

A Visualization System for Binary Rooted Trees

M. Forghani¹, P. Vasev², V. Averbukh³

majid.forqani@gmail.com|vasev@imm.uran.ru|averbukh@imm.uran.ru

¹Ural Federal University, Yekaterinburg, Russia;

²IMM UB RAS, Yekaterinburg, Russia;

³IMM UB RAS & Ural Federal University, Yekaterinburg, Russia.

A system VirTree has been designed for visualization of ultrametric phylogenetic trees, which creates a three dimensional view and also provides a virtual reality presentation. The system can validate clusters in tree and adding a new dimension gives advantage to visualize each cluster independently. In addition to customization of 2D & 3D displays, a virtual reality presentation is provided in system through Viewlang. Viewlang is a web environment of three-dimensional visualization that has a high development speed due to QML language. The virtual reality can dramatically increase the information content of visualization and provides a wide range of view to see the general picture of tree with details while due to limited size of monitor screen this is not possible for large trees.

Keywords: ultrametric phylogenetic tree, visualization, virtual reality, Viewlang.

1. Introduction

Sequence alignments and phylogenetic trees can provide information about the evolutionary history of the proteins that comprise a family [7]. The evolutionary history is achieved based upon similarity and differences in genetic characteristics. Phylogenetic trees represent a backbone for various other biological studies and it is essential to have state-of-the-art tools for their visualization, customization and interpretation [4]. Due to growth of genetics databases, the visualization and customization of phylogenetic tree becomes difficult especially for large trees. The aim of tree visualization is to display complex evolutionary relationships between species to graphically human-readable presentation.

An ultrametric tree is a weighted rooted tree that all of its leaves are at the same depth. Each leaf represents a species and is labeled with unique taxon while two nodes can be joined together in the tree by common ancestor. In fact, each branch defines a relationship between sequences of proteins. A rooted phylogenetic tree is a directed tree with a unique node that is most recent ancestor of all leaves [2]. The distance between a node and its ancestor is evolutionary distance and represents how distant the node (taxa) is from its most recent ancestor.

Traditionally, Phylogenetic trees have been represented as two dimensional (flat) diagrams and the difficulties of interpreting of such tree are growing in size and complexity [3]. Trees can be displayed in a number ways, e.g., circular radial, phylogram and dendrogram. There are several tools for drawing phylogenetic trees both online and standalone, e.g. TREEVIEW, PHYLO_WIN, TREECON, FIGTREE, SEEVIEW, iTOL.

The goal of visualization is to leverage existing scientific methods by providing new scientific insight through visual methods. To visualize in 3-dimensional, tree is displayed in a circular (radial) mode. This form of displaying is particularly useful for the visualization of mid-sized trees (up to several thousand leaves) [4]. A high-resolution 3D phylogenetic tree can be constructed using parameters such as isoelectric focusing, molecular weight and immuno-crossreactivity or any other relevant attribute of a character or trait that gives discrete phylogenetic relationships [5].

As an example of 3D visualization of phylogenetic tree, 3D Phylogeny Explorer (3DPE) is a novel phylogeny tree viewer that maps trees onto three spatial axes (species on the X-axis; paralogs on Z; evolutionary distance on Y), making interpretation of the tree visually intuitive [4]. In addition to 3DPE, Hughes et al. [1] had introduced the idea of visualizing phylogenetic trees in three dimensional hyperbolic space with the Walrus graph visualization tool. With Walrus, it becomes

possible to visualize and navigate phylogenetic trees with more than 100,000 nodes.

In current paper, a system for 3D visualization of ultrametric tree is presented. In this tree, all leaves have the same distance from root. Tree is displayed in circular (radial) mode and by default all arc lengths between each two neighbor leaves are equal. The program gives two views 2D and 3D. In general, using a relevant attribute of a character or trait of leaves such as isoelectric focusing, can led to a complex visualization that is not easy to interpret while VirTree defines Z distance between two leaves equal to arc length between them. This allows bettering visualization of circular tree and tree can be rotated and displayed freely.

Virtree system consists of two parts. The first one is written as a Matlab program that computes coordination of visualization. The second part is based on the virtual reality presentation of results by Viewlang system. Viewlang is the technology of creating interactive 3D-applications for the Internet that has a high development speed due to QML language and supports virtual reality and mobile devices [6, 8]. Virtual Reality (VR) environments are actively used to provide leaping into a new quality of cognitive visualization. Virtual environments are characterized by such features as egocentric point-of-views and user-centered often multi-sensory interactions. The VR environments are dynamic, rather than static. The user's experience of the virtual world may combine a visual channel with the auditory or kinesthetic feedbacks. The immersion and sense of presence (feeling of "being there") are factors which defines VR. The sense of presence distinguishes VR from "traditional" 3D Computer Graphics. Users "immersed" in VR control the graphics output. Thus presence and immersion effects are integral parts of VR. The VR loses its most remarkable ability without them. Similar to color and animation, presence and immersion open a new dimension in visualization. Color and animation directly extend informativness of visualization, but presence and immersion create conditions for another perception of information. In scientific visualization, as in applying VR in industrial applications, it must be looked for new ways to not only increase the quality of interpretation, but also to control the visualization processes.

2. VirTree System

As mentioned above, Virtree consists of two parts. The first part is a GUI Matlab program that allows user to customize the visualization and produce a file for Viewlang representation. A 2D plot displays two-dimensional (flat) circular tree while 3D one shows the same tree but with additional dimension Z (see Fig.1). Final 2D and/or 3D customized graphical representation of phylogenetic trees can be saved separately.

There is an option to validate clusters in phylogenetic tree based upon “cluster” command for phytree object in Matlab [9]. When clusters are declared, uniform color will be assigned to a cluster automatically. For better visualization of clusters, within-cluster leaves distance can be decreased by ratio and this allows to increase between-clusters distance (see Fig. 2.). The compaction ratio slider controls how within-cluster leaves be close together. Similarly, the compaction also affects the Z distance between and within clusters. There are also additional options to show trunk and labels, control the markers size and select color of markers.

Phylogenetic trees can be loaded from a file in the standard Newick format. The GUI program does not include alignment or tree reconstruction methods. It is a tool for visualization. It provides an output JSON file that has information of 3D plot presentation.

Virtual reality provides a wide range of view comparing with monitor screen. Using Viewlang web-based presentation (that is second part of system), tree can be visualized in VR by JSON file. This provides a new quality of visualization and enables the user to view a general picture while the details are simply visible. In fact, Viewlang allows a direct virtual reality presentation within Matlab.

As an example, an ultrametric tree has been visualized in VirTree (see Fig.3.). VR provides us with a wider view space that allows user change some visualization parameters such as the thickness of the tree, the scaling parameter of the tree, the size of the label leaf. In order to see an individual leaf or branch, the user just needs to go closer to it by passing through VR environment. That is much more convenient than zoom in /out in 2D format. In other words, VR visualization is more flexible than 2D presentation.

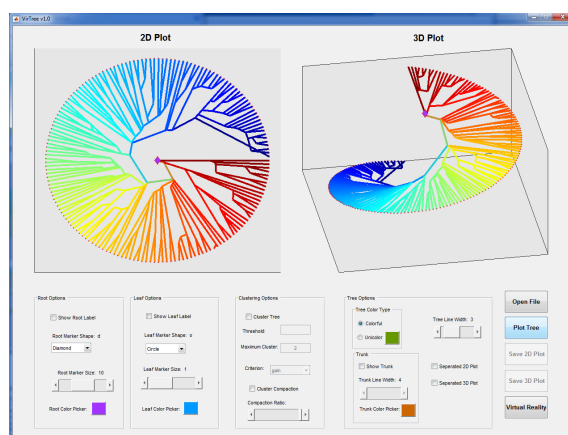


Figure 1. General representation of tree visualization.

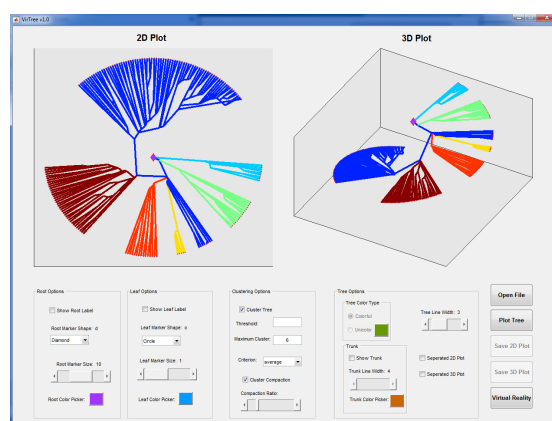


Figure 2. Cluster representation of tree from Fig.1.

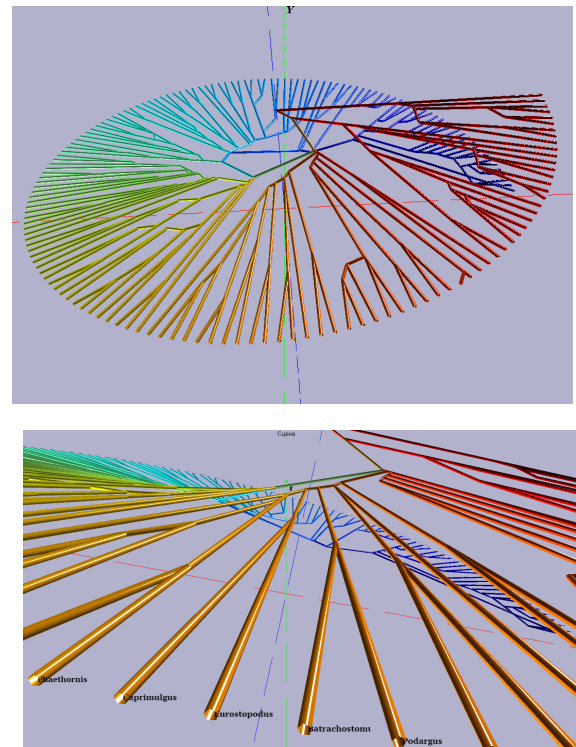


Figure 3. Virtual reality presentation of tree.

3. Conclusion

In summary, Virtree allows user to customize the tree visualization. It supplies 2D and 3D circular tree presentation. Moreover, system can validate tree clusters and additional dimension Z provides better recognition of clusters in plot. The Z dimension also makes a distance between leaves labels for easy to read them. Addition virtual reality capabilities will be used to enable a wide range of view. Immersion into three-dimensional space and easy navigation in it should provide researchers with more opportunities for studying the structure of trees. In future work, the system will be improved by adding alignment and tree reconstruction modules that allows working with FASTA format and different types of tree (root and unrooted). In addition, the interface in virtual reality will be more interactive that provides tree customization in virtual space directly.

4. References

- [1] Hughes, Timothy, Young Hyun, and David A. Liberles. Visualising very large phylogenetic trees in three dimensional hyperbolic space. // BMC bioinformatics 5.1 (2004): 48.
- [2] Huson, Daniel H., and Celine Scornavacca. Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. // Systematic biology 61.6 (2012): 1061-1067.
- [3] Kim, Namshin, and Christopher Lee. Three-Dimensional Phylogeny Explorer: Distinguishing paralogs, lateral transfer, and violation of" molecular clock" assumption with 3D visualization.// BMC bioinformatics 8.1 (2007): 213.
- [4] Letunic, Ivica, and Peer Bork. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation.// Bioinformatics 23.1 (2006): 127-128.
- [5] Milner, M., et al. Constructing 3D phylogenetic trees. // Current Science (2003): 1471-1478.
- [6] V Vasev P. Three-dimensional visualization in a web based environment based on Qml declarative description // International (47th all-Russian) youth school-conference,

Yekaterinburg, January 31 - February 6, 2016.
<http://data2.lact.ru/fl/s/0/299/basic/1605/992/vasev-sopromat-2016-1.pdf?t=1482676145>

- [7] Zhai, Yufeng, Jason Tchieu, and Milton H. Saier Jr. A web-based Tree View (TV) program for the visualization of phylogenetic trees." *Journal of molecular microbiology and biotechnology* 4.1 (2002): 69-70.
- [8] <http://viewlang.ru>.
- [9] <https://www.mathworks.com/help/bioinfo/ref/clusterphytref.html>.

About the authors

Vladimir L. Averbukh, Ph.D., head of the researcher's section in the Institute of Mathematics and Mechanics, Ural Branch of the Russian Academy of Sciences, associate professor, Ural Federal University. E-mail averbukh@imm.uran.ru.

Pavel F. Vasev, researcher in the Institute of Mathematics and Mechanics, Ural Branch of the Russian Academy of Sciences. E-mail vasev@imm.uran.ru.

Majid A. Forghani, researcher in the Ural Federal University. E-mail majid.forqani@gmail.com.