Spatial Weighting for Bag-of-Features

Authors: Marcin Marszałek, Cordelia Schmid
Presented by: Brendan Younger

Better Bags-of-Features

- Better Kernels
  - Pyramid Match Kernel, Grauman & Darrell, 2005
  - Mercer Kernels, Lyu, 2005
- Interest Point Detection
  - Distinctive features from keypoints, Lowe, 2004
- Localization
  - Combined segmentation & categorization, Liebe et. al., 2004

Better Bags-of-Features

- Better Kernels
  - Okay, but still no spatial information
- Interest Point Detection
  - Uses Hough transform, so restricted set of shapes
- Localization
  - Finds “interesting” parts okay, but can’t fill in the rest

Spatial weighting

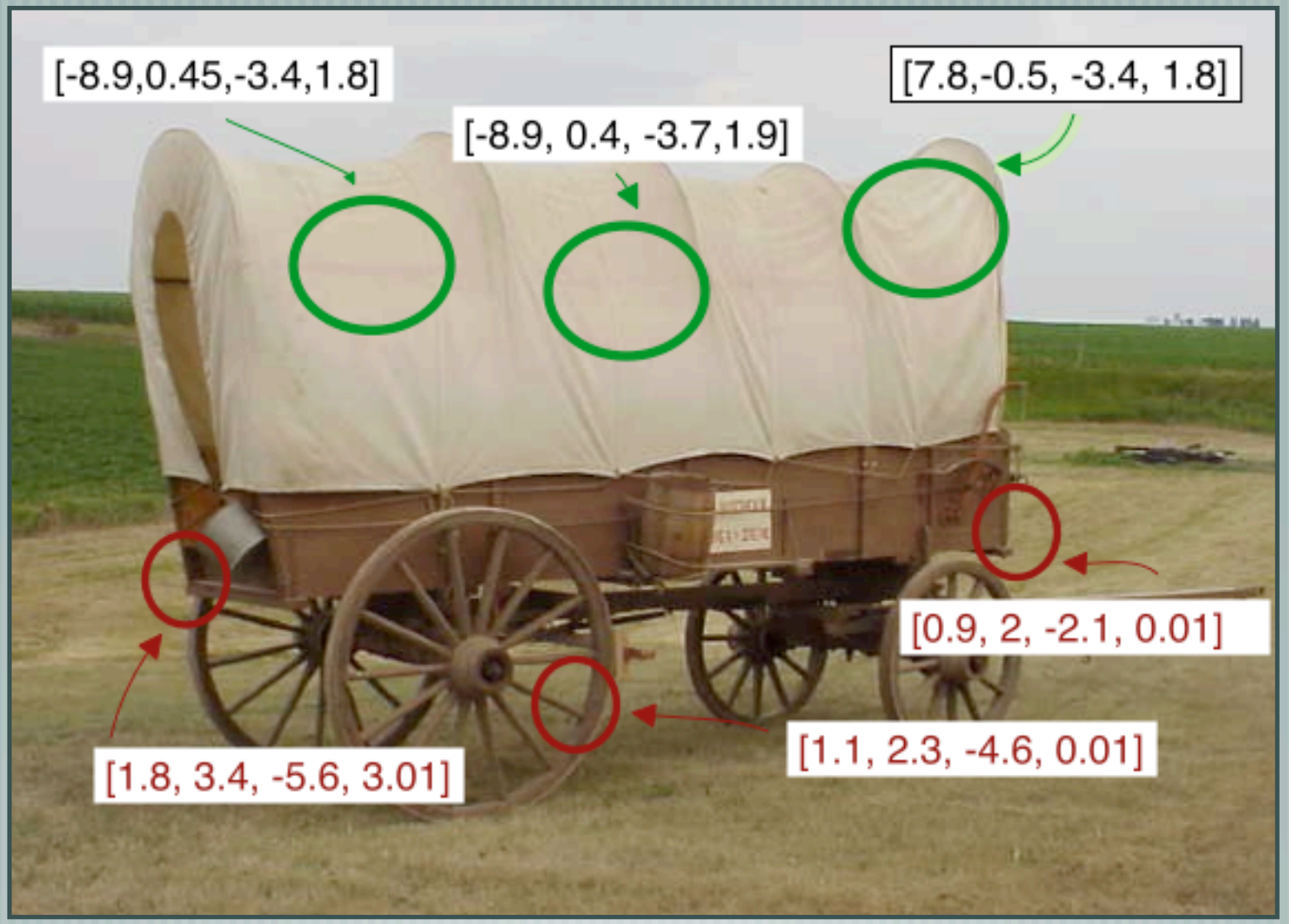
Features help other features in their neighborhood

Overview of Classification

- Interest-point detection
- SIFT descriptor at each interest point
- Find nearest descriptors in vocabulary
- Create segmentation image based on segmentations from training set
- Weight each feature with segmentation image
- Build histogram of features and use SVM to classify

Local Features

- Corner Regions: Harris-Laplace Detector (HS)
- "Blob"-like Regions: Laplacian Detector (LS)
- 128-dimensional Descriptor: Lowe's SIFT
  - Normalized descriptor for illumination invariance
- Vocabulary: K-means clustering; 1,000 features
  - Classification is insensitive to choice of vocabulary

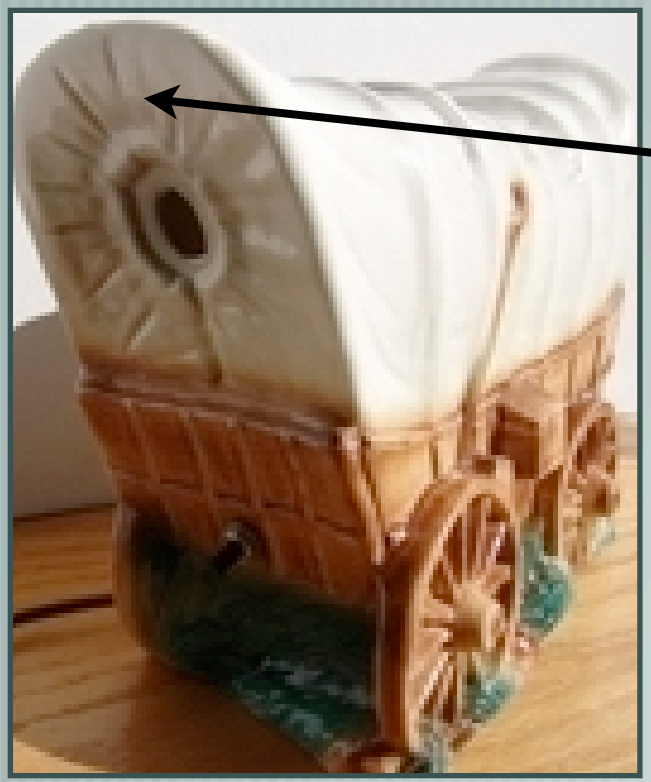

[-8.9,0.45,-3.4,1.8]
[-8.9, 0.4, -3.7,1.9]
[7.8,-0.5, -3.4, 1.8]
[0.9, 2, -2.1, 0.01]
[1.1, 2.3, -4.6, 0.01]
[1.8, 3.4, -5.6, 3.01]

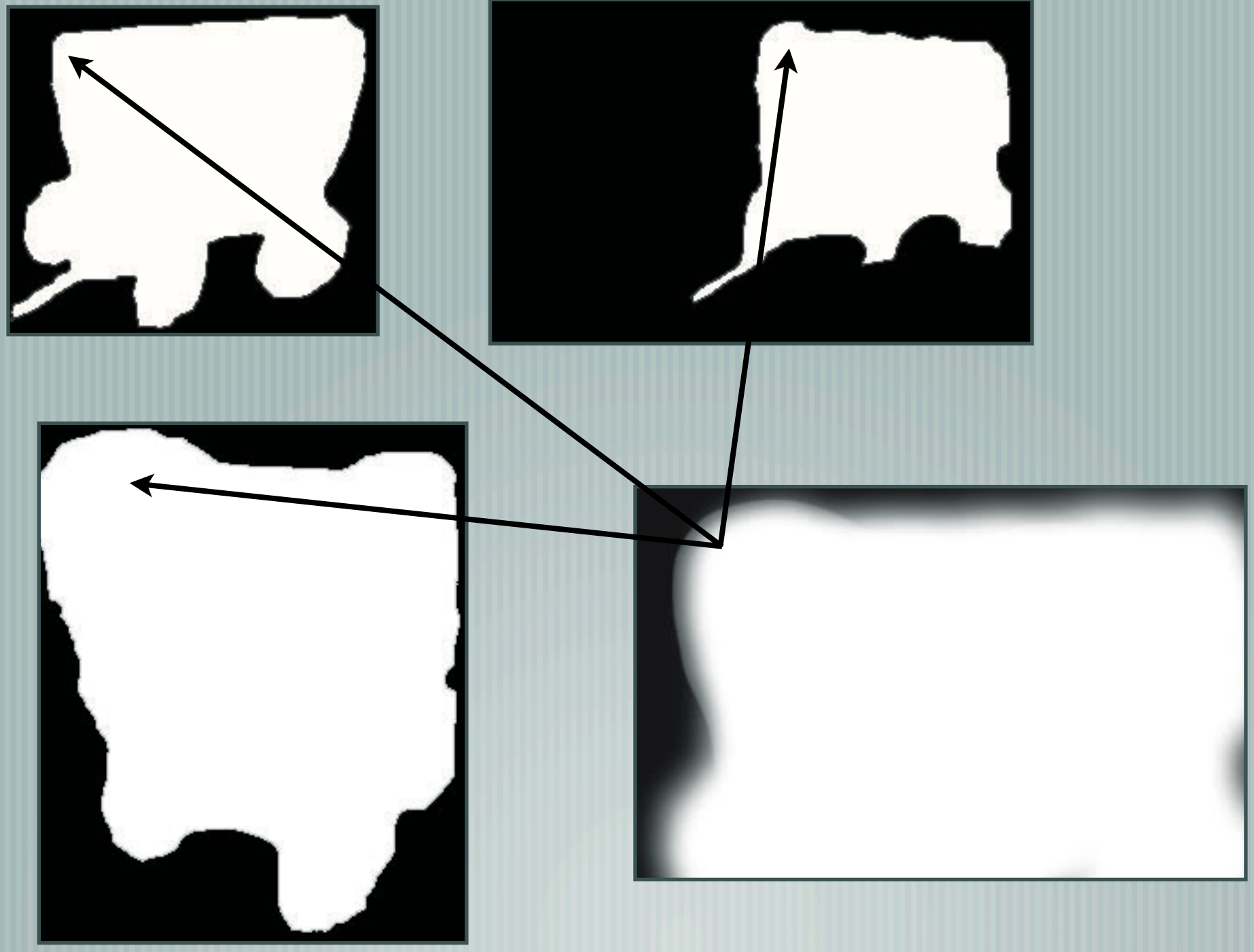
Segmentation

For each feature in test image:

- Find nearby features in training data
- Match location and orientation of both features
- Blur segmentation of training image based on distance between features
- Add blurred segmentation to computed segmentation

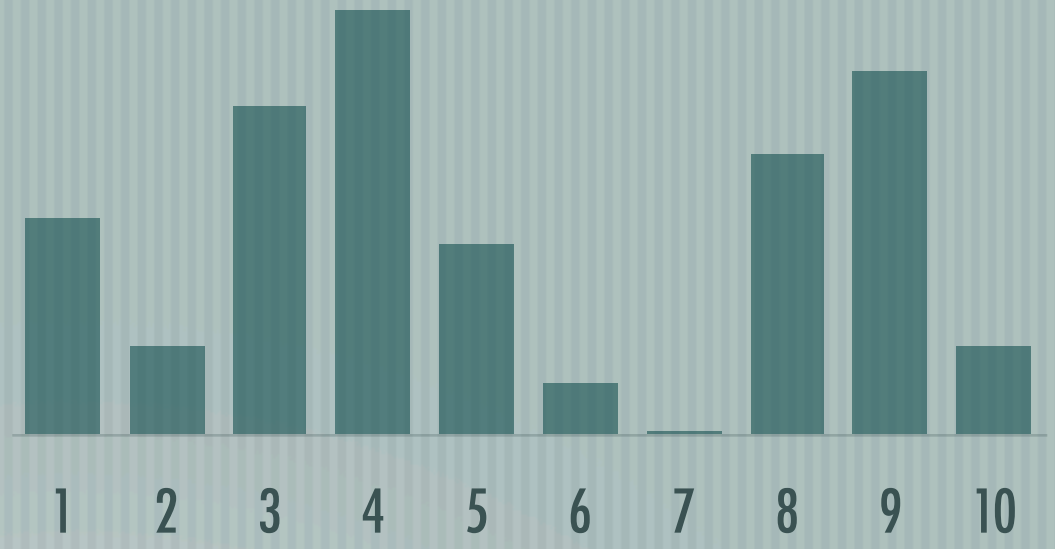
Lather; rinse; repeat





Histograms and Classification

- Using the segmentation, weight each feature
- Place features in the bucket of the nearest vocabulary feature
- Apply the class-specific SVM to the histogram



$$\sum_i \pm 1 \; w_i \; e^{\frac{1}{2A}\sum_n \frac{(h_{in}-h_{jn})^2}{h_{in}+h_{jn}}}$$

Object Localization

- Computed a segmentation just to classify?
- Use the segmentation to localize and object
- Improve the localization by re-running the algorithm
- Each time, there are fewer background features to blur the segmentation