CPSC 533Y: Term Project Proposal

Guanxiong Chen
The University of British Columbia
2336 Main Mall, Vancouver, BC
gxchen@cs.ubc.ca

Abstract

I plan to improve Luo's method for diffusion probabilistic model (DPM)-based 3D point cloud generation. Previous work has shown that DPM-based models suffer from unsatisfactory log-likelihood on 2D image generation tasks, and whether the same problem exists for 3D point cloud generation is still unclear. I aim to answer this question; moreover, I would augment an existing method that boosts log-likelihood on 2D image generation with Luo's method, and prove that the augmentation is effective.

1. Motivation and Goals

1.1. Motivation

Significance. 3D point cloud generation is an activate area of computer vision research with a multitude of applications in robotics, from autonomous driving [10] to seafloor investigation [2].

Related work. In the past few years, researchers have proposed many methods for this task. As pointed out in [8], one approach to generation demonstrated in early methods such as [1,3] is to treat point clouds as an $N \times 3$ matrix, then apply generative models such as Variational Autoencoders (VAE) [6] or Generative Adversial Networks (GAN) [4] to generate samples. Another approach is to treat point clouds as samples from a probablistic distribution [7, 8, 12].

Luo and Hu [8] took inspiration from the probabilistic distribution approach and the diffusion model in thermodynamics [11], modelling point clouds as particles in space that diffuse from an orderly shape to a Gaussian-noise distribution due to contact with a heat source. Then the authors modelled point cloud generation as a reversed diffusion process, in which particles move from the noise distribution back to their original places in the shape. The method has achieved generation quality on-par with that of existing methods such as [1] and [12] when evaluated with metrics including minimum matching distance (MMD) and 1-NN Classifier Accuracy (1-NNA) [8].

The problem. The diffusion probabilistic model (DPM) itself however has not been well-adapted to learning tasks: in [5], the authors pointed out that for 2D image generation tasks, DPM could not achieve log-likelihood as competitive as other likelihood-based models. Also, in [8] the authors did not discuss whether DPM has achieved better log-likelihood for point cloud generation. Since log-likelihood is an important metric for generative models, [9], we need to ensure that DPM is competitive in this regard for point cloud generation.

Nichol and Dhariwal [9] improved the original DPM to achieve competitive log-likelihood by changing the variance of the reverse diffusion kernel from a fixed hyperparameter to a parameter that is learnable over a small range. They demonstrated that with learned variance, the resulting negative log-likelihood (NLL) became much lower for 2D image generation.

1.2. Goals

Under the assumption that unsatisfactory log-likelihood would also be a problem for 3D point cloud generation, I would like to augment Nichol and Dhariwal's work with Luo's method, and show that:

1. Nichol and Dhariwal's method can also improve the log-likelihood of DPM for 3D point cloud generation significantly;
2. DPM can achieve log-likelihood competitive against other state-of-the-art likelihood methods for point cloud generation.

Specifically, I plan to modify Luo's training set-up so variance in the reverse diffusion kernel is also a trainable parameter.

2. Expected Outcome

The expected outcome is to demonstrate the two contributions listed in 1.2 by showing that the NLL yielded from Luo's method on point cloud generation is lower than previous methods, and is much lower than before switching the variance to a trainable parameter.

References

[1] Panos Achlioptas, Olga Diamanti, Ioannis Mitliagkas, and Leonidas Guibas. Learning representations and generative models for 3d point clouds. In *International conference on machine learning*, pages 40–49. PMLR, 2018. 1

[2] Hyeonwoo Cho, Byeongjin Kim, and Son-Cheol Yu. Auv-based underwater 3-d point cloud generation using acoustic lens-based multibeam sonar. *IEEE Journal of Oceanic Engineering*, 43(4):856–872, 2017. 1

[3] Matheus Gadelha, Rui Wang, and Subhransu Maji. Multiresolution tree networks for 3d point cloud processing. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 103–118, 2018. 1

[4] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in neural information processing systems*, 27, 2014. 1

[5] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *arXiv preprint arXiv:2006.11239*, 2020. 1

[6] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013. 1

[7] Roman Klokov, Edmond Boyer, and Jakob Verbeek. Discrete point flow networks for efficient point cloud generation. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXIII 16*, pages 694–710. Springer, 2020. 1

[8] Shitong Luo and Wei Hu. Diffusion probabilistic models for 3d point cloud generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2837–2845, 2021. 1

[9] Alex Nichol and Prafulla Dhariwal. Improved denoising diffusion probabilistic models. *arXiv preprint arXiv:2102.09672*, 2021. 1

[10] Kun Qian, Zhaoyuan He, and Xinyu Zhang. 3d point cloud generation with millimeter-wave radar. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(4):1–23, 2020. 1

[11] Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In *International Conference on Machine Learning*, pages 2256–2265. PMLR, 2015. 1

[12] Guandao Yang, Xun Huang, Zekun Hao, Ming-Yu Liu, Serge Belongie, and Bharath Hariharan. Pointflow: 3d point cloud generation with continuous normalizing flows. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 4541–4550, 2019. 1