Integrating Image Clustering and Codebook Learning

Pengtao Xie and Eric Xing

{pengtaox,epxing}@cs.cmu.edu

School of Computer Science
Carnegie Mellon University
Outline

- Motivation
- Double Layer Gaussian Mixture Model
- Spatially Coherent Double Layer Gaussian Mixture Model
- Experiments
- Conclusions
Outline

- Motivation
- Double Layer Gaussian Mixture Model
- Spatially Coherent Double Layer Gaussian Mixture Model
- Experiments
- Conclusions
Motivation

- Image clustering and codebook learning are closely related and can mutually benefit each other.
Motivation

- Better codebook contributes to better clustering results.
Motivation

- Cluster labels can guide codebook learning.
Motivation

- Existing approaches perform them separately.
- **Goal:** develop a unified framework to perform two tasks simultaneously to make them mutually promote each other.
Outline

- Motivation
- Double Layer Gaussian Mixture Model
- Spatially Coherent Double Layer Gaussian Mixture Model
- Experiments
- Conclusions
Double Layer Gaussian Mixture Model
Outline

- Motivation
- Double Layer Gaussian Mixture Model
- Spatially Coherent Double Layer Gaussian Mixture Model
- Experiments
- Conclusions
Spatially Coherent Double Layer Gaussian Mixture Model
Spatially Coherent Double Layer Gaussian Mixture Model

Using MRF to enforce spatial coherence
Outline

- Motivation
- Double Layer Gaussian Mixture Model
- Spatially Coherent Double Layer Gaussian Mixture Model
- Experiments
- Conclusions
Experimental Setup

• Datasets
  • 15-Scenes: 4485 images, 15 classes
  • Caltech-101: 9144 images, 101 classes
  • 16x16 dense patches, 128-dimensional SIFT descriptor

• Baselines
  • K-means (KM), normalized cut (NC), joint scene and object model (JSOM), latent Dirichlet allocation (LDA)

• Evaluation Metrics
  • Accuracy (AC)
  • Normalized Mutual Information (NMI)
Accuracy on 15-Scenes

Accuracy (%) on 15-Scenes Dataset

Codebook Size

Accuracy (%)
Accuracy on Caltech-101

Accuracy (%) on Caltech-101 Dataset

Accuracy (%)

Codebook Size

KM
NC
JSOM
LDA
DLGMM
SC-DLGMM
Normalized Mutual Information (NMI) on 15-Scenes Dataset

NMI (%) on 15-Scenes Dataset

Codebook Size

NMI (%) on 15-Scenes Dataset

KM
NC
JSOM
LDA
DLGMM
SC-DLGMM
Normalized Mutual Information (NMI) on Caltech-101

NMI (%) on Caltech-101 Dataset

- KM
- NC
- JSOM
- LDA
- DLGMM
- SC-DLGMM

Codebook Size
Outline

- Motivation
- Double Layer Gaussian Mixture Model
- Spatially Coherent Double Layer Gaussian Mixture Model
- Experiments
- Conclusions
Conclusions

- Image clustering and codebook learning are closely related.
- We propose Double Layer Gaussian Mixture Model to perform two tasks jointly.
- We propose Spatially Coherent DLGMM to incorporate spatial coherence.
- Experiments on two datasets demonstrate the effectiveness of two models.
Thank you!
Questions?