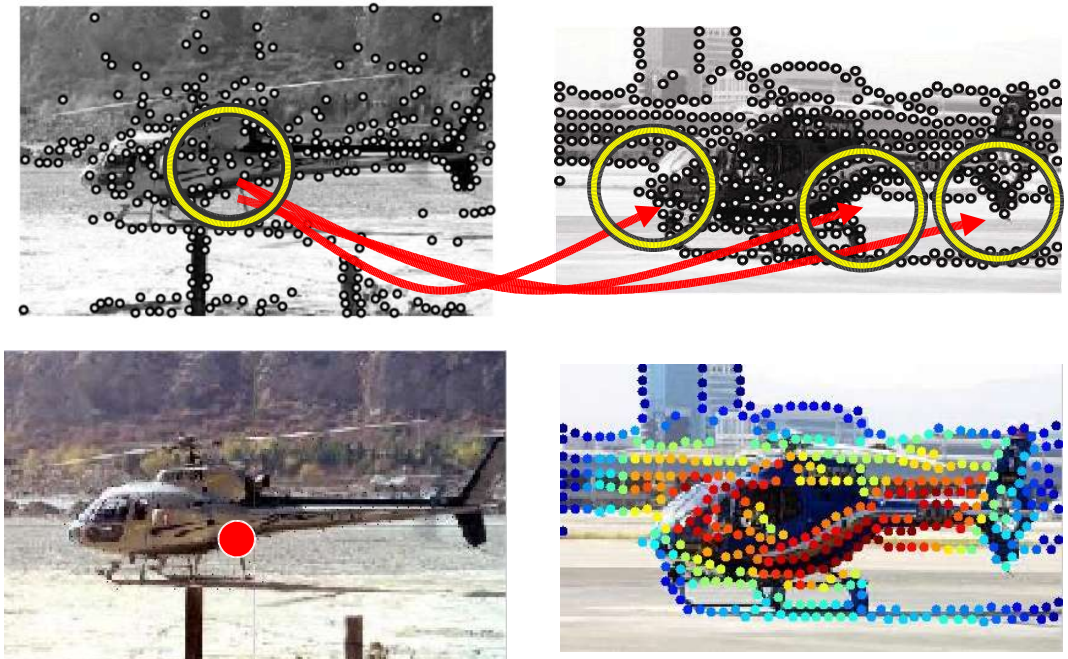
- Use exemplars for recognition
- Compare query image and each exemplar using shape cues



Finding similar patches

Exemplar

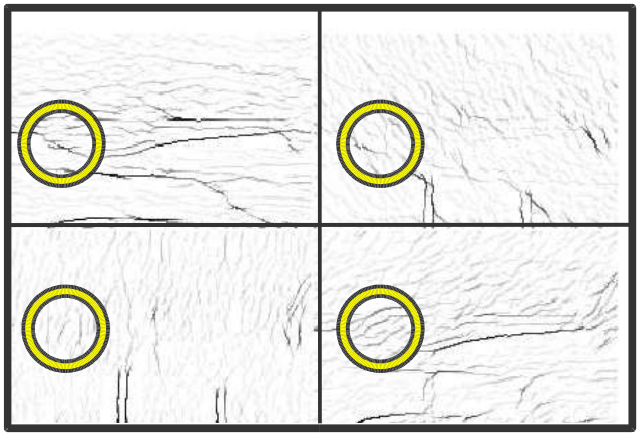
Query



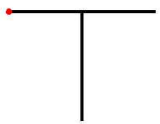
Geometric Blur (Local Appearance Descriptor)

[Berg & Malik, CVPR'01]

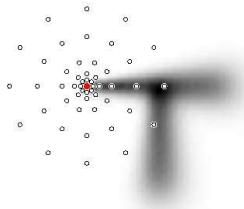
Compute sparse channels from image



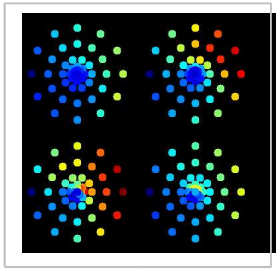
Extract a patch in each channel



Idealized signal



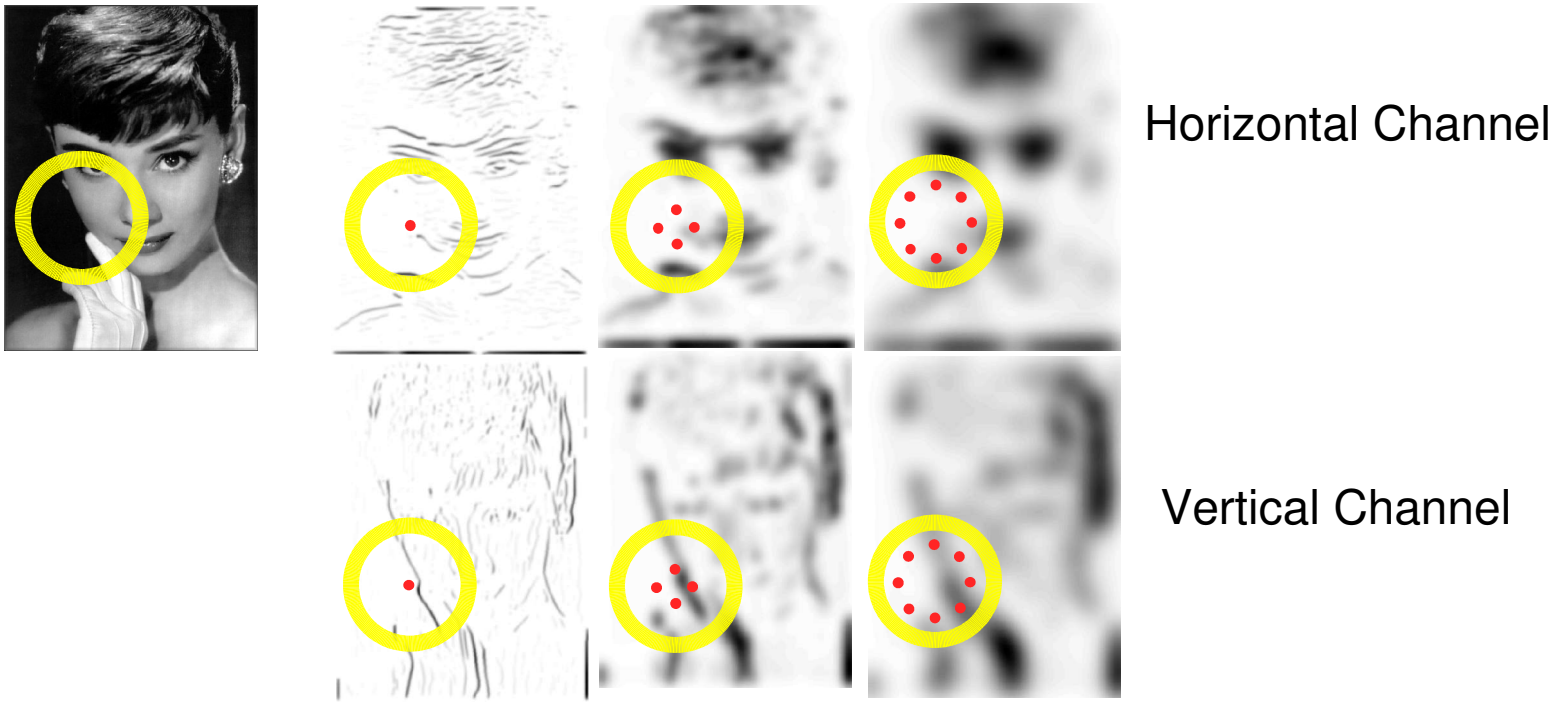
Apply spatially varying blur and sub-sample



Descriptor is robust to small affine distortions

GB in Practice

- In practice compute discrete blur levels for whole image and sample as needed for each feature location.



[Berg, Berg & Malik, CVPR'05]

Comparing Images

- Sample 200 GB features from edge points
- Dissimilarity from A to B is

$$D(A \rightarrow B) \propto \sum_i \min_j ||F_i^A - F_j^B||^2$$

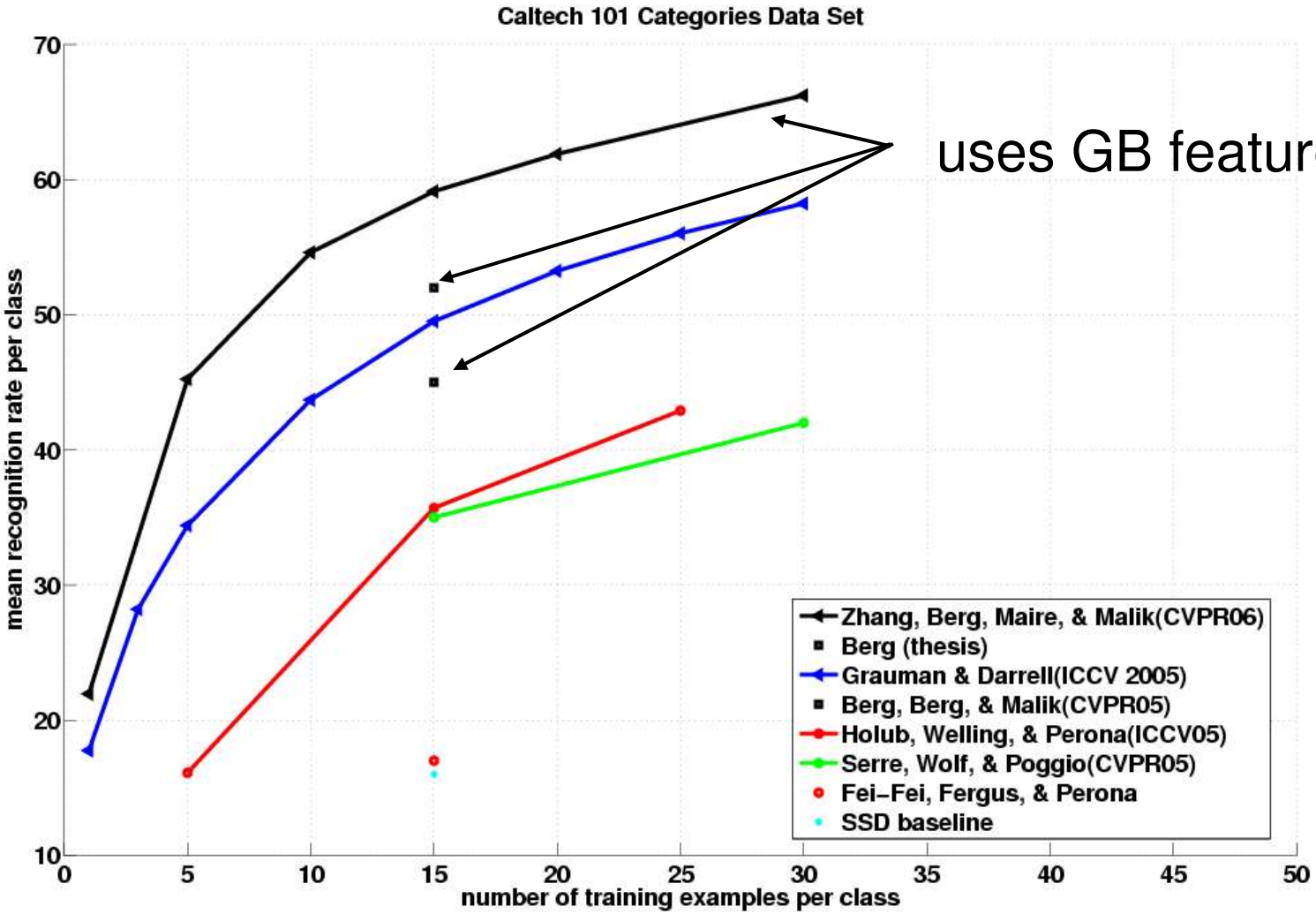
where the F_x are the GB features.

Caltech 101 Dataset

- Object Recognition Benchmark
- 101 Categories:
  - Stereotypical pose
  - Little clutter
  - Objects centered
  - One object per image

[Zhang, Berg, Maire & Malik, CVPR'06]

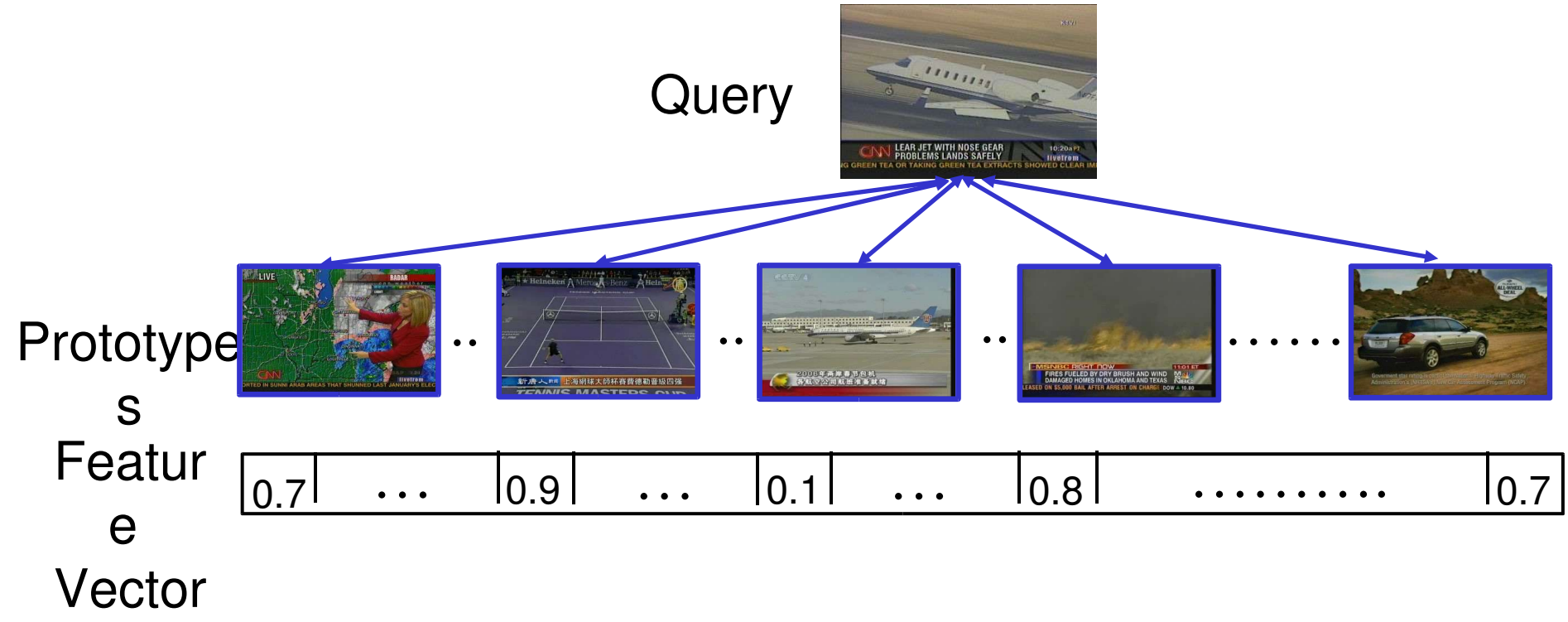
Caltech 101 Results



uses GB features

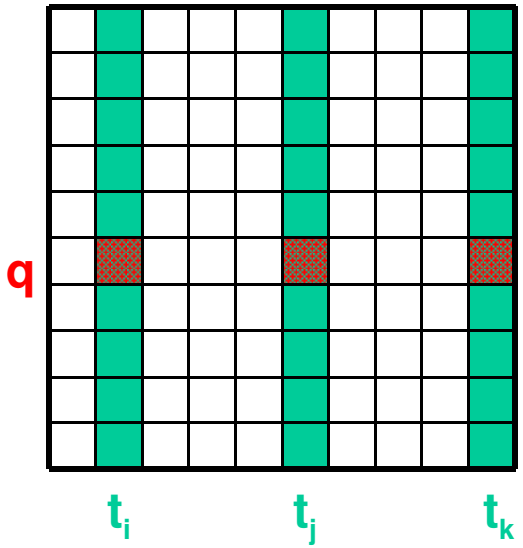
Primal features for SVM

- Compare to 50 prototypes from each class
- Use distances as feature vector for an SVM

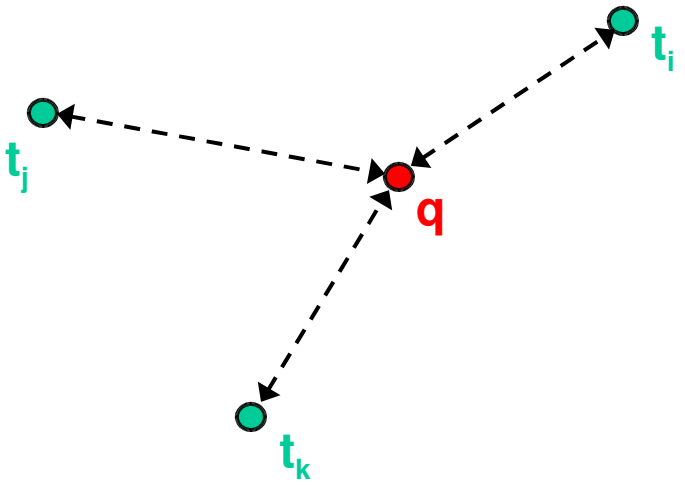


SVM features interpretation

- Slices of the Kernel Matrix:



- Fixed-points in a higher dimensional vector space:



SVM Specifics

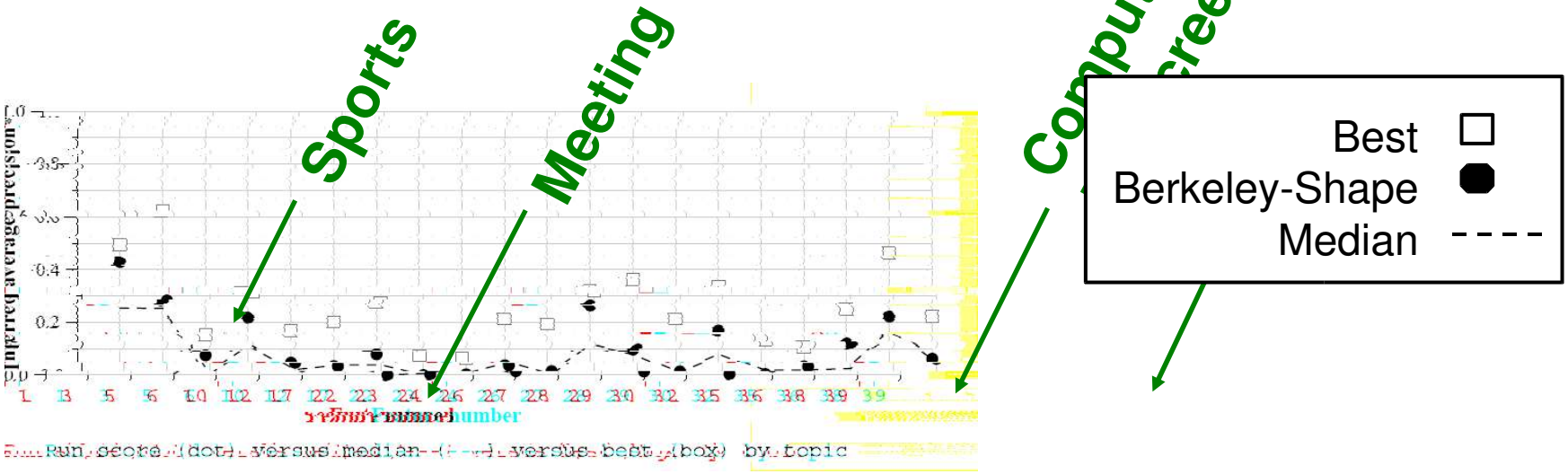
- SVM^{light} package
- Same parameters for all categories:
  - Linear kernel
  - Default regularization parameter
  - Asymmetric cost doubling the weight of positive examples

Results ’06

Results ’05

Berkeley-Shape mAP = 0.38

Best ’05 (IBM) mAP = 0.34



mAP = 0.11

Limitations

- Several objects per image:

- Features do not capture:
  - Different Scales
  - Color

Conclusions

- Shape is an important cue for object recognition.
- System that uses shape features only can have competitive performance.
- Shape features are orthogonal to features used in the past.

Thank You!

petrov@eecs.berkeley.edu