Institute for Computer Graphics and Vision
Graz University of Technology
Graz
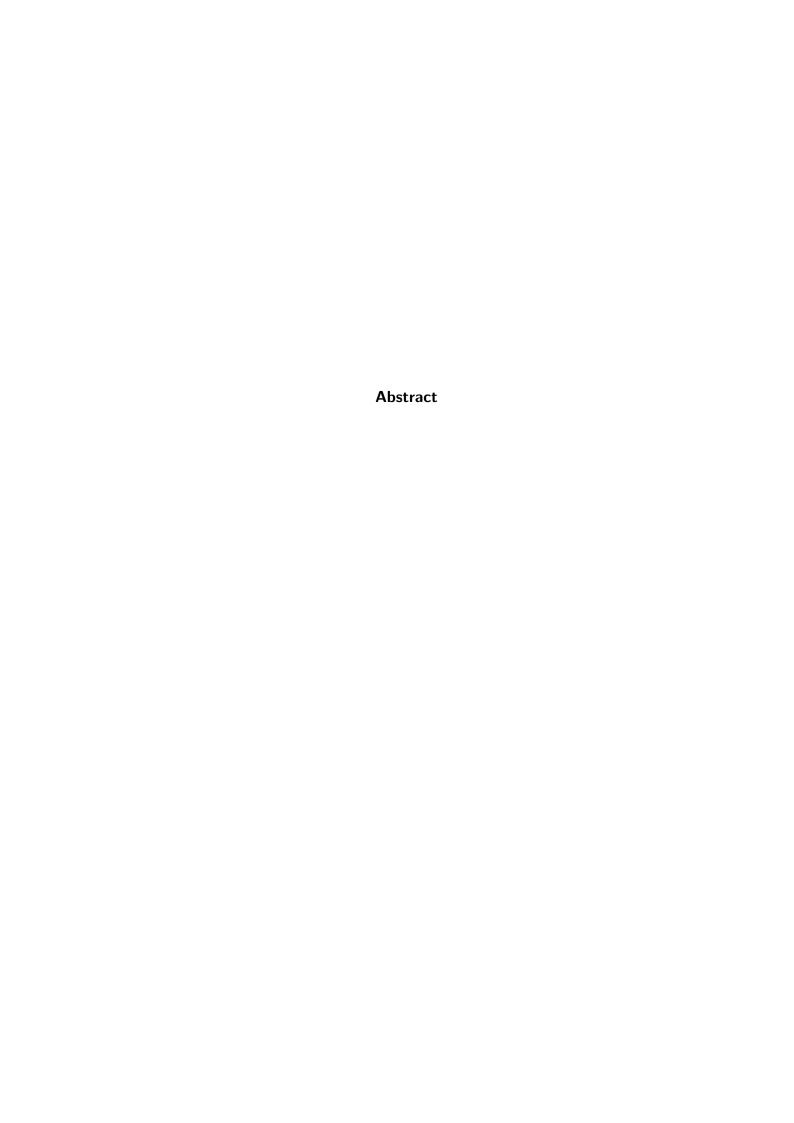
Metabolic Pathways Influencing Gene-Expression Analysis
Thesis paper

Marc Streit

mstreit@icg.tugraz.at

February 2007

Supervision:
Univ. Prof. DI Dr. techn. Dieter Schmalstieg
DI Micheal Kalkusch

Abstract

Ich versichere hiermit, diese Arbeit selbständig verfaßt, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und mich auch sonst keiner unerlaubten Hilfsmittel bedient zu haben.

I hereby certify that the work presented in this diploma thesis is my own and that work performed by others is appropriately cited.

Acknowledgements

Contents

Chapter 1

Introduction

Chapter 2

Problem Definiton

Chapter 3

Related Work

3.1 Biomedical Databases

3.1.1 Entrez

3.1.2 Kyoto Encyclopedia of Genes and Genomes (KEGG)

3.1.3 Gene Ontology (GO)

3.1.4 Biomedical Identification Numbers

3.2 Metabolic Pahtway Visualization

3.2.1 Medical Pathways

3.2.2 State-of-the-art Frameworks

3.3 Gene Expression Visualization

3.3.1 Gene-Expression Analysis

3.3.2 State-of-the-art Frameworks

3.4 Information Visualization Methods

3.4.1 Multiple Views

3.4.2 Focus + Context

3.4.3 Linking & Brushing

3.4.4 Semantic Zooom

3.4.5 Neighborhood Visualization

3.5 Application of Gene-Expression Data on Metabolic Pathways

3.6 Pathway Visualization influencing Gene-Expression Analysis

Chapter 4

System Architecture

4.1 Overall Design

4.2 Data Management

4.2.1 Pathway Data

4.2.2 Set / Storage / Virtual Array

4.3 Graphical User Interface (GUI)

4.4 Data Update Mechanism

4.5 Selection Handling

Chapter 5

Implementation

5.1 Used Technologies

5.1.1 SWT

5.1.2 JOGL

5.1.3 JGraph

5.2 Data Loading

5.3 Visualization Techniques

5.3.1 2D Pathway Implementation

5.3.1.1 Pathway Switching

5.3.1.2 Hierarchical Pathways

5.3.1.3 Neighborhood Visualization

5.3.2 OpenGL Pathways

5.3.2.1 Pathway Texture Overlay

5.3.2.2 Pathway Linking

5.3.2.3 Pathway Element Picking

5.3.2.4 Hierarchical Display Lists

5.3.2.5 Layered Pathways

5.3.2.6 Panel View

Chapter 6

Results

Chapter 7

Conclusions

Chapter 8

Future Work

List of Figures

List of Tables

Bibliography

[Agr06]   AGRAWALA, Maneesh: Software Design Patterns for Information Visualization. In: *IEEE Transactions on Visualization and Computer Graphics* 12 (2006), Nr. 5, S. 853–860. `http://dx.doi.org/http://dx.doi.org/10.1109/TVCG.2006.178`. – Student Member-Jeffrey Heer. – DOI http://dx.doi.org/10.1109/TVCG.2006.178. – ISSN 1077–2626

[BDS04]   BRANDES, Ulrik ; DWYER, Tim ; SCHREIBER, Falk: Visualizing Related Metabolic Pathways in Two and a Half Dimensions (Long Paper). In: LIOTTA, Giuseppe (Editor): *Graph Drawing, Perugia, 2003*, Springer, 2004, S. 111–122

[BFP⁺04]   BRANDENBURG, Franz J. ; FORSTER, Michael ; PICK, Andreas ; RAITNER, Marcus ; SCHREIBER, Falk: BioPath - Exploration and Visualization of Biochemical Pathways. In: JÜNGER, Michael (Editor) ; MUTZEL, Petra (Editor): *Graph Drawing Software*. Springer, 2004, S. 215 – 236

[BKPT05]   BRESLIN, Thomas ; KROGH, Morten ; PETERSON, Carsten ; TROEIN, Carl: Signal transduction pathway profiling of individual tumor samples. In: *BMC Bioinformatics* 6 (2005), S. 163

[BR01]   BECKER, Moritz ; ROJAS, Isabel: A Graph Layout Algorithm for Drawing Metabolic Pathways. In: *Bioinformatics* 17 (2001), Nr. 5, S. 461–467

[BWK00]   BALDONADO, Michelle Q. W. ; WOODRUFF, Allison ; KUCHINSKY, Allan: Guidelines for using multiple views in information visualization. In: *AVI '00: Proceedings of the working conference on Advanced visual interfaces*. New York, NY, USA : ACM Press, 2000. – ISBN 1–58113–252–2, S. 110–119

[CKP⁺04]   CHUNG, Hee-Joon ; KIM, Mingoo ; PARK, Chan H. ; KIM, Jihoon ; KIM, Ju H.: ArrayXPath: mapping and visualizing microarray gene-expression data with integrated biological pathway resources using Scalable Vector Graphics. In: *Nucleic Acids Res* 32 (2004), Jul, Nr. Web Server issue, S. W460–W464. `http://dx.doi.org/10.1093/nar/gkh476`. – DOI 10.1093/nar/gkh476

[CPH⁺05]   CHUNG, Hee-Joon ; PARK, Chan H. ; HAN, Mi R. ; LEE, Seokho ; OHN, Jung H. ; KIM, Jihoon ; KIM, Jihun ; KIM, Ju H.: ArrayXPath II: mapping and visualizing microarray gene-expression data with biomedical ontologies and integrated biological pathway resources using Scalable Vector Graphics.

In: *Nucleic Acids Res* 33 (2005), Jul, Nr. Web Server issue, S. W621–W626. `http://dx.doi.org/10.1093/nar/gki450`. – DOI 10.1093/nar/gki450

[DBD⁺02] Demir, E. ; Babur, O. ; Dogrusoz, U. ; Gursoy, A. ; Nisanci, G. ; Cetin-Atalay, R. ; Ozturk, M.: PATIKA: an integrated visual environment for collaborative construction and analysis of cellular pathways. In: *Bioinformatics* 18 (2002), Nr. 7, S. 996–1003

[DGC⁺04] Dogrusoz, Ugur ; Giral, Erhan ; Cetintas, Ahmet ; Civril, Ali ; Demir, Emek: A Compound Graph Layout Algorithm for Biological Pathways. In: Pach, János (Editor): *Graph Drawing, New York, 2004*, Springer, 2004, S. pp. 442–447

[DJ03] Dickerson JA, Blom K Reinot A Lie J Cruz-Neira C Wurtele ES. . ”.” September. In p.: Using Virtual Reality to Understand Complex Metabolic Networks. In: *Atlantic Symposium on Computational Biology and Genomic Information Systems and Technology, In Press*, 2003

[GHJV95] Gamma, Erich ; Helm, Richard ; Johnson, Ralph ; Vlissides, John: *Design patterns: elements of reusable object-oriented software.* Boston, MA, USA : Addison-Wesley Longman Publishing Co., Inc., 1995. – ISBN 0–201–63361–2

[GTHC02] Grosu, Paul ; Townsend, Jeffrey P. ; Hartl, Daniel L. ; Cavalieri, Duccio: Pathway Processor: a tool for integrating whole-genome expression results into metabolic networks. In: *Genome Res* 12 (2002), Jul, Nr. 7, S. 1121–6. `http://dx.doi.org/10.1101/gr.226602`. – DOI 10.1101/gr.226602

[HDLT05] Hibbs, M. A. ; Dirksen, N. C. ; Li, K. ; Troyanskaya, O. G.: Visualization methods for statistical analysis of microarray clusters. In: *BMC Bioinformatics* 6 (2005). `http://dx.doi.org/10.1186/1471-2105-6-115`. – DOI 10.1186/1471–2105–6–115. – ISSN 1471–2105

[HK04] Hauser, Helwig ; Kosara, Robert: Interactive Analysis of High-Dimensional Data using Visualization. In: *Workshop on Robustness for High-dimensional Data (RobHD 2004)*. Vorau, Austria, May 2004

[HWH04] Harris, Bobby ; Warner, Rob ; Harris, Robert: *The Definitive Guide to Swt and Jface.* APress, 2004. – ISBN 1590593251

[KAJR02] Kurhekar, Manish P. ; Adak, Sudeshna ; Jhunjhunwala, S. ; Raghupathy, K.: Genome-Wide Pathway Analysis and Visualization Using Gene Expression Data. In: *Pacific Symposium on Biocomputing*, 2002, S. 462–473

[KG00] Kanehisa, M. ; Goto, S.: KEGG: kyoto encyclopedia of genes and

genomes. In: *Nucleic Acids Research* 28 (2000), January, Nr. 1, S. 27–30. – ISSN 0305–1048

[KHG03] KOSARA, Robert ; HAUSER, Helwig ; GRESH, Donna L.: An Interaction View on Information Visualization. In: *State-of-the-Art Proceedings of EUROGRAPHICS 2003 (EG 2003)*, 2003, S. 123–137

[Kin04] KINCAID, Robert: VistaClara: an interactive visualization for exploratory analysis of DNA microarrays. In: *SAC '04: Proceedings of the 2004 ACM symposium on Applied computing.* New York, NY, USA : ACM Press, 2004. – ISBN 1–58113–812–1, S. 167–174

[KP94] KARP, P. ; PALEY, S. M.: Automated Drawing of Metabolic Pathways. In: LIM, H. (Editor) ; CANTOR, C. (Editor) ; ROBBINS, R. (Editor): *Third International Conference on Bioinformatics and Genome Research*, 1994

[KR93] KARP, Peter D. ; RILEY, Monica: Representations of Metabolic Knowledge: Pathways. In: *Proceedings of the 1st International Conference on Intelligent Systems for Molecular Biology*, AAAI Press, 1993. – ISBN 0–929280–47–4, S. 207–215

[KSH04] KOSARA, R. ; SAHLING, G. ; HAUSER, H.: Linking scientific and information visualization with interactive 3D scatterplots. In: *12th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG)*, 2004, S. 133–140

[Lin02] LINDROOS, S.G.E.: Visualizing metabolic pathways: comparative genomics and expression analysis. In: *Proceedings of the IEEE* Bd. 90, 2002, S. 1793 – 1802

[MW95] MARTIN, Allen R. ; WARD, Matthew O.: High Dimensional Brushing for Interactive Exploration of Multivariate Data. In: *VIS '95: Proceedings of the 6th conference on Visualization '95.* Washington, DC, USA : IEEE Computer Society, 1995. – ISBN 0–8186–7187–4, S. 271

[NBK+99] NAKAO, Mitsuteru ; BONO, Hidemasa ; KAWASHIMA, Shuichi ; KAMIYA, Tomomi ; SATO, Kazushige ; GOTO, Susumu ; KANEHISA, Minoru: Genome-scale Gene Expression Analysis and Pathway Reconstruction in KEGG. In: *Genome Informatics* Bd. 10, 1999, S. 94–103

[PC06] PELLEGRINI, M. ; COKUS, S.: Pathway Analysis of Microarray Data. In: *In Press* (2006)

[RC02] ROJDESTVENSKI, Igor ; COTTAM, Michael: Visualizing Metabolic Networks in VRML. In: *6th International Conference on InformationVisualisation* 00 (2002), S. 175. `http://dx.doi.org/http:`

//doi.ieeecomputersociety.org/10.1109/IV.2002.1028774. – DOI http://doi.ieeecomputersociety.org/10.1109/IV.2002.1028774. – ISSN 1093–9547

[RCB⁺05] ROBERTSON, George ; CZERWINSKI, Mary ; BAUDISCH, Patrick ; MEYERS, Brian ; ROBBINS, Daniel ; SMITH, Greg ; TAN, Desney: The Large-Display User Experience. In: *IEEE Comput. Graph. Appl.* 25 (2005), Nr. 4, S. 44–51. http://dx.doi.org/http://dx.doi.org/10.1109/MCG.2005.88. – DOI http://dx.doi.org/10.1109/MCG.2005.88. – ISSN 0272–1716

[Roj03] ROJDESTVENSKI, Igor: Metabolic pathways in three dimensions. In: *Bioinformatics* 19 (2003), Nr. 18, S. 2436–2441

[Sch03a] SCHREIBER, Falk: Comparison of metabolic pathways using constraint graph drawing. In: *APBC '03: Proceedings of the First Asia-Pacific bioinformatics conference on Bioinformatics 2003.* Darlinghurst, Australia, Australia : Australian Computer Society, Inc., 2003. – ISBN 0–909–92597–6, S. 105–110

[Sch03b] SCHREIBER, Falk: Visual Comparison of Metabolic Pathways. In: *Journal of Visual Languages and Computing* 14 (2003), Nr. 4, S. 327–340

[Shn96] SHNEIDERMAN, Ben: The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In: *VL '96: Proceedings of the 1996 IEEE Symposium on Visual Languages*, IEEE Computer Society, 1996. – ISBN 081867508X

[TSB04] TANG, Diane ; STOLTE, Chris ; BOSCH, Robert: Design choices when architecting visualizations. In: *Information Visualization* 3 (2004), Nr. 2, S. 65–79. http://dx.doi.org/http://dx.doi.org/10.1057/palgrave.ivs.9500067. – DOI http://dx.doi.org/10.1057/palgrave.ivs.9500067. – ISSN 1473–8716

[TSKS05] TANAY, Amos ; STEINFELD, Israel ; KUPIEC, Martin ; SHAMIR, Ron: Integrative analysis of genome-wide experiments in the context of a large high-throughput data compendium. In: *Molecular Systems Biology* 1 (2005), March, Nr. 1, S. msb4100005–E1–msb4100005–E10. http://dx.doi.org/10.1038/msb4100005. – DOI 10.1038/msb4100005

[WF94] WARE, Colin ; FRANCK, Glenn: Viewing a Graph in a Virtual Reality Display is Three Times as Good as 2D Diagram. In: *VL*, 1994, S. 182–183

[WGS00] WOLF, Detlef ; GRAY, Christopher P. ; SAIZIEU, Antoine de: Visualising Gene Expression in Its Metabolic Context. In: *Briefings in Bioinformatics* 1 (2000), Nr. 3, S. 297–304

[Wol05]   WOLFF, David: Using openGL in Java with JOGL. In: *J. Comput. Small Coll.* 21 (2005), Nr. 1, S. 223–224

[XYC05]   XU, Zhigen ; YAN, Yusong ; CHEN, Jim X.: OpenGL Programming in Java. In: *Computing in Science and Engg.* 7 (2005), Nr. 1, S. 51–55. `http://dx.doi.org/http://dx.doi.org/10.1109/MCSE.2005.20`. – DOI http://dx.doi.org/10.1109/MCSE.2005.20. – ISSN 1521–9615

[YEW+05]   YANG, Yuting ; ENGIN, Levent ; WURTELE, Eve S. ; CRUZ-NEIRA, Carolina ; DICKERSON, Julie A.: Integration of metabolic networks and gene expression in virtual reality. In: *Bioinformatics* 21 (2005), Nr. 18, S. 3645–3650. `http://dx.doi.org/http://dx.doi.org/10.1093/bioinformatics/bti581`. – DOI http://dx.doi.org/10.1093/bioinformatics/bti581. – ISSN 1367–4803

[YWCND06]   YANG, Yuting ; WURTELE, Eve S. ; CRUZ-NEIRA, Carolina ; DICKERSON, Julie A.: Hierarchical visualization of metabolic networks using virtual reality. In: *VRCIA '06: Proceedings of the 2006 ACM international conference on Virtual reality continuum and its applications.* New York, NY, USA : ACM Press, 2006. – ISBN 1–59593–324–7, S. 377–381

Index

bla, 15