

# Unsupervised Learning of Shape-Motion Patterns for Objects in Urban Street Scenes

Dirk Klostermann  
 dirk.klostermann@rwth-aachen.de  
 Aljoša Ošep  
 osepa@vision.rwth-aachen.de  
 Jörg Stückler  
 stueckler@vision.rwth-aachen.de  
 Bastian Leibe  
 leibe@vision.rwth-aachen.de

Computer Vision Group,  
 Visual Computing Institute,  
 RWTH Aachen University  
 Aachen, Germany

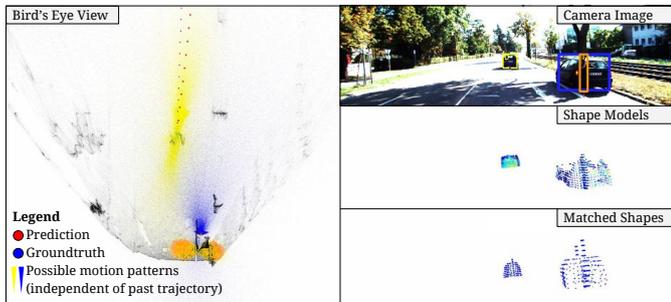


Figure 1: Possible future motion predictions based on shape-motion patterns (SMPs). In this scene, a car (yellow) is driving forward, another car (blue) is parked on the right and next to the parked car a pedestrian (orange) is waiting. Our model predicts possible motion behaviour even for static objects. Note that in this situation our model can indicate that the standing pedestrian may walk onto the road.

Analysing and predicting the movement of objects is a vital ability for self-driving cars and autonomous mobile robots. Such systems need to be able to foresee potential collisions and also react to possibly harmful situations. We propose an unsupervised approach to learn typical motion patterns of object categories from example data. In our approach, object categories are not limited to predefined classes. Instead, our approach provides a categorization into objects with similar shapes and trajectories.

Current research on motion prediction can be structured into dynamical, maneuver-based, and complex interaction-aware models [3]. Maneuver-based methods, such as ours (e.g. [1, 2, 5, 9]) find patterns in previously observed trajectories to predict the future evolution of the trajectory. Compared to previous maneuver-based methods, we additionally distinguish the observed objects by their shape, which allows us to assign object category-specific motion patterns. While most related work in the research area of object categorization focuses on supervised methods (e.g. [4, 6]), only a small number of approaches tackles the semi-/unsupervised categorization of objects [7, 8]. Based on noisy stereo data we cluster objects based on shape and trajectory information in an unsupervised way.

We propose to describe objects in a hierarchical approach, as visualized in Fig. 2: (1.) First, we run a tracker on the training set and gather training data by sampling the trajectories of the tracked objects. A training example contains a shape model of the object together with its trajectory. (2.) We cluster the instances from the training set based on their shape. The shape clusters  $c_S \in \mathcal{C}_S$  represent categories of objects, differentiating view points within an object class. (3.) Various instances in each shape cluster can have different motion models, e.g. a car can be parked or drive with a large velocity. Hence, we cluster the trajectories  $m \in \mathcal{M}(c_S)$  of each shape cluster  $c_S$  to obtain shape-specific trajectory clusters  $c_{M|c_S} \in \mathcal{C}_{M|c_S}$ . (4.) Each trajectory cluster in a shape forms one shape-motion pattern (SMP)  $p = (c_S, c_{M|c_S})$ . The trajectories in the SMP can be described by the Gaussian distribution on the past and future positions relative to their current position.

Using our learned model, we classify novel object shapes and trajectories into one of the learned SMPs. (5.) For this, we gather shape and trajectory information from the same kind of tracker that we use in the training phase. (6.) By comparing the integrated shape of the new instance to all cluster centers we extract a subset of shape centers which are

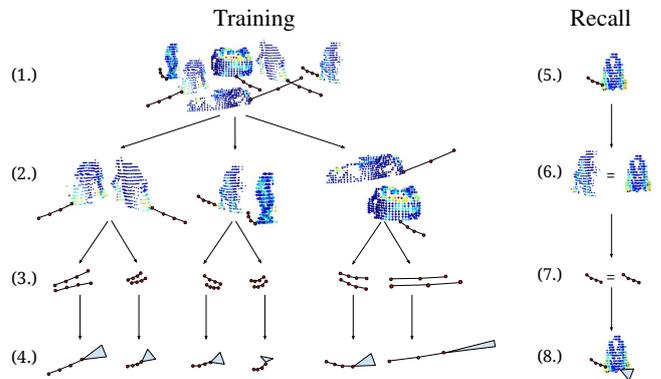


Figure 2: Motion prediction pipeline based on hierarchical clustering.

most similar to the observed object.

$$C_S(s) := \left\{ c_S \in \mathcal{C}_S \mid d(s, \hat{s}(c_S)) < \lambda \min_{c'_S \in \mathcal{C}_S} d(s, \hat{s}(c'_S)) \right\}, \quad (1)$$

where  $\hat{s}(c_S)$  is the exemplar shape of cluster  $c_S$  and parameter  $\lambda$  controls the number of closest shape clusters. (7.) We compare all motion models from this subset of shape clusters to the observed trajectory and select the most similar one. (8.) Based on the selected motion model we predict the future motion. Furthermore, we use the subset of motion models from the previous step to infer a probability distribution describing where the object could move based on its shape.

We demonstrate in our experiments that our approach outperforms Kalman filter based predictions. Furthermore, we demonstrate that reasoning on shape can increase prediction performance w.r.t. trajectory-only maneuver-based methods, especially for the less frequent object categories e.g. pedestrians whose trajectory models differ significantly from the shape-independent trajectory models. In addition, we demonstrate qualitatively that our method can predict possible future motions for static objects (see Fig. 1) and thus foresee potentially dangerous situations.

- [1] S. Calderara, A. Prati, and R. Cucchiara. Mixtures of von mises distributions for people trajectory shape analysis. *IEEE Trans. on Circuits and Systems for Video Technology*, 21(4):457–471, 2011.
- [2] J. Joseph, F. Doshi-Velez, A. S. Huang, and N. Roy. A bayesian nonparametric approach to modeling motion patterns. *Autonomous Robots*, 31(4):383–400, 2011.
- [3] S. Lefèvre, D. Vasquez, and C. Laugier. A survey on motion prediction and risk assessment for intelligent vehicles. *Robomech Journal*, 1(1), 2014.
- [4] D. Mitzel, J. Diesel, A. Ošep, U. Rafi, and B. Leibe. A fixed-dimensional 3d shape representation for matching partially observed objects in street scenes. In *ICRA*, 2015.
- [5] V. Romero-Cano, J. Nieto, G. Agamennoni, et al. Unsupervised motion learning from a moving platform. In *Intel. Vehicles Symp.*, 2013.
- [6] L. Spinello, R. Triebel, and R. Siegwart. Multiclass multimodal detection and tracking in urban environments. *IJRR*, 29(12):1498–1515, 2010.
- [7] Y. Tao, R. Triebel, and D. Cremers. Semi-supervised online learning for efficient classification of objects in 3d data streams. In *IROS*, 2015.
- [8] A. Teichman, J. Levinson, and S. Thrun. Towards 3d object recognition via classification of arbitrary object tracks. In *ICRA*, 2011.
- [9] J. Wiest, M. Höffken, U. Kresel, and K. Dietmayer. Probabilistic trajectory prediction with gaussian mixture models. In *Intel. Vehicles Symp.*, pages 141–146, 2012.