Automatic molecular tour creation: a study

V. Larroque[1], M. Maria[1], S. Mérillou[1], M. Montes[2]

[1]Laboratoire XLIM, UMR CNRS 7252, Université de Limoges, France
[2]GBCM - Laboratoire Génomique, Bioinformatique et Chimie Moléculaire, EA7528, Conservatoire National des Arts et Métiers, Hésam Université, France

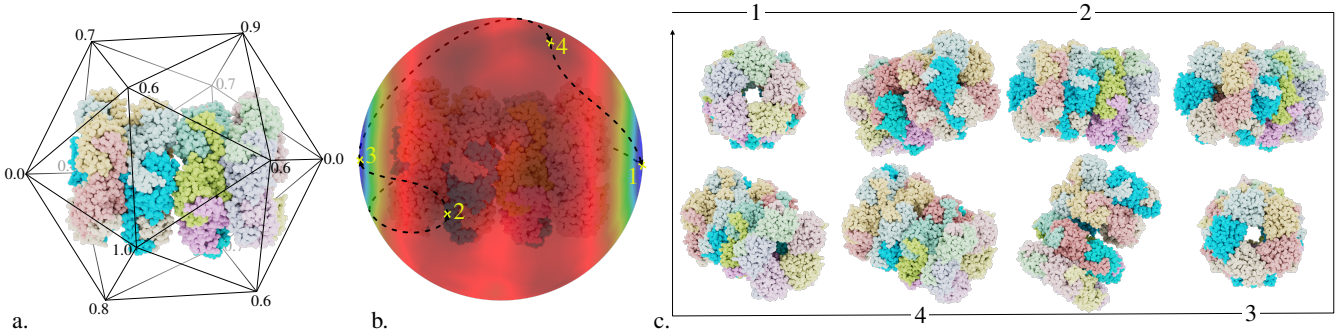


Figure 1: *Pipeline for creating molecular tours: (a) illustrates the icosahedron used to sample viewpoints; (b) shows the molecular tour's path; (c) displays the key (numbered) and intermediary viewpoints.*

Abstract

Molecular system visualization is a difficult task even for experts as molecules can contain millions of atoms. Our goal is to create a tool to improve the preliminary study of molecules by automatically creating a tour of the interesting viewpoints around them. Since we noticed limited research specific to molecular visualization, we analyzed and adapted methods from the general field. Our preliminary study shows that our molecular tour is able to smoothly present key information automatically.

CCS Concepts

• ***Computing methodologies*** → ***Perception;*** • ***Human-centered computing*** → ***Scientific visualization;***

1. Introduction

Computer simulations are essential for progress in drug design. For example, they help in the selection of potential drug candidates, according to their inhibitory effect on the target. Biochemists all over the world use molecular visualization software to support their research by allowing them to explore molecular systems with millions of atoms in real-time. Molecules can be represented in various ways, including bond-centric models, surface models, or "cartoon" representation. Each displays different information useful to scientists and should be considered when looking for interesting features. The state-of-the-art of Kozlíková *et al.* [KKF*17] provides a comprehensive overview of molecular visualization techniques.

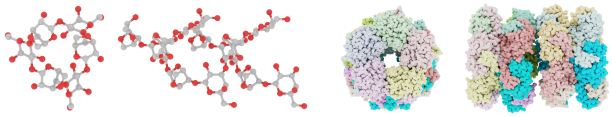
The geometric complexity of molecules makes exploration difficult, even for experts: we aim to provide a tool that will help the preliminary study of molecules by automatically creating a tour of the interesting viewpoints around them.

2. Related work

Viewpoint selection for molecular visualization has not received much attention: Vázquez *et al.* [VFSL06] proposed an information theory-based method to select viewpoints of molecules displayed using the "Ball and Stick" representation. Doulamis *et al.* [DCMP10] built on Vázquez's work by training a non-linear classifier using input from domain experts. Heinrich *et al.* [HVH*16] also extended Vázquez's method for the "cartoon" representation and provided a study with experts and non-experts on the definition of best viewpoints for molecular visualization. A good viewpoint, according to Vázquez and Heinrich, shows many atoms and bonds, exposing most of the molecule, or shows a geometrical configuration of the molecule (see figure 2 for examples).

Due to limited research specific to molecular visualization, our work consisted in studying general methods based on the surveys of Secord *et al.* [SLF*11] and Bonaventura *et al.* [BFS*18] for molecular systems.

DOI: 10.2312/egp.20231024



(a) *Alignment [VFSL06]* **(b)** *"Tunnel" [HVH*16]*

Figure 2: *Examples of two specific geometric configurations.*

3. Scoring and selection of viewpoints

The goal is to find the best viewpoints around a molecule: we first must select the candidate viewpoints around it. To do so, we use the recursive discretization of an icosahedron (figure 1.a), to select regularly spaced points on the bounding sphere of the molecule. The camera is placed on each of these points, looking at the molecule center, to compute an image. Several metrics will then use this image to compute a score based on various properties visible on it. Finally, this score allows us to choose the best viewpoints, which have either the lowest or highest scores. Metrics are defined as $VQ(v) = F_m(v)$ where $VQ(v)$ is the quality score of a viewpoint v computed with the $F_m(v)$ metric function.

4. Viewpoint sphere and molecular tour

To build the molecular tour, we run each metric on each selected viewpoint (figure 1.a). These vertices are then projected onto an image resulting in a heatmap of the scores that can be displayed on a sphere, thereby providing the viewpoint sphere (figure 1.b) and easy access to score data for any point on it. Researchers can visualize zones of interest using the viewpoint sphere itself, which are represented in blue (low score) and red (high score).

We filtered the extrema using the heatmap and the scores to find the highest and lowest values representing the best viewpoints. The molecular tour is built with the selected points by creating a path between them; the camera will orbit the molecule's bounding sphere, looking at its center, from one point to another. Figure 1.b depicts an example path while figure 1.c shows the viewpoints of each numbered point. The tour will alternate between low and high scores, giving a complete overview of the molecule and allowing them to quickly identify essential information without any input.

5. Results

We analyzed and adapted the 6 categories of metrics presented in [SLF*11] and [BFS*18] to molecular visualization. We conducted a preliminary study on 10 molecules to detect some geometric configurations (figure 2), and we made the following observations:

- Area metrics suffer from the fact that they only compute a score based on the visible area, ignoring critical information such as the molecule topology;
- Silhouette metrics use more information about the general topology to provide a score, but most of the internal topology information is lost when only the silhouette is used;
- Depth metrics is the most interesting category because depth conveys a molecule's complexity and topology, providing the best results for geometrical configuration research;
- Stability metrics are also interesting because they compare the similarity of a viewpoint with its neighbors which can be used to detect features such as the "tunnel" or an alignment;
- Surface curvature metrics may work in the general field, but there is no clear way to define what the curvature of a molecule is;
- Semantic metrics are currently designed to score features of real objects and are thus inapplicable to molecular visualization.

This study shows that the molecular tour can show both a viewpoint exposing most of the molecule and a geometrical configuration exposing the viewpoint for most metrics. The study also revealed that most metrics are unreliable. Figure 3 shows a plot of the precision of some to detect a "tunnel" on their highest (red) or lowest (blue) score, ranging from reliable (± 1) to unreliable (0).

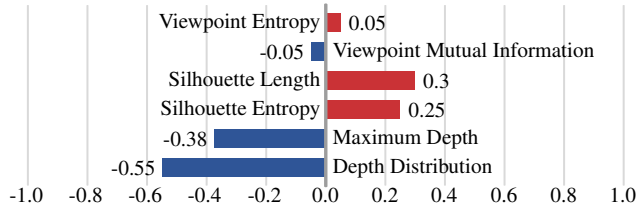


Figure 3: *Precision of some metrics from [SLF*11] and [BFS*18].*

6. Conclusion and future work

This preliminary study shows that our tool can generate tours of interesting viewpoints around molecules but cannot find specific geometrical configurations when requested. A larger study could provide more insight into our tool's true capabilities, but it is certain that new specific metrics for molecules could provide benefits.

Acknowledgments

This work is supported by the European Research Council Executive Agency under grant numbers 640283 and 101069190 and a fellowship from Qubit Pharmaceuticals.

References

[BFS*18] BONAVENTURA X., FEIXAS M., SBERT M., CHUANG L., WALLRAVEN C.: A Survey of Viewpoint Selection Methods for Polygonal Models. *Entropy 20*, 5 (2018), 370. 1, 2

[DCMP10] DOULAMIS N., CHRONIS E., MIAOULIS G., PLEMENOS D.: Personalized View Selection of 3D Molecular Proteins. In *Intelligent Computer Graphics*. Springer Berlin Heidelberg, 2010, pp. 211–227. 1

[HVH*16] HEINRICH J., VUONG J., HAMMANG C. J., WU A., RITTENBRUCH M., HOGAN J., BRERETON M., O'DONOGHUE S. I.: Evaluating Viewpoint Entropy for Ribbon Representation of Protein Structure. *Computer Graphics Forum 35*, 3 (June 2016), 181–190. 1, 2

[KKF*17] KOZLÍKOVÁ B., KRONE M., FALK M., LINDOW N., BAADEN M., BAUM D., VIOLA I., PARULEK J., HEGE H.-C.: Visualization of Biomolecular Structures: State of the Art Revisited. *Computer Graphics Forum 36*, 8 (2017), 178–204. 1

[SLF*11] SECORD A., LU J., FINKELSTEIN A., SINGH M., NEALEN A.: Perceptual Models of Viewpoint Preference. *ACM Trans. Graph. 30*, 5 (Oct. 2011), 1–12. 1, 2

[VFSL06] VÁZQUEZ P.-P., FEIXAS M., SBERT M., LLOBET A.: Real-time automatic selection of good molecular views. *Computers & Graphics 30*, 1 (2006), 98–110. 1, 2