LONG BEACH
CALIFORNIA
June 16-20, 2019

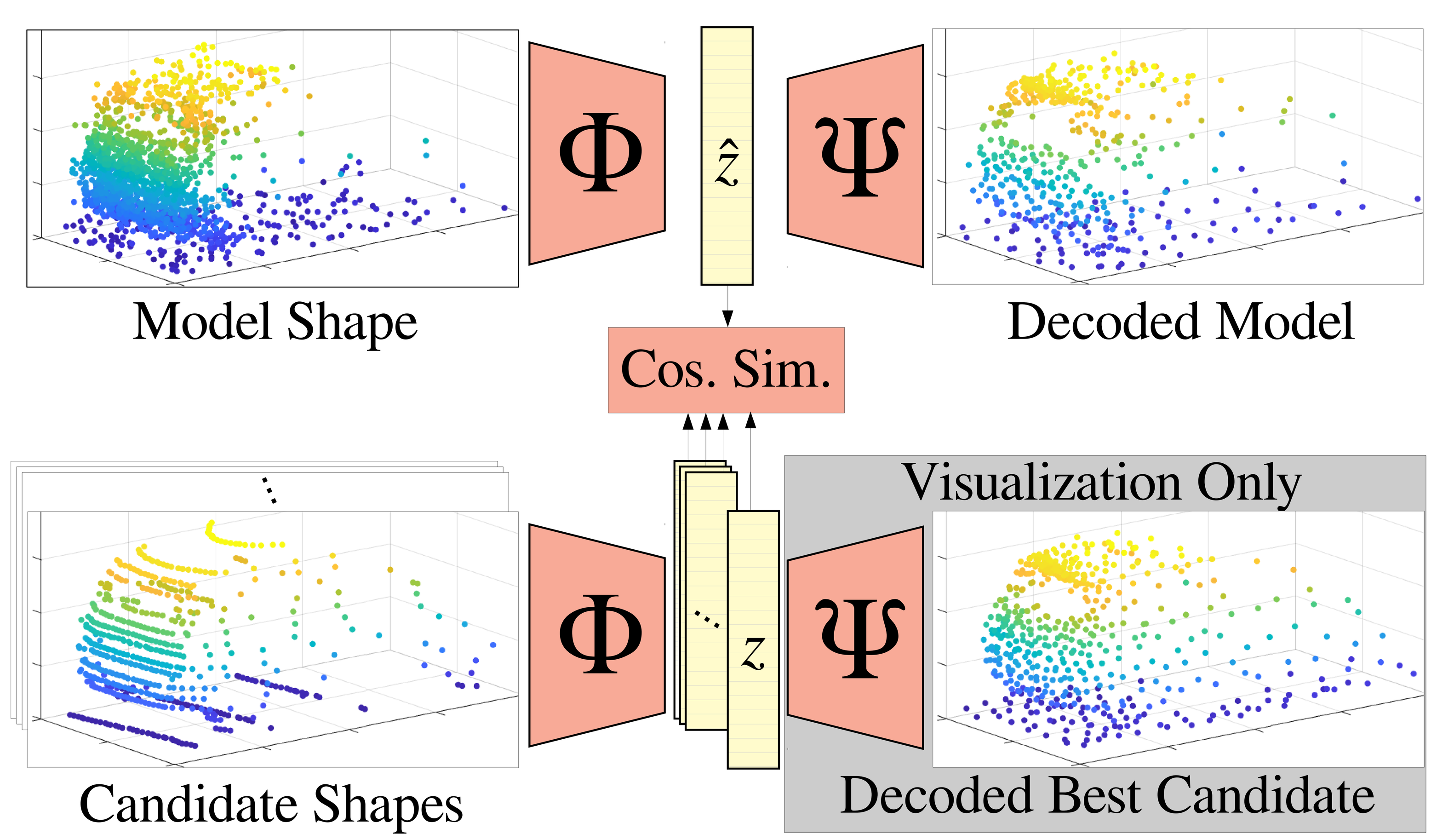
Leveraging Shape Completion for 3D Siamese Tracking

Silvio Giancola*, Jesus Zarzar* and Bernard Ghanem
King Abdullah University of Science and Technology (KAUST)

https://github.com/SilvioGiancola/ShapeCompletion3DTracking

Abstract

- We propose the **first 3D Siamese tracker** applied to point clouds rather than images.
- We **regularize** the 3D Siamese network's latent space such that it resembles the latent space of a **shape completion** network.
- We show that imbedding semantic information for the shape leads in **better discrimination and tracking**.

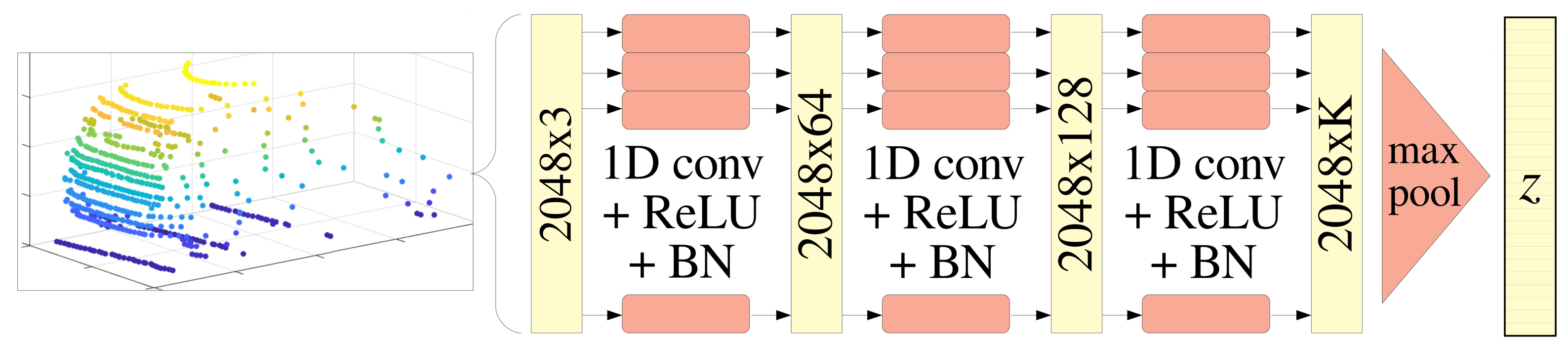


Motivation

- 3D LIDAR perception can provide crucial geometrical information for urban navigation at extended range.
- LIDARs are more reliable in challenging environment like low light and weather conditions (rain, fog,…).
- However their sparse representation are challenging to process and require cumbersome processing.

Methodology

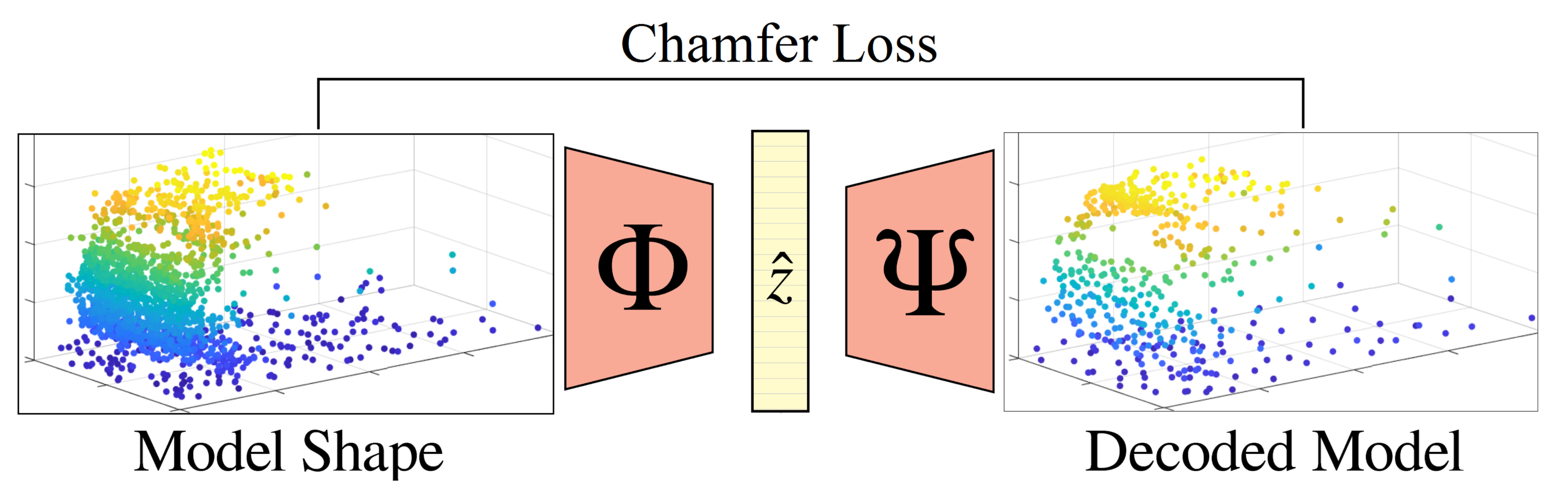
- **3D Siamese Tracking:** A K-dimensional latent vector is extracted from point clouds using an encoder composed of three layers of **1-D CNN + ReLU + BN** along with a final max pooling.



- We train our Siamese network to **regress** the cosine similarity between model and candidates to the **Gaussian distance** of the GT positions.

$$\mathcal{L}_{tr} = \frac{1}{n} \sum_{\mathbf{x}} \Big(CosSim\big(\phi(\mathbf{x}), \phi(\hat{\mathbf{x}})\big) - \rho\big(d\,(\mathbf{x}, \hat{\mathbf{x}})\big) \Big)^2$$

- **Shape Regularization:** We decode the latent shape features to **auto-encode the model point cloud**, regressing its coordinate with two FC layers. We train our shape regularization with a **chamfer loss** and enforce the latent vectors to hold meaningful and useful semantic information for discrimination.



$$\mathcal{L}_{comp} = \sum_{\hat{\mathbf{x}}_i \in \hat{\mathbf{x}}} \min_{\tilde{\mathbf{x}}_j \in \tilde{\mathbf{x}}} \|\hat{\mathbf{x}}_i - \tilde{\mathbf{x}}_j\|_2^2 + \sum_{\tilde{\mathbf{x}}_j \in \tilde{\mathbf{x}}} \min_{\hat{\mathbf{x}}_i \in \hat{\mathbf{x}}} \|\hat{\mathbf{x}}_i - \tilde{\mathbf{x}}_j\|_2^2$$

Experimental results

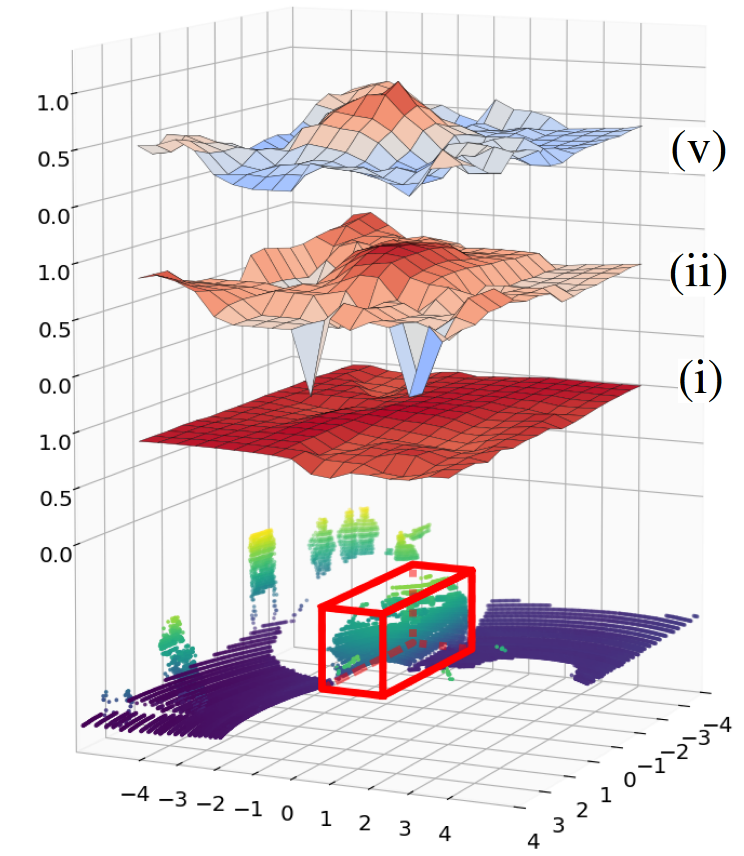
- We show quantitative improvement using the shape completion regularization in both the OPE metric for tracking performances and the completeness metric for shape completion.
  - **Tracking:** The shape completion regularization helps improving both **tracking metrics by 3%.**

Ablation	Success	Precision
(i) Before Training (Random)	39.06	41.79
(ii) Pre-trained on ShapeNet	44.54	49.38
(iii) Ours – Completion only	65.36	70.62
(iv) Ours – Tracking only	73.96	78.68
(v) Ours – $\lambda_{comp}@1e^{-6}$	**76.94**	**81.38**

  - **Completion:** Tracking the vehicle helps improving its **completion metrics by 5%.**

Method	(iii)	(iv)	(v)	[15]	[50]
Comp. [m]	0.188	0.690	**0.179**	0.130	0.078

  - **Speed:** Encoding 147 candidates occurs in only **1.8ms**.
  - **Qualitative Results:** Solving for both tasks improves the quality of the shape discrimination for tracking along the **search space manifold**.



Acknowledgments: This publication is based upon work supported by the King Abdullah University of Science and Technology (KAUST) Office of Sponsored Research (OSR) under Award No. 2017-266-1