Nesti-Net: Normal Estimation for Unstructured 3D Point Clouds using Convolutional Neural Networks

Technion Israel Institute of Technology
CAD Lab, Faculty of Mechanical Engineering
ISL Lab, Faculty of Computer Science

Yizhak Ben-Shabat, Michael Lindenbaum, Anath Fischer

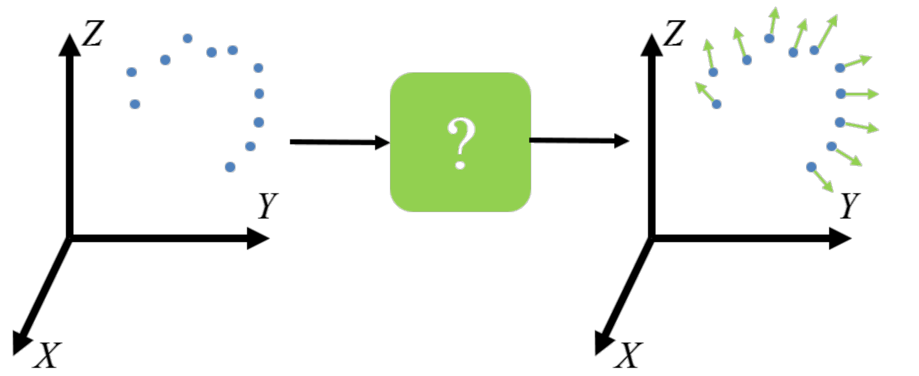
CVPR JUNE 16-21, 2019 LONG BEACH • CALIFORNIA

Motivation

- 3D point clouds represent the geometry of scanned objects and environments.
- Normal estimation is a fundamental problem in shape analysis.
- The normal vector provides local geometric information and can improve performance of different tasks e.g. segmentation, classification, and surface reconstruction.

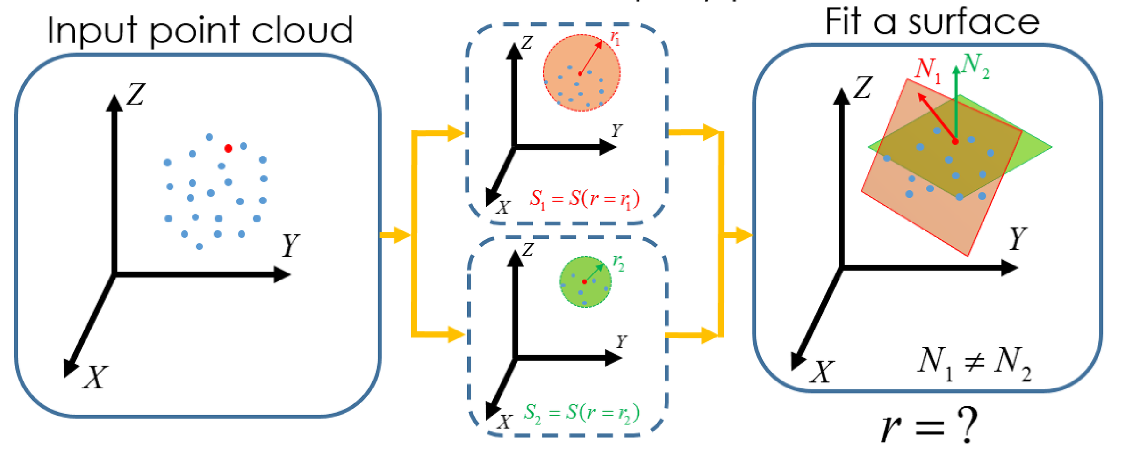
Goal

For each point, estimate the normal vector of the underlying surface



Challenges

- Point clouds are hard to process using CNNs (**unstructured**).
- Point clouds are **Noisy**
- Point clouds suffer from **density variations**
- Unknown scale (size of suitable neighborhood)



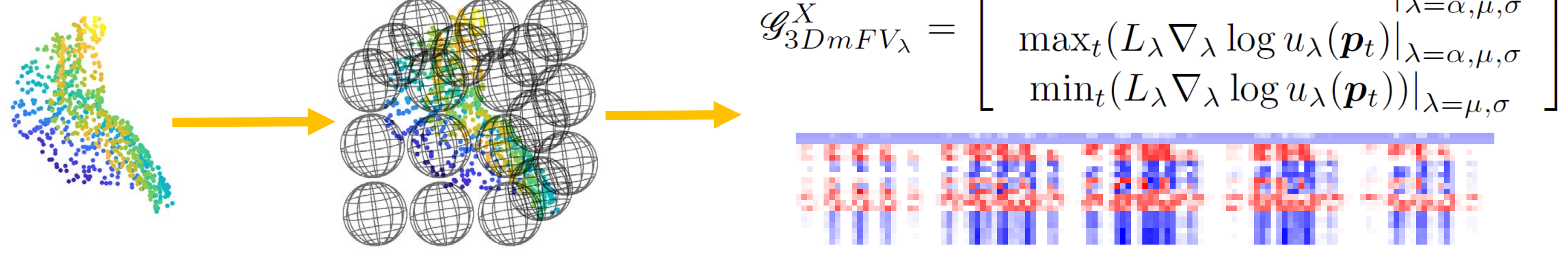
Can we learn the best radius ?

Previous work – 3D modified Fisher Vectors (3DmFV) [1]

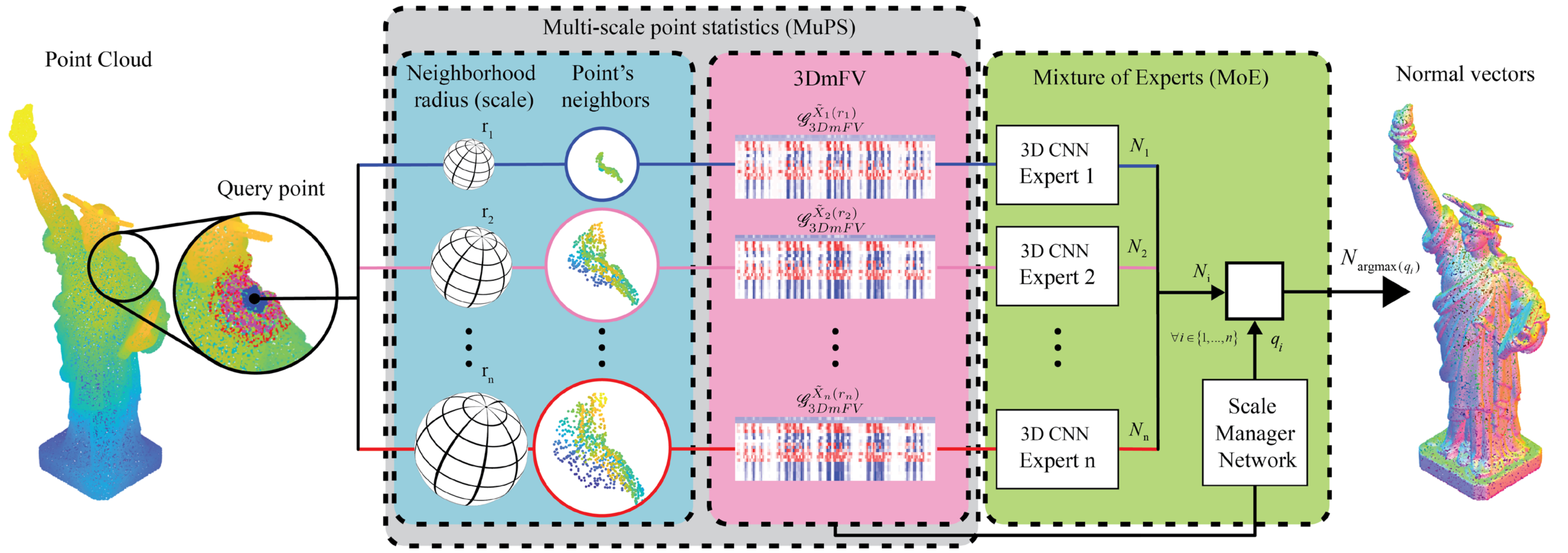
- Specify a special GMM: uniform size and Gaussians center on a grid.
- Compute the derivatives of each point w.r.t model parameters.
- **Advantages**: Asymptotically optimal [2], good classification performance [1,3], structured, size and order invariant.

$$u_\lambda(\mathbf{p}) = \sum_{k=1}^{K} w_k u_k(\mathbf{p})$$

$$\mathcal{G}^X_{3DmFV_\lambda} = \begin{bmatrix} \sum_{t=1}^{T} L_\lambda \nabla_\lambda \log u_\lambda(\mathbf{p}_t)\big|_{\lambda=\alpha,\mu,\sigma} \\ \max_t(L_\lambda \nabla_\lambda \log u_\lambda(\mathbf{p}_t))\big|_{\lambda=\alpha,\mu,\sigma} \\ \min_t(L_\lambda \nabla_\lambda \log u_\lambda(\mathbf{p}_t))\big|_{\lambda=\mu,\sigma} \end{bmatrix}$$



Approach



The Multi-scale Point Statistics (MUPS) representation

- Extract local subsets of neighboring points using different radii (scales).
- Normalize the subset to fit in a unit sphere.
- Compute the 3D modified Fisher Vector (3DmFV) representation for each scale.
- Concatenate

$$\mathcal{G}^{\mathbf{p}}_{MuPS} = \left(\mathcal{G}^{\tilde{X}_1(r_1)}_{3DmFV}, \ldots, \mathcal{G}^{\tilde{X}_n(r_n)}_{3DmFV}\right)$$

The Mixture of Experts (MOE) network for scale prediction

- Expert networks [4] receive different scales as input and estimates a normal vector.
- A scale manager network receives the MUPS as input and estimates which expert is best.
- The final normal vector estimation is chosen according to the managers prediction.
- The training loss is given by

$$L = \sum_{i=1}^{n} q_i \cdot D_N = \sum_{i=1}^{n} q_i \frac{\|N_i \times N_{GT}\|}{\|N_i\| \cdot \|N_{GT}\|}$$

Quantitative Experimental Results

		Ours	PCPnet	Jet	Jet ms	PCA	PCA ms	HoughCNN	Improve(%)
PGP5	(a)	**0.39**	0.24	0.29	0.31	0.28	0.30	0.33	**19.6**
	(b)	**0.77**	0.65	0.63	0.74	0.62	0.73	0.60	**3.05**

Table 1. Normal estimation PGP5 performance on PCPnet dataset for subsets containing (a) sharp features (b) smooth surfaces.

	Our Nesti-Net	PCA			Jet			PCPNet		HoughCNN	
scale	ms	small	med	large	small	med	large	ss	ms	ss	ms
Clean	**6.99**	8.31	12.29	16.77	7.60	12.35	17.35	9.68	9.62	10.23	10.02
Augmentations	**12.41**	21.97	16.25	18.87	21.95	16.29	19.02	14.56	14.34	16.9	17.37

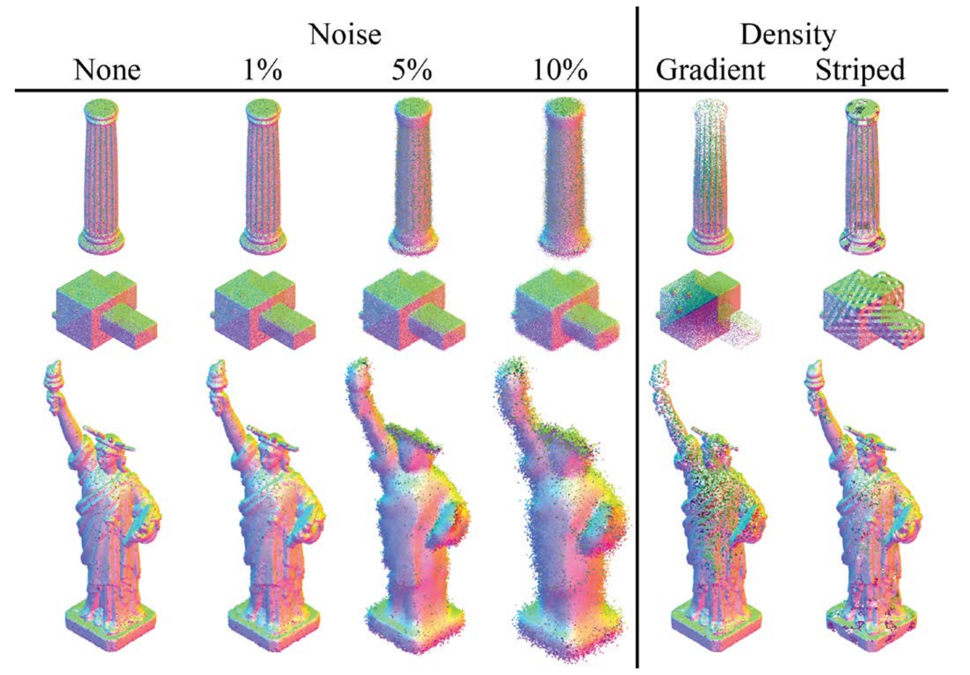
Table 2. Normal estimation RMS results on PCPnet dataset compared to state of the art for synthetic (clean) and augmented (Gaussian noise and density variation) point clouds

	ss		ms	ms-sw	NestiNet	
scale	0.01	0.05	0.01 0.05	0.01 0.05	0.01 0.05	0.01 0.03 0.05
Clean	9.32	12.73	10.83	7.88	7.76	**6.99**
Augmentations	23.91	16.14	16.11	13.55	12.97	**12.41**

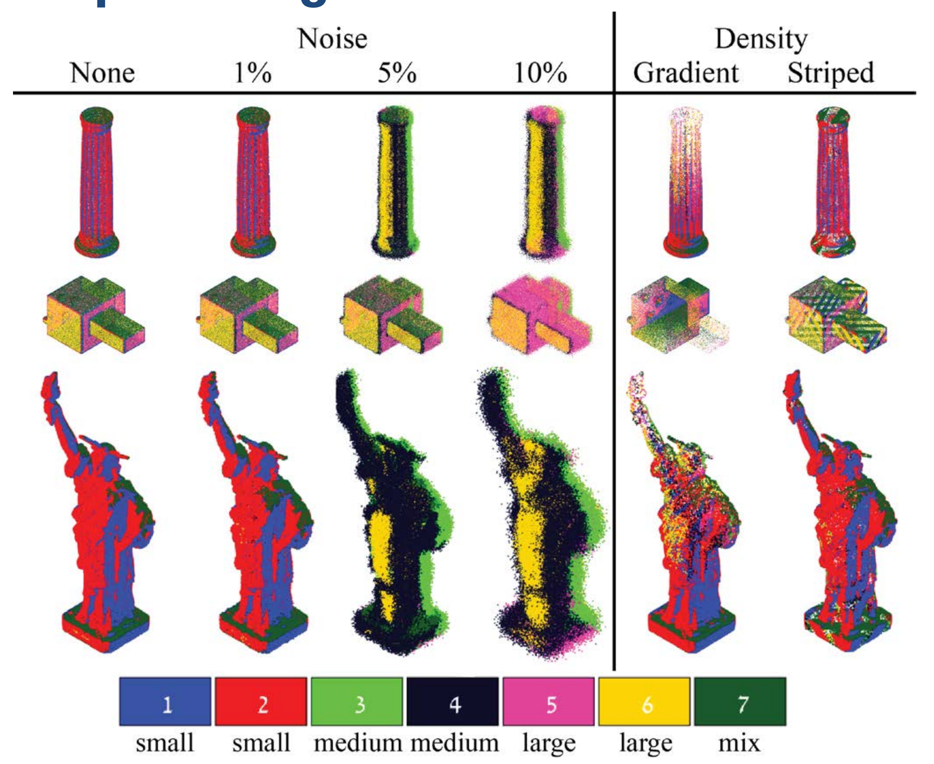
Table 3. Normal estimation RMS results on PCPnet dataset for Nesti-Net ablations: single-scale (ss), multi-scale (ms), multi-scale with switching (ms-sw) and multi-scale with mixture of experts (Nesti-Net)

Qualitative Experimental Results

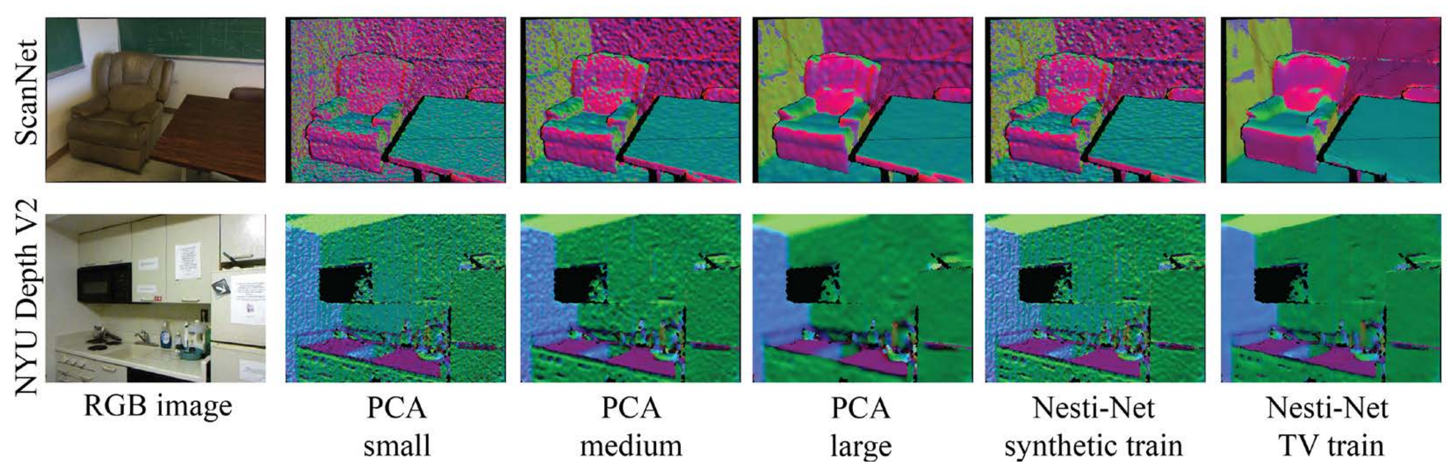
Normal vector visualization



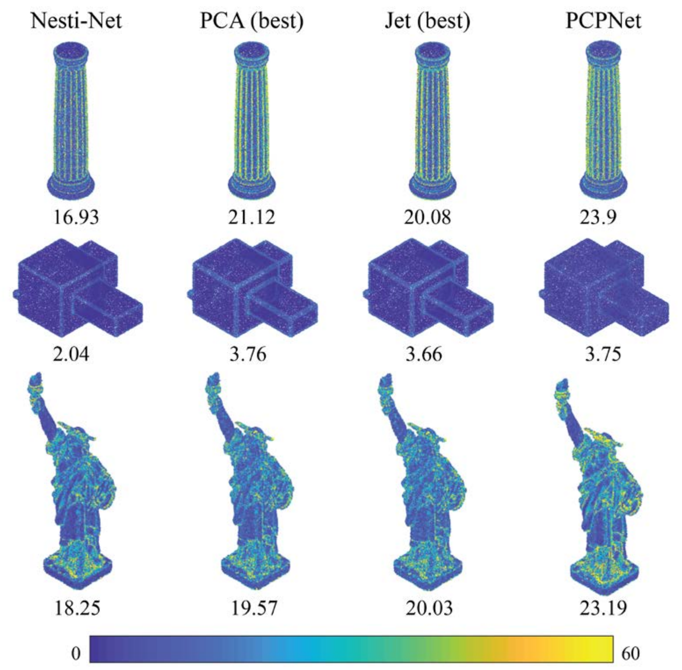
Expert assignment visualization



Normal vector visualization for scanned data



Error visualization



Contribution

- A new state-of-of-the-art solution for normal estimation in unstructured 3D point clouds.
- A local point representation which can be used as input to a CNN (MUPS).
- Unsupervised scale selection using mixture of experts (MOE).

Project website

References

[1] Y. Ben-Shabat, M. Lindenbaum, and A. Fischer. 3DmFV: Three-dimensional point cloud classification in real-time using convolutional neural networks. IROS and IEEE RA-L, 3(4):3145-3152, 2018.
[2] T. Jaakkola and D. Haussler. Exploiting generative models in discriminative classifiers. In Advances in Neural Information Processing Systems, pages 487–493, 1999.
[3] J. Sanchez, F. Perronnin, T. Mensink, and J. Verbeek. Image classification with the Fisher vector: Theory and practice. IJCV, 105(3):222-245, 2013.
[4] R. A. Jacobs, M. I. Jordan, S. J. Nowlan, and G. E. Hinton. Adaptive mixtures of local experts. Neural Computation, 3(1):79-87, 1991.