Zernikes3D: A new tool to study molecular motions through CryoEM particles

D. Herreros [a], R.R Lederman [b], J. Krieger [c], A. Jiménez-Moreno [a], M.Martínez [a], D. Myska [d], D. Strelak [a,e], J. Filipovic [d], I. Bahar [c], C.O.S. Sorzano [a], J.M. Carazo [a]

[a] Centro Nacional de Biotecnologia-CSIC, C/ Darwin 3, Cantoblanco, Madrid 28049, Spain, [b] Department of Statistics and Data Science, Yale University, New Haven, Connecticut, USA, [c] Department of Computational and Systems Biology, University of Pittsburgh, Pennsylvania, USA, [d] Institute of Computer Science, Masaryk University, Botanicka´ 68a, 60200 Brno, Czech Republic, and [e] Faculty of Informatics, Masaryk University, Botanicka´ 68a, 60200 Brno, Czech Republic.

Introduction

The study of continuous macromolecuar flexibility is pushing a paradigm shift in cryoEM, no longer restricted to the analysis of a few stable states. As a result, the topic is currently attracting much attention (CryoDrgn [1], GMM [2], Manifold Embedding [3]…)

In this work, we extend a new tool called Zernikes3D [4] (able to estimate deformation fields to reproduce molecular motions) to work with CryoEM particle images. The estimated per-particle 3D deformation fields allow obtaining a visual representation of the conformational landscape of a molecule. In addition, deformation fields can be applied to correct the motion blur artefacts in CryoEM maps

An additional application of the Zernikes3D is its capacity to help in understanding relationships among volumes or structural models, being able to work with the two types of data indistinctly and project them in the same reduced representation space

Methods

- Displacement finding problem:

$$\max_{g_L} \rho\left(I(r) - P_H\left(V\left(r + g_L(r)\right)\right)\right) + \lambda_1 \int \|g_L(r)\|^2$$

- Basis properties:
  - Closed under rotations. This property helps avoiding artifacts related to over deformations along projection directions

$$Ag_L(A^{-1}r) = \sum_n \sum_l \sum_m A \begin{pmatrix} \alpha^x_{l,n,m} \\ \alpha^y_{l,n,m} \\ \alpha^z_{l,n,m} \end{pmatrix} Z_{l,n,m}(A^{-1}r)$$

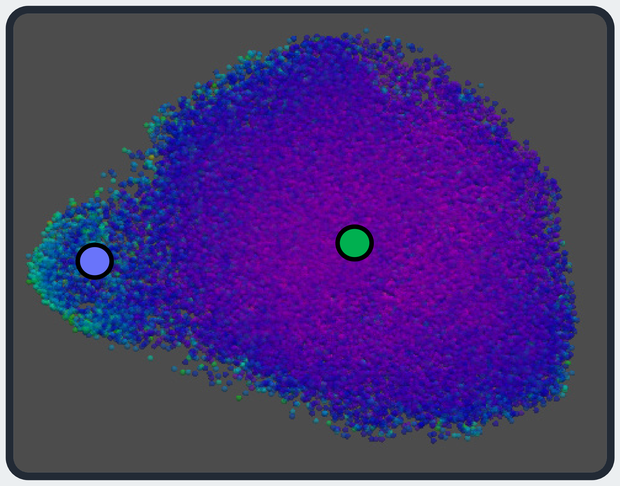
  - Coefficient $\alpha_{l,n,m}$ scaling to match different resolutions
  - Combined embeddings (volumes, particles, structures…)
- "Undoing" per–particle conformational changes (canonical volume) → Zernikes3D + ART (ZART):

$$V_R^{k+1} = V_R^k + \lambda(P_H(V(r + g_L(r)) - I(P_H(r)))$$

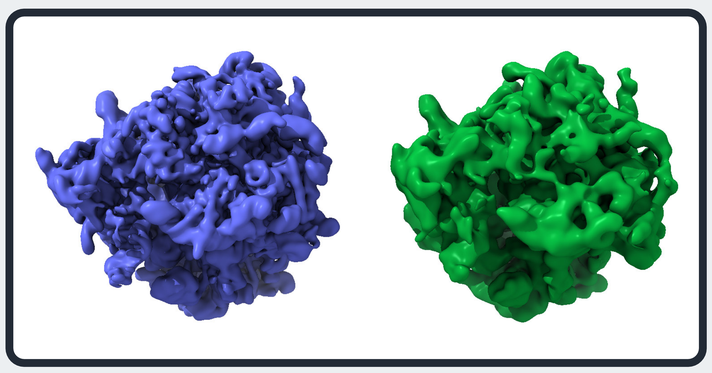
Results

We show results only with a data set due to space limitations, but current applications extend to other systems (SARS CoV2 spike and Her2). The system shown here has already been used in other publications, acting as a "test data set"

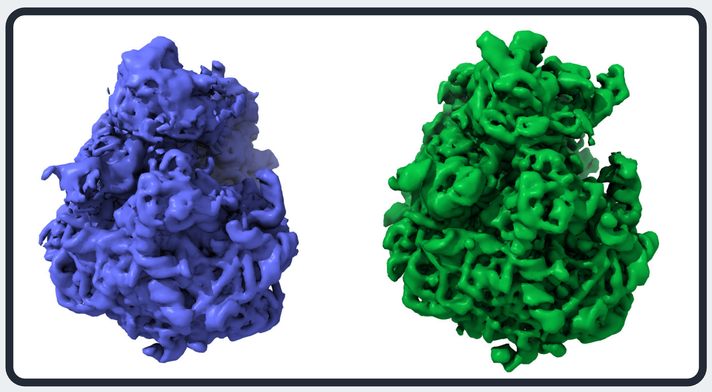
Zernikes3D coefficient space for the EMPIAR 10028 dataset (UMAP)
Two clear states are well differentiated (green and blue dots) representing the Pf80S rotation



Map visualization of the unrotated (blue) and rotated states (green) reconstructed from the particles selected from the Zernikes3D UMAP space



New ZART reconstruction method (green) compared to CryoSparc reconstruction (blue). Green map shows an improvement of both, moving and still areas of the ribosome



Conclusions

- The Zernikes3D basis can be effectively applied to study per-particle conformational changes by the estimation of 3D deformation fields
- For each particle, a set of Zernikes3D coefficients is estimated, which can be afterwards represented in a low-dimensional mapping showing the conformational landscape of a molecule
- Maps, structures, and particles can be simultaneously projected to the Zernikes3D space
- Projection artefacts can be completely avoided by rotating the coefficient space and cancelling the basis components associated with the projection direction
- The deformation fields can be applied to "undo" a particle conformation to reconstruct a motion-corrected map

References

1. E.D. Zhong, et al. CryoDRGN: reconstruction of heterogeneous cryo-EM structures using neural networks. **Nature Methods, 18: 176-185 (2021)**

2. M. Chend, S. Ludtke. Deep learning-based mixed-dimensional Gaussian mixture model for characterizing variability in cryo-EM. **Nature Methods, 18: 930-936 (2021)**

3. J. Frank, A. Ourmazd. Continuous changes in structure mapped by manifold embedding of single-particle data in cryo-EM. **Methods, 100: 61-67 (2016)**

4. D. Herreros, et al. Approximating deformation fields for the analysis of continuous heterogeneity of biological macromolecules by 3D Zernike polynomials. **IUCR J, 8: 992-1005 (2021)**