

Vergleichende Visualisierung von Moleküloberflächen durch ähnlichkeitsbasiertes Clustering

(Comparative Visualization of Molecular Surfaces using Similarity-based Clustering)

Das Teilprojekt D.4 des Sonderforschungsbereichs 716 beschäftigt sich mit der Visualisierung von Biomolekülen, insbesondere Proteinen. Verschiedene Verfahren zur Visualisierung und Analyse von Proteinen sind in das am Institut entwickelte Visualisierungsprogramm MegaMol [1] integriert¹.

Eine für die Analyse der Proteindaten wichtige Darstellung sind Moleküloberflächen wie beispielsweise die in Abbildung 1 gezeigte *Solvent Excluded Surface* [2] (SES). Die SES ist die Kontaktfläche eines Proteins bezüglich eines bestimmten Lösungsmittelmoleküls, welches durch eine Kugel approximiert wird. Dadurch kann beispielsweise die Erreichbarkeit von Bindungsstellen am Protein analysiert werden. In einer vorangegangenen Arbeit wurde ein Verfahren entwickelt, welches die SES auf ein zweidimensionales Bild projiziert, also sozusagen eine Landkarte des Moleküls erstellt (*Molecular Surface Map* [3], siehe Abbildung 2). Dadurch können die Eigenschaften verschiedener Proteine einfach und schnell verglichen werden, da auf der 2D-Karte keine Verdeckung besteht.

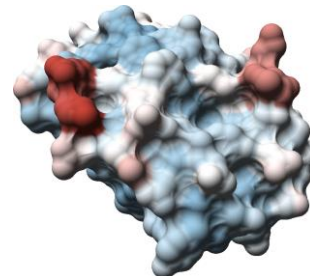


Abbildung 1: Solvent Excluded Surface eines Proteins.

Aufgabenbeschreibung:

Im Rahmen dieses Projektes sollen die vorhandenen Molecular Surface Maps für eine vergleichende Visualisierung verwendet werden. Die 2D-Bilder sollen zuerst nach Ähnlichkeit geclustert werden. Dazu soll ein geeignetes bildbasiertes Clusteringverfahren gewählt werden, denkbar wären beispielsweise das von Kolesar et al. [4] verwendete Verfahren, welche auf *Image Moments* basiert. Zur Überprüfung des bildbasierten Ähnlichkeitsvergleichs sollen geeignete Verfahren zum Vergleich von 3D Moleküldaten ausgewählt und verwendet werden (z.B. [5]–[8]). Die Ergebnisse der beiden ähnlichkeitsbasierten Clusterings sollen geeignet visualisiert werden. Denkbar wäre hier beispielsweise ein an die Graphen-Visualisierung angelehntes *Force Directed Layout*. Die Visualisierung soll es dem Benutzer ermöglichen, sowohl einen Überblick über den Datensatz zu bekommen (z.B. in dem für jedes Cluster nur ein repräsentativer Datensatz angezeigt wird) als auch alle Datensätze innerhalb eines ausgewählten Clusters zu betrachten (z.B. auf einem Raster angeordnet, siehe [9]). Für den Vergleich von großen Datensätzen mit sehr vielen Proteinen ist eine Darstellung auf einem großen, hochauflösenden Display wie der VISUS Powerwall von Vorteil. Die vergleichende Visualisierung soll deshalb nicht nur mit Maus und Tastatur, sondern auch mit dem am VISUS vorhandenen optischen Trackingsystem möglich sein.

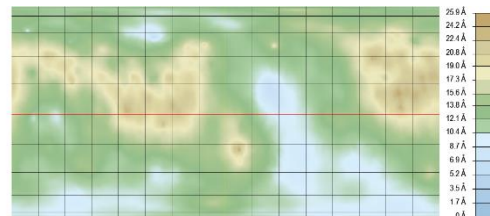
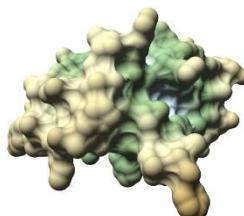


Abbildung 2: Solvent Excluded Surface eines Proteins (links) und die entsprechende Molecular Surface Map (rechts).

Voraussetzungen und Bedingungen:

Die vergleichende Visualisierung soll in MegaMol entwickelt werden, wodurch eine Verwendung auf der VISUS Powerwall ohne weitere Anpassungen möglich ist. Die Implementierung der Programmteile im MegaMol-Framework müssen in C/C++ erfolgen. Die Graphikprogrammierung erfolgt in OpenGL/GLSL. Module zur Visualisierung von 3D-Daten und 2D-Daten sind in MegaMol vorhanden und können im Rahmen des Projektes genutzt oder als Referenz verwendet werden. Für die Vorverarbeitung der Bilder und 3D-Daten sollen soweit möglich bestehende externe Programme und Bibliotheken eingesetzt werden.

Betreuer:

Dr. Michael Krone (michael.krone@visus.uni-stuttgart.de)
M.Sc. Florian Friß (florian.friess@visus.uni-stuttgart.de)

Prüfer:

Prof. Dr. Thomas Ertl (ertl@vis.uni-stuttgart.de)

Bearbeiter:

1. Shoma Kaiser
2. Enis Spahiu
3. Wilhelm Buchmüller
4. Damir Ravlija

Anfangsdatum:

01.12.2017

¹ <http://www.megamol.org/>

Literatur:

- [1] S. Grottel, M. Krone, C. Müller, G. Reina, and T. Ertl, "MegaMol - A Prototyping Framework for Particle-based Visualization," *IEEE Trans. Vis. Comput. Graph.*, vol. 21, no. 2, pp. 201–214, 2015.
- [2] M. F. Sanner, A. J. Olson, and J.-C. Spehner, "Reduced Surface: An Efficient Way to Compute Molecular Surfaces," *Biopolymers*, vol. 38, no. 3, pp. 305–320, 1996.
- [3] M. Krone *et al.*, "Molecular Surface Maps," *IEEE Trans. Vis. Comput. Graph.*, vol. 23, no. 1, pp. 701–710, 2017.
- [4] I. Kolesár, J. Byska, J. Parulek, H. Hauser, and B. Kozlíková, "Unfolding and Interactive Exploration of Protein Tunnels and their Dynamics," in *Eurographics Workshop on Visual Computing for Biology and Medicine*, 2016.
- [5] A. Via, F. Ferrè, B. Brannetti, and M. Helmer-Citterich, "Protein surface similarities: a survey of methods to describe and compare protein surfaces," *Cell. Mol. Life Sci. CMLS*, vol. 57, no. 13–14, pp. 1970–1977, 2000.
- [6] C. Hofbauer, H. Lohninger, and A. Aszódi, "SURFCOMP: A Novel Graph-Based Approach to Molecular Surface Comparison," *J. Chem. Inf. Comput. Sci.*, vol. 44, no. 3, pp. 837–847, 2004.
- [7] L. Sael, D. La, B. Li, R. Rustamov, and D. Kihara, "Rapid comparison of properties on protein surface," *Proteins Struct. Funct. Bioinforma.*, vol. 73, no. 1, pp. 1–10, 2008.
- [8] D. La *et al.*, "3D-SURFER: Software for High-throughput Protein Surface Comparison and Analysis," *Bioinformatics*, vol. 25, no. 21, pp. 2843–4, 2009.
- [9] C. Müller, M. Krone, K. Scharnowski, G. Reina, and T. Ertl, "On the Utility of Large High-Resolution Displays for Comparative Scientific Visualisation," in *International Symposium on Visual Information Communication and Interaction (VINCI)*, 2015, vol. 8, pp. 131–136.