Dirichlet Process Mixture Models for Object Category Recognition

- Diploma Thesis -

TECHNISCHE UNIVERSITÄT DARMSTADT

Multimodale interaktive Systeme

Sandra Ebert

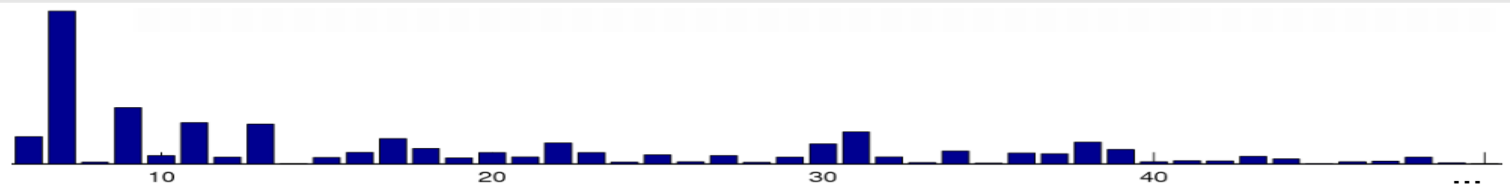
Advisors: Michael Stark, Mario Fritz and Bernt Schiele

Introduction

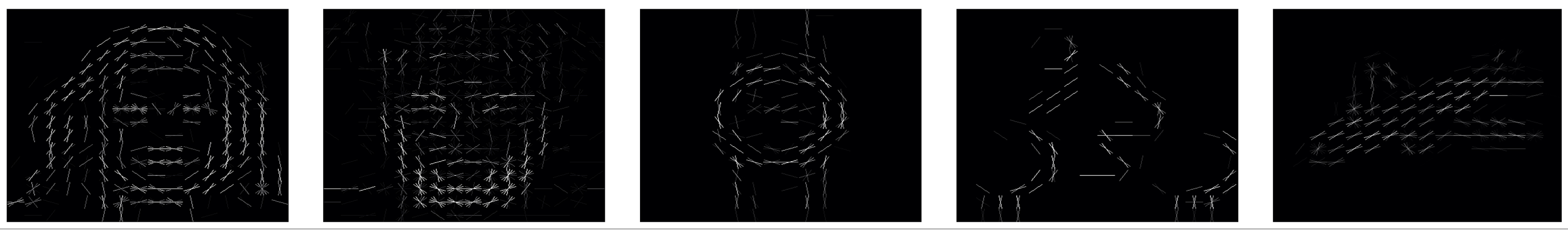
- Exploration of Dirichlet Processes in different models to learn more about the trade-off between computational complexity and quality of the model selection
- Unclear how to choose the hyperparameter and how the complexity evolves when multiple Dirichlet Processes (DP) are combined
- Comparison of different Dirichlet Process Mixture Models on different data sets
- Propose several heuristics to improve the performance and the quality

Topic representation

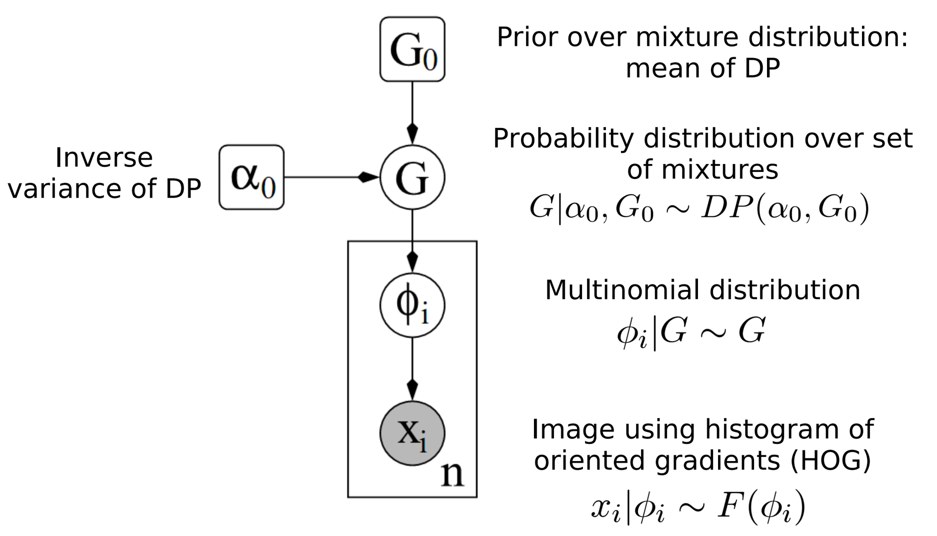
- Feature reduction with topics: each gradient is assigned to one topic
- Topics are typical gradient distributions, e.g. female or male face



Histogram of **O**riented **G**radients (HOG)

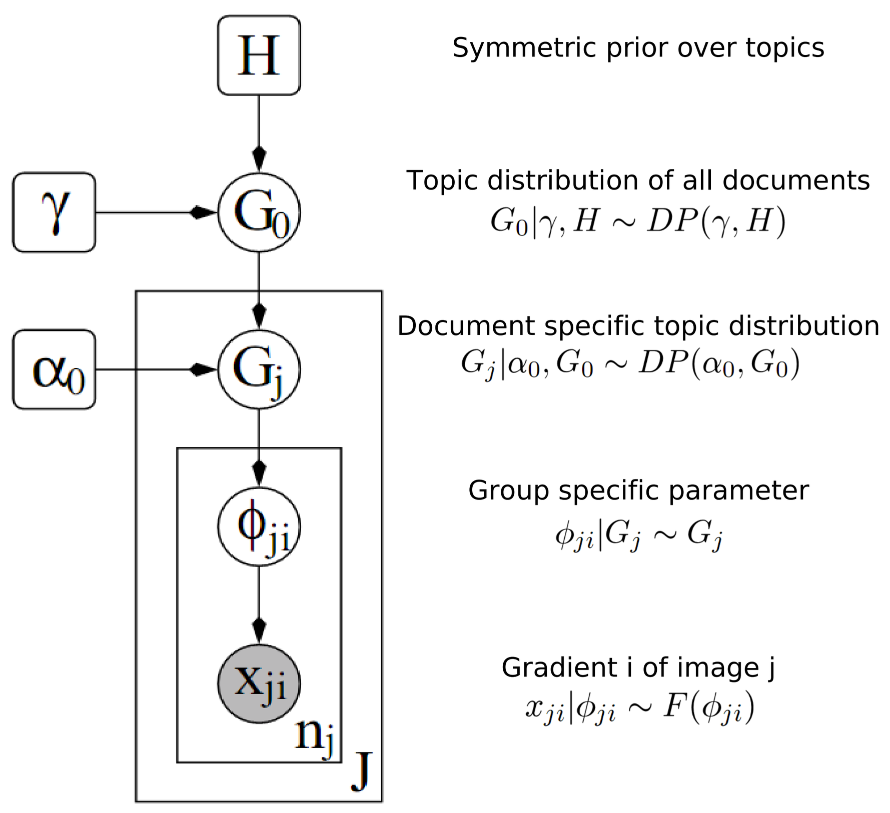


Dirichlet Process Mixture Model



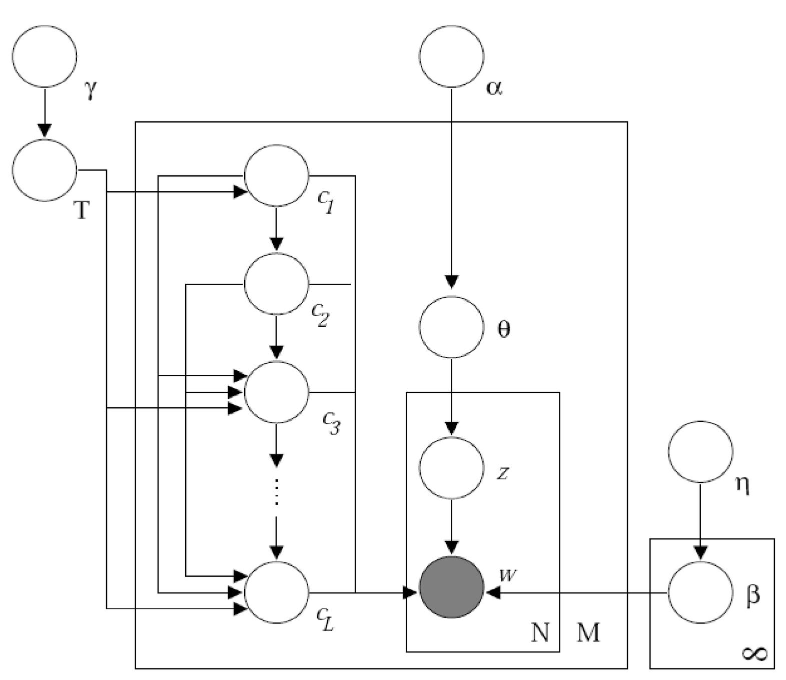
- Dirichlet Process to determine the number of mixtures (clusters)
- Images using histogram of oriented gradients (HOG)
- 16x16x9 Features as multinomial distribution
- Log-likelihood to describe the quality of mixture assignment
- Comparison with k-means clustering

Hierarchical Dirichlet Process Model



- Topic model: counterpart of Latent Dirichlet Allocation [Blei03] to reduce feature dimensions
- Multiple Dirichlet Processes as hierarchy
- Global DP for topic distribution over all documents
- Local DP for each document for document specific topic distribution
- k-means clustering on new feature representation
- Comparison with Latent Dirichlet Allocation and k-means Clustering [Fritz08]

Nested Chinese Restaurant Process



- Topic model: to learn topics and hierarchy of topics (tree structure)
- Root topic is most general
- Leaf topic is most specific
- Each document is assigned to one path
- Dirichlet Process for number of paths

Heuristics:

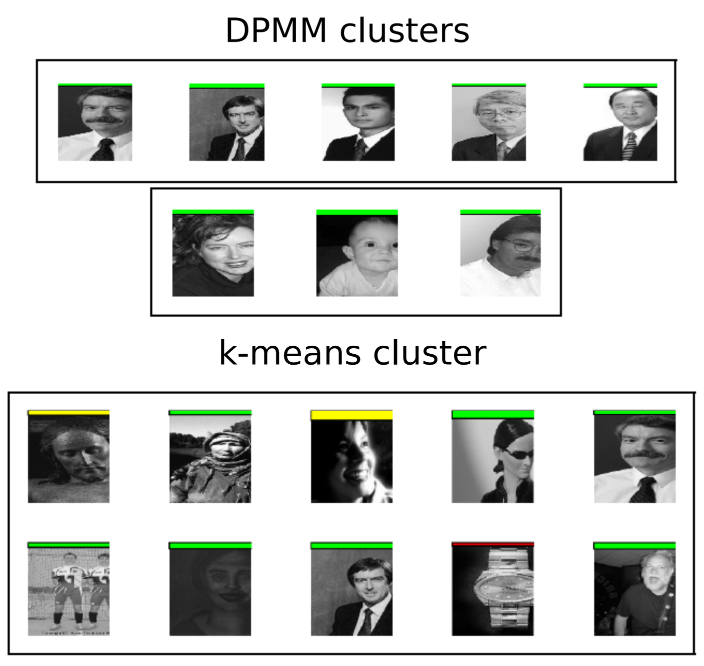
1) Good starting point with LDA and topic assignment
2) Aggregation of similar paths independent of the topic placement in the path with random redistribution of gradients

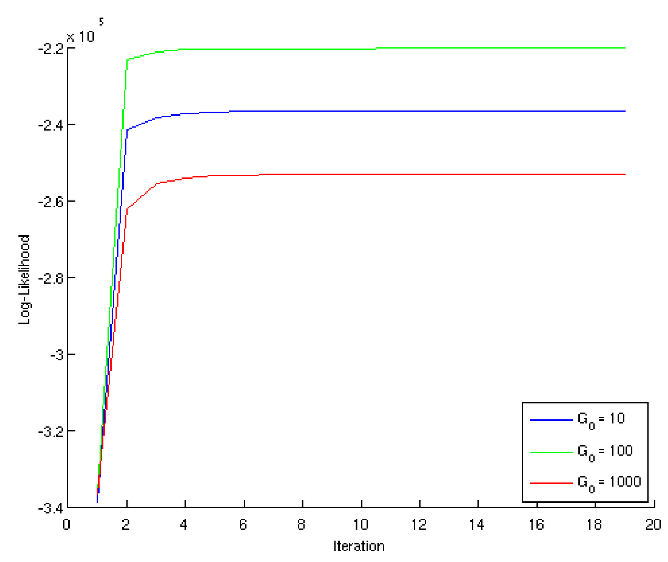
Results

Datasets:

- Google image re-ranking: 7 queries; follow [Fergus05] and sort clusters in descending order of size, and report precision at 15% recall; comparison with k-means clustering and different k from 10 to 300 clusters
- Caltech 101: report purity in the cluster

Dirichlet Process Mixture Models on Google image re-ranking



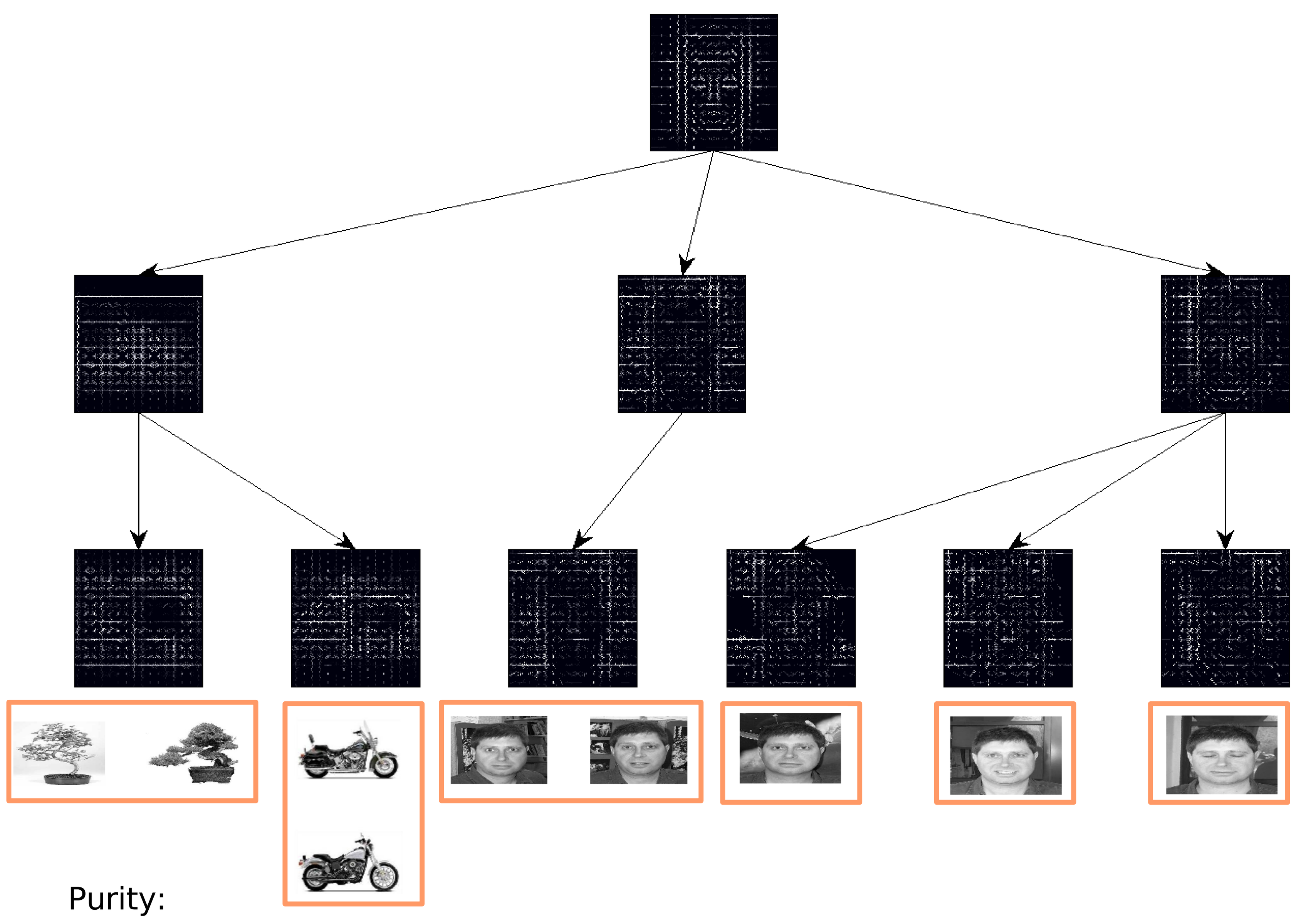


Precision at 15% recall

category	DPMM	k-means
airplane	1.000	0.824
cars rear	0.892	0.866
face	0.457	0.272
guitar	0.333	0.253
leopard	0.459	0.381
motorbike	0.643	0.689
wrist watch	1.000	0.971
average	0.680	0.564

Nested Chinese Restaurant Process on Caltech 101

Partial topic hierarchy with both heuristics



Purity:

71.4%	90.0%	100%	100%	100%	100%

Conclusion

- Cluster are very specific (people with suits, square clocks etc.)
- Single Dirichlet Process (DPM, NCRP): fast convergence, hyperparameter easy to estimate and quite robust
- Multiple Dirichlet Processes (HDPM): slow convergence and results depend heavily on hyperparameters

Literature

D.M. Blei, T.L. Griffiths, M.I. Jordan, and J.B. Tenenbaum. Hierarchical topic models and the nested chinese restaurant process. In NIPS, 2003.

M. Fritz, and B. Schiele. Decomposition, discovery and detection of visual categories using topic models. In CVPR, 2008.

J. Sivic, B.C. Russell, A. Zisserman, W.T. Freeman, and A.A. Efros. Unsupervised discovery of visual object class hierarchies. In CVPR, 2008.

Y.W. Teh, M.I. Jordan, M.J. Beal, and D.M. Blei. Sharing clusters among related groups: Hierarchical Dirichlet processes. In NIPS, 2005.