IEEE Pacific Visualization 2010
Caleydo: Design and Evaluation of a Visual Analysis Framework for Gene Expression Data in its Biological Context

Alexander Lex, Marc Streit, Ernst Kruijff, Dieter Schmalstiege
Graz University of Technology, Austria
Help life science experts to understand function of genes

Help to understand biological processes and the graphs depicting them (pathways)

Help to identify marker genes for targeted treatment

Help understand the interdependencies between a concrete patient’s gene expression regulation and pathways
Gene Expression Data

- Measures how much of a gene’s functional product is produced
- One float value per gene, up to 30,000 float values per experiment

http://en.wikipedia.org
Pathway Data

- Models of cellular processes
- Subdivision into small graphs
- Widely-used public databases, containing about 800 pathways
  - KEGG (www.genome.ad.jp/kegg)
  - BioCarta (www.biocarta.com)
- Nodes are enzymes / genes
- Hand-crafted
- Encoded meta-information
- Available as XML + images
Web sources

- Gene databases like Entrez Gene
- Pathway meta-information
- Publications e.g. from PubMed
Caleydo was developed in the course of three projects that included funding for both life- and computer scientists

Close collaboration with several doctors and biologists

All want to understand their microarray data better
Goal: understand pathways and their interdependencies

- Pathways can be nested
- Pathways can share nodes
- Gene expression data helps understand a pathway
Goal: Find patterns in gene expression data that allow to reason about different conditions

- Knowledge about clinical factors essential
- Pathways can provide biological context for findings
Visualizing Gene Expression Data
Heat Maps

- Common representation
- Color coded gene regulation
- Clustered, including Dendrograms

[Seo2002] [Eisen1998]
Hierarchical Heat Map
Parallel Coordinates
The Bucket

Visualizing Pathway Interdependencies
To understand pathways it is necessary to:

- view the pathway itself
- view related pathways and identify interdependencies
- view meta-information
- view expression data for the pathway

Boils down to a multiple-view problem
How to manage these views?

Classical approach:
Multiple windows / tabs
New Concept: The Bucket

- Multiple layers:
  - Bottom → Focus
  - Walls → First level context
  - Rim → Second level context
The Bucket
Open Questions

- Does the Bucket help find related information compared to list-based methods?
- Is the distortion in the Bucket a problem?
- Do the visual links facilitate finding related items?
- Does a multiple-display setup (increased display space) have an effect on the results?
Setup

- 4 conditions
- 2 principal tasks (following the use-cases)
- 1 task to evaluate visual links
- 12 experts as participants (4 senior, 4 PhD, 4 master)

- 16 questions
- Open discussion
Hypotheses

- H1: The Bucket performs better than the list-based mode.
- H2: Multi-screen performs better than single screen, both in list-based and Bucket-mode.
- H3: The visual links are a significant aid in the identification of relevant information.
Results for Conditions

- Significant results (p<.001) for Bucket vs. list (for 7 out of 11 questions):
  - improved quality of context
  - ability to detect, compare and relate information

- Significant results (p<.05) for single vs. multi (for 4 out of 11 questions):
  - improved quality of context
  - relate information
Results for Bucket and Visual Links

- Graphics distortion in the Bucket was rated negligible
- Visual links were rated very helpful
Results Summary

- H1 (Bucket better than list): supported
- H2 (multiple better than single displays): not supported
- H3 (visual links help): supported
Conclusion

- Caleydo is implemented in Java and OpenGL
  - Runs on Linux and Windows systems
- Freely available for academic use at [www.caleydo.org](http://www.caleydo.org)
- Videos on: [www.youtube.com/user/caleydotugraz](http://www.youtube.com/user/caleydotugraz)
- Talk to me or Marc if you would like a demo

Thank you for your attention!
References


