3D Object Class Geometry Modeling with Spatial Latent Dirichlet Markov Random Fields

{HANCHEN.XIONG, SANDOR.SZEDMAK, JUSTUS.PIATER }@UIBK.AC.AT

Background

In computer vision research, there exists an obvious gap between 2D appearance modeling and 3D geometry modeling in terms of interpretation and representation abilities, and it has been advocated that robust 3D geometry modeling is highly desirable. Motived by this gap and desire, we put forward a novel 3D object class geometry model in the light of state-of-the-art techniques developed in machine learning and computer graphics.

The basic underlying principle of our modeling is that different instances of the same class should share similar 3D structure of composing parts, although their parts can slightly vary from one instance to another.

Contribution

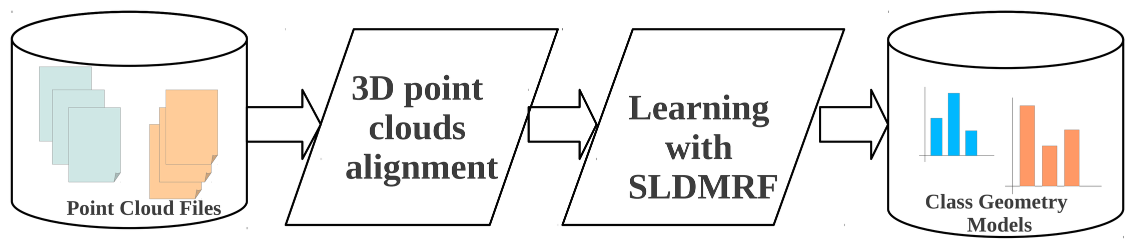
A novel part-based geometry model (Spatial Latent Dirichlet Markov Random Fields) for 3D object classes:

- constructs discrete vocabulary for 3D shapes;
- extends latent Dirichlet allocation (LDA) by strategically constructing a Markov random field (MRF) on the part labels, which enhancing the spatial coherence for part segmentation;
- exhibits superior semantic interpretation and discriminative ability in model classification to LDA and other related models.

Learning Framework

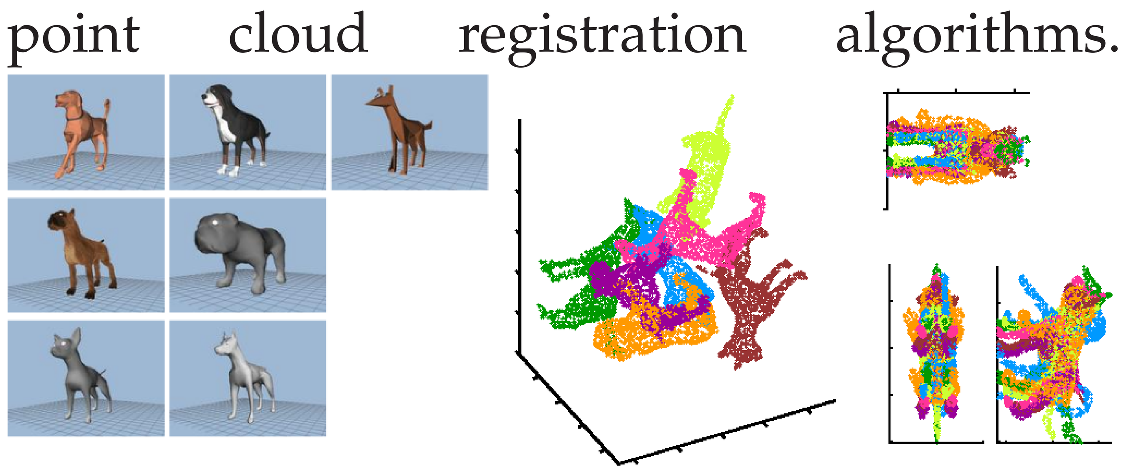
3D object shapes are represented by point cloud data (PCD).

- First, for each class, different PCDs of object instances are aligned using point cloud registration methods.
- Secondly, spatial latent Dirichlet Markov Random fields (SLDMRF) are constructed and learned for each category based on the aligned instance point clouds.



Alignment

The alignment of different instances of the same class is achieved with point cloud registration algorithms.



Inference and Learning

A hybrid Gibbs sampler:

$$q_{\mathrm{LDA}}(z_m^{(n)}=k)\cdot q_M(z_m^{(n)}=k)\cdot q_c(z_m^{(n)}=k)$$

where $q_{\mathrm{LDA}}(z_m^{(n)}=k)$ is the collapsed Gibbs sampler of LDA,

$$q_M(z_m^{(n)}=k) \propto \frac{\exp\left(\sum_{(z_j,z_m^{(n)})} g(z_j,z_m^{(n)}=k)\right)}{\sum_h \exp\left(\sum_{(z_j,z_m^{(n)})} g(z_j,z_m^{(n)}=h)\right)}$$

is the Gibbs sampler based on the Markov random field, and

$$q_c(z_m^{(n)}=k) \propto \frac{\mathcal{N}(\hat{w}_m^{(n)};\boldsymbol{\mu}_l^{(k)},\boldsymbol{\Lambda}_l^{(k)})}{\sum_h \mathcal{N}(\hat{w}_m^{(n)};\boldsymbol{\mu}_l^{(h)},\boldsymbol{\Lambda}_l^{(h)})}$$

is a collapsed Gibbs sampler derived from the information of all part positions. The parameters $\{\boldsymbol{\pi}_m\}_{m=1}^M$, $\{\boldsymbol{\theta}_k\}_{k=1}^K$, $\{c_k\}_{k=1}^K$ can be read out once Gibbs sampling converges.

Quantitative Results

Representational capabilities of models are tested based on classification, comparison between GM (which is currently most popular model in practice) and SLDMRF:

Classifiation with SLDMRF

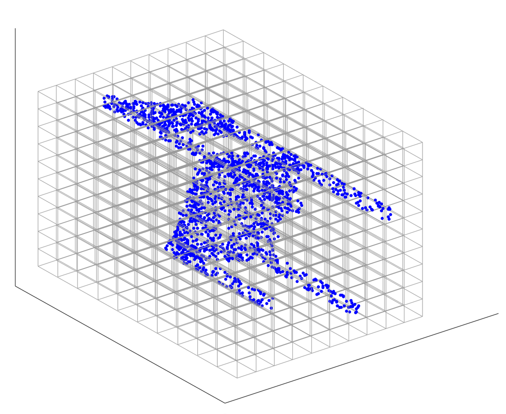
Ground Truth	airplane	bike	car	dog	motorcycle
airplane	0.92	0.02	0.04	0.01	0.01
bike	0.01	0.71	0.01	0.03	0.25
car	0.00	0.02	0.80	0.03	0.07
dog	0.02	0.05	0.12	0.70	0.10
motorcycle	0.01	0.10	0.03	0.03	0.82

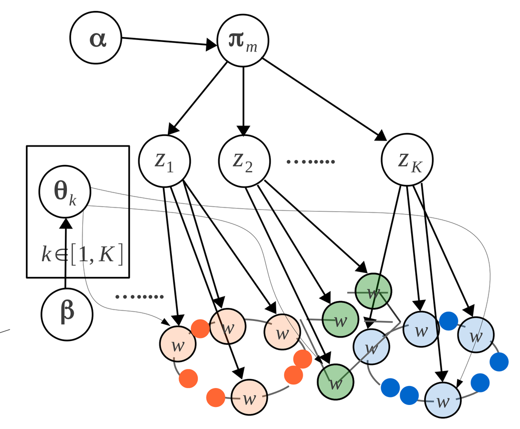
Classifiation with GM

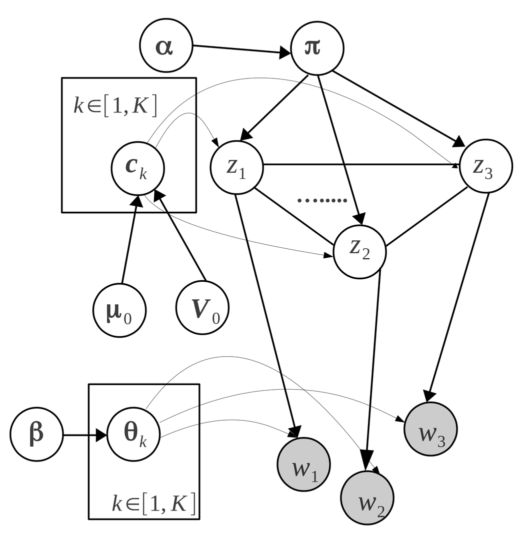
Ground Truth	airplane	bike	car	dog	motorcycle
airplane	0.85	0.04	0.06	0.03	0.02
bike	0.05	0.56	0.07	0.10	0.22
car	0.06	0.04	0.65	0.10	0.05
dog	0.02	0.07	0.20	0.61	0.10
motorcycle	0.01	0.17	0.06	0.05	0.71

Spatial Latent Dirichlet Markov Random Fields

3D vocabulary construction and LDA [1] modeling for object class geometry. $\boldsymbol{\pi} = \{\pi_m\}_{m=1}^M$ denote part weights in each object class, and $\boldsymbol{\theta} = \{\theta_k\}_{k=1}^K$ describe the spatial distribution of 3D words. $\boldsymbol{\alpha}, \boldsymbol{\beta}$ are hyper-parameters.

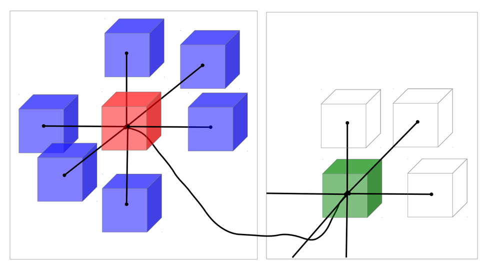






SLDRMF

SLDMRF models the positions of all parts with parameters $\mathbf{c}_k$ such that 3D words that share the same label k are likely to be close; **SLDMRF** constructs Markov random fields, in which the links are intra-connections between neighboring words (between red and blue) and inter-connections between corresponding words (between red and green) across different objects. Potential functions on links are defined based on the compatibility between corresponding orientation vectors.



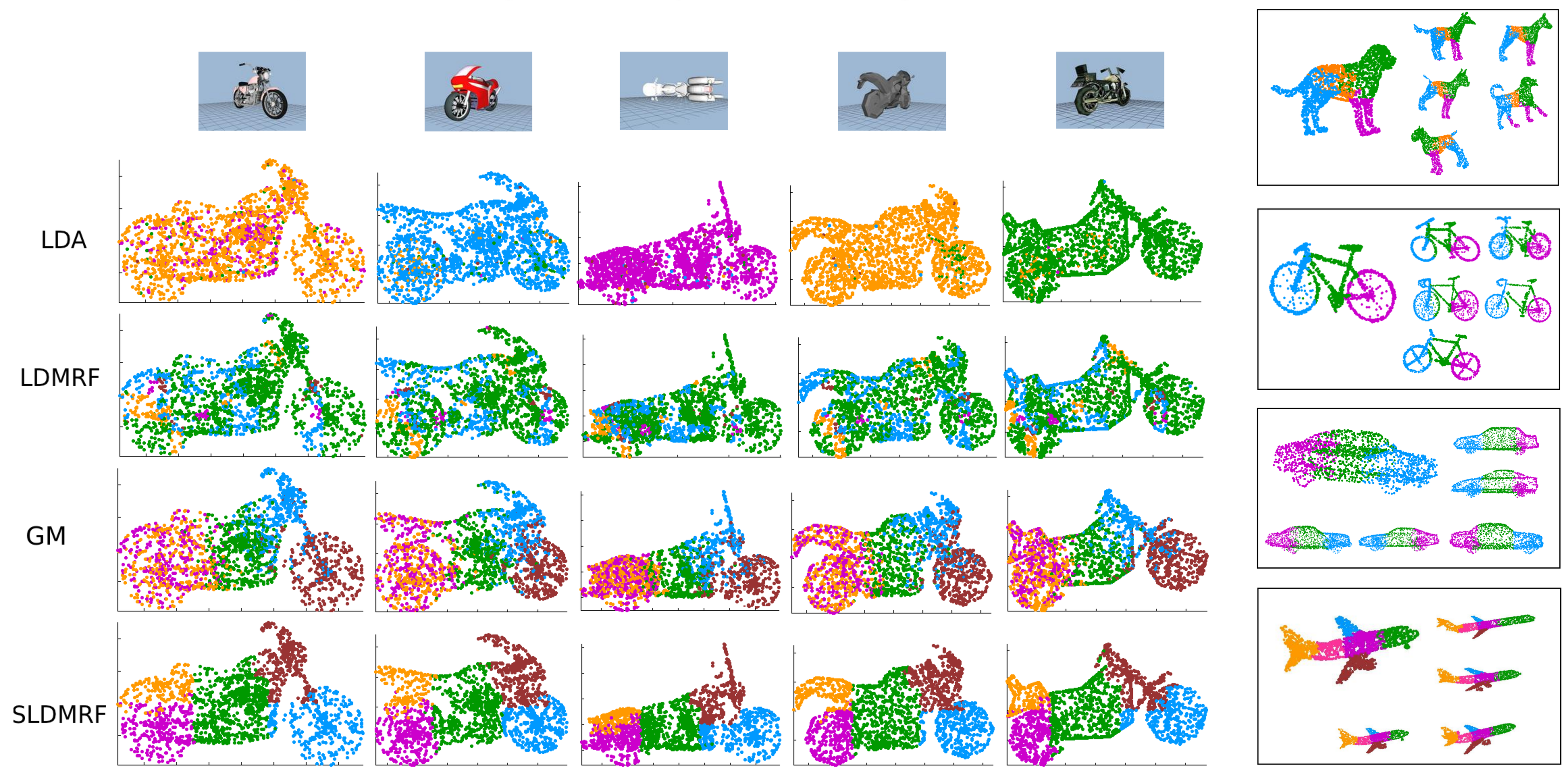
connections:

$$g(z_i,z_j) = \exp(|\langle o_i,o_j\rangle|)$$

where o_i is the orientation vector

Qualitative Results

For comparison, LDA [1], LDMRF [2] and Gaussian Mixtures (GM) models are tested on the same data for 3D object class geometry modeling:



Acknowledgement

This research has received funding from the European Community's Seventh Framework Programme FP7/2007-2013 under grant agreement no. 270273, Xperience.

References

(1) Blei, D., Ng, A., Jordan, M.: Latent Dirichlet Allocation. Journal of Machine Learning Research 3, 993 to 1022 (2003)

(2) Mackey, L.: Latent Dirichlet Markov Random Fields for Semi-supervised Image Segmentation and Object Recognition (2007)