



DBTA Workshop on Information Retrieval:
Algorithms and Systems for Text and Multimedia Retrieval

PLSA-BASED APPROACHES TO IMAGE ANNOTATION

Dr. Florent Monay, Idiap Research Institute

Dr. Daniel Gatica-Perez, Idiap Research Institute

The problem

- In general, image captions are:

- Incomplete (subjectivity)

How to **expand** existing annotations?

- Inexistent

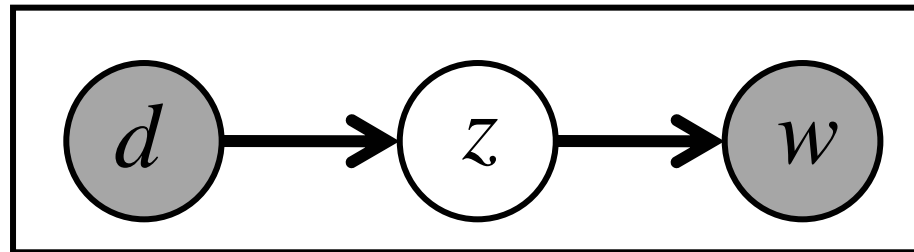
How to **propose** an annotation given an image?

Outline

- Probabilistic Latent Semantic Analysis
- Annotation expansion with PLSA
- Modeling images and their caption with PLSA
- Retrieval demo (Corel)
- Conclusion

Probabilistic latent semantic analysis

- Observations: document and word pairs (d_i, w_j) where $d_i \in \{d_1, \dots, d_N\}$ and $w_j \in \{w_1, \dots, w_M\}$
- A document is a mixture of hidden aspects z_k that generated each word w_j ($z_k \in \{z_1, \dots, z_K\}$)



Thomas Hofmann, *Unsupervised Learning by Probabilistic Latent Semantic Analysis*, Machine Learning Journal, 42(1), 2001.

Posterior word probability

$$P(w_j | d_i) = \sum_{k=1}^K P(w_j | z_k) P(z_k | d_i)$$

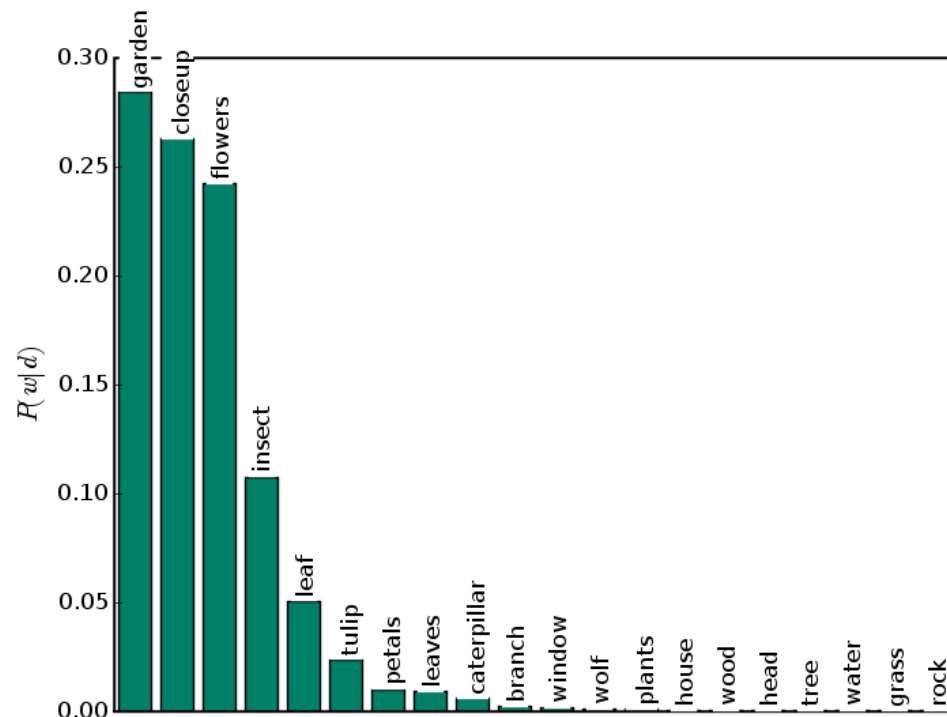
- Parameters:
 - ▣ A document is defined by $P(z_k | d_i)$
 - ▣ An aspect is a mixture of words $P(w_j | z_k)$
- Unsupervised parameter estimation with
Expectation Maximization

Outline

- Probabilistic Latent Semantic Analysis
- Annotation expansion with PLSA
- Modeling images and their caption with PLSA
- Retrieval demo
- Conclusion

Annotation expansion

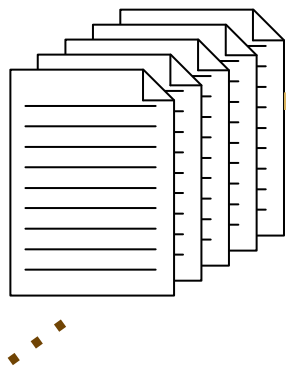
- Goal: estimate the probability of all words in the vocabulary $P(w_j | d_{new})$ from a given ann



closeup
flowers
garden

Annotation expansion with PLSA

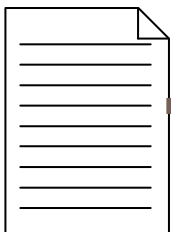
Image caption
Training set



$$P(z|d)$$

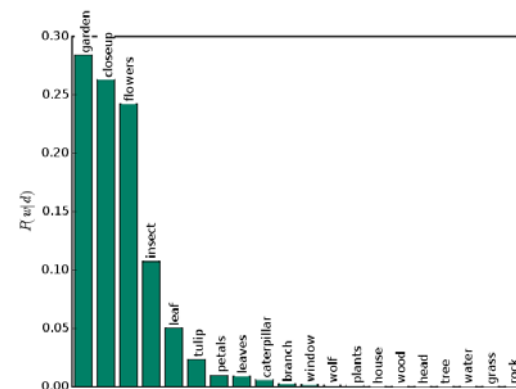
$$P(w|z)$$

New caption

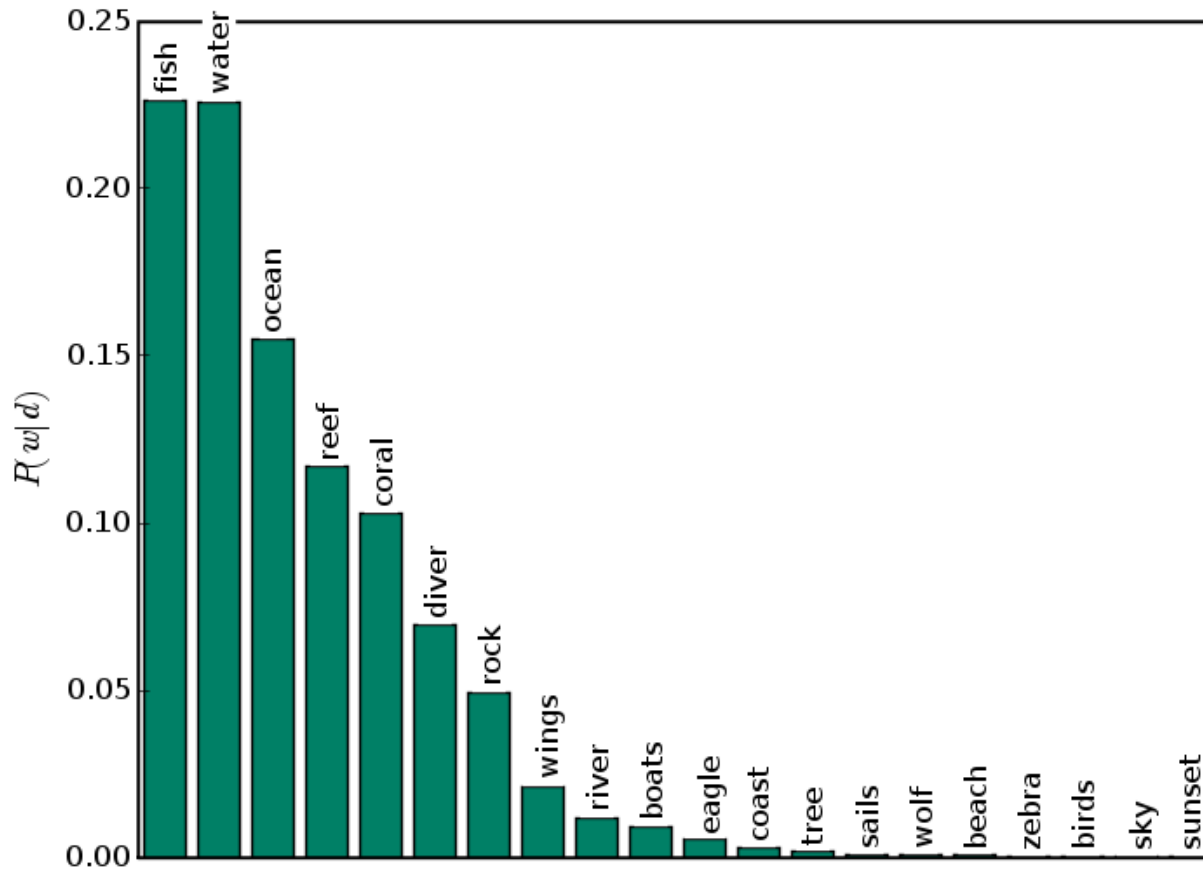


$$P(z|d_{\text{new}})$$

$$P(w|d_{\text{new}})$$

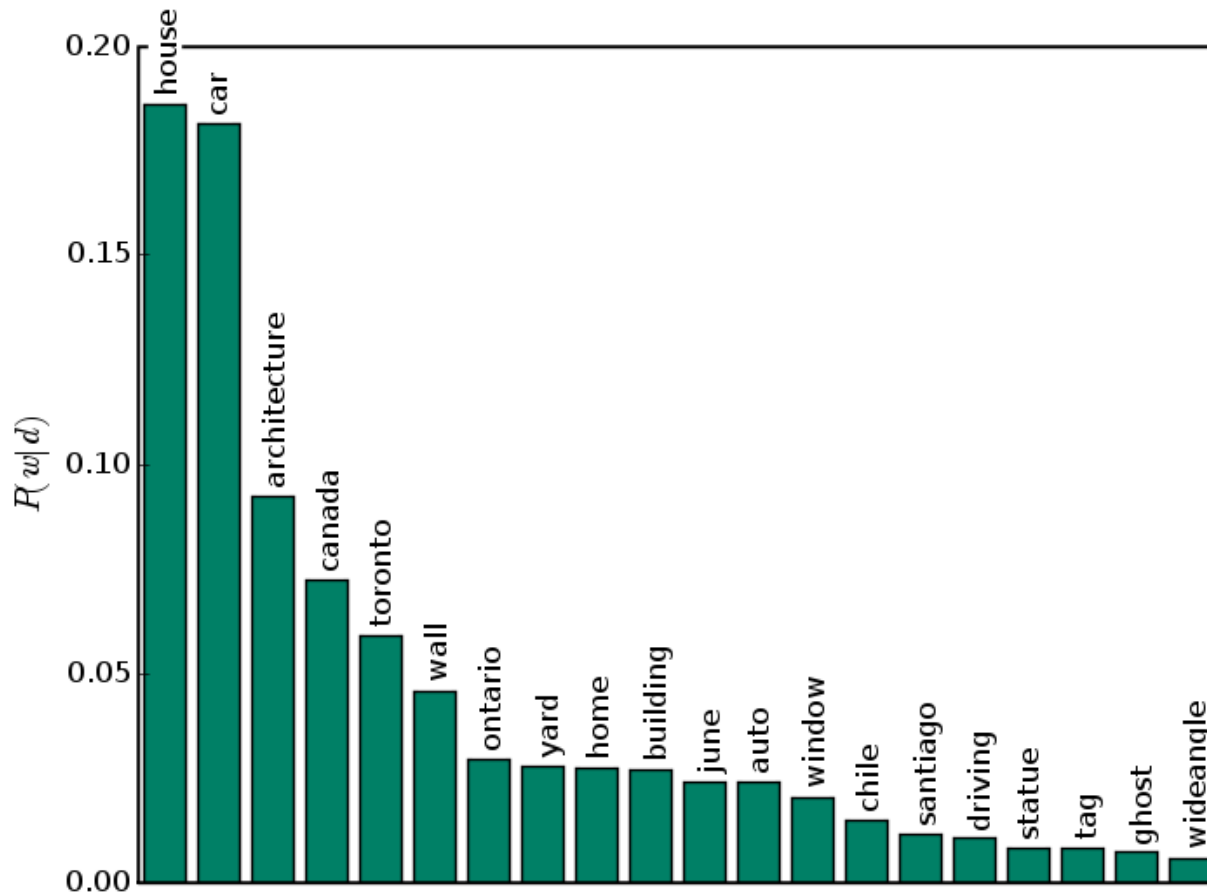


Example: Corel data



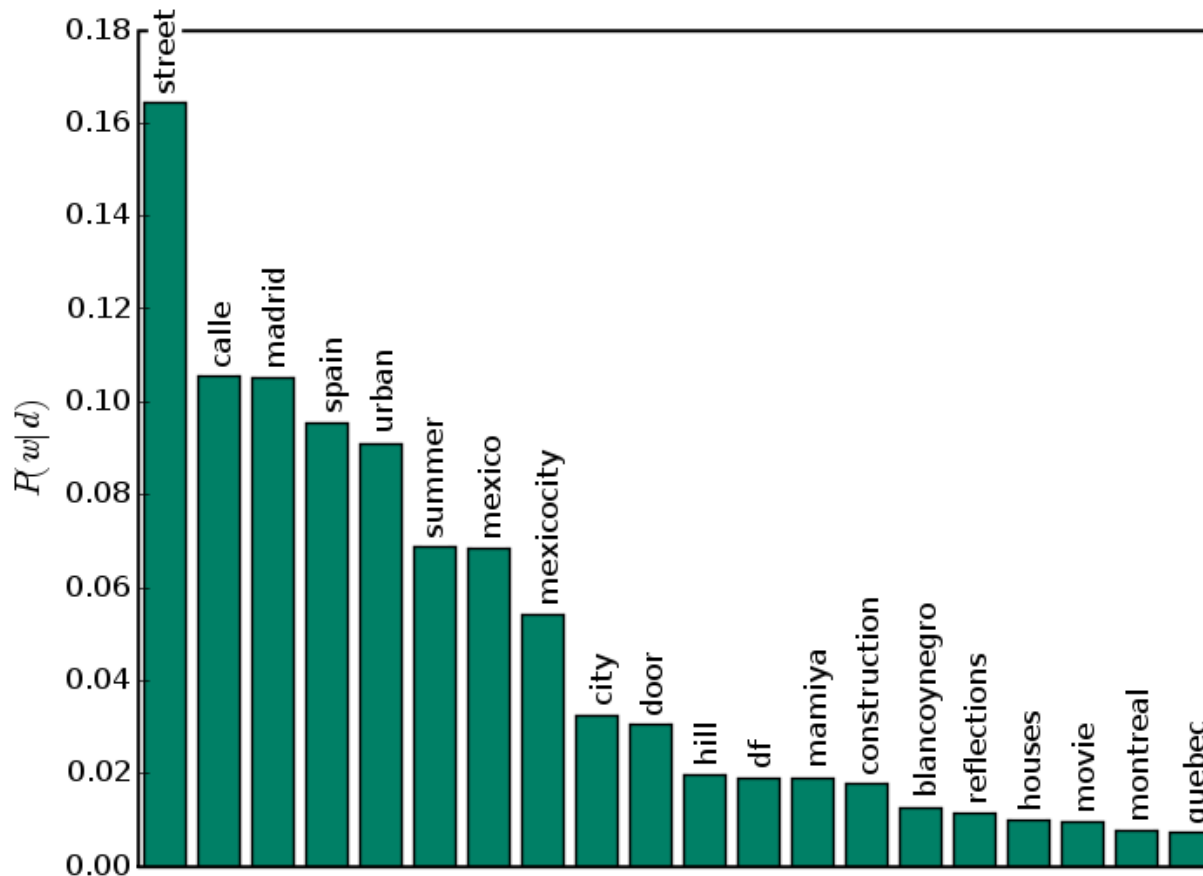
fish
reef
water

Example: Flickr data



office
toronto
architecture
house

Example: Flickr data



door
calle
mexico
urban
street
mexicocity

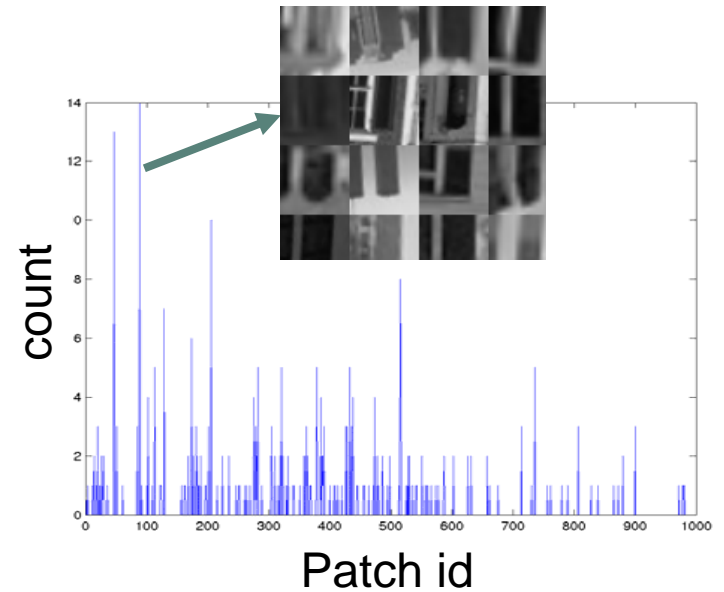
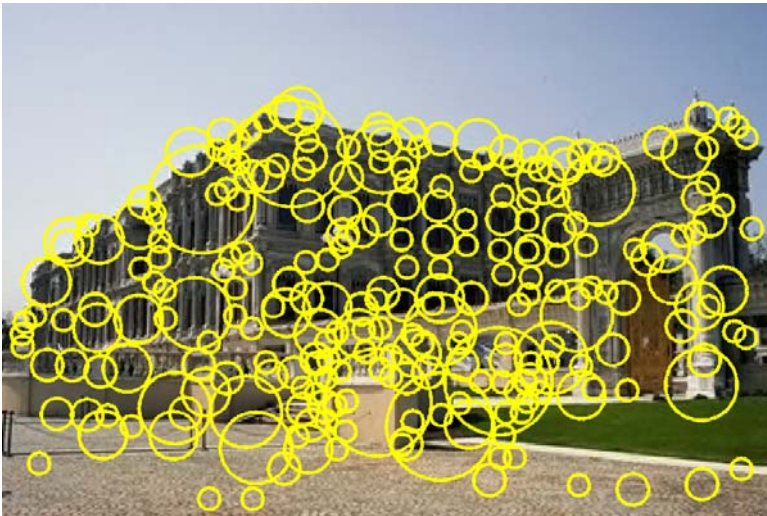
Outline

- Probabilistic Latent Semantic Analysis
- Annotation expansion with PLSA
- Modeling images and their caption with PLSA
- Retrieval demo
- Conclusion

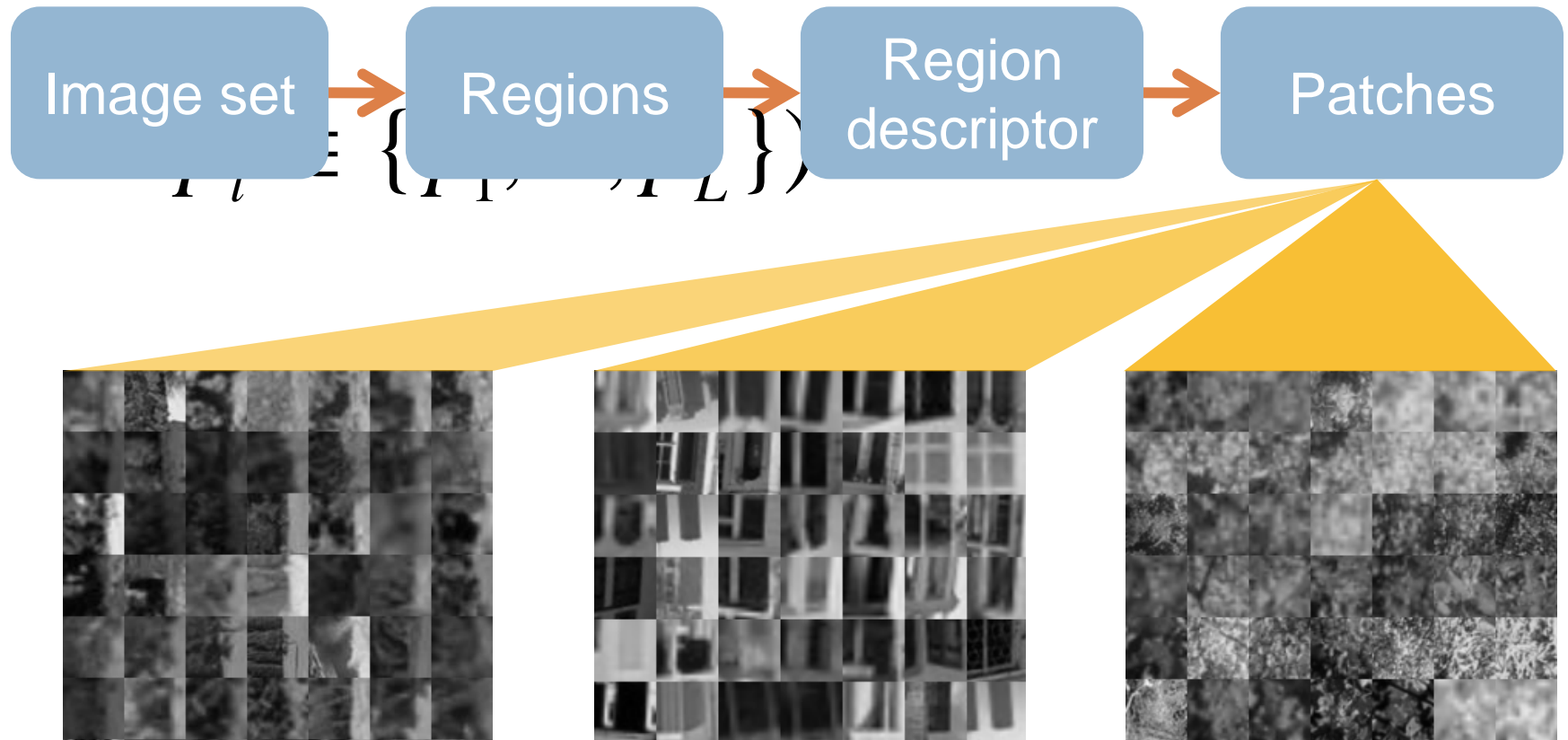
Image representation

- Caption: set of words, Image: set of patches

$$p_l \in \{p_1, \dots, p_L\}$$



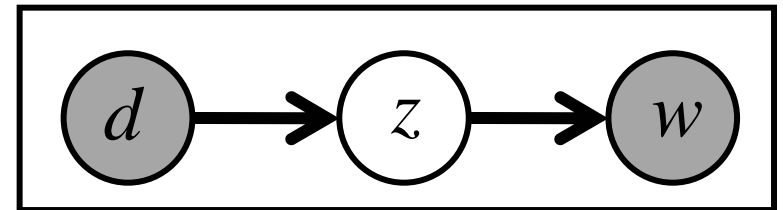
Construction of a patch vocabulary



Modeling images and their caption

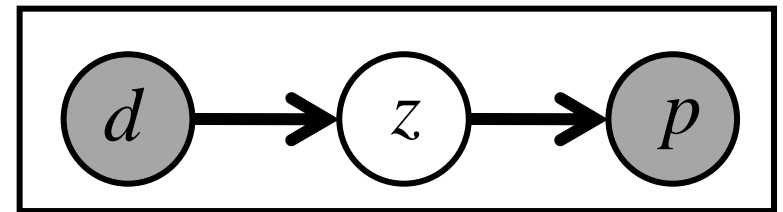
- Observations: image-patch and image-word

pairs (a_i, p_l) and (d_i, w_j)



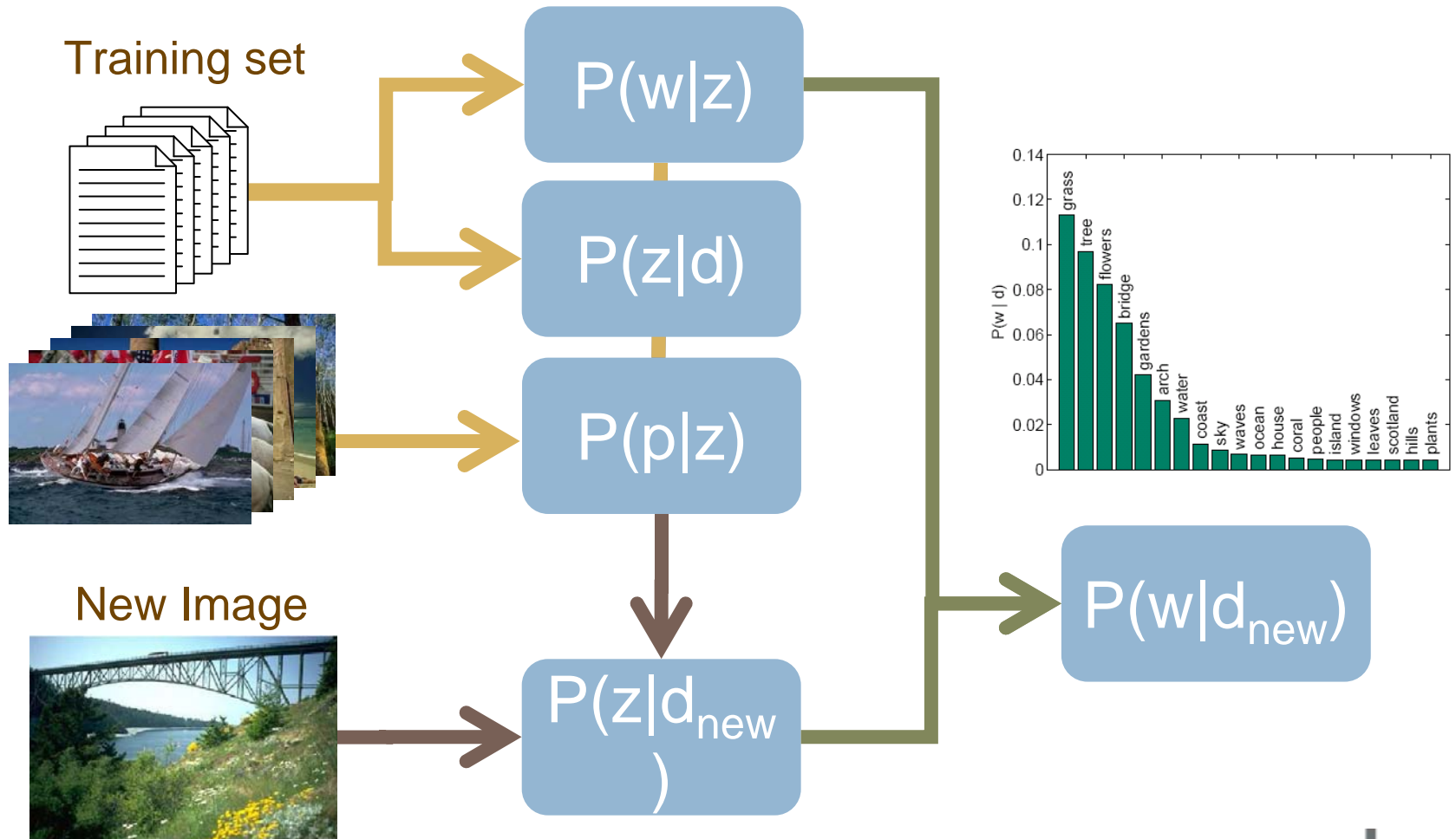
- New parameters:

$P(p_l | z_i)$



- An image is defined by a single aspect distribution $P(z_k | a_i)$

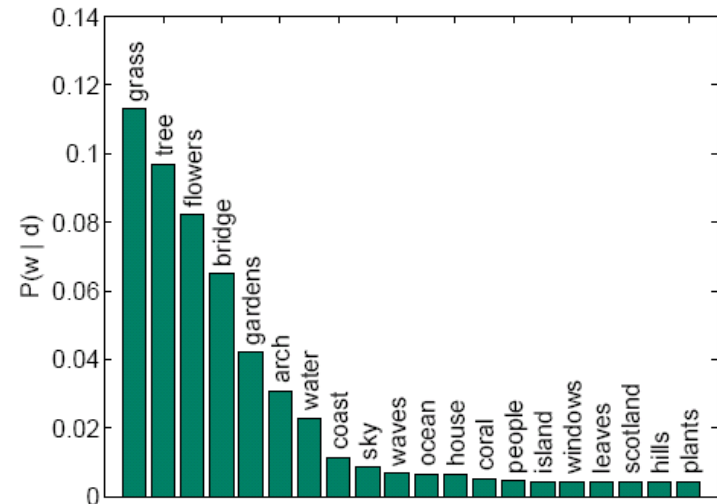
PLSA for image annotation



Outline

- Probabilistic Latent Semantic Analysis
- Annotation expansion with PLSA
- Modeling images and their caption with PLSA
- Retrieval demo
- Conclusion

Image annotation examples



- Image retrieval based on this probabilistic annotation? [see demo](#)

Conclusion

- Illustration of **annotation expansion** with PLSA
- Annotation of images from their bag of patches representation
- To do:
 - ▣ Measure the annotation expansion performance
 - ▣ Experiment the system for real case scenarios (Flickr dataset)
- Questions?

Related publication

- Kobus Barnard, Pinar Duygulu, Nando de Freitas, David Forsyth, David Blei, and Michael I. Jordan, *Matching Words and Pictures*, Journal of Machine Learning Research, Vol 3, 2003.
- Florent Monay and Daniel Gatica-Perez, *Modeling semantic aspects for cross-media image indexing*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 29, 2007.