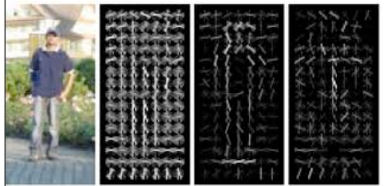
Descriptors for CV

2017



T. Chateau

Content

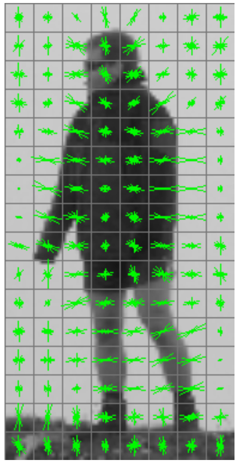
Introduction:

Image descriptors

- Descriptor: a high level representation of an image or video.
- Descriptor: a vector
- Useful for:
  - geometric based matching
  - image retrieval,
  - object recognition and categorization,
  - ....

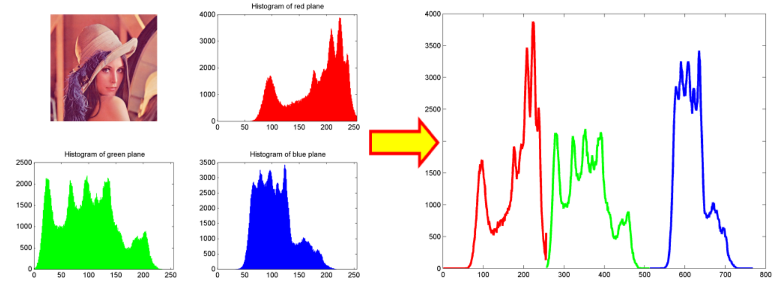
Introduction:

Dense and Sparse descriptors

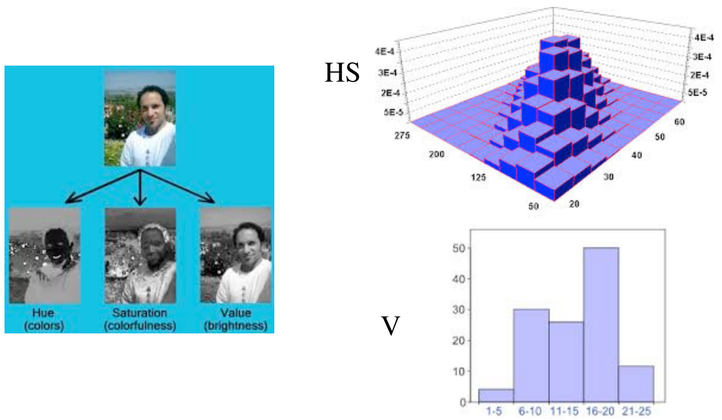


Some dense Descriptors

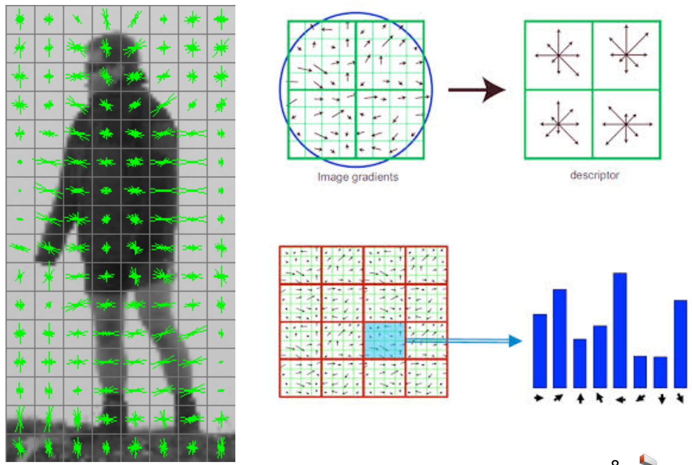
Color RGB histogram



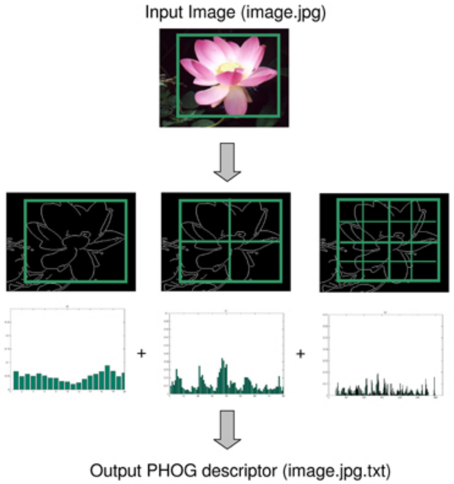
Color HS+V histogram



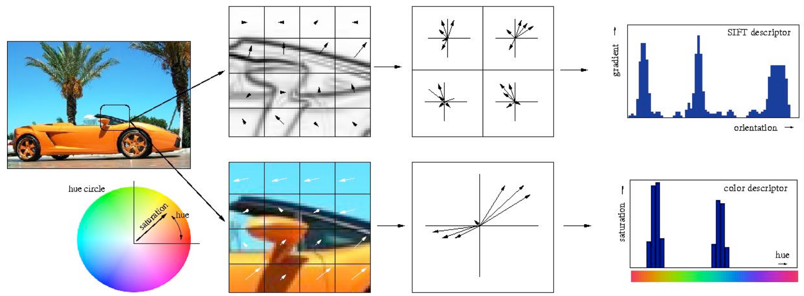
Histogram of oriented gradients (HOG)



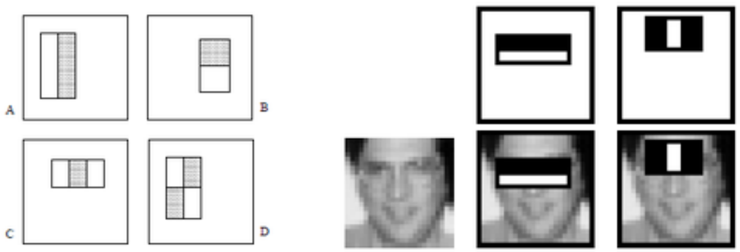
Histogram of oriented gradients (HOG)



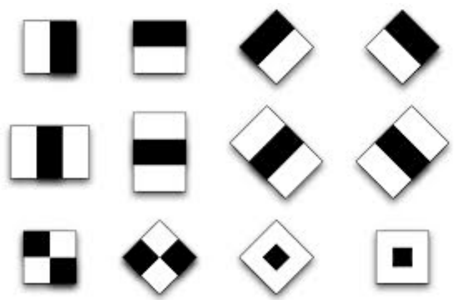
Combining histograms



Haar Wavelets

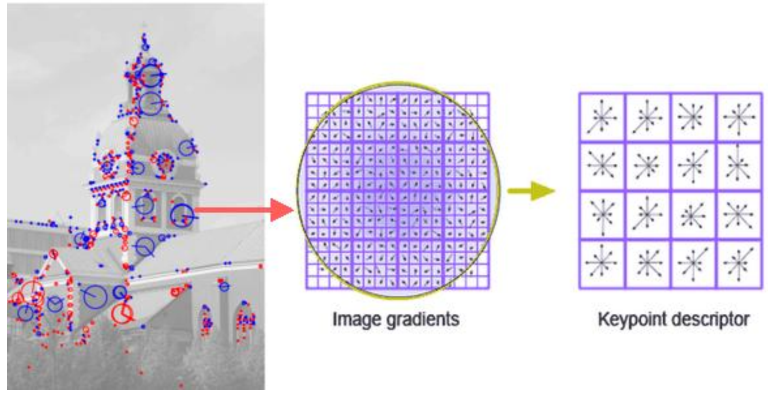


Haar Wavelets



Some Sparse Descriptors

SIFT

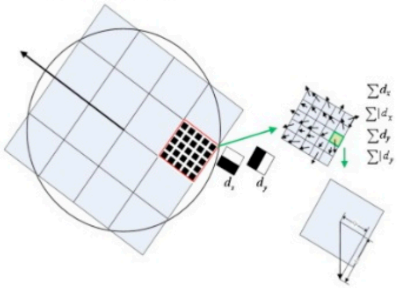


SURF

Speeded Up Robust Features

- Approximate derivatives with Haar wavelets
- Exploit integral images

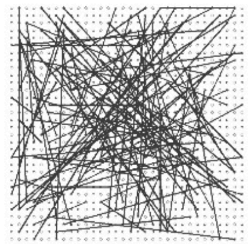
Citations:
4500 (2012)



BRIEF

Binary Robust Independent Elementary Features

- Random selection of pairs of intensity values.
- Fixed sampling pattern of 128, 256 or 512 pairs.
- Hamming distance to compare descriptors (XOR).



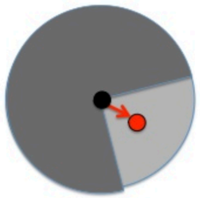
Citations:
149 (2012)

ORB

Oriented FAST and Rotated BRIEF

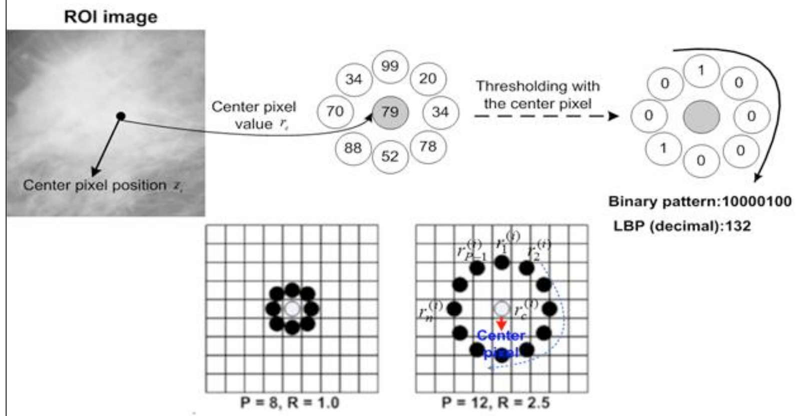
- Add rotation invariance to BRIEF
- Orientation assignment based on the intensity centroid

Citations:
43 (2012)



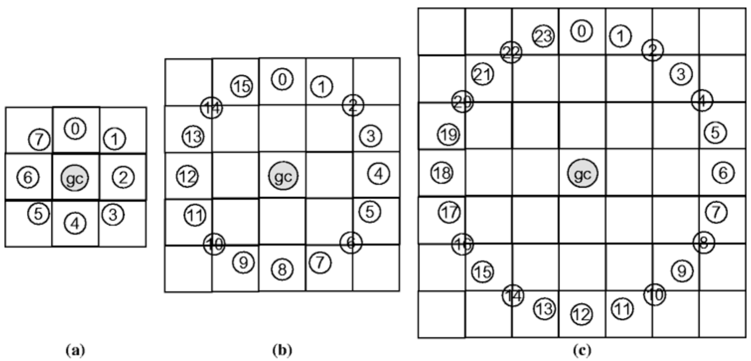
T. Chateau

LBP: local binary pattern



T. Chateau

LBP: local binary pattern



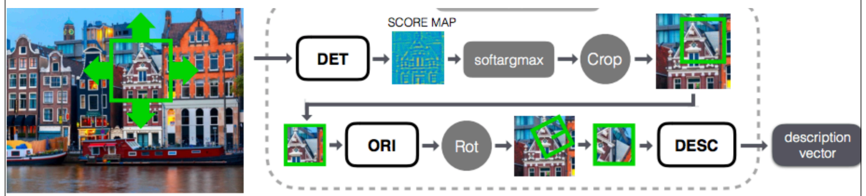
Notes: (a) P = 8, R = 1; (b) P = 16, R = 2; (c) P = 24, R = 3

T. Chateau

LIFT: Learned Invariant Feature Transform

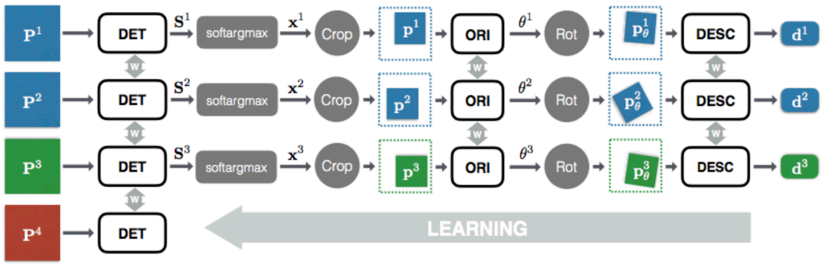
Kwang Moo Yi*[,1], Eduard Trulls*[,1], Vincent Lepetit[2], Pascal Fua[1]

[1]Computer Vision Laboratory, Ecole Polytechnique Fédérale de Lausanne (EPFL)
[2]Institute for Computer Graphics and Vision, Graz University of Technology
{kwang.yi, eduard.trulls, pascal.fua}@epfl.ch, lepetit@icg.tugraz.at

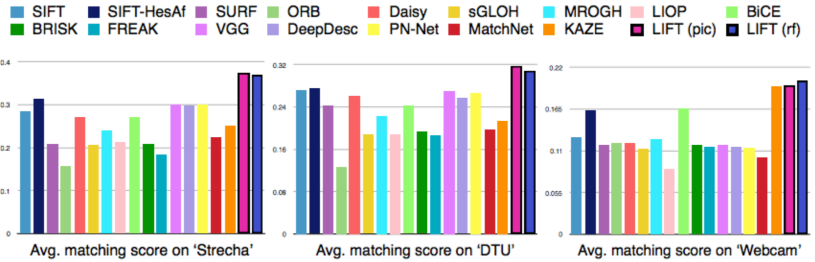


T. Chateau

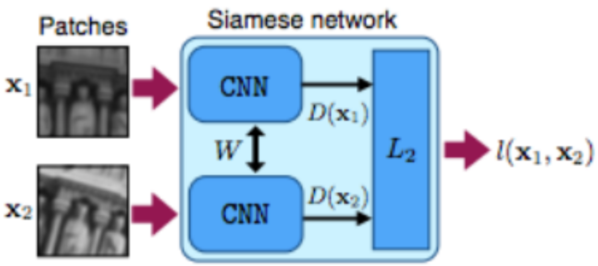
LIFT: Learned Invariant Feature Transform



LIFT: Learned Invariant Feature Transform

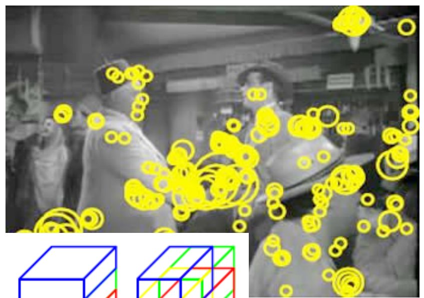


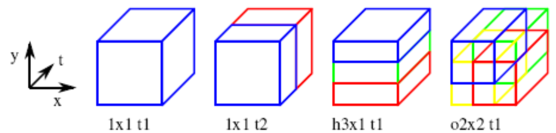
Deep Neural network descriptor for matching



Video based descriptors (HOG/HOF)

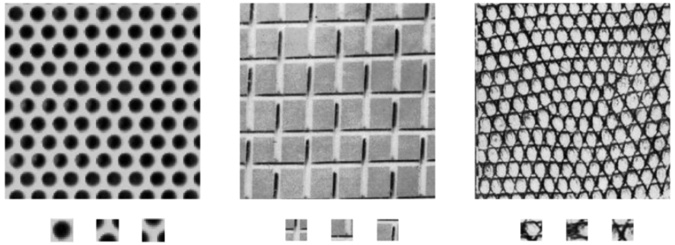
$$\mu = g(.;\sigma_i^2;\tau_i^2) * \begin{pmatrix} L_x^2 & L_xL_y & L_xL_t \\ L_xL_y & L_y^2 & L_yL_t \\ L_xL_t & L_yL_t & L_t^2 \end{pmatrix}$$





Bag of Words (bag of features models)

Origin 1: texture recognition

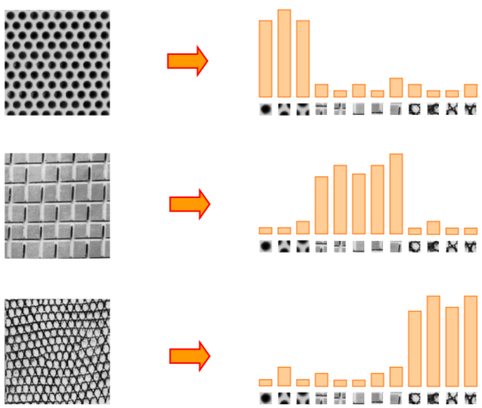


Julesz, 1981; Cula & Dana, 2001; Leung & Malik 2001; Mori, Belongie & Malik, 2001; Schmid 2001; Varma & Zisserman, 2002, 2003; Lazebnik, Schmid & Ponce, 2003

Content

Bag of Words (bag of features models)

Origin 1: texture recognition

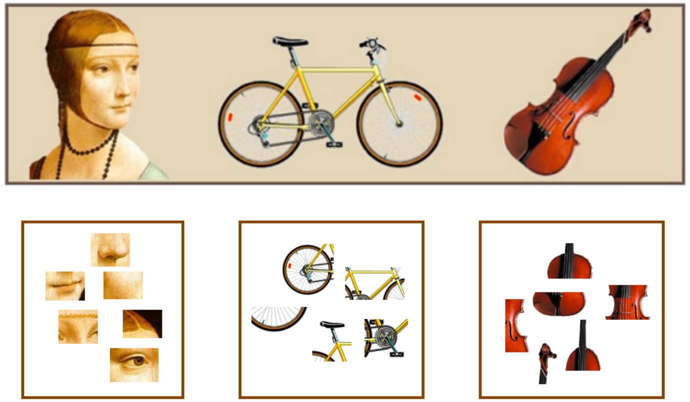


Bag of Words (bag of features models)

Origin 2: text analysis (Frequency of words of a dictionary, Salton & McGill (1983))

Bag of Words

1. extract features

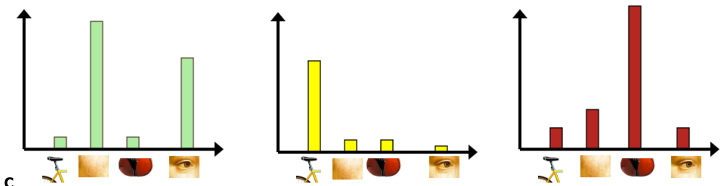


Bag of Words

2. learn visual vocabulary

Bag of Words

3. represent images by frequencies of « visual words »



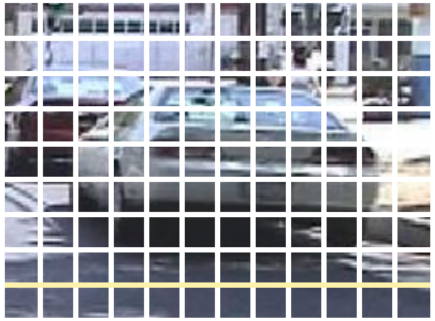
Bag of features for image classification

Given an image to classify:
1. Extract features
2. Quantize features using visual vocabulary
3. Represent images by frequencies of “visual words”
4. Estimate the class using a previously learn classifier (eg. SVM, KNN, AdaBoost,…)

Bag of features for image classification

1. Feature extraction

Regular grid
– Vogel & Schiele, 2003
– Fei-Fei & Perona, 2005

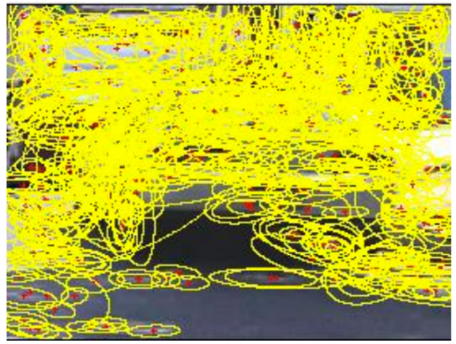


Bag of features for image classification

1. Feature extraction

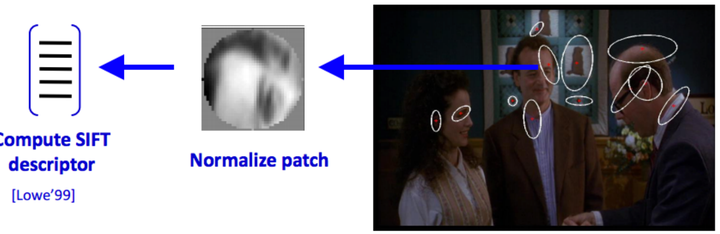
Regular grid
– Vogel & Schiele, 2003
– Fei-Fei & Perona, 2005

Interest point detector
– Csurka et al. 2004
– Fei-Fei & Perona, 2005
– Sivic et al. 2005



Bag of features for image classification

1. Feature extraction

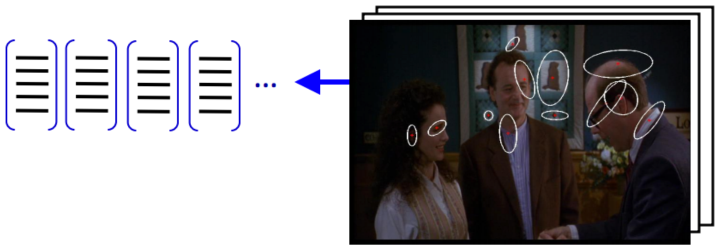


Detect patches
[Mikojaczyk and Schmid '02]
[Mata, Chum, Urban & Pajdla, '02]
[Sivic & Zisserman, '03]

Slide credit: Josef Sivic

Bag of features for image classification

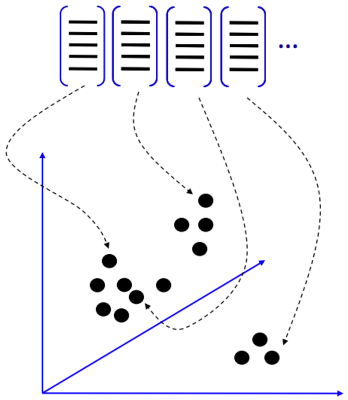
1. Feature extraction



Slide credit: Josef Sivic

Bag of features for image classification

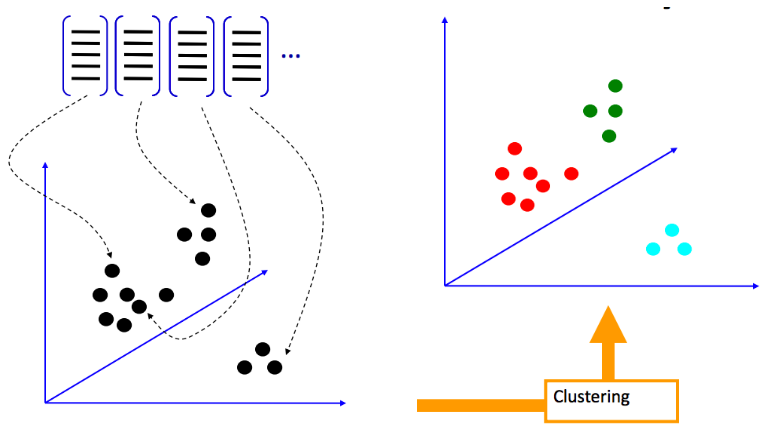
2. Learning the visual vocabulary



Slide credit: Josef Sivic

Bag of features for image classification

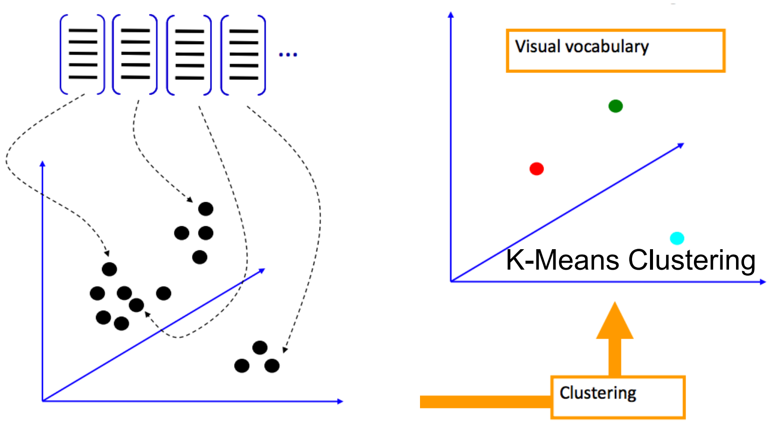
2. Learning the visual vocabulary



Slide credit: Josef Sivic

Bag of features for image classification

2. Learning the visual vocabulary



Slide credit: Josef Sivic

Bag of features for image classification

2. Learning the visual vocabulary

The codebook is used for quantizing features
– A vector quantizer takes a feature vector and maps it to the index of the nearest codevector in a codebook
– Codebook = visual vocabulary
– Codevector = visual word

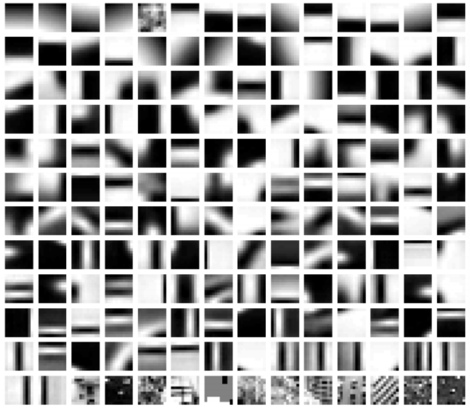
Bag of features for image classification

2. Learning the visual vocabulary

The codebook is used for quantizing features
– A vector quantizer takes a feature vector and maps it to the index of the nearest codevector in a codebook
– Codebook = visual vocabulary
– Codevector = visual word

Bag of features for image classification

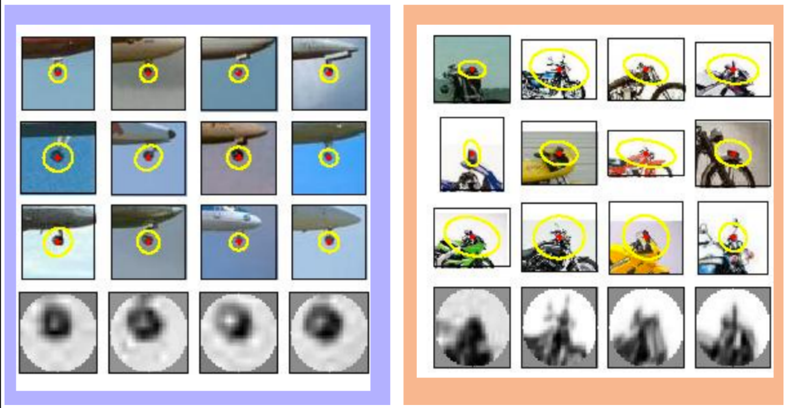
2. Learning the visual vocabulary
exemple of visual vocabulary



Fei-Fei et al. 2005

Bag of features for image classification

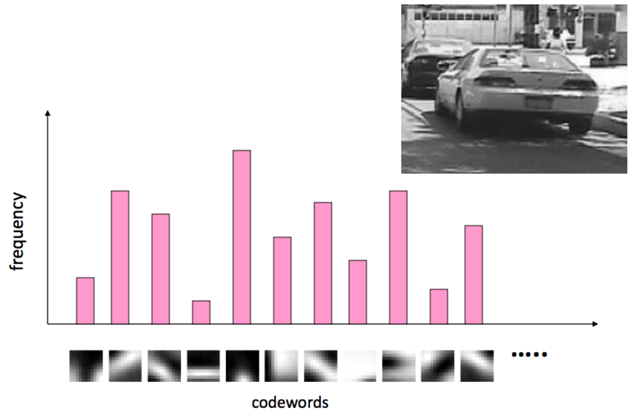
2. Learning the visual vocabulary
exemple of visual words



Sivic et al. 2005

Bag of features for image classification

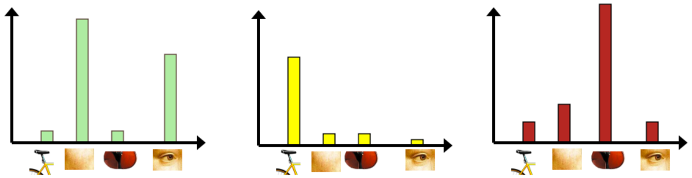
3. Image representation
exemple of codewords



Bag of features for image classification

4. Image classification

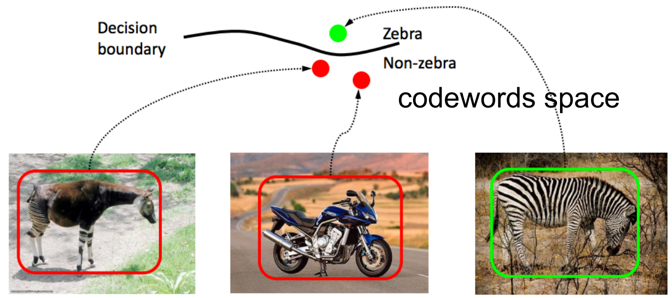
- Given the bag-of-features representations of images from different classes, how do we learn a model for distinguishing them?



Bag of features for image classification

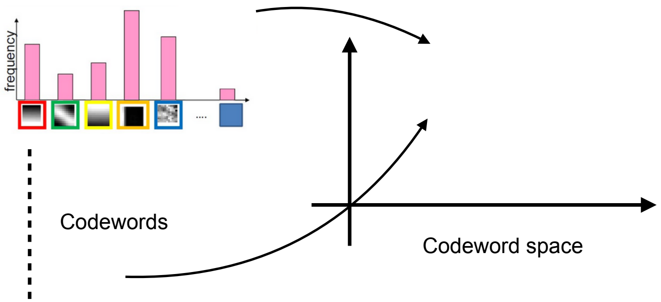
4. Image classification: discriminative method

- Learn a decision rule (classifier) assigning bag-of-features representations of images to different classes



Bag of features for image classification

4. Image classification: discriminative method



We need a to define a distance between codewords

Bag of features for image classification

4. Image classification: discriminative method

Minskowski-form distance:

$$D(X,Y) = \left(\sum_i |f(i;X) - f(i;Y)|^p\right)^{1/p}$$

CHI2 distance:

$$D(X,Y) = \sum_i \frac{(f(i;X) - \hat{f}(i))^2}{\hat{f}(i)},$$

$$\hat{f}(i) = [f(i;X) + f(i;Y)]/2$$

We need a to define a distance between codewords

Bag of features for image classification

4. Image classification: discriminative method

Kullback-Leiber divergence

$$D(X,Y)=\sum_i f(i;X)\log\frac{f(i;X)}{f(i;Y)},$$

Battacharyya distance

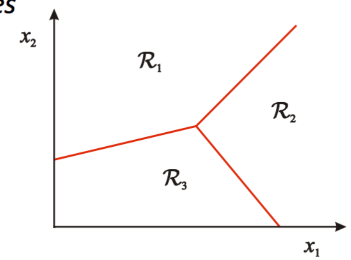
$$BC(p,q)=\sum_{x\in X}\sqrt{p(x)q(x)}$$

We need a to define a distance between codewords

Bag of features for image classification

4. Image classification: discriminative method

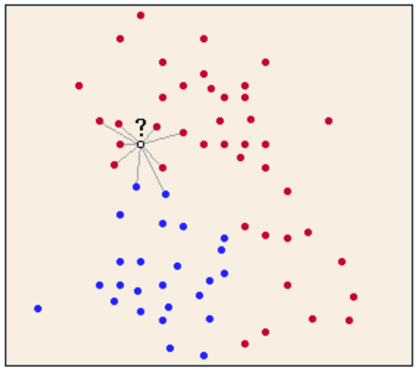
- Assign input vector to one of two or more classes
- Any decision rule divides input space into *decision regions* separated by *decision boundaries*



Bag of features for image classification

4. Image classification: discriminative method

KNN: K- Nearest Neighbors



Bag of features for image classification

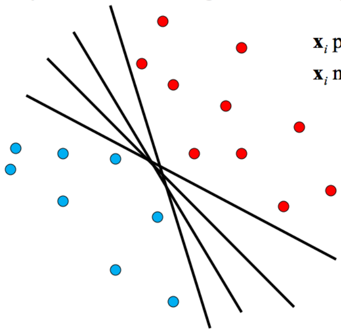
4. Image classification: discriminative method

linear classifiers

- Find linear function (*hyperplane*) to separate positive and negative examples

$\mathbf{x}_i$ positive : $\quad \mathbf{x}_i \cdot \mathbf{w} + b \geq 0$

$\mathbf{x}_i$ negative : $\quad \mathbf{x}_i \cdot \mathbf{w} + b < 0$



Which hyperplane is best?

Slide: S. Lazebnik

Bag of features for image classification

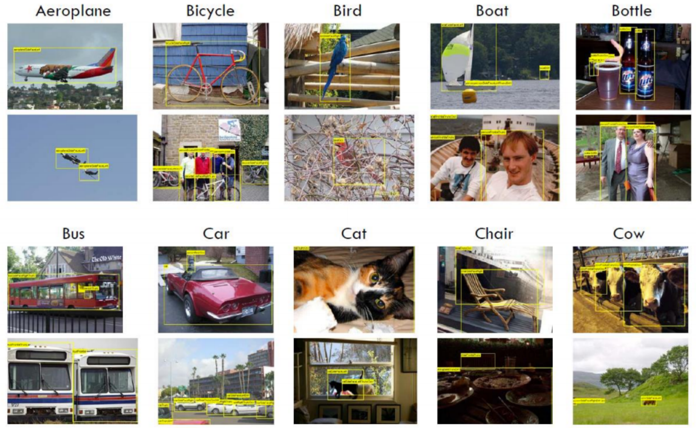
4. Image classification: discriminative method

and much more …

SVM, …

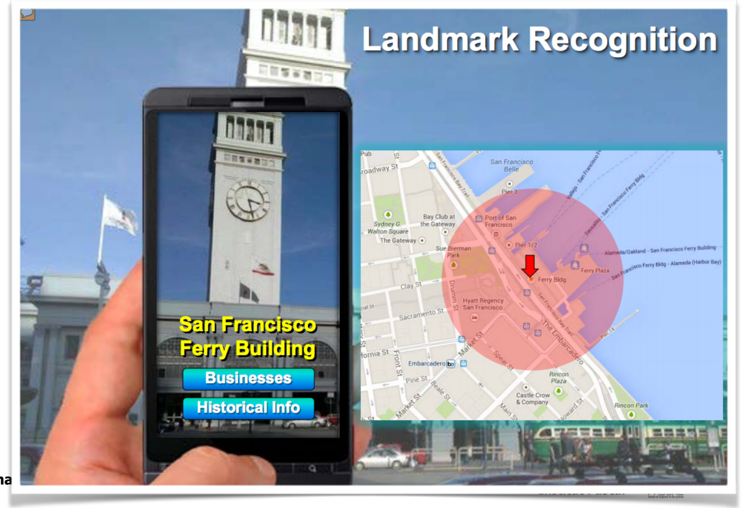
Bag of features for image classification

Results from Pascal VOC Challenge 2010



Bag of features for image classification

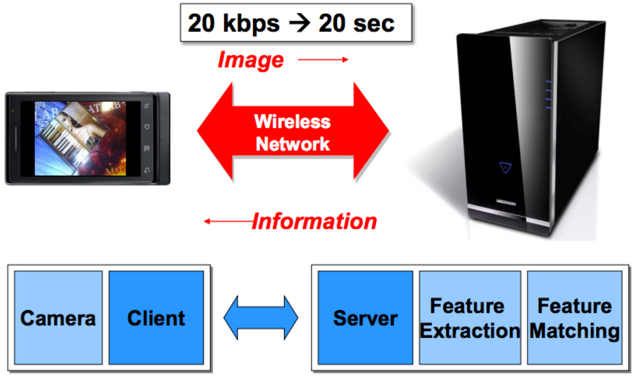
AR Applications



Bag of features for image classification

AR Applications

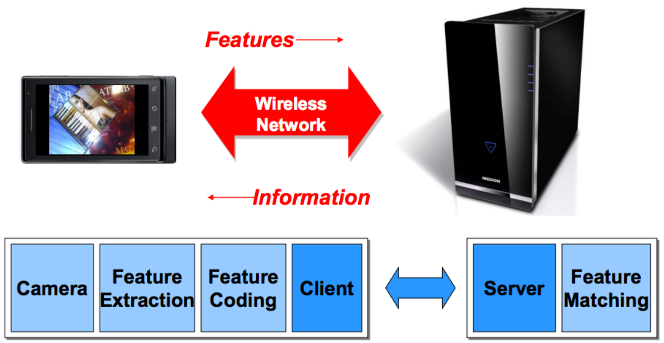
Architecture A: Send Image



Bag of features for image classification

AR Applications

Architecture B: Send Features



Bag of features for image classification

AR Applications

Architecture C: Features on Mobile Device