EE367 Computational Imaging - Project Proposal
3D Gaussian Splatting for Volume Reconstruction in Cryo-EM

Nikolaus Dräger (06765586)

Winter 2024

1 Motivation

In the field of cryogenic electron microscopy (cryo-EM), the reconstruction of a three-dimensional model from two-dimensional images represents a pivotal step in molecular structure analysis. This project proposes an innovative approach by employing 3D Gaussian Splatting [1], a technique primarily explored in volumetric rendering, to improve the cryo-EM reconstruction process. Current methodologies in cryo-EM, while effective, face limitations in terms of reconstruction accuracy and computational efficiency. Through the application of 3D Gaussian Splatting, we aim to address these challenges by potentially improving the speed and precision of the reconstruction process.

2 Related Work

The field of cryo-electron microscopy (cryo-EM) has seen remarkable advancements thanks to the application of machine learning in 3D reconstruction. A notable development is cryoDRGN by Zhong et al. [2, 3], which utilizes Variational Autoencoders (VAEs) [4] for improved molecular structure reconstruction. In a related attempt, Zhong et al. [5] also developed Cryofold, which employed Gaussian-based methods for cryo-EM reconstruction. While promising, Cryofold faced challenges, particularly with experimental validation and limitations in modeling certain types of heterogeneity, underscoring the need for further innovation in this space.

Alongside this, Mildenhall et al. [6] introduced Neural Radiance Fields (NeRFs), a deep fully-connected neural network approach for 3D scene reconstruction. NeRFs have shown impressive results in image quality, but they fall short in terms of training and inference speeds. NeRFs have shown impressive results in image quality, but they fall short in terms of training and inference speeds. This gap has led to innovative approaches like Plenoxels by Fridovich-Keil et al. [7] and, more relevant to this project, 3D Gaussian Splatting by Kerbl et al. [1]. Both methods abandoned the neural network in favor of different methods for representing the latent space. The latter method leverages anisotropic Gaussians and stands out for its significantly faster inference and training times, making it a strong candidate for cryo-EM applications.

3 Methodology

We define $V : \mathbb{R}^3 \to \mathbb{R}$ as the electrostatic potential of a protein or molecule. I_i are two-dimensional images and thus represent projections of the potential onto a 2D plane. C_i is the point spread function, the Fourier transform of which we call the contrast transfer function (CTF). The image formation model is then given as:

$$I_i(x, y) = C_i * \int_t V(\Phi_i \begin{pmatrix} x \\ y \\ t \end{pmatrix}) \, dt + \eta_i, \tag{1}$$

where $\Phi_i \in SO(3)$ is a rotation of the coordinates $(x, y, t)^T$ and η_i is a noise term [8].

In the Gaussian mixture model used for 3D Gaussian splatting, the potential at a coordinate $\boldsymbol{r} = (x, y, z)^T$ is written as the sum of normal distributions

$$V(\boldsymbol{r}) = \sum_k \mathcal{N}(\boldsymbol{r}, \boldsymbol{\mu}_k, \Sigma_k). \tag{2}$$

μ_k defines the mean of a distribution and Σ_k the covariance matrix. We introduce $\theta = \{\mu_k, \Sigma_k\}$ as the set of means and covariance matrices defining our potential such that $V \in \{V_\theta, \theta \in \Theta\}$.

The problem then becomes a non-convex optimization problem in which we solve for

$$\arg\min_{\theta} \|I_i - \Gamma(V_\theta, C_i, \boldsymbol{\phi}_i)\|_2^2, \tag{3}$$

where I_i is an image from the training dataset and Γ is renders a 2D image for a given potential V_θ, together with the CTF C_i and a set of coordinates $\boldsymbol{\phi}_i$. Intuitively, we want to align the rendered or reconstructed images as closely as possible with the actual images from the training dataset. The CTF will for the first stage of the project be largely ignored to reduce the complexity of the task.

4 Milestones

1. **Projection:** The first step is the implementation of the Γ function for rendering or projecting 2D images for a given potential V_ϕ and for given coordinates $\boldsymbol{\phi}_i$.

2. **Potential:** The next step is the construction of the potential from a set of means and covariance matrices $\theta : \theta \to V_\theta$.

3. **Sanity Check:** We then create a synthetic dataset $\{I_i = \Gamma(V_{\theta^*}, \boldsymbol{\phi}_i)\}$ from a given potential V_θ. To check whether or not the algorithm works, we use Eq. 3 to compute θ and compare $\theta = \theta^*$. For simplicity, this project will not include any work with real cryo-EM data but will rely only on synthetic data for the proof of concept.

4. **Evaluation:** As a final step, we want to know how well the algorithm performs. To do this, we initialize $\theta_0 = \theta^* + \varepsilon$, where ε is a noise term. We then, again, run the minimization problem from Eq. 3 and check whether $\theta = \theta^*$.

Miscellaneous Tools for data visualization will likely be needed and implemented throughout the project.

Future Work To preserve a reasonable scope for this project, we will only consider scenarios where $\boldsymbol{\phi}_i$ is known. This is called a *refinement task*. In the future, it is desirable to lift this constraint and also solve the problem for unknown $\boldsymbol{\phi}_i$ which is then called *ab initio* reconstruction. To support this, one can utilize different strategies for regularization. Physics-based regularizers leverage knowledge about the underlying problem. For example, atoms in a molecule cannot be too close or too far apart. In contrast, a data-driven approach could use a denoiser for regularization.

References

[1] Bernhard Kerbl et al. "3D Gaussian Splatting for Real-Time Radiance Field Rendering". In: *ACM Transactions on Graphics* 42.4 (2023).

[2] Ellen D Zhong et al. "CryoDRGN: reconstruction of heterogeneous cryo-EM structures using neural networks". In: *Nature methods* 18.2 (2021), pp. 176–185.

[3] Ellen D Zhong et al. "Reconstructing continuous distributions of 3D protein structure from cryo-EM images". In: *arXiv preprint arXiv:1909.05215* (2019).

[4] Diederik P Kingma and Max Welling. "Auto-encoding variational bayes". In: *arXiv preprint arXiv:1312.6114* (2013).

[5] Ellen D Zhong et al. "Exploring generative atomic models in cryo-EM reconstruction". In: *arXiv preprint arXiv:2107.01331* (2021).

[6] Ben Mildenhall et al. "Nerf: Representing scenes as neural radiance fields for view synthesis". In: *Communications of the ACM* 65.1 (2021), pp. 99–106.

[7] Sara Fridovich-Keil et al. "Plenoxels: Radiance fields without neural networks". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022, pp. 5501–5510.

[8] Amit Singer. "Mathematics for cryo-electron microscopy". In: *Proceedings of the International Congress of Mathematicians: Rio de Janeiro 2018*. World Scientific. 2018, pp. 3995–4014.