Hierarchical Visualization of Metabolic Networks using Virtual Reality

Yuting Yang[1,4] Eve Syrkin Wurtele[2,4] Carolina Cruz-Neira[1,3,4] Julie A. Dickerson[1,4]

[1]Department of Electrical and Computer Engineering

[2]Department of Genetics, Development, and Cell Biology

[3]Department of Industrial and Manufacturing Systems Engineering

[4]Virtual Reality Applications Center

Iowa State University

Abstract

MetNetVR combines graph layouts in 3D space, computer graphics, and VR technologies for interactive visualization of high dimensional metabolic networks. This paper highlights the use of hierarchical visualization that captures the hierarchical relationships in metabolic networks and the use of detail-on-demand interactions to help users understand both the global relationships and local details of a large network simultaneously. Complex interactions are enabled by MetNetVR Tweek, an auxiliary 2D GUI program which communicates with MetNetVR in real time. These interactions control graph theoretic operations, layout management, and visual metaphors.

CR Categories: H.5.1 [INFORMATION INTERFACES AND PRESENTATION]: Multimedia Information Systems—Artificial realities; I.3.7 [COMPUTER GRAPHICS]: Three-Dimensional Graphics and Realism—Virtual reality; I.3.8 [COMPUTER GRAPHICS]: Applications

Keywords: metabolic network, interactive visualization, graph layout, computer graphics rendering, virtual reality

1 Introduction

A major challenge in the post-genomic era is to understand how interactions among molecules in a cell determine its form and function. Experimental data derived by metabolomic, proteomic, and transcriptomic analysis technologies have the potential to give biologists vast amounts of valuable data on the status of metabolic network interactions. Understanding the structure and status of complex metabolic networks without graphics is very difficult.

Metabolic network visualization is an application of graph drawing. A metabolic network is modeled as a directed graph, G (N, E), where N is the set of nodes representing molecules in the network; E is the set of reaction edges connecting the nodes. The process of graph drawing includes three stages [Kamada 1989]: modeling, layout, and rendering. In the stage of modeling, a graph is extracted from the information. The stage of layout assigns positions to nodes and edges in two-dimensional (2D) or three-dimensional (3D) space. The stage of rendering produces an image of the graph. Most of current metabolic network visualization systems exploit 2D space in the stage of layout or produce flowchart-like display in the stage of rendering [Karp and Paley 1994; Becker and Rojas 2001; Dickerson et al. 2001; Demir et al. 2002; Schreiber 2002; Schreiber 2003; Shannon et al. 2003]. Layouts in 2D space are limited due to lack the display space.

Layouts in 3D space may be able to create a more effective information workspace for the study of large metabolic networks. The extra dimension gives greater flexibility for placing nodes and edges in the network. 3D computer graphics uses computers to generate photorealistic images for a real world or a synthetic world. Rendering in computer graphics features various material properties, such as color, transparency, and texture. Some researchers have started to use 3D space and computer graphics technologies to display metabolic networks. Brandes [Brandes et al. 2003] focuses on comparing pathways across species. Each pathway is drawn in one layer; then the layers are ordered in the third dimension such that the most similar pathways are adjacent. Rojdestvenski [Rojdestvenski 2003] statically models metabolic networks as graphs in 3D space using Virtual Reality Modeling Language (VRML).

Usually resultant images of computer graphics rendering are displayed in 2D media, such as monoscopic computer monitors or paper. Virtual reality (VR) extends the display into the third dimension by introducing stereoscopic display. VR also uses advanced devices, such as head and wand tracking devices and various input devices, to enable richer interactions than conventional computers.

MetNetVR builds on above technologies to visualize high dimensional metabolic networks (more than 1,000 nodes) in 3D space in an interactive way [Yang et al. 2005]. Multiple graph layout methods are implemented to automatically decide node and edge positions. The network is then visually represented by a scene. Shapes and material properties, such as color, transparency, and texture, are used as visual metaphors indicating basic properties of nodes and edges. VR technologies enable the interaction to change the viewpoint position for rendering and other interactions such as selecting. MetNetVR also provides a solution for linking text-rich information that is important to biologists but difficult to visualize, such as chemical reaction stoichiometry, sources of information, and synonyms, to network nodes and edges. Each node and edge is linked to a record in the database of text-based annotation data. MetNetVR Tweek, a Java program with a 2D Graphical User Interface (GUI),

VRCIA 2006, Hong Kong, 14–17 June 2006.
© 2006 ACM 1-59593-324-7/06/0006 $5.00

communicates with MetNetVR and displays the detailed text information.

MetNetVR is implemented in C++ and uses two APIs, VRJuggler (http://www.vrjuggler.org/) and OpenSG (http://www.opensg.org/). VRJuggler is a flexible development platform for VR applications [Bierbaum et al. 2001]. It enables MetNetVR to run in a range of platforms without changing source codes, from a fully immersive CAVE [Cruz-Neira et al. 1993] to a conventional desktop. OpenSG is an OpenGL based API for graphics scene construction and rendering [Reiners et al. 2002]. MetNetVR Tweek is implemented in C++ and Java and uses the Tweek API [Hartling et al. 2002], which uses the Common Object Request Broker Architecture (http://www.corba.org/) for the cross-language communication. MetNetVR Tweek runs on a portable tablet computer and connects to MetNetVR wirelessly when the latter runs in an immersive CAVE. It can also run on the same desktop as MetNetVR when the latter runs on a conventional desktop. Figure 1 shows MetNetVR running in a CAVE and MetNetVR Tweek running on a tablet computer.

A new method to exploit hierarchical relationships contained in metabolic networks is recently added. Very few metabolic network visualization programs reveal or make use of hierarchical relationships. [Dogrusoz et al. 2004] shows abstractions for some parts of a network statically by using the geometric inclusion in 2D space. In MetNetVR, hierarchical relationships or quasi-hierarchical relationships, together with detail-on-demand interactions, help users to understand global relationships and local details of a large network simultaneously. A menu is added in MetNetVR Tweek that contains complex interaction commands beyond the existing wand based interactions. These commands enable functions in MetNetVR to manipulate network visualization, such as graph theory operations, layout management, and visual metaphor controls.

2 Hierarchical Relationships in Metabolic Networks and Detail-on-Demand Method

Many metabolic networks contain some hierarchical relationships or quasi-hierarchical relationships. For example, a metabolic network may include several pathways. A pathway contains molecules. The relationship between pathways and molecules is a quasi-hierarchical relationship. One pathway involves many molecules. Some molecules take part in more than one pathway. Another example is the grouping of molecules in a network according to their location within the cell or sub-cellular compartments. One compartment may contain multiple molecules. Molecules that appear in different compartments are modeled as different nodes. These hierarchical relationships show different levels-of-detail in the metabolic network (Figure 2). A graph model containing both adjacent relationships and hierarchical relationships is called a *compound graph* [Sugiyama and Misue 1991]. A metabolic network is called a *compound network* if it is represented as a compound graph. It is called a *standard network* if represented as a graph. A compound network is reduced to a standard network if all pathway (or compartment) nodes and hierarchical relationships are removed.

Displaying the whole graph representing a large standard network shows the overall structure of the network. The drawback is that details such as node and edge names are too small to be seen. Zooming into a part of the network and panning to other parts show local details but make it difficult to see the global structure.

The detail-on-demand method uses hierarchical relationships in a compound network to allow users to dynamically change the visual level-of-detail. Figure 2 shows how the detail-on-demand method reduces the number of displayed nodes and still maintains the correct relationships in a compound network.

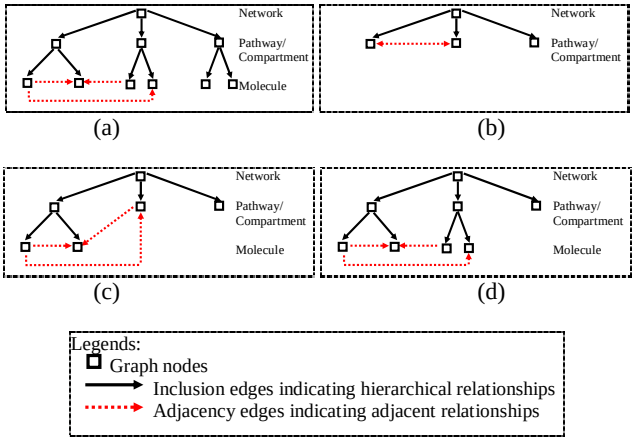


Figure 2. (a) A compound network contains hierarchical and adjacent relationships. (b) The snapshot in the level of least details. Adjacency edges in more detailed levels are converted to this level. (c) Expand the leftmost pathway (or compartment) node in (b), the adjacency edges are converted to the currently deepest level accordingly (d) Expand one more node

A quasi-hierarchical relationship is a loose hierarchical relationship where a child may belongs to more than one parent in the hierarchy. A quasi-hierarchical relationship converts to a hierarchical relationship using duplication. In Figure 3(a), molecule 'A' appears in two pathways. In Figure 3(b), 'A' is duplicated so that each pathway has a copy. There is a duplicate edge between two copies of 'A'. The duplicate edges are treated the same as the adjacency edges in the detail-on-demand method discussed above.

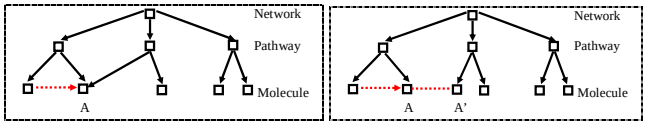


Figure 3. Node duplication converts a quasi-hierarchical relationship to a hierarchical relationship

MetNetVR implements hierarchical relationships as geometric inclusions in 3D space. This is represented by displaying all the visual representations of molecules as bounded by the visual representation of the pathway (or compartment) they belong to.

The detail-on-demand interaction is enabled by an abstract interface for a tracked wand-like input device, which supports both real and simulated wands. A ray (the red line in Figure 4) is cast from the wand (the green gadget in Figure 4) according to its position and orientation. When the ray intersects with the visual representation of a pathway (or compartment), a button click expands the pathway, i.e., molecules in the pathway appear. When the ray intersects any one of molecules in the expanded pathway, a button click will shrink the pathway back into the pathway node.

Figure 4 shows some snapshots in the exploration of a network in Arabidopsis from the MetNet Database [Wurtele et al. 2003] using hierarchical relationships and detail-on-demand interactions. There are 572 molecules and 648 reaction edges among molecules in this network. The molecules in the network belong to the following sub-cellular compartments: plastid, mitochondrion, nucleus, cytosol, plastid stroma, and unknown. The network contains three pathways, Acetyl-CoA biotin, starch degradation, and starch synthesis. The thickness of an edge between two compartments or between a compartment and a molecule indicates the density of the edges (between molecules) crossing these two compartments.

These still snapshots are caught when MetNetVR runs in a desktop with Windows Operation System (OS). The original snapshots are available at (http://www.vrac.iastate.edu/research/sites/metnet/VRCIA/). The video clip (video-1.mpg) shows the exploration of the same network. The video clip is caught when MetNetVR runs in a desktop with Linux OS. Both of them are in the monoscopic mode for the sake of viewing without any aids.

3 MetNetVR Tweek

MetNetVR Tweek was designed to display text-rich information for molecules and reaction edges in metabolic networks. The information of each molecule node and reaction edges is displayed in a scrollable list [Yang et al. 2005]. Wand-based interactions include moving forward and backward along the wand orientation, rotating right and left along the center of the network, and selecting the nodes. All other functions are called through menus in MetNetVR Tweek. The main interaction commands in the MetNetVR Tweek menu fall into the following categories:

- Compound Network Visualization
  - Hierarchical Relationship Management. This function toggles the exploration method between different hierarchical relationships.

- Standard Network Visualization (exploring the network without taking hierarchical relationships into consideration)
  - Layout Management
    - Switch among different layouts for the network including a GEM-3D layout and a weighted GEM-3D layout [Yang et al. 2005] and a recently added multilevel force directed layout[Gajer et al. 2000]. The weighted GEM-3D layout gives a weight to each reaction edge according to its biological type. Other two layouts are pure graph theoretic layouts. The multilevel force directed layout is much faster to calculate than the other two and suitable for large networks.
    - Edge length adjustment for each layout. This function likes a zoom in/out function. However, it keeps the molecule node size fixed to maintain the label's readability. Small edge lengths are suitable for the overview of a network because they keep the network volume small. Large edge lengths work best for local details as they reduce the overlapping of node labels.
  - Reactions of Interest Management:
    - Select a metabolite (metabolite is a type of molecule) or a gene (gene is another type of molecule) from the node list
    - Generate (the scene for) reactions of interest focusing on the selected metabolite or the selected gene [Yang et al. 2005]
  - Sub-network Management:
    - Select a molecule of any type from the node list and set the step size.
    - Generate the sub-network composed of the selected molecule and its neighbors within the distance of the step size. Switch display between the whole network and the sub-network.

- Network Management.
  - Toggle between the standard network and the compound network. In MetNetVR, both the standard graph model and the compound graph model exist for a metabolic network. They come from the same data source. The standard network filters out pathway (or compartment) nodes and all hierarchical relationships. The adjacent relationships among molecules are identical between a standard network and a compound network with all the compartments expanded. The layout of the compound network groups the molecules belonging to one compartment together. The layouts of the standard network only look at graph connectivity. Users can toggle between these representations.

- Visual Metaphor (controlling the visual representations for both the standard network and the compound network).
  - Color Management: select colors for different types of nodes and different types of reaction edges(Figure 5). Different colors for different types of nodes and edges help users to easily recognize the type of specific nodes and edges.
  - Geometry Management. This function toggles molecule nodes or reaction edges on and off to increase readability. For example, the user can temporarily turn off edges to read a node label because one or two letters may be blocked by the incoming/outgoing edges of this node. The user can also temporarily turn off node labels for a better vision when trying to find a cycle structure in the network.

- Navigation Management
  - Reset: move back to the initial position
  - Select a specific node in the node list and move to the selected node

4 Discussion and Conclusions

3D computer graphics is used for visualizing abstract information like metabolic networks. Metabolic networks are modeled as graphs whose nodes and edges represent biological properties. Material properties in computer graphics, which were originally used for photorealistic rendering of the real world, are effective visual metaphors for basic biological properties.

Metabolic networks usually have a large number of nodes and edges. 3D space, rendering from different viewpoints, and stereoscopic display help to increase the size of understandable networks. For example, two edges may look like they cross each other from one viewpoint in a monoscopic display. The apparent crossing will disappear when viewed from another viewpoint, or the depth difference shown by the stereoscopic display helps distinguish the edges.

Besides enabling binocular disparity through stereoscopic display, VR has the option of head tracking to enable the rendering of the metabolic network from different and continuous viewpoints. Moving the viewpoint and fixing the network position have the same effect as fixing the viewpoint and moving the network position in an opposite way. Both of them bring about motion parallax, which is helpful to understand the complex structures of high dimensional metabolic networks. VR also features tracked 3D input devices such as wands to enable movement and find intersections along the device orientation. Intersecting with visual representation to pick a node is of great importance for interactive visualization. For example, in the detail-on-demand method described in Section 2, user needs to pick the pathway (or compartment) node in order to expand it. 3D input devices can be simulated by mice and keyboards with full six degrees of freedom in conventional desktops. However, controlling the position and orientation in 3D space using mice and keyboards requires considerable practice.

The detail-on-demand interactions enable different hierarchical views of a metabolic network. 3D geometric inclusions, together with the transparency material property and stereoscopic display, are effective to represent hierarchical relationships in the network. Hierarchical visualization reduces the amount of information displayed at one time and helps users understand both the global relationships and local details of a large network simultaneously. The detail-on-demand interactions, 3D geometric inclusions, and the layout method support the hierarchical structure of multiple levels. The current networks contain 3 levels, i.e. the network, the pathways (or compartments), and the molecule nodes.

Text-rich information display methods and complex interaction methods are needed for an effective visualization. In traditional 2D visualization systems, it is very natural to add popup windows and menu items for this purpose. In a 3D computer graphics based visualization systems for metabolic networks, an auxiliary 2D GUI program can be used to display text-rich and domain-related information and enable interactions to manipulate visualization.

Acknowledgements

Funding for this project is provided by grants from the National Science Foundation in the Arabidopsis 2010 and the Information Technology Research Program (Grant Number: 0219366). We are grateful to Jie Li and Pan Du for developing and providing the XML data files of metabolic networks. This work was performed at the Virtual Reality Application Center, Iowa State University.

References

Becker, M. Y. and Rojas, I. 2001. A graph layout algorithm for drawing metabolic pathways. *Bioinformatics* ***17***(5): 461-467.

Bierbaum, A., Just, C., Hartling, P., Meinert, K., Baker, A. and Cruz-Neira, C. 2001. VR Juggler: A Virtual Platform for Virtual Reality Application Development. *IEEE Virtual Reality Conference*, Yokohama, Japan, 89-96.

Brandes, U., Dwyer, T. and Schreiber, F. 2003. Visualizing Related Metabolic Pathways in Two and a Half Dimensions. *11th International Symposium on Graph Drawing (GD '03)*, Springer-Verlag, LNCS 2912, 111-122.

Cruz-Neira, C., Sandin, D. J. and Defanti, T. A. 1993. Surround-Screen Projection-based Virtual Reality: The Design and Implementation of the CAVE. *Proceedings of ACM SIGGRAPH 1993*, ACM Press / ACM SIGGRAPH, Computer Graphics Proceedings, Annual Conference Series, ACM, 135-142.

Demir, E., Babur, O., Dogrusoz, U., Gursoy, A., Nisanci, G., Cetin-Atalay, R. and Ozturk, M. 2002. PATIKA: An Integrated Visual Environment for Collaborative Construction and Analysis of Cellular Pathways. *Bioinformatics* ***18***(7): 996-1003.

Dickerson, J. A., Cox, Z., Wurtele, E. S. and Fulmer, A. W. 2001. Creating Metabolic and Regulatory Network Models using Fuzzy Cognitive Maps. *North American Fuzzy Information Processing Conference (NAFIPS)*, Vancouver B.C., 2171-2176.

Dogrusoz, U., Giral, E., Cetintas, A., Civril, A. and Demir, E. 2004. A Compound Graph Layout Algorithm for Biological Pathways. *12th International Symposium on Graph Drawing*, NYC, USA

Gajer, P., Goodrich, M. and Kobourov, S. 2000. A fast multi-dimensional algorithm for drawing large graphs. *8th Symposium on Graph Drawing*, 211-221.

Hartling, P., Bierbaum, A. and Cruz-Neira, C. 2002. Tweek: Merging 2D and 3D Interaction in Immersive Environments. *6th World Multiconference on Systemics, Cybernetics, and Informatics*, Orlando, Florida, 1-5.

Kamada, T. 1989. *Visualizing Abstract Objects and Relations*. Singapore, World Scientific.

Karp, P. D. and Paley, S. M. 1994. Automated drawing of metabolic pathways. *3rd Int. Conf. on Bioinformatics and Genome Research*, 225-238.

Reiners, D., Vo, G. and Behr, J. 2002. OpenSG: Basic concepts. *OpenSG Symposium*, Darmstadt

Rojdestvenski, I. 2003. Metabolic pathways in three dimensions. *Bioinformatics* ***19***(18): 2436-2441.

Schreiber, F. 2002. High quality visualization of biochemical pathways in BioPath. *Silico Biology* ***2***(2): 59-73.

Schreiber, F. 2003. Comparison of metabolic pathways using constraint graph drawing. *Conferences in Research and Practice in Information Technology Series*, Australian Computer Society, Inc, Adelaide, Australia, Proceedings of the First Asia-Pacific bioinformatics conference on Bioinformatics 2003, 105-110.

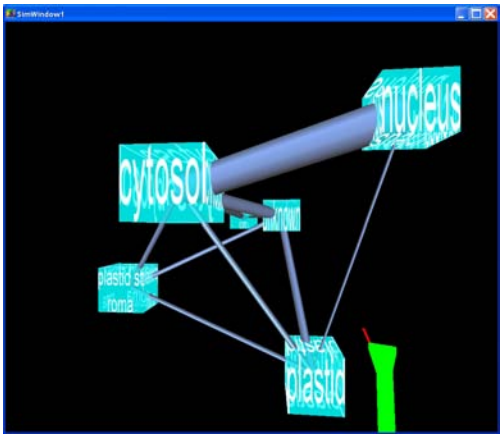
Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B. and Ideker, T. 2003. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Research* ***13***: 2498–2504.

Sugiyama, K. and Misue, K. 1991. Visualization of structural information: Automatic drawing of compound digraphs. *IEEE Trans. Systems, Man and Cybernetics*, ***21***(4): 876-892.

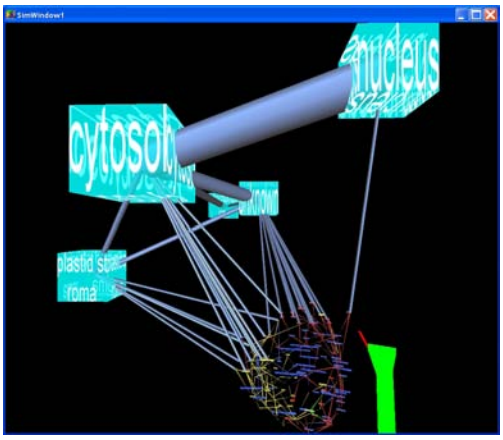
Wurtele, E. S., Dickerson, J. A., Cook, D., Hofmann, H., Li, J. and Diao, L. 2003. MetNet: software to build and model the biogenetic lattice of Arabidopsis. *Comp Funct Genom* ***4***: 239-245.

Yang, Y., Engin, L., Wurtele, E. S., Cruz-Neira, C. and Dickerson, J. A. 2005. Integration of metabolic networks and gene expression in virtual reality. *Bioinformatics* ***21***(18): 3645-3650.

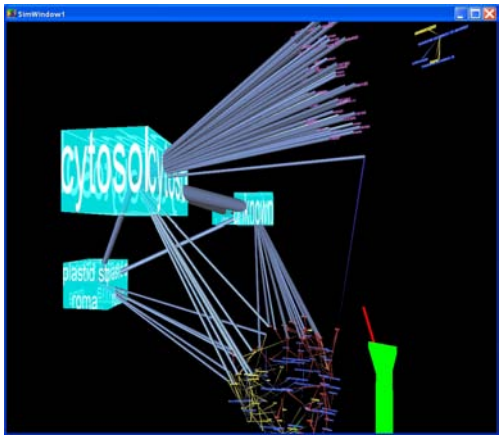
Figure 1. A user navigating through the metabolic network in a CAVE. MetNetVR Tweek running in the tablet PC communicates with the visualization program wirelessly and displays the text rich information of the network.



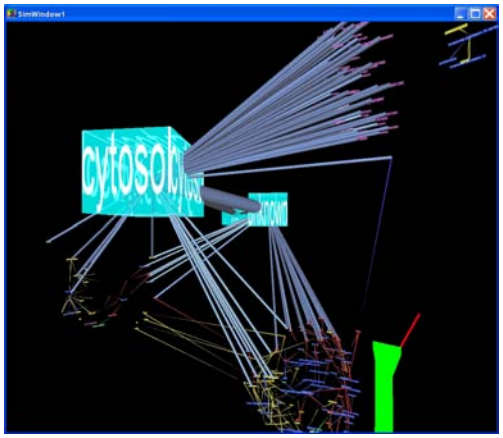
(a) A network whose molecules in the network belong to six compartments



(b) Expand the plastid compartment. After the expansion, the representation of plastid (the translucent boundary box) is turned off to show the nodes and edges inside.



(c) Expand the nucleus compartment.



(d) Expand the plastid stroma compartment.

Figure 4. Snapshots illustrating the exploration of a metabolic network using hierarchical relationships and detail-on-demand interactions.

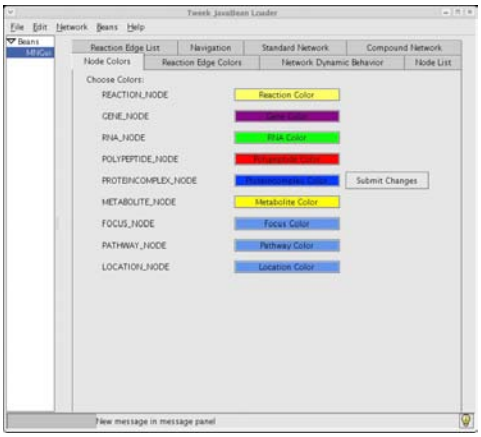


Figure 5. Menu for interaction commands in MetNetVR Tweek.