Visualizing Protein Structures in Virtual Interactive Interface

Kamran Ahsan
Department of Computer Science, Federal Urdu University of Arts, Science, & Technology, Karachi, Pakistan
Email: kamran.ahsan@fuuast.edu.pk

Muhammad Shahzad
Department of Computer Science, Usman Institute of Technology, Karachi, Pakistan

***Abstract*—In biomedical data set, macromolecular structures such as DNA and proteins have huge multifaceted bio molecular units containing many atoms and residues in big strands of amino acids. From the few decades, different visualizations techniques have been introduced to display these data sets by using several software tools such as Kinemage. Every type of former representation ranging from primary structure to quaternary structure have certain inadequacy that limits the scientists to examining the protein interaction because of protein's size and complexity. Comparing these type different representations helps researchers in their investigation of relationships between structures and functions of protein. This paper studies the major needs for virtual intelligent environment and elaborates the high performance simulation service for bio molecular visualization.**

***Index Terms*—protein structures, computational biology, intelligent interface, virtual reality**

I. Introduction

One of most complex and large molecules that plays different vital roles in the body is Protein. It is consisting of one or more polypeptide chain of 20 different amino acids. Every protein in any living organism performs variety of functions. Some proteins involved in muscle contraction and movement, some facilitates in speeding the chemical reactions, some act as defenders against viruses, bacteria and other foreign intruders, some helps in the movement of molecule from one place to another around the body, some protein bridging the tissues, some work as storage for amino acids and some protein helps in coordination of bodily activities [1]. Protein functions depend on its presence in the chain of amino acids and which is formed from the nucleotide sequence of their genes. Since protein function directly depends to its structure and some molecules are highly complex, so it is difficult to predict features in a traditional 2D visualization. Therefore primary, secondary, tertiary and quaternary structures [2] of proteins are being used to reduce the complexities while scientific investigation (Fig. 1).

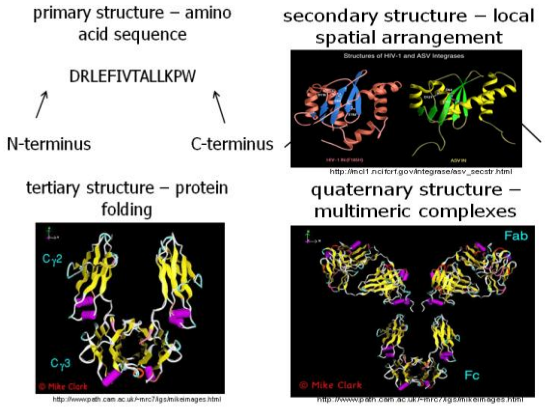

Figure 1. Computer generated protein structure

Fig. 1 is computer generated and can be developed using different software tools like Kinemage [3], which was the first software tool used for visual representation of the molecular structure. Kinemage is a viewing language. Scripting language is used to create 'kin' files. 3D features of structures can be focused easily using kin files. Associated labels or colors may be used to distinguish these structures. Associated labels can be rotated with the help o f a mouse cursor. Apart from visualizing macromolecular structures it may be used to view 3D data. Kin provides the facility to rotate the images in real time, cutting down the display image according to the use choice so that user can visualize the parts in which it is more interested. RasMol [4], developed by Roger Sayle, is a molecular graphics visualization tool, used for visualization of structure protein data sets. It provides pre-compiled binaries for different platform such as PC, SGI, Macintosh and x86 Linux. It is an open source, stand-alone program. MolScript [5] is a program which translates protein structure of Protein Data Bank format into geometric representation. In this PDB code can be used instead of complete file name. MolScript works on user entered file. User needs to specify the coordinate fie, points of object to be rendered and the rigorous look of the object. There are many other commercial software applications for the computer simulation of molecular structure.

Dynamic interaction with the protein and developing complex 3D structures is not possible using the tools

Manuscript received June 6, 2014; revised October 26, 2014.

doi: 10.12720/jomb.4.5.422-425

mentioned above. Virtual interactive interface and environment make possible the detailed inspection and comparison of the related molecular structure. They provide different quality as compared to standard 2D methods available in most computer generated rendering and text books. Immersive display technology, user centered projection, stereoscopic and multimodal user interfaces provide the facility of understanding of complex data sets.

In spite of the fact that the last decade has observed significance advances in the field of human computer interface, only a very small number of groups collected the respective novelties. In this paper we try to smooth the path for other researchers through presenting corresponding advances.

This paper has been structured as follows: In the upcoming section a general review of related work has been described. In section III we highlighted current research projects. Section IV comprises of a few useful recommendations and suggestions exhaustively. In the last section some conclusions and an outlook for future work is presented.

II. Related Work

Different objectives and approaches were adopted by research community for visualizing complex 3D protein structure in virtual interactive interface. Some aspects and visualization processes have attain extra ordinary progress. CAVE1, build in 1991, is known as the first interactive environment. It showed full scale (3m3). Research from different parts of the world constructed hundreds of CAVE and alternatives during the last two decades. The upcoming part of this section discusses the examples for efficacious work in different related areas.

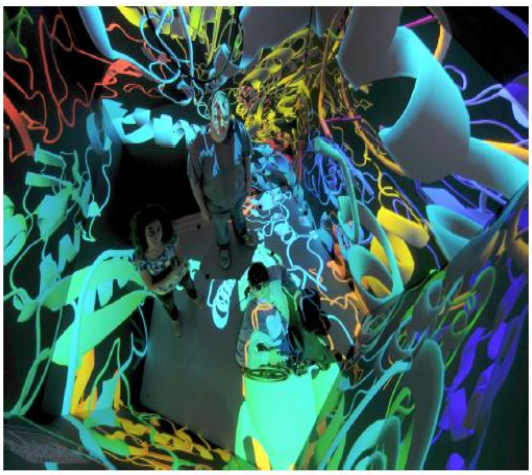

Figure 2. The StarCAVE from above, looking down on a RNA protein rendering. The still camera taking the picture is not being tracked so the perspective is skewed, but this image shows the floor as well as the walls and shows some of the effects of vignetting and abnormally severe off-axis viewing [7].

A. StarCAVE

StarCAVE is a projected virtual reality (Fig. 2) room consisting of 5 walls and one floor. Two softwares are supported to drive the StarCAVE. One is OpenCover, and the other is EVL's Electro. The StarCAVE can be run on any software which is capable of driving a multiscreen and multi node environment, upper and lower displays, and pentagonal arrangement of the walls. They used ROCKS-based operating system. To display protein in the virtual 3D environment, the user can download one of the 50000 protein structure from the PDB data bank server. The PDB files are converted into VRML (Virtual Reality Modeling Language) file format. Conversion to VRML format is done by using PyMOL software. Hence the converted file is used by visualization software. The application provides the s the facility to indicate between numerous visualization modes for example cartoon and surface. It can load multiple proteins at a time. Alignment of protein and display the corresponding amino acid sequence is also possible. A large number of protein structures [6] can be visualized simultaneously due to the high resolution and its surrounded display capacity. By using StarCAVE scientist can find similarities between proteins to reach the results which do not exist before [7].

B. Virtual Reality System for PyMOL:

PyMOL is a molecular visualization system on an open source foundation. Vertalis, an advanced visualization company, developed virtual reality (VR) software plug-in for PyMOL. In this system there are both screen-based VR tracking and ability to interact with the protein structure via tracked hand held device to PyMOL (Fig. 3). With this system, user can jumps into a 3D display of complex biochemical structures. Users have facility to touch, feel and see the molecules from different angle. It helps researchers and students to understand and examine the complex molecule structures from all sides and from within the molecule itself, which can help to identify the relationship between the structures and functions [8].

Figure 3. An innovative new ways interaction of users [8]

C. CAVEman:

This project is constructed in order to make visual map for the diseases which have genetic components such as Schizophrenia, Alzheimer and cancer. High-resolution digital atlas of a human body with medical data related to specific diseases is integrated through CAVE. CAVE is an immersive virtual reality environment. The software for this project is java based and delivers complete portability across visualization. This project creates an

atomically atlas of the human body and data mapping mechanism. Disease process and the effect of involvement of drugs on these processes can be visualized from the next generation 4D output of the software [9].

III. Critical Analysis

One of the major needs for virtual intelligent environment is high performance simulation center, which offers a global simulation service for biomolecular visualization. This kind of center should open to all industrial players and in particular, academics institution. One of the most popular resources, providing these services is offered by University of Groningen. They have visualization cluster of 8 high-end HP workstations (xw8600) in a master-slave setup (Fig. 4) which is the driving force behind the Reality Cube (Fig. 5) and the Reality Theatre (Fig. 6) [9]. For the more sophisticated system, it should be extended with respect to the multimodal human computer interface part, refinement and extension in the interaction techniques, and investigation for the contribution to a faster knowledge acquisition by performing user studies is also needed for future work.

Figure 4. The 8-node visualization cluster [9]

Figure 5. Understanding chemical structures using the reality cube [9]

Figure 6. Visualizing molecular structures in the reality theatre [9]

Application of immersive stereoscopic virtual environment in the field of integration between gene expression and metabolic networks is another challenging and one of the open research opportunities.

This is the new way to explore metabolism that definitely helps researchers for insight observation and investigation in their studies. One of related work has been done in [10] (Fig. 7), which uses MetNet3D, virtual reality system for the exploration of such kind of research. Animated network graph are used to explore gene function and the common behavior of genes in metabolic network that has been taken from MetNetDB. MetNetDB is a database of Arabidopsis metabolic networks. But still advancement is needed specially pathway layout problem. In order to keep user mental model compact, part of the graph is adopted and the rest of the graph is not taken. In result this creates an open research area in which someone can apply a complex animation method on the whole graph, which usually can be entertained using high performance computing.

(a) (b)

Figure 7. (a) Navigating through the metabolic network in the virtual environment (b) Interacting with the same application in a desktop [10].

IV. Discussion

Virtual Interactive environment makes the system more effective as users do not feel confusion while investigating the results during the experiment. Especially in case of biological information at molecular level where massive information is associated with every element.

The user can enjoy more representative and realistic feeling and more natural interaction through the use of immersive stereoscopic VR. Immersive stereoscopic VR can display protein sequence data around the user. By this method user feels that it is operating the data as operating real objects.

Stereoscopic VR have the prospective to assist and support people to comprehend the relationships between the protein structure and its function. Visualization and navigation in three dimensional spaces make it potential to explore huge amount of dataset than in two dimensional spaces. It also simplifies the process of knowledge perceiving and learning procedures.

High performance simulation center for offering global simulation services for industrial usage and academic research as discussed in previous section will be vital step

in the development of this emerging field. All available software for simulation should be there so that researchers can integrate with their hardware resources. The advantage of high performance computing in multimodal human interface which was lacking in the past research, will definitely increase usability: the weaknesses of one modality are balance by the strengths of another.

Another challenge for the developers of immersive stereoscopic virtual environment for exploring metabolism is to enabling complex animation support to whole displayed protein sequence at a same time, which is easily entertained with the help of high performance computing. No such virtual environment system exists till now fulfilling this kind of whole animation service. It would definitely help researchers for analyzing whole effect simultaneously.

V. Conclusion

In summary, we explore basics of computational biology which wrap the literature that is prerequisite for the deep understanding of virtual environment. Three most popular research works like StarCAVE, virtual Reality system for PyMOL and CAVEman that have been explored. StarCAVE integrate the PDB data bank server with its application to display structures in 3D. In the second research work, there are two kinds of facilities: screen-based VR tracking and tracking with hand held device. The aim of CAVEman is to develop a tool that can display information about certain diseases like cancer, diabetes, and Alzheimer in an interactive virtual environment. We have also discussed some future work that will probably non-trivial for solving latest problems.

Acknowledgement

The authors highly appreciate the support of research group "Information & Communication Technology Trends" at Federal Urdu University, and International Center for Chemical and Biological Sciences, HEJ University of Karachi, for their useful suggestions and comments. We are also thankful to Prof. Dr. Kamran Azim from the University of Karachi for their valuable suggestions and useful recommendations in writing this piece of work.

References

[1] J. N. Onuchic and P. Wolynes, "Theory of protein folding," *Current Opinion in Structural Biology*, vol. 14, pp. 70–75, 2004.
[2] J. Setubal and J. Meidanis, *Introduction to Computational Molecular Biology*, PWS Publishing Company, Boston, 1997.
[3] D. C. Richardson and J. S. Richardson, "The kinemage: A tool for scientific communication," *Protein Science*, vol. 1, pp. 3–9, 1992.
[4] R. A. Sayle and E. J. Milner-White, "Rasmol: Biomolecular graphics for all," *Biochem. Science*, vol. 20, pp. 374–376, 1995.
[5] P. J. Kraulis, "Molscript: A program to produce both detailed and schematic plots of protein structures," *Journal of Applied Crystallography*, vol. 24, pp. 946–950, 1991.
[6] E. Moritz and J. Meyer, "Interactive 3D protein structure visualization using virtual reality," in *Proc. Fourth IEEE Symposium on Bioinformatics and Bioengineering*, 2004, pp. 503-507.
[7] T. A. DeFanti, G. Dawe, D. J. Sandin, J. P. Schulze, *et al*., "The StarCAVE, a third-generation CAVE and virtual reality OptIPortal," *Future Generation Computer Systems*, vol. 25, pp. 169–178, 2009.
[8] Science Daily: Virtual reality gives insight on protein structures. [Online]. Available: http://www.sciencedaily.com/releases/2010/07/100722160114.htm
[9] Visualisation Cluster. [Online]. Available: http://www.rug.nl/science-and-society/centre-for-information-technology/research/hpcv/faciliteiten/visualisationcluster
[10] Y. Yang, L. Engin, E. Syrkin Wurtele, C. Cruz-Neira, and J. A. Dickerson, "Integration of metabolic networks and gene expression in virtual reality," *Bioinformatics*, vol. 21 no. 18, pp. 3645–3650, 2005.

Kamran Ahsan has PhD in Computer Science and an MS in Mobile Computer Systems from Staffordshire University. He has MCS in Computer Science from University of Karachi & MBA from Sweden. Kamran is an Assistant Professor at Federeal Urdu University of Arts, Science and Technology, in Pakistan. Dr. Ahsan is a formar researcher and lecturer in FCET (Faculty of Computing, Engineering and Technology) and, web researcher in Centre for Ageing and Mental Health, Staffordshire University, UK for 7 years. He has published many research papers so far. He has involved in several research funding (UK) includes KTP, NHS Trust, Innovation Funds etc. He is Visiting Faculty at University of Karachi. He is a consultant to businesses in IT applications, software development and web tools. His research interests are mobile technology applications in healthcare including knowledge management.