Lecture 13: Visual Bag of Words

Visual bag of words: applications

Juan Carlos Niebles and Jiajun Wu

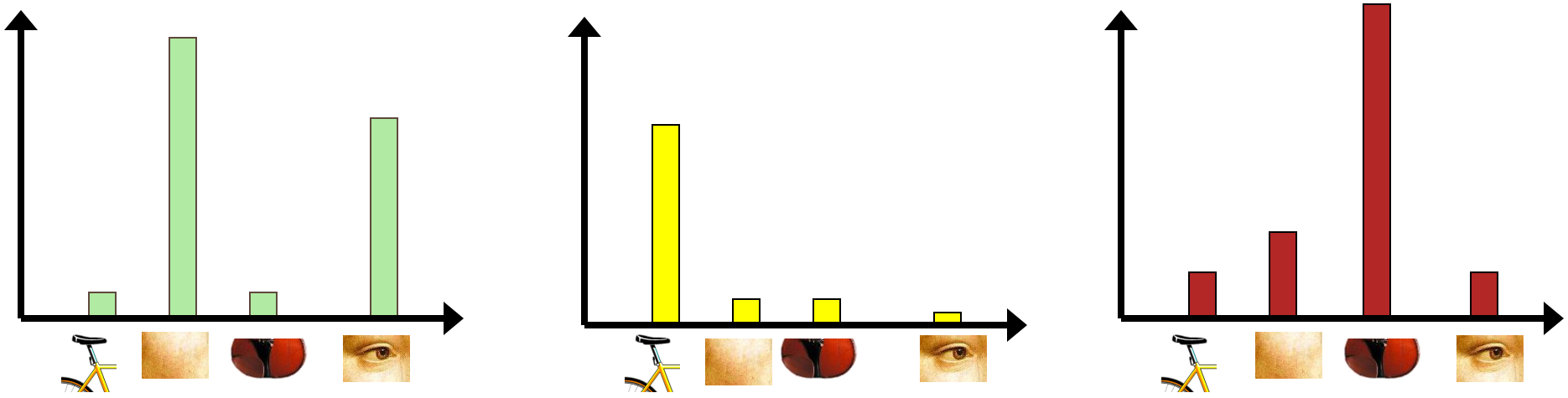
CS131 Computer Vision: Foundations and Applications

What will we learn today?

- Visual bag of words: applications
  - Image search
  - Action recognition

Image classification

- Given the bag-of-features representations of images from different classes, how do we learn a model for distinguishing them?



Uses of the BoW representation

- Treat as the input feature vector for a standard classifier
  - e.g k-nearest neighbors, support vector machine

- Cluster BoW vectors over image collection
  - Discover visual themes

What will we learn today?

- Visual bag of words: applications
  - Image search
  - Action recognition

Large-scale image search

11,400 images of game covers
(Caltech games dataset)

- Bag-of-words models have been useful in matching an image to a large database of object *instances*

how do I find this image in the database?

Large-scale image search

Build the database:

- Extract features from the database images
- Learn a vocabulary using k-means (typical k: 100,000)
- Compute *weights* for each word
- Create an inverted file mapping words → images

Weighting the words

- Just as with text, some visual words are more discriminative than others

the, and, or vs. ***cow, AT&T, Cher***

- the bigger fraction of the documents a word appears in, the less useful it is for matching
  - e.g., a word that appears in *all* documents is not helping us

Large-scale image search

- Cons:
  - performance degrades as the database grows

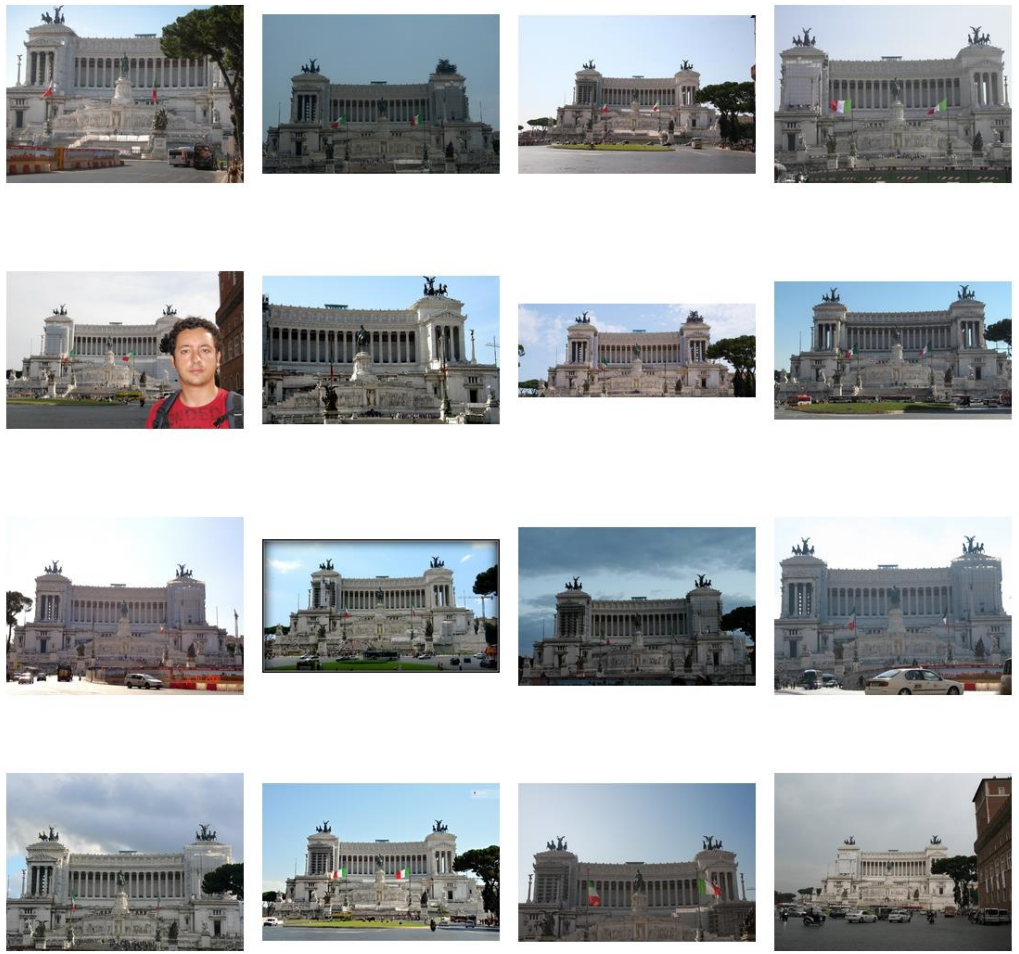
Large-scale image search

- Pros:
  - Works well for CD covers, movie posters
  - Real-time performance possible

real-time retrieval from a database of 40,000 CD covers

Nister & Stewenius, **Scalable Recognition with a Vocabulary Tree**

Example bag-of-words matches



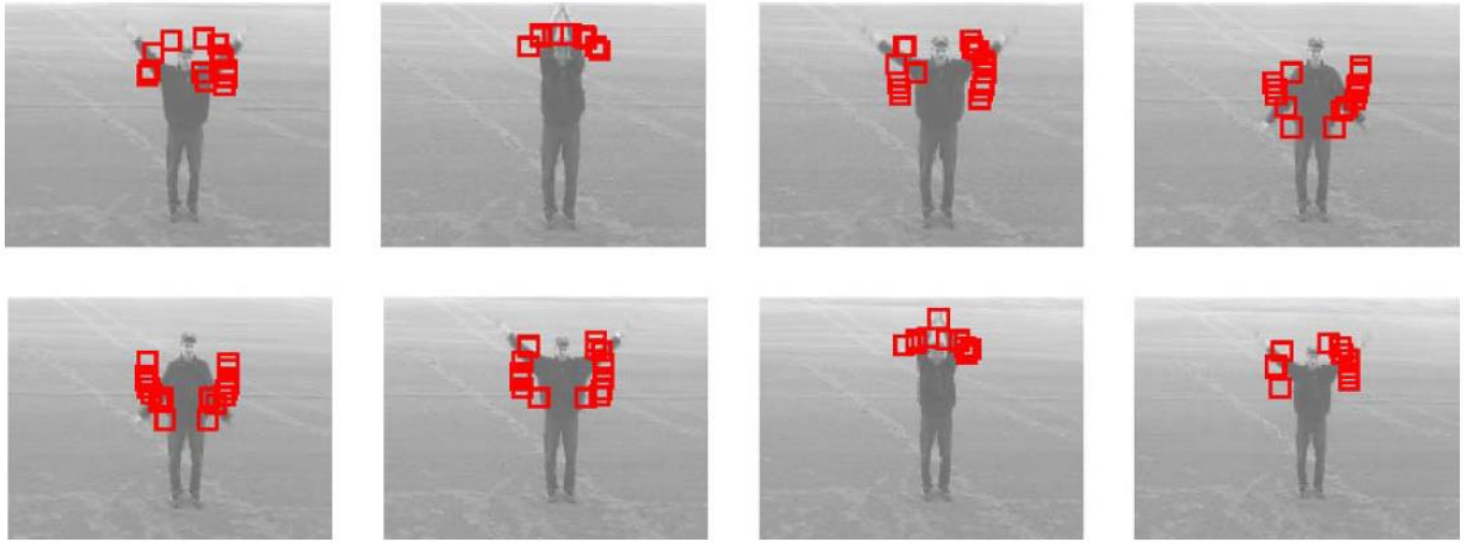
Example bag-of-words matches

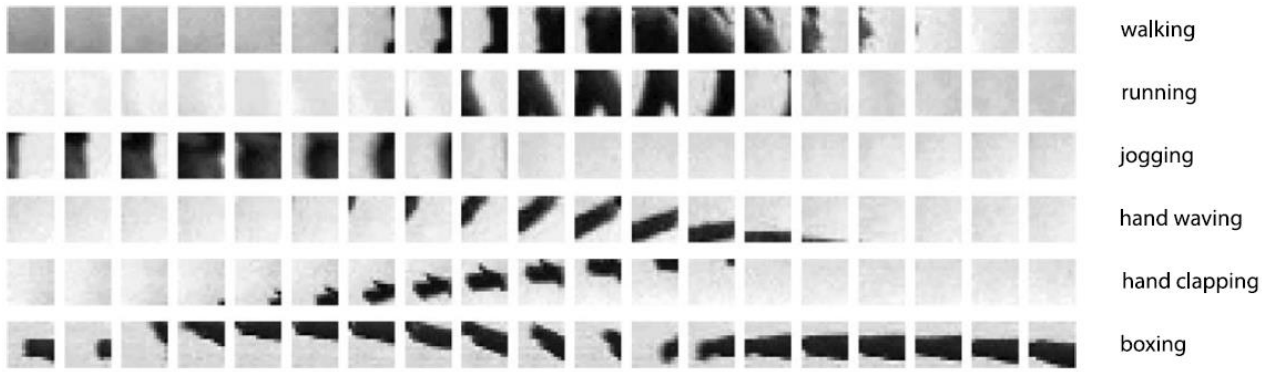
What will we learn today?

- Visual bag of words: applications
  - Image search
  - Action recognition

Bags of features for action recognition

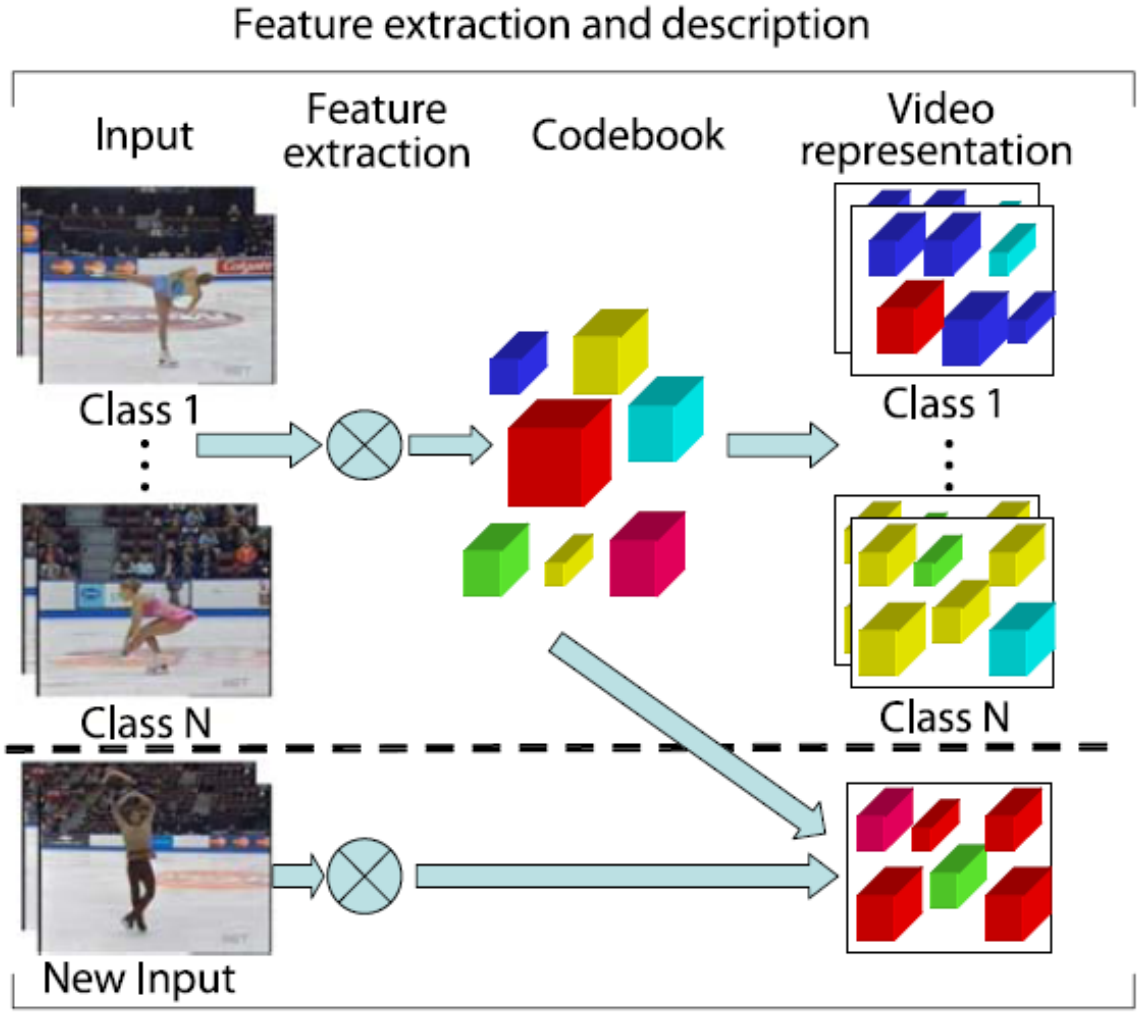
Space-time interest points





Juan Carlos Niebles, Hongcheng Wang and Li Fei-Fei, **Unsupervised Learning of Human Action Categories Using Spatial-Temporal Words**, IJCV 2008.

Bags of features for action recognition



Juan Carlos Niebles, Hongcheng Wang and Li Fei-Fei, **Unsupervised Learning of Human Action Categories Using Spatial-Temporal Words**, IJCV 2008.

Summary

- Visual bag of words: applications
  - Image search
  - Action recognition