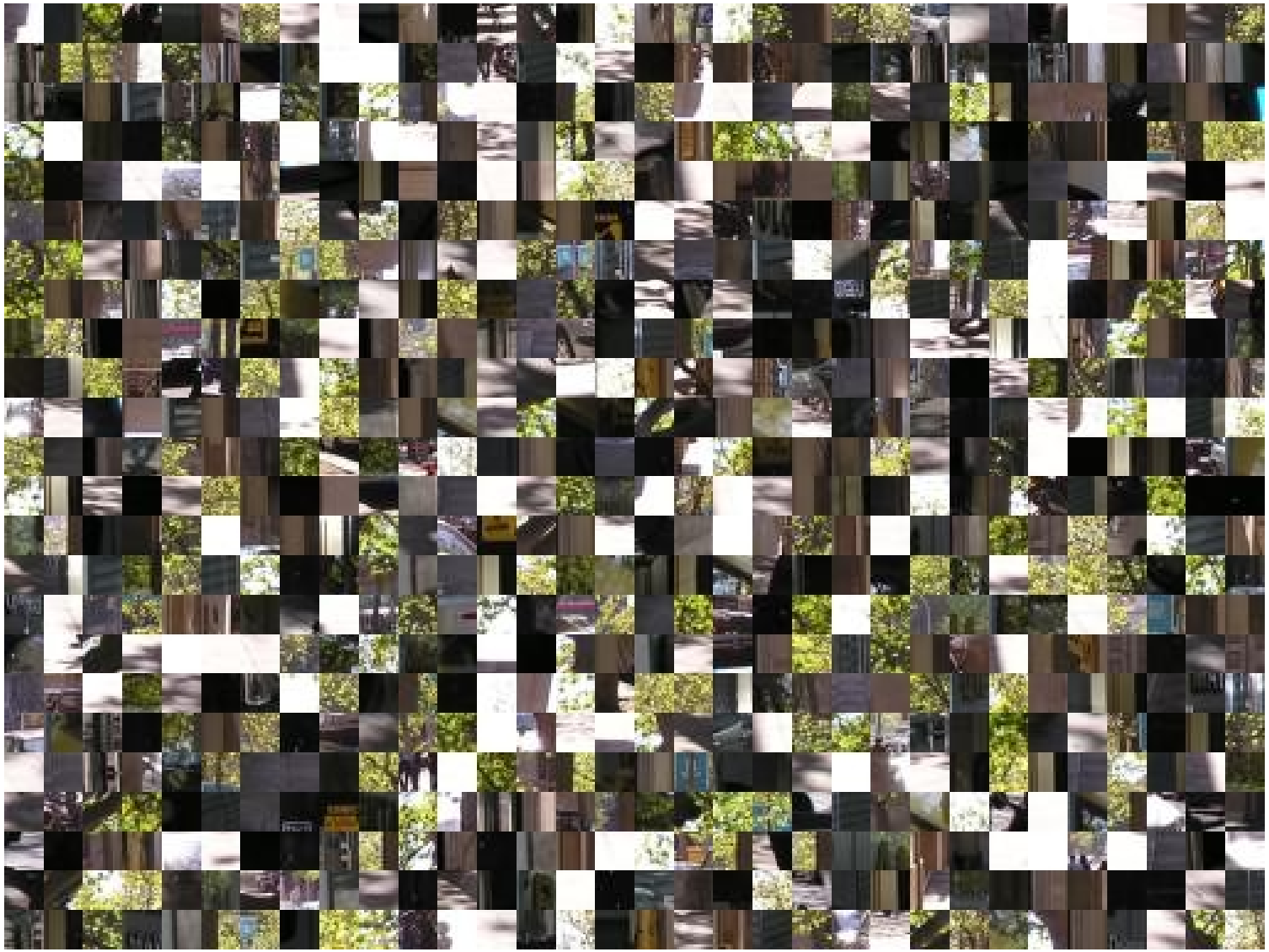
Spatial Relationships in Image Understanding

Roman Stanchak

CMSC 828G Link Mining

05/13/2008

Image as a Bag of Words

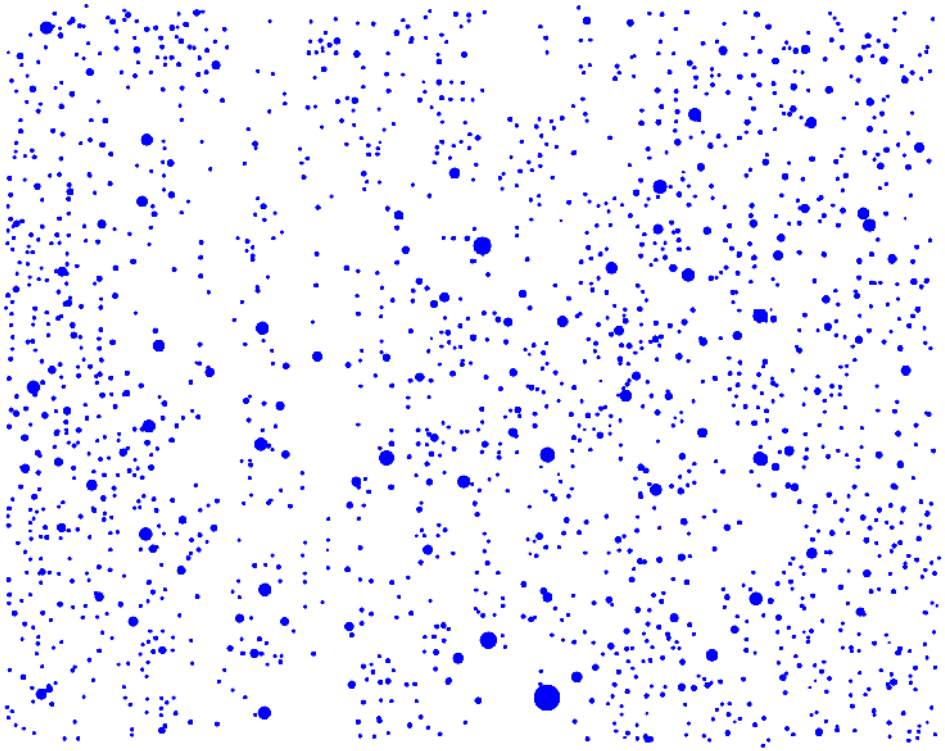


Spatial Context is Important

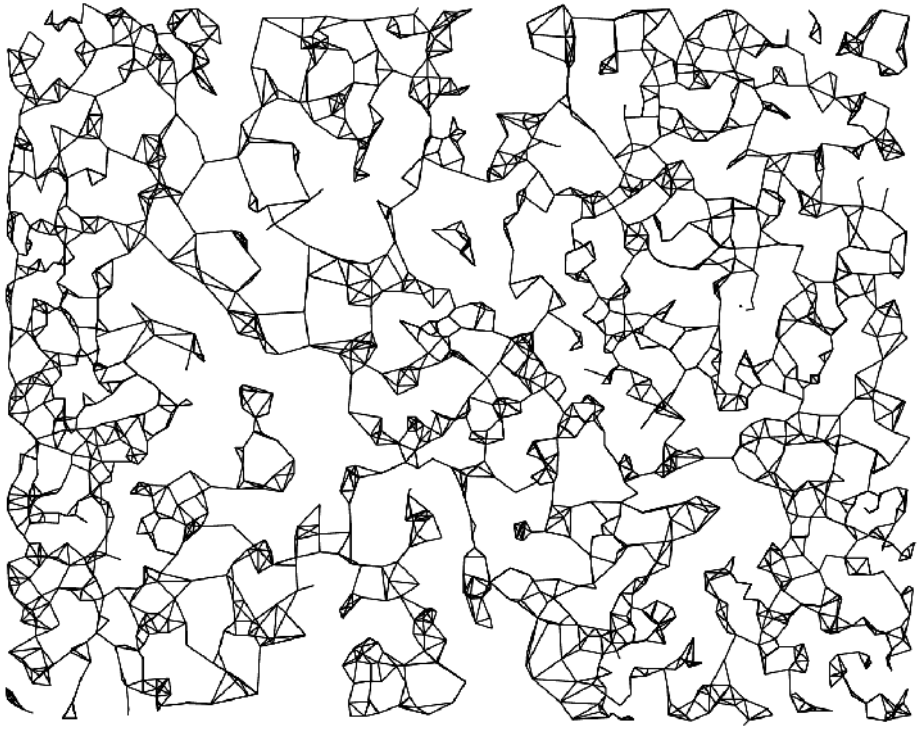
Visual Words

- Salient Points in Image
- Descriptor of Local Appearance

Inducing a Graph

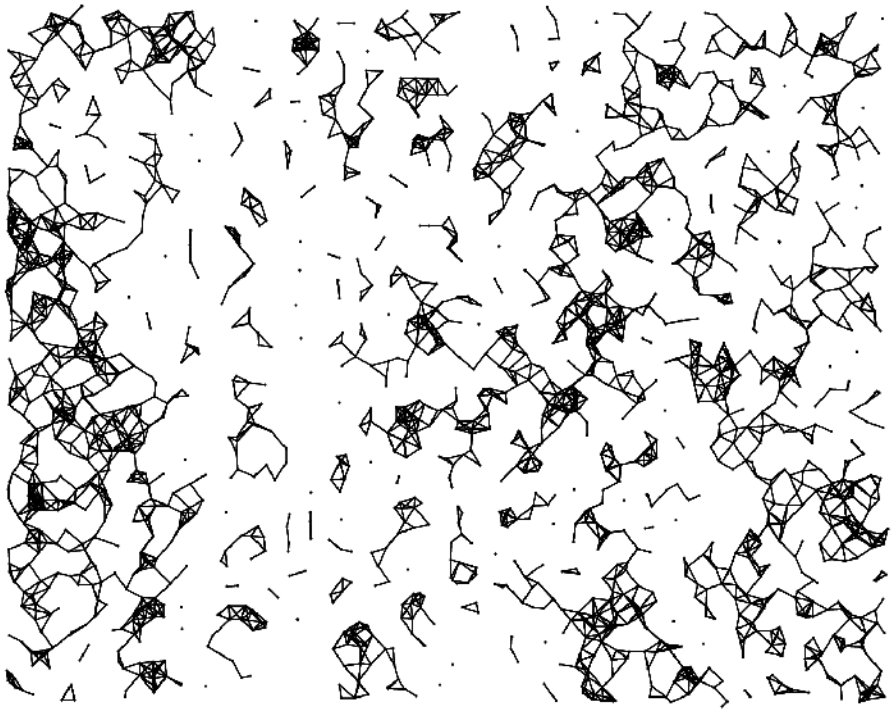


Visual Word Positions

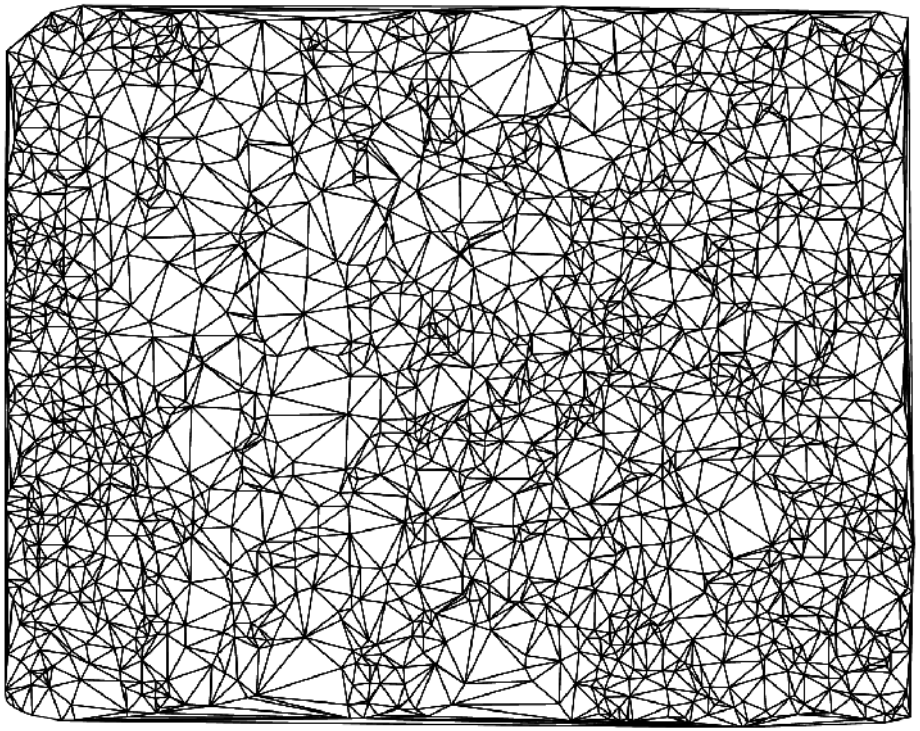


K-Nearest Neighbors (K=5)

Inducing a Graph

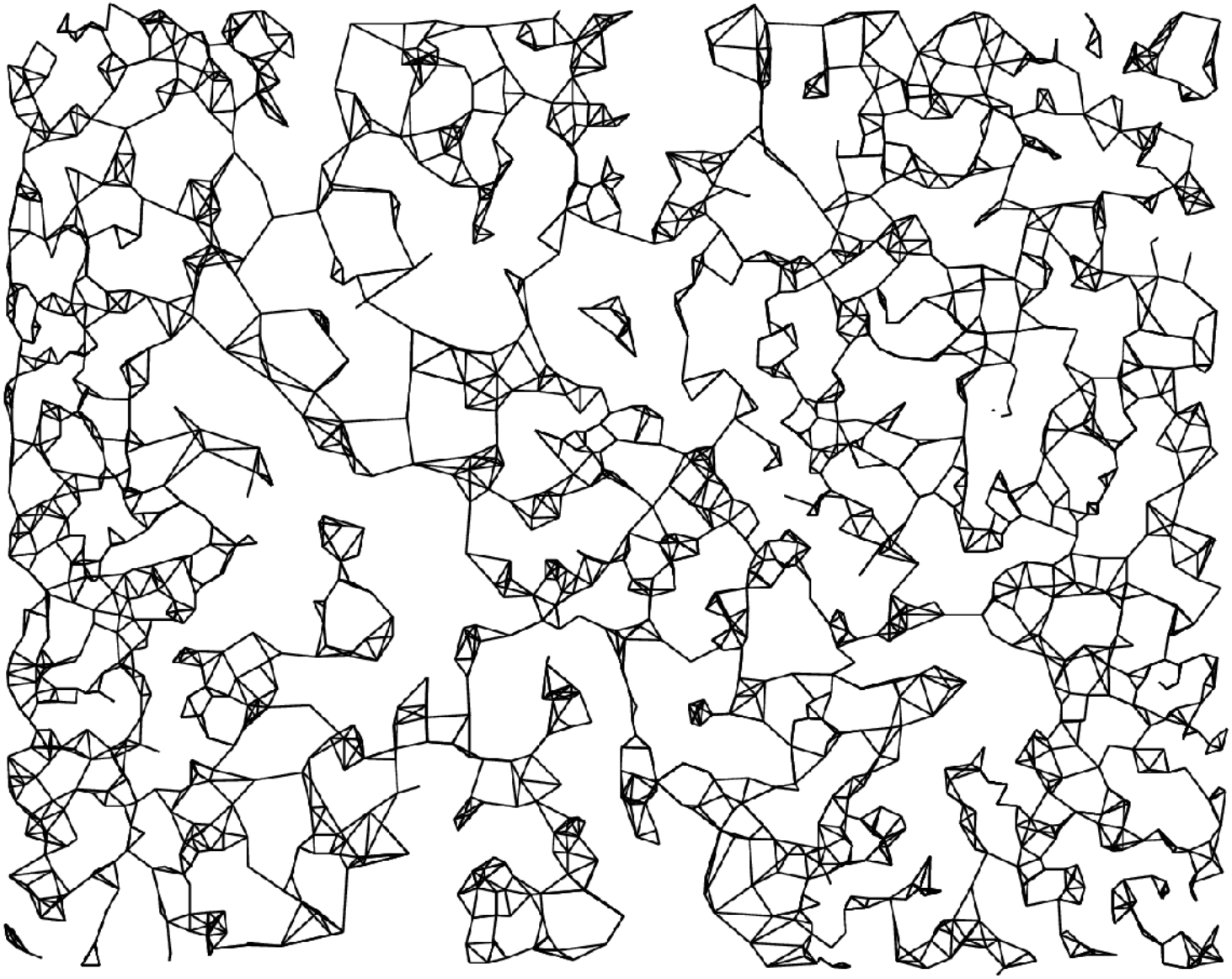


Fixed Radius Neighbors
(R=10)

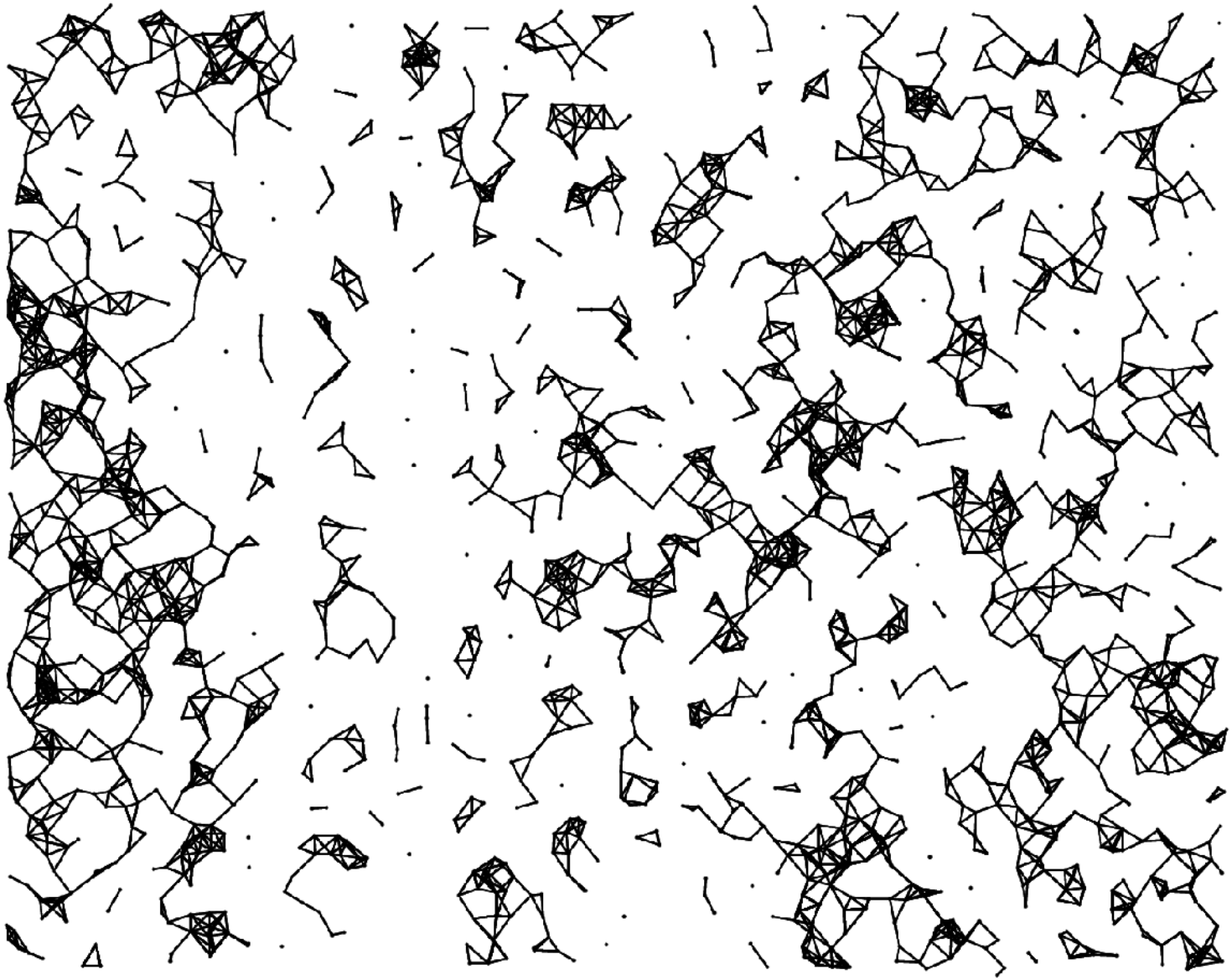


Delaunay Triangulation

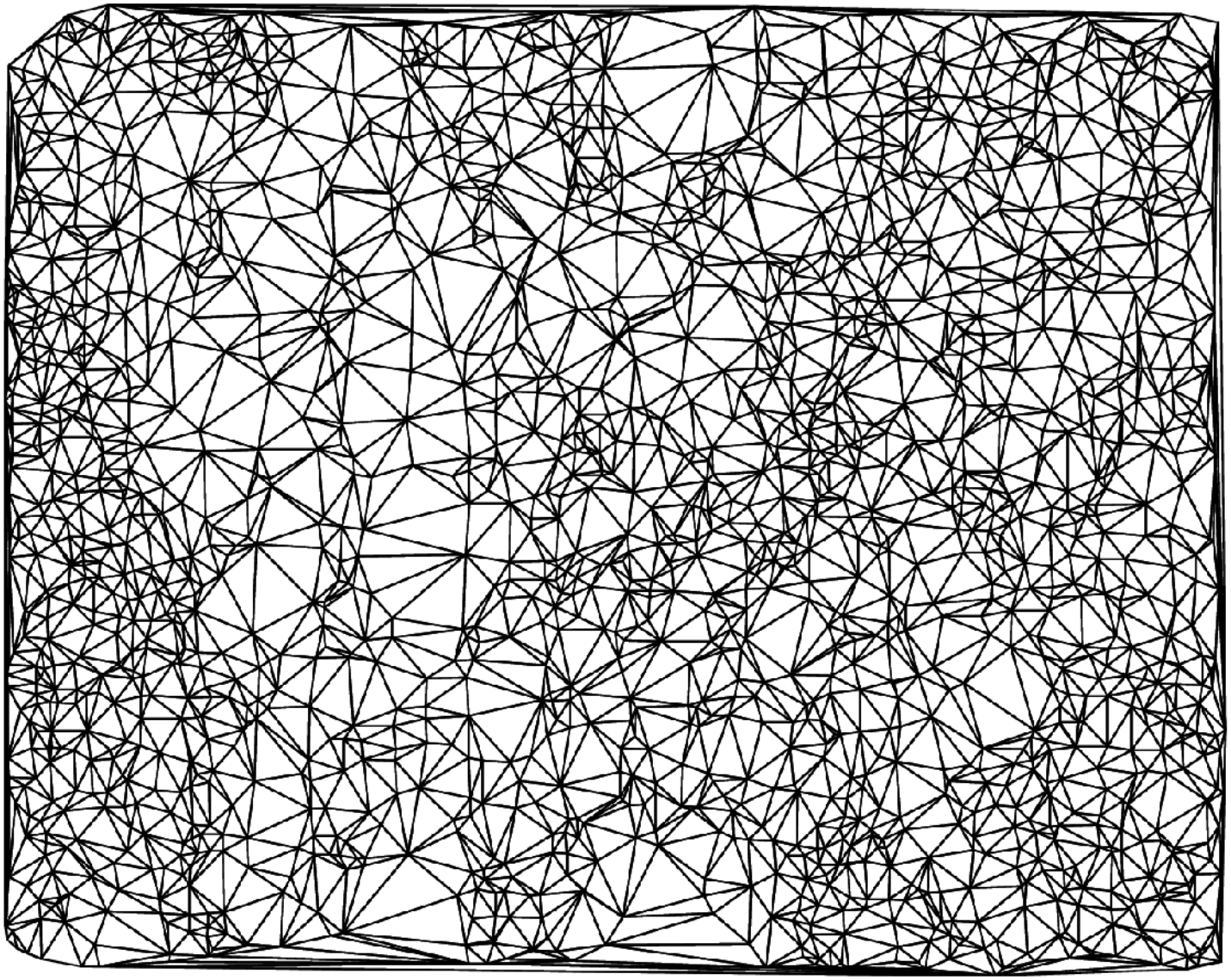
K-Nearest Neighbors



Fixed-Radius Neighbors



Delaunay Triangulation



LabelMe

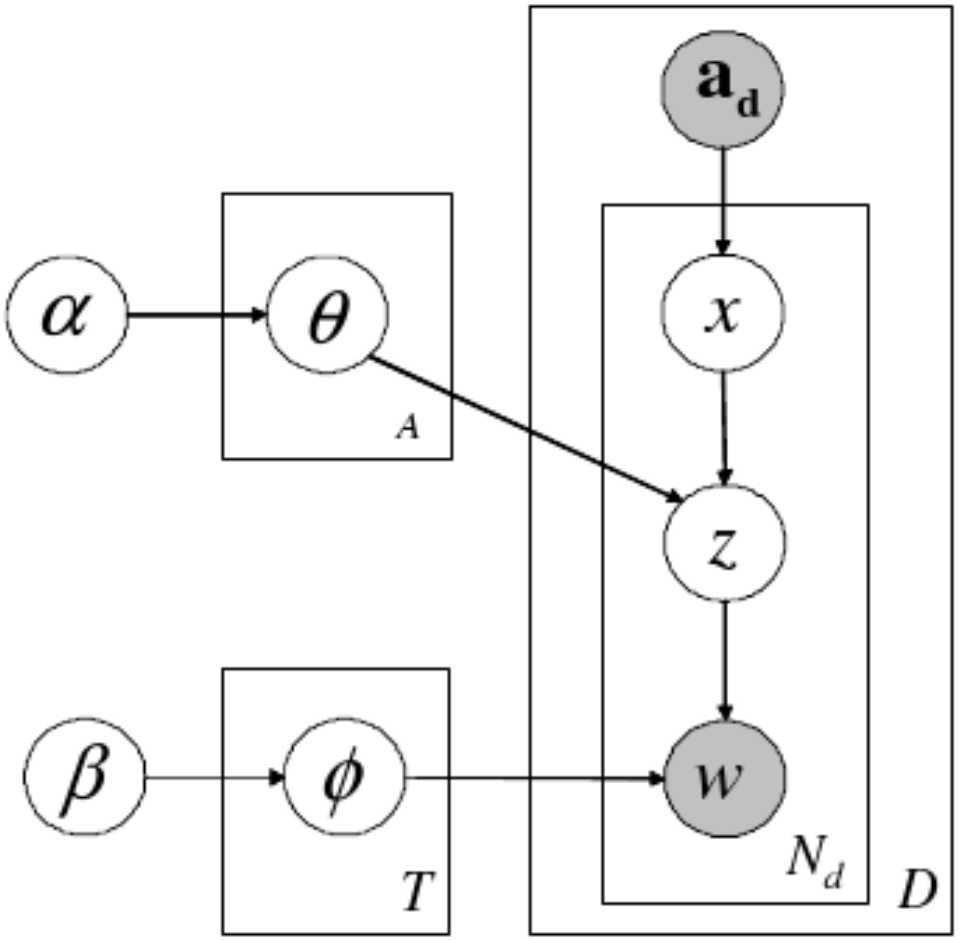
(Russell et al. 2007)

- Images w/ user annotations
  - 162,988 Images (43,175 annotated)
- Benchmark Data
  - 2920 Train/1033 Test Images

Author-Topic Model

(Rosen-Zv et al.i 2004)

- Image
  - Document D
- Image Patch
  - Word w
- Image Annotation
  - Author x
- Topic z

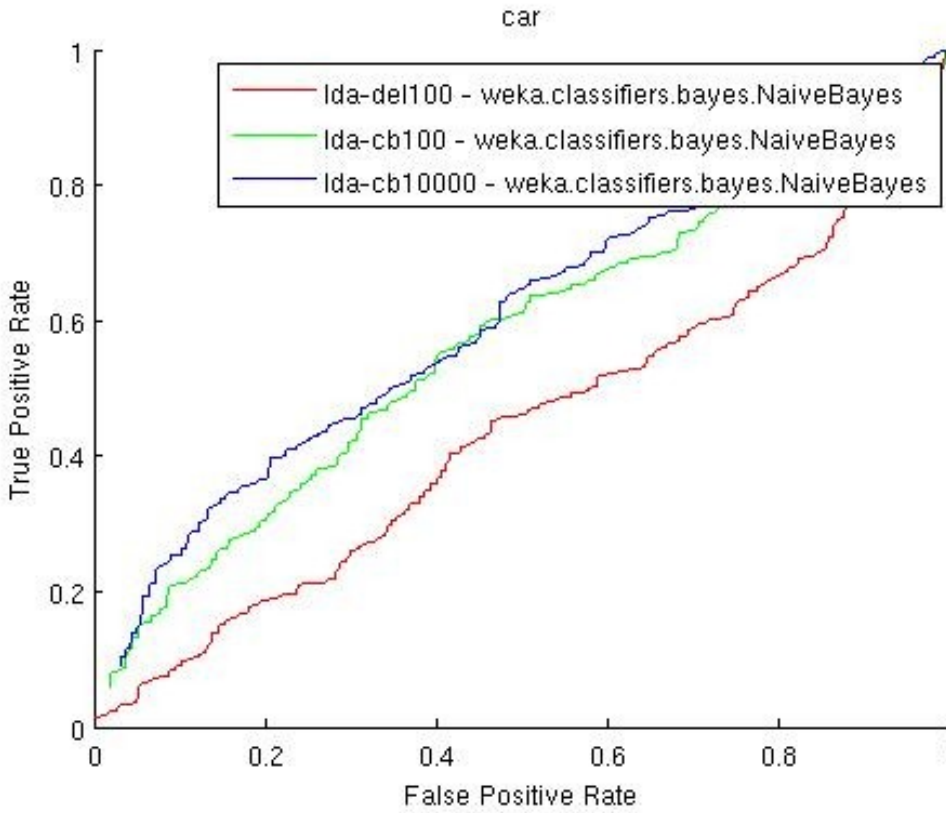


Experimental Setup

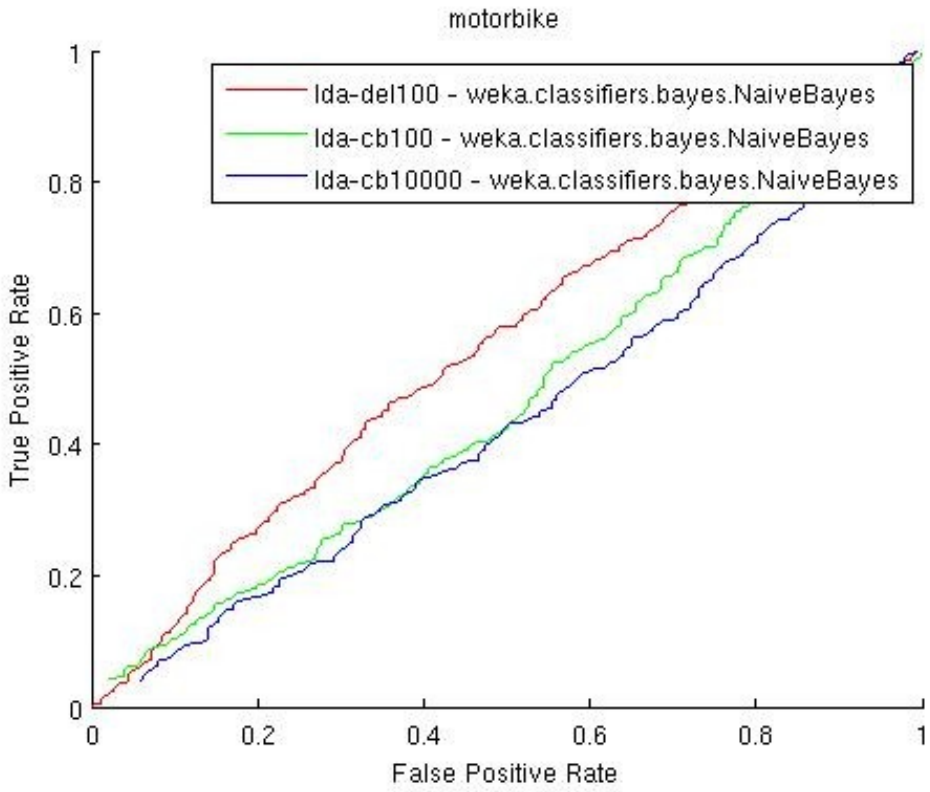
- 10000 visual codewords
- 100 visual codewords
- 5000 relational codewords
  - KNN, Fixed Radius, Delaunay
- LDA and Author Topic Model
  - Low-dimensional representation -> Classifier
- Binary Classification Task
  - Does the image contain object X?

Initial Results

ROC Curves for two object categories (Naïve Bayes classifier)



Delaunay feature does poorly



Some improvement, but still barely better than random

LDA is reported here, Author-Topic Model did not complete in time for this Poster

Conclusions & Future Work

- Gibbs Sampling is Slow
- Poor classification accuracy
  - Difficult data? Wrong model? More Gibbs iterations?
- Integrate neighbor properties into generative model