Discovering objects and their location in images

Sivic, Russell, Efros, Zisserman and Freeman

Ulf Martin Blom

UT Austin

April 9, 2009

Outline

Main points in the paper:

Outline

Main points in the paper:

- Seeks to find the object categories in *unlabelled* images

Outline

Main points in the paper:

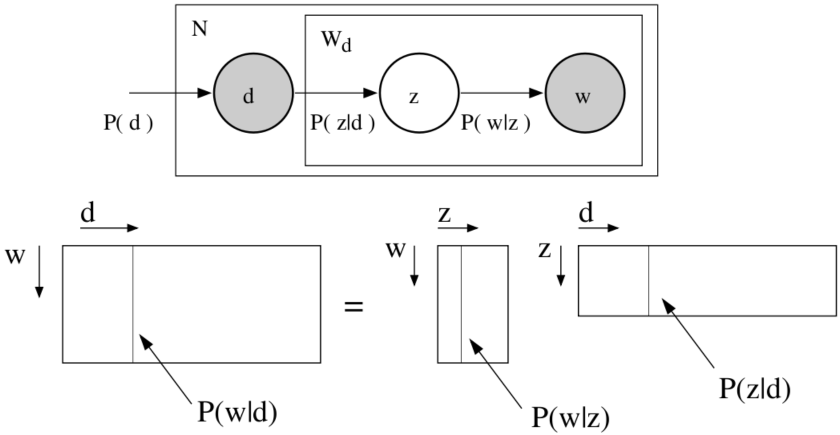
- Seeks to find the object categories in *unlabelled* images
- Inspired and analogous to pLSA in text

Outline

Main points in the paper:

- Seeks to find the object categories in *unlabelled* images
- Inspired and analogous to pLSA in text
- Includes doublets for clearer representation

pLSA



Dictionary

Text	**Vision**
Document	Image
Word	Vector quantized SIFT
Topic	Category

Words words words

Take a subset of the data.

Words words words

Take a subset of the data.

- Find affine covariant regions.

Words words words

Take a subset of the data.

- Find affine covariant regions.
- Calculate SIFT descriptors on rescaled ellipses.

Words words words

Take a subset of the data.

- Find affine covariant regions.
- Calculate SIFT descriptors on rescaled ellipses.
- Cluster by k-means.

Nonnegative Matrix Factorization

- Early, similar, work by Lee and Seung (Nature 1999).

Nonnegative Matrix Factorization

- Early, similar, work by Lee and Seung (Nature 1999).
- More CS-y, matrix factorization interpretation.

Nonnegative Matrix Factorization

- Early, similar, work by Lee and Seung (Nature 1999).
- More CS-y, matrix factorization interpretation.
- Minimizes distance between matrix and its (low dimensional) factorization.

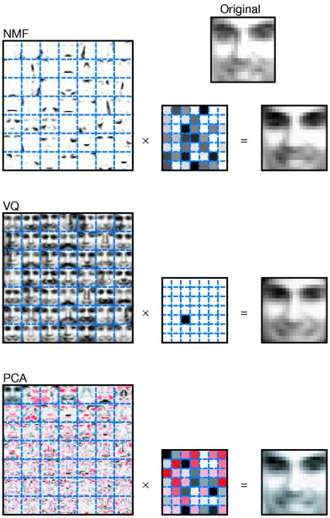


Figure: Nonnegative matrix factorization compared to vector quantization and PCA (Nature 401, 788-791. 21 October 1999)

Experiments

Sivic et al. investigate three areas:

Experiments

Sivic et al. investigate three areas:

- Topic Discovery

Experiments

Sivic et al. investigate three areas:

- Topic Discovery
- Classification

Experiments

Sivic et al. investigate three areas:

- Topic Discovery
- Classification
- Object Detection

Topic Discovery

Infer the distributions via maximum likelihood.

Expectation-Maximization of

$$L = \prod_{i=1}^{M} \prod_{j=1}^{N} P(w_i | d_j)^{n(w_i, d_j)},$$

w.r.t. the factorization

$$P(w_i | d_j) = \sum_{k=1}^{K} P(z_k | d_j) P(w_i | z_k).$$

Segmentation

We can also calculate the probability of topics given words, by using Bayes rule.

$$P(z_k|w_i, d_j) = \frac{P(w_i|z_k)P(z_k|d_j)}{\sum_{l=1}^{K} P(z_l|d_j)P(w_i|z_l)}$$

So, for a given word in a particular document, we can calculate the probability that it comes from a given topic.

Code

Code

- Oldish code, causes some matlab problems.

Code

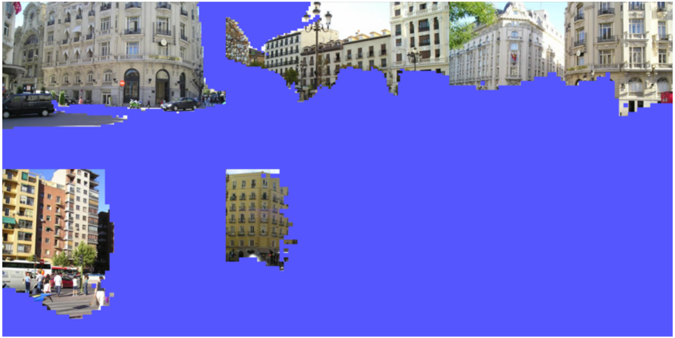
- Oldish code, causes some matlab problems.
- Eats all you memory.

Code

- Oldish code, causes some matlab problems.
- Eats all you memory.
- Exists in 64-bit version for large datasets.

Sample Topics, LabelMe $k = 5$

2



LabelMe $k = 5$

4

SALAD BAR
1200
JORDAN ST
BREAKFAST & LUNCH
ONE WAY
383-7705
Keno
CINES
ONE WAY
ONE WAY

LabelMe $k = 5$

3

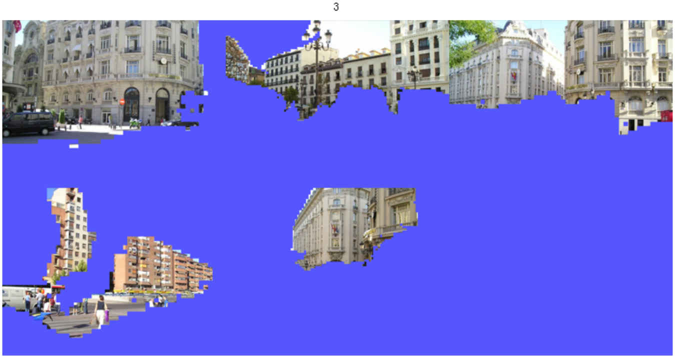
Different Categories

What happens when we vary the number of categories?

LabelMe, $k = 10$

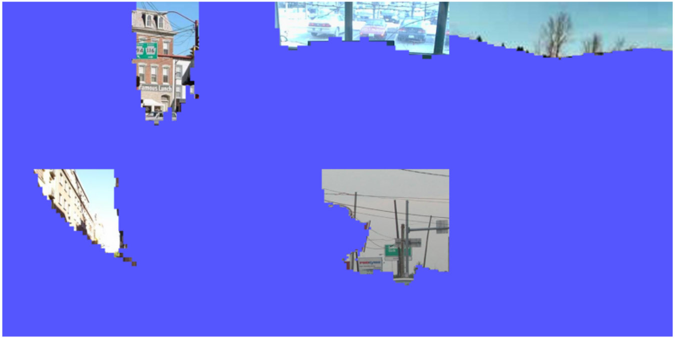
2

LabelMe, $k = 10$



LabelMe, $k = 10$

4



LabelMe, $k = 10$

5

LabelMe, $k = 10$

6

LabelMe, $k = 10$

7

LabelMe, $k = 10$

8

Slightly different data!

What happens when we learn on a slightly different dataset?

LabelMe, $k = 10$

3

LabelMe, $k = 10$

4

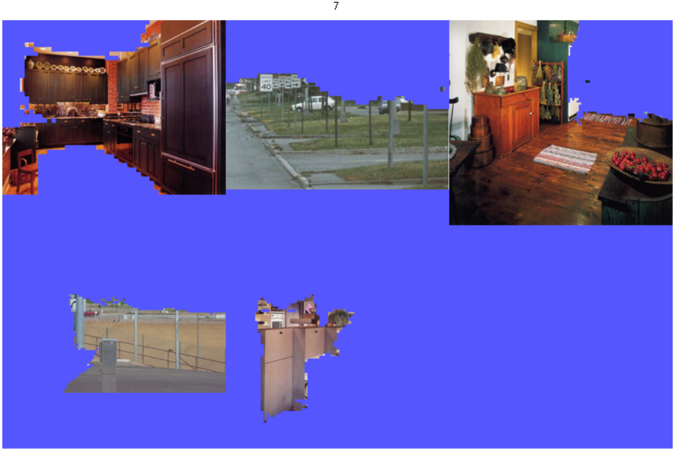
LabelMe, $k = 10$

5

LabelMe, $k = 10$

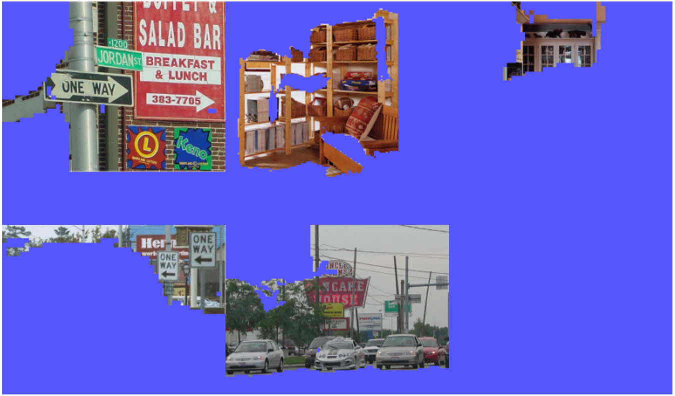
6

LabelMe, $k = 10$

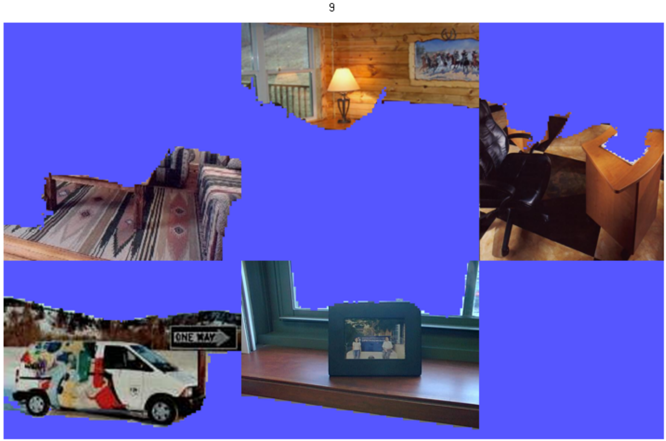


LabelMe, $k = 10$

8



LabelMe, $k = 10$



LabelMe, $k = 10$

Discussion

- When should we use many categories, when should we use few?
- What kind of advantages compared to supervised learning?
- Do people use doublets in text?