RPSRNet: End-to-End Trainable Rigid Point Set Registration Network using Barnes-Hut 2^D-Tree Representation

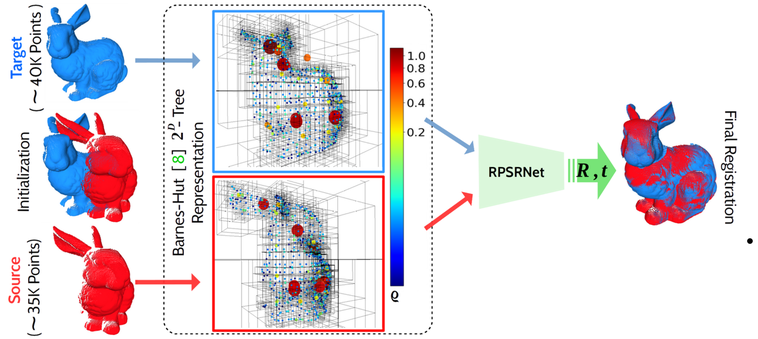
Sk Aziz Ali , Kerem Kahraman, Gerd Reis, Didier Stricker
TU Kaiserslautern, German Research Center for Artificial Intelligence (DFKI GmbH)

CVPR 2021 VIRTUAL JUNE 19-25

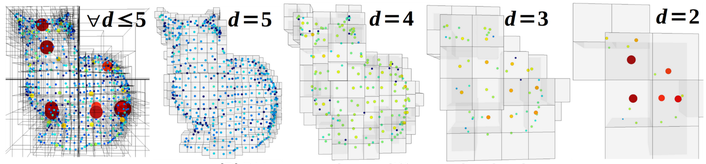
Overview

RPSRNet is a novel end-to-end trainable deep neural network for rigid point set registration with state-of-the-art accuracy and run-time –

(i) uses a novel Barnes-Hut -tree representation for the input point sets,
(ii) inference speed of 12-15 ms on a pair of point clouds as large as 250K
(iii) robust to noise, data disturbances, indoor and outdoor scenes



Transformation Estimation



(B) BH-tree nodes at different depths d

multi-scale sum of mean-squared distance errors between the CoMs

$$\mathbf{U}(\mathbf{R}, \mathbf{t}, \tau^{\mathbf{X}}, \tau^{\mathbf{Y}}) = \sum_{d} \sum_{l,\hat{l}} \varrho^{y-}_{d,l}\, \varrho^{x-}_{d,\hat{l}} \left\| \left(\mathbf{R}\mu^{y}_{d,l} + t \right) - \mu^{x}_{d,\hat{l}} \right\|_2^2$$

BH-trees | Depth | Node Labels | Inverse Node Densities | Center of Masses of nodes from $\tau^{\mathbf{X}}, \tau^{\mathbf{Y}}$

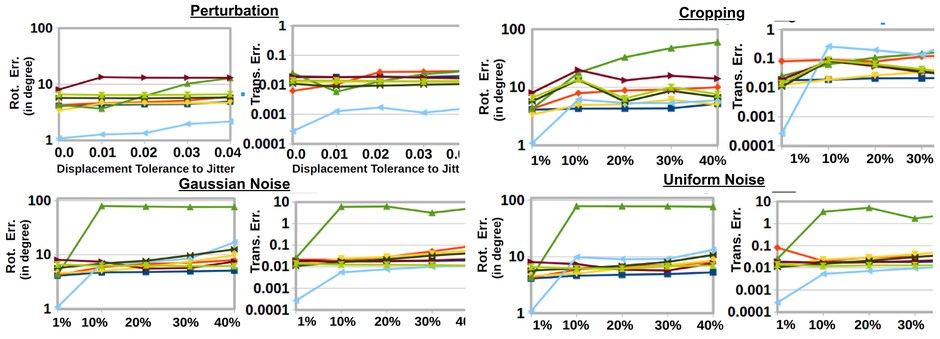
- RPSRNet is a deep-learning framework to estimate **R, t** of $\mathbf{U}(\mathbf{R}, \mathbf{t}, \tau^{\mathbf{X}}, \tau^{\mathbf{Y}})$:
  - Obtain hierarchical features $\mathcal{F}_{\tau^{\mathbf{Y}}}$ as $\mathcal{F}_{\mathbf{M}^{\mathbf{Y}}} \circ \mathcal{F}_{\rho^{\mathbf{Y}-}}$ & $\mathcal{F}_{\tau^{\mathbf{X}}}$ as $\mathcal{F}_{\mathbf{M}^{\mathbf{X}}} \circ \mathcal{F}_{\rho^{\mathbf{X}-}}$
  - Learn the contextual residuals $\phi(\mathcal{F}_{\tau^{\mathbf{X}}}, \mathcal{F}_{\tau^{\mathbf{Y}}})$ and $\phi(\mathcal{F}_{\tau^{\mathbf{Y}}}, \mathcal{F}_{\tau^{\mathbf{X}}})$
  - Apply differential SVD on $S = \mathrm{SoftMax}([\mathcal{F}_{\tau^{\mathbf{X}}} + \phi(\mathcal{F}_{\tau^{\mathbf{X}}}, \mathcal{F}_{\tau^{\mathbf{Y}}})]\,[\mathcal{F}_{\tau^{\mathbf{Y}}} + \phi(\mathcal{F}_{\tau^{\mathbf{Y}}}, \mathcal{F}_{\tau^{\mathbf{X}}})]^{\mathrm{T}})$

Quantitative Results

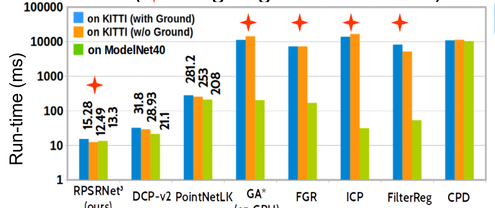
- Evaluation on **KITTI LiDAR seq. 00 to 07**. Each cell (gray →with ground points) in the table denotes the RMSE on angular and transnational deviations from GT.

Seq.	CPD [7] φ_{rmse}, Δt_{rmse}	GA* [6] φ_{rmse}, Δt_{rmse}	FGR [13] φ_{rmse}, Δt_{rmse}	ICP [3] φ_{rmse}, Δt_{rmse}	FilterReg [4] φ_{rmse}, Δt_{rmse}	DCP-v2 [10] φ_{rmse}, Δt_{rmse}	PointNetLK [1] φ_{rmse}, Δt_{rmse}	RPSRNet[1] (ours) φ_{rmse}, Δt_{rmse}	RPSRNet[3] (o φ_{rmse}, Δt_{rm}
00	4.99, 1.12	4.82, 1.09	4.77, 0.82	4.72, **0.80**	4.77, 0.80	4.87, 1.07	5.44, 1.27	4.48, 1.01	**3.30**, 0.81
	4.91, 1.39	*4.78, 0.93*	*4.78, 0.83*	*4.70, 0.84*	*4.92, 0.88*	*4.78, 0.80*	*6.11, 1.27*	*4.56, **0.73***	***3.31**, 1.0*
01	3.18, 1.54	2.99, 1.74	3.06, 1.10	3.06, 2.27	3.02, 0.89	3.0, 0.95	4.50, 1.33	2.79, 1.28	**2.13, 0.48**
	3.15, 1.88	*3.06, 2.42*	*2.88, 2.42*	*2.83, 2.41*	*2.77, 2.41*	***1.80**, **0.68***	*5.22, 1.18*	*3.04, 1.38*	*2.0, 1.03*
02	3.87, 1.28	3.67, 0.98	3.71, 1.02	3.63, 1.0	3.42, 0.69	3.52, 0.76	4.21, 1.0	3.69, 0.73	**2.66, 0.6**
	3.0, 1.11	*3.71, 1.09*	*3.66, 0.9*	*3.66, 0.91*	*3.81, 1.1*	*3.52, **0.53***	*5.77, 0.96*	*3.73, 1.10*	***2.88**, 1.11*
03	0.38, 0.88	0.34, 0.72	0.31, 0.59	0.14, 0.56	0.14, 0.49	0.24, 0.66	0.90, 0.81	0.14, 0.72	**0.10, 0.41**
	0.23, 0.58	*0.18, 0.43*	*0.20, 0.74*	*0.13, 0.91*	*0.35, 0.78*	*0.18, **0.3395***	*2.1, 1.01*	*0.16, 0.85*	***0.08**, 0.68*
04	2.58, 1.09	2.64, 1.12	2.65, 1.06	2.64, 1.07	1.97, 0.89	2.20, 1.37	3.88, 1.31	2.74, 0.55	**1.11, 0.50**
	2.77, 0.91	*2.64, 0.85*	*2.64, 1.09*	*2.65, 1.11*	*2.11, 1.18*	*2.07, 1.01*	*4.86, 1.27*	*2.28, 0.38*	***1.19**, **0.21***
05	3.81, 0.79	3.41, 0.7135	3.29, **0.4142**	2.95, 0.75	3.16, 0.62	2.01, 0.42	4.09, 0.79	3.09, 0.87	**1.91, 0.71**
	3.0, 0.69	*3.30, 0.64*	*3.27, 0.80*	*2.8, 1.12*	*3.89, 0.99*	*3.15, **0.63***	*4.2, 1.12*	*3.35, 1.06*	***1.11**, 0.91*
06	4.67, 0.96	4.13, 0.85	4.04, 0.87	3.64, 1.20	4.04, 0.88	3.02, 0.69	3.08, 0.99	4.01, 0.64	**3.0, 0.50**
	2.85, 1.28	***1.55**, 1.02*	*4.13, 1.21*	*3.26, 1.32*	*4.03, 1.29*	*3.74, **0.48***	*4.87, 0.97*	*2.64, 0.48*	*3.94, 0.69*
07	4.89, 1.0	4.39, 0.78	4.46, 0.91	4.41, 1.03	4.11, 0.99	4.48, 1.22	6.04, 1.44	4.06, 0.81	**3.58, 0.61**
	*4.31, **0.71***	*4.34, 0.77*	*4.46, 0.88*	*4.37, 0.95*	*4.22, 0.99*	*4.45, 1.58*	*8.21, 1.81*	*4.45, 1.08*	***3.11**, 1.07*

- Evaluation on **ModelNet40** data, Transformation errors averaged over validation set

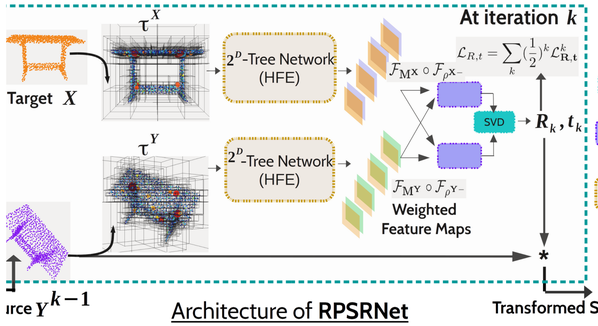


- Run-time (✦ using original Point Sizes)

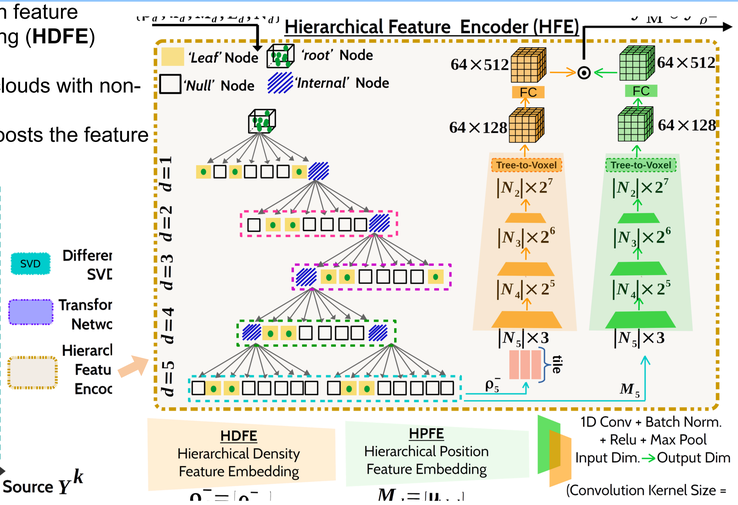


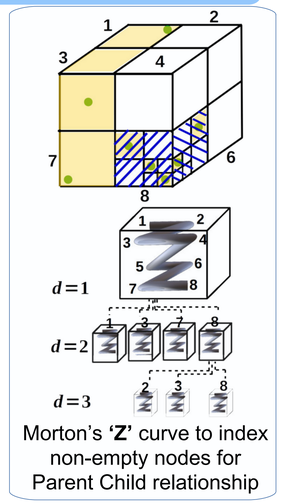
RPSRNet Architecture & Hierarchical Feature Embedding

- The HFE module of RPSRNet uses -- 1. hierarchical position feature
- embedding (**HPFE**) 2. hierarchical density feature embedding (**HDFE**)
- Apply late-fusion between **HPFE** and **HDFE**
- learned-features become homogeneous for the input point clouds with non-
- uniform point sampling densities (e.g., LiDAR scans)
- iterative transformation refinement module of our network boosts the feature
- matching accuracy in the intermediate stages



Architecture of **RPSRNet**





Morton's '**Z**' curve to index non-empty nodes for Parent Child relationship

More Details

Acknowledgment
This work was partially supported by the project **VIDETE (01IW18002)** of the German Federal Ministry of Education and Research (BMBF).