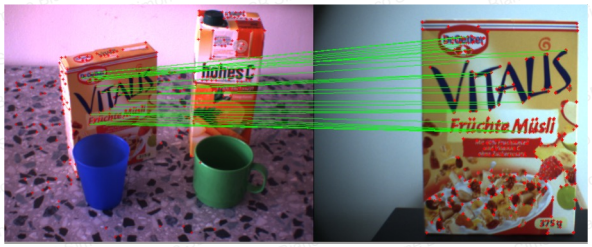
Bag of Visual-Words

76

Small Recap



SIFT + RANSAC → Instance Recognition

SIFT + Bag of Words → Object Recognition

→ Global image descriptor

→ more...

77

Bag-of-words: motivation

President George W. Bush Speech in 2001

ADDRESS TO THE JOINT SESSION OF THE 107TH CONGRESS
UNITED STATES CAPITOL
WASHINGTON, D.C. SEPTEMBER 20, 2001

Mr. Speaker, Mr. President Pro Tempore, members of Congress, and fellow Americans: In the normal course of events, Presidents come to this chamber to report on the state of the Union. Tonight, no such report is needed. It has already been delivered by the American people. We have seen it in the courage of passengers, who rushed terrorists to save others on the ground — passengers like an exceptional man named Todd Beamer. And would you please help me to welcome his wife, Lisa Beamer, here tonight. We have seen the state of our Union in the endurance of rescuers, working past exhaustion. We have seen the unfurling of fl ags, the lighting of candles, the giving of blood, the saying of prayers — in English, Hebrew, and Arabic. We have seen the decency of a loving and giving people who have made the grief of strangers their own. My fellow citizens, for the last nine days, the entire world has seen for itself the state of our Union — and it is strong. Tonight we are a country awakened to danger and called to defend freedom. Our grief has turned to anger, and anger to resolution. Whether we bring our enemies to justice, or bring justice to our enemies, justice will be done. I thank the Congress for its leadership at such an important time. All of America was touched on the evening of the tragedy to see Republicans and Democrats joined together on the steps of this Capitol, singing "God Bless America." And you did more than sing; you acted, by delivering $40 billion to rebuild our communities and meet the needs of our military. Speaker Hastert, Minority Leader Gephardt, Majority Leader Daschle and Senator Lott, I thank you for your friendship, for your leadership and for your service to our country. .. (to be continued)

- We want to represent the **topic** of the document in a **compact** way
- **Orderless** document representation: frequencies of words from a dictionary
  Salton & McGill (1983)

78

Bag-of-words: motivation

US Presidential Speeches Tag Cloud
`http://chir.ag/projects/preztags/`

79

Bag-of-words: motivation

80

Bag-of-words: motivation

81

Bag of visual words (Bag of features)

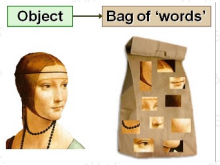
82

Bag-of-words

- Fei Fei et al. [1]
- From text analysis
- Document representation as a bag of important keywords
- An object is represented as a bag of visual words (patches)
- Allows the recognition of a large collection of objects

Text

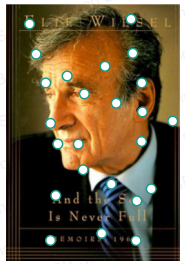
Images



[1] Fei-Fei, Li, and Pietro Perona. "A bayesian hierarchical model for learning natural scene categories." *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*. Vol. 2. IEEE, 2005.

83

Bag-of-words



84

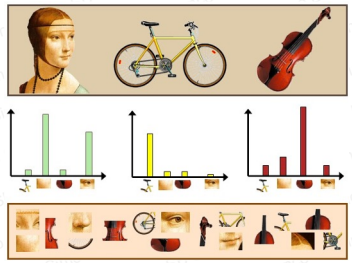
Bag-of-words

a. Extract local features

b. Learn "visual vocabulary"

c. Quantize local features using visual vocabulary

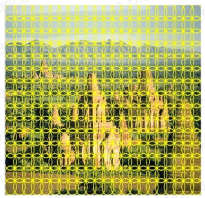
d. Represent images by frequencies of "visual words"



85

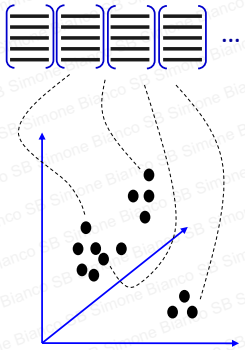
Local features extraction

- Sample patches and extract descriptors



86

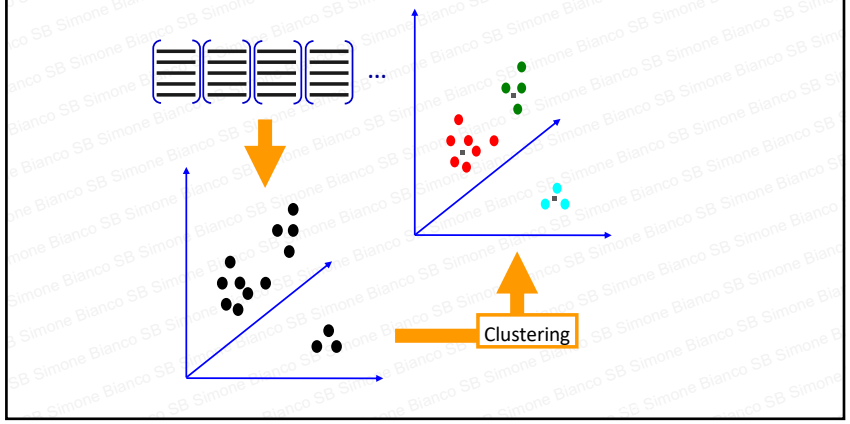
Learning the visual vocabulary (1/2)



Extracted descriptors from the training set

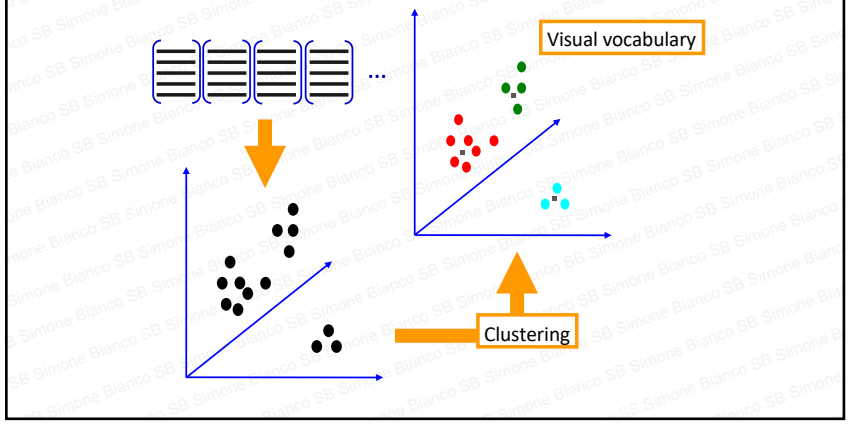
87

Learning the visual vocabulary (2/2)



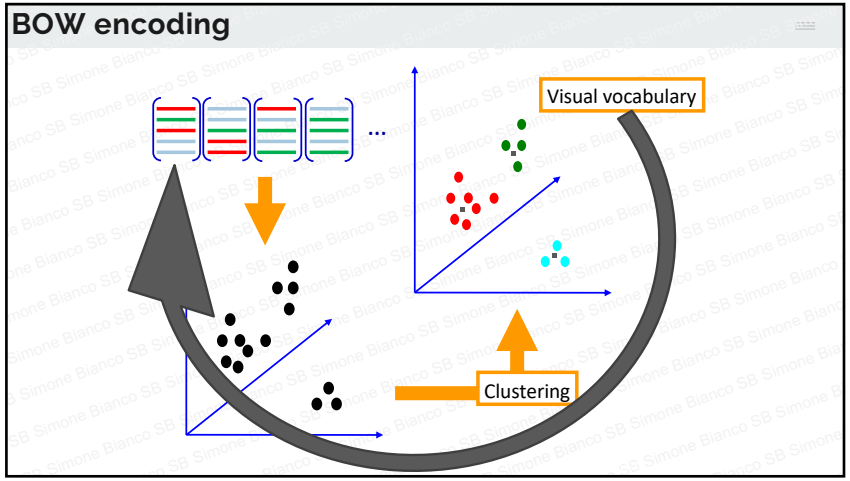
88

Learning the visual vocabulary (2/2)



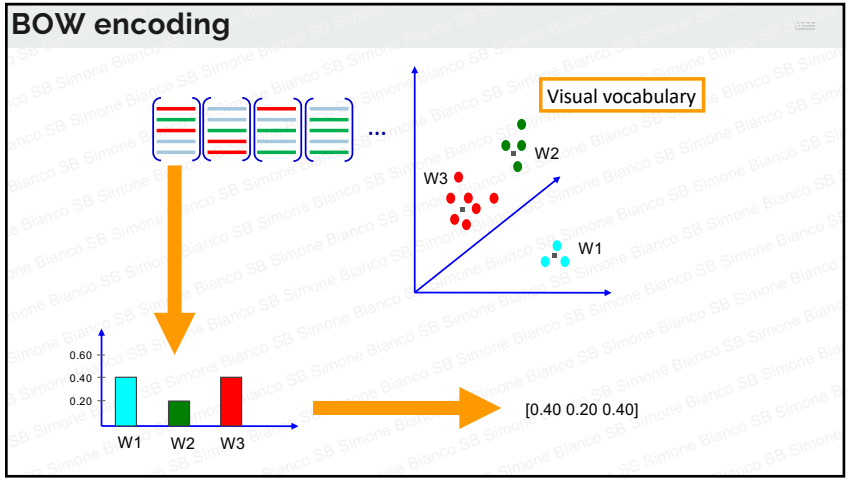
89

BOW encoding



90

BOW encoding

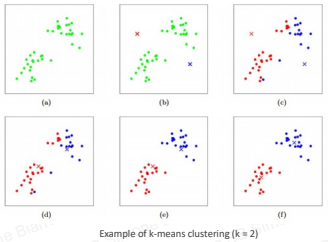


91

Review: K-means clustering

- Want to minimize sum of squared Euclidean distances between features $\mathbf{x}_i$ and their nearest cluster centers $\mathbf{m}_k$

Algorithm: $D(X,M) = \sum_{\text{cluster } k} \sum_{\substack{\text{point } i \text{ in} \\ \text{cluster } k}} (\mathbf{x}_i - \mathbf{m}_k)^2$



Example of k-means clustering (k = 2)

- Randomly initialize K cluster centers
- Iterate until convergence:
  - Assign each feature to the nearest center
  - Recompute each cluster center as the mean of all features assigned to it

92

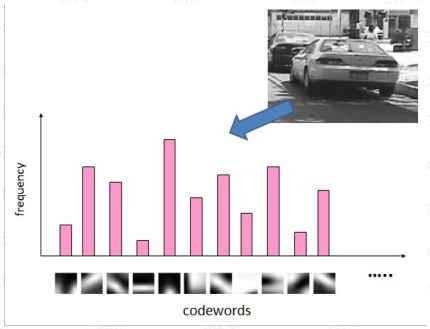
Example visual vocabulary

93

Bag-of-words

Test:

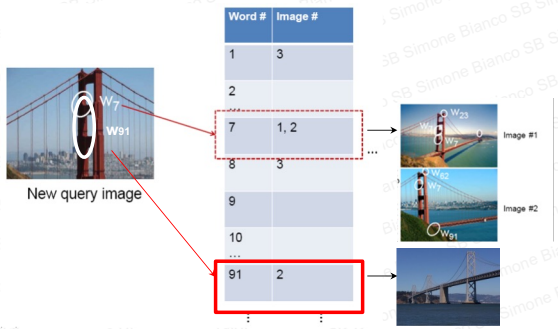
1. Extract descriptors from the image
2. For each descriptor extracted compute its nearest neighbor in the dictionary
3. Build a histogram of length k where the i'th value is the frequency of the i'th dictionary word
4. Classify the histogram with a classifier (e.g. Nearest Neighbor, SVM or Naive Bayes)



94

Applications of Bag of words

Inverted file index and bags of words similarity



95

Bag-of-words

Works with **Deformable Objects**!
Example: tracking

96

Bag-of-words

Works with **Object Classes**!
Example: dogs/parrots classification

97