Eurographics/ IEEE-VGTC Symposium on Visualization 2008
A. Vilanova, A. Telea, G. Scheuermann, and T. Möller
(Guest Editors)

Volume 27 (2008), Number 3

Navigation and Exploration of Interconnected Pathways

M. Streit[1] and M. Kalkusch[1] and K. Kashofer[2] and D. Schmalstieg[1]

[1]Institute for Computer Graphics and Vision, Graz University of Technology, Austria
[2]Institute of Pathology, Medical University of Graz, Austria

Abstract

Visualizing pathways, i. e. models of cellular functional networks, is a challenging task in computer assisted biomedicine. Pathways are represented as large collections of interwoven graphs, with complex structures present in both the individual graphs and their interconnections. This situation requires the development of novel visualization techniques to allow efficient visual exploration. We present the Caleydo framework, which incorporates a number of approaches to handle such pathways. Navigation in the network of pathways is facilitated by a hierarchical approach which dynamically selects a working set of individual pathways for closer inspection. These pathways are interactively rendered together with visual interconnections in a 2.5D view using graphics hardware acceleration. The layout of individual graphs is not computed automatically, but taken from the KEGG and BioCarta databases, which use layouts that life scientists are familiar with. Therefore they encode essential meta-information. While the KEGG and BioCarta pathways use a pre-defined layout, interactions such as linking+brushing, neighborhood search or detail on demand are still fully interactive in Caleydo. We have evaluated Caleydo with pathologists working on the determination of unknown gene functions. Informal experiences confirm that Caleydo is useful in both generating and validating such hypotheses.
Even though the presented techniques are applied to medical pathways, the proposed way of interaction is not limited to cellular processes and therefore has the potential to open new possibilities in other fields of application.

Categories and Subject Descriptors (according to ACM CCS): H.5.2 [Information Interfaces and Presentation]: User Interfaces and J.3 [Computer Applications]: Life and Medical Sciences

1. Introduction

In biomedicine, so-called pathways are used to model cellular processes and to describe their states and transformations. An example of a fundamental pathway is the *Citric Acid Cycle* where a series of chemical reactions perform a conversion of fats, proteins and carbohydrates to energy. Pathways are represented as directed, attributed graphs: Nodes represent enzymes or chemical compounds, and edges represent reactions or signals. It is characteristic for pathways that nodes can be involved several times within the same pathway, or in multiple pathways. One enzyme can for instance catalyze multiple reactions. Nodes can also represent complete pathways, leading to a recursive definition of the graph. A good definition of pathways as graphs is given in [KS07].

Traditionally, pathways are printed on large posters. The design of such a poster constitutes a huge effort, and the result is widely used in the community. Figure 1 shows the popular metabolic network published by Roche Applied Science [Mic99]. Important about these hand-crafted visualizations is the encoded meta-information in the layout. The layout conveys structure that is not evident from the graph topology or graph attributes. Illustrators use grouping or annotations (cf. figure 2) to convey meta-information that is essential for understanding and memorizing the pathways.

Posters are still widely-used as an inexpensive, static way of presenting pathway networks. However, the handling of huge posters is obviously tedious. Pathway databases such as KEGG (Kyoto Encyclopedia of Genes and Genomes) [Kan06] and BioCarta (`http://www.biocarta.com`) are therefore quickly replacing traditional paper based representations. One example can be seen in figure 2, which is thought to be behind *Long Term Depression*.

KEGG and BioCarta decompose the huge connected networks into smaller pieces relating to one individual pathway.

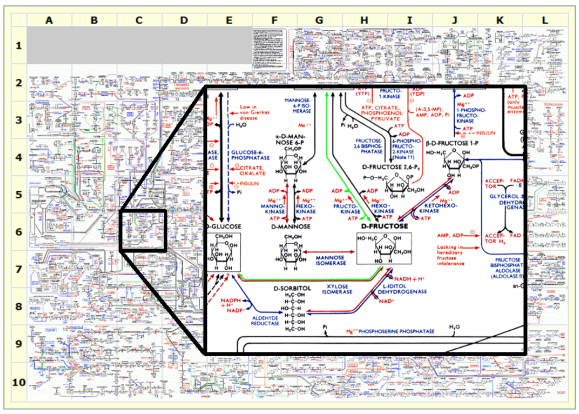

Figure 1: *Section of the Roche Metabolic Network. The graph is hand-routed. The poster comes with a booklet which contains an index that provides the mapping from entity to a sector on the poster.*

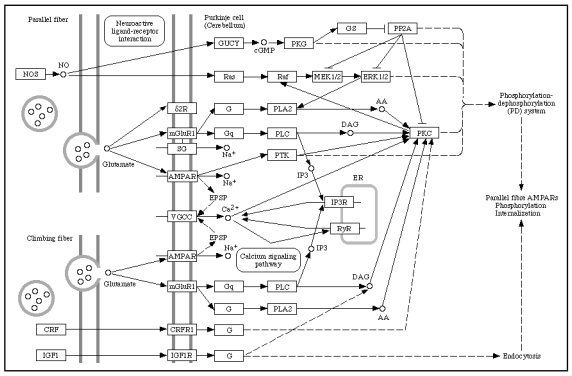

Figure 2: *KEGG pathway map modeling Long Term Depression. Remarkable are the structures in the background picture of the graph which provide important meta-information.*

The graph information of such a pathway is encoded in explicit form, but meta-information is still only available in raster image format and can only be used if the layout is retained. Moreover, online browsing of the pathway databases is complicated by the fact that nodes can occur multiple times within one pathway or in multiple related pathways. Therefore researchers must typically consider a working set of pathways at once, including the interconnections between pathways. Traditional visual exploration techniques, such as the hyperlinked navigation, do not support 1:n connections very well.

Our contribution is a set of techniques for fast and easy navigation inside the network of connected graphs. Caleydo is an interactive system for visual exploration of metabolic as well as signaling pathways. It respects the meta-information available in the hand-crafted pathway layouts but turns the static pathway layouts into fully interactive representations, which can be manipulated through techniques such as linking+brushing, neighborhood search or detail on demand. We address the problem of browsing through the hierarchical graph definition and managing the working set of relevant pathways by a set of hierarchically linked views. A study carried out by [SND05] identified the visualization of interconnectivities as a significant requirement for pathway research. We propose a 2.5D representation of the pathway working set, which makes inter-pathway connections evident and maximizes the use of screen real estate while avoiding clutter. While interaction and presentation retain the simplicity of 2D, we use 3D graphics acceleration to assemble the final visualization. Initial trials with pathologists indicate that our approach is well suited to typical problems of pathway analysis, such as finding missing or invalid relations in the network or the determination of unknown gene functions.

2. Related Work

2.1. Visualization of Pathways

A straightforward approach for the visualization of pathways is treating the whole network as one huge graph. Because of the high degree of interconnectedness, [Roj03] suggests the visualization of the pathway network as a 3D graph. *MetNetVR* [YWCND06] goes a step further by presenting a hierarchical visualization in a 3D CAVE environment. These 3D approaches suffer from the common problem that the user can easily get confused in a purely abstract, but complex 3D environment. In particular there is no obvious three-dimensional subspace in the data that can be used for natural organization of the domain. Moreover, Virtual Reality environments such as a CAVE are expensive. We believe that more constrained 2.5D methods are better suited to organize pathways. We were inspired by the work presented in [Dwy05, BDS04] where similar pathways are stacked on top of each other to visually differentiate them. However, their methods do not provide solutions for showing relations among the layers. In contrast, our approach uses the additional relation afforded by 2.5D to link multiple ordered pathway layers.

The high degree of data complexity makes it indispensable to use some sort of multiple coordinated views [NS00]. *SimVis* [PKH04, DGH03] is a good example for a framework that successfully applies the concept of linking+brushing between several views. A trend-setting strategy of connecting data entities by drawing edges between several views was presented in [SA06]. These concepts work well for a small number of connected views. However, since each pathway graph is a separate view, the space for placing the views side-by-side on the screen is the restricting factor. These considerations led us to adopt a 2.5D layout where the additional ordinal dimension can be used to manage and arrange the view representations.

Since KEGG and BioCarta use nested pathways (i. e., pathways as nodes inside pathways), the problem of visualizing multiple pathways can also be interpreted as a problem of browsing hierarchical graphs. [KS07] addresses this issue by combining multiple pathways in a single network that supports interactive level of detail change by expanding and collapsing the pathways from/to single nodes. While the system partly solves the users needs in terms of navigation, the relations between the graphs depicted by connecting lines still get cluttered easily.

2.2. Automated Layouting of Pathways

Independently of whether the chosen graph visualization techniques operate in 2D or 3D, a layout for the graph needs to be provided. Considering the complexity of the graphs, this task is a challenging research topic on its own. One of the first attempts to dynamically model metabolic pathways was done by Karp et al. starting in 1994 [KP94]. As pathway networks became more and more diversified and complex over time, graph drawing approaches needed to be employed and enhanced. In 2001 Becker et al. proposed an algorithm, that builds on the ideas of Karp and enhances them by including topological structure [BR01].

Even with a perfect layout, handling visualizations of huge graphs is challenging. A very large and dense representation is overwhelming and does often not allow easy discrimination of focus and context. Our work with life scientists revealed that they manage the complexity by memorizing patterns of well known pathways. Frequently pathways are immediately recognized by the user because of a particular layout, such as circularly arranged nodes. The *MetaViz* tool proposed by Bourqui et al. [Bou07] that creates a metabolic network using multiple pathways takes this fact into account. Moreover this valuable work also addresses the duplication problem of nodes by clever clustering and overlapped drawing of the graphs. However, recently contextual information such as localization of the compound in the cell or cellular structures have found their way into pathway layouts. For example, consider the pathway in figure 2, which denotes the cell border as two bold horizontal lines. BioCarta even draws pathways in a cartoon like style (cf. part 3 in 3). This trend towards a massive incorporation of meta-information is a strong argument against the application of automatic drawing and layouting in pathway networks.

2.3. Gene Mapping

Since the human genome as well as the genome of many other species is completely sequenced, the main thrust of scientific effort has moved towards the identification of gene functions [Pel01]. One used tool for the identification of gene functions are DNA microarrays [SSDB95]. A microarray analyzes a full genome profile from a tissue sample for a certain point in time. These so-called gene-expressions rate the activation of every analyzed gene with a numeric value. Microarray analysis can only measure transcription of genes and not the absolute expression value. Therefore the measured values are relative values that can only be compared to one another.

Pathways operate on the level of enzymes which are encoded by one ore more genes. The relation of enzymes to genes makes it important to display the pathway together with the current regulation of its encoding genes. Standard gene mapping approaches such as [LA02, TMK03, Mle05] present a vector of color coded gene-expressions, one for every gene associated with an enzyme. Commercial genetic data mining tools such as *GeneSpring* (Agilent Technologies, Inc., USA) and *PathwayStudio* (Ariadne Genomics, Inc., USA) use a static mapping of gene-expression data onto pathways. However, these approaches do not support the interactive visualization of relations among several pathways.

3. Methods

In this section we describe a combination of methods that allow efficient interactive exploration of pathway databases. The user operates with objects on two levels of granularity, individual pathways with a recognizable layout, and individual enzymes, which are manipulated to find identical or adjacent enzymes. In the following sections we describe the handling of the layout, the manipulation of pathways with the "jukebox" concept, the neighborhood visualization and the gene-expression mapping tools.

All methods presented in this paper are integrated in the Caleydo visualization framework (`http://www.caleydo.org`). The system is written in Java and supports various types of 2D and 3D representations. 3D rendering is implemented using the Java OpenGL library (JOGL). The framework supports linked views as well as system-wide synchronized interaction.

3.1. Superimposing Pathway Textures

Hand-crafted, annotated pathways with high recognition value are readily available in the KEGG and BioCarta databases. Rather than discarding this information in favor of a computer-generated layout, we use the pathway textures directly as an image-based background. Interactive representations of the nodes are overlaid so that the blended results have a uniform appearance. The pathway visualization is rendered as a flat 3D model with OpenGL, applying the pre-filtered pathway image as a texture to the ground plane. This allows seamless scaling of the visualization as well as arbitrary 2D and 3D transformations, which we employ in the jukebox as described in the next section.

3.2. Jukebox View

For the navigation inside the set of interconnected pathways we propose a four-stage concept:

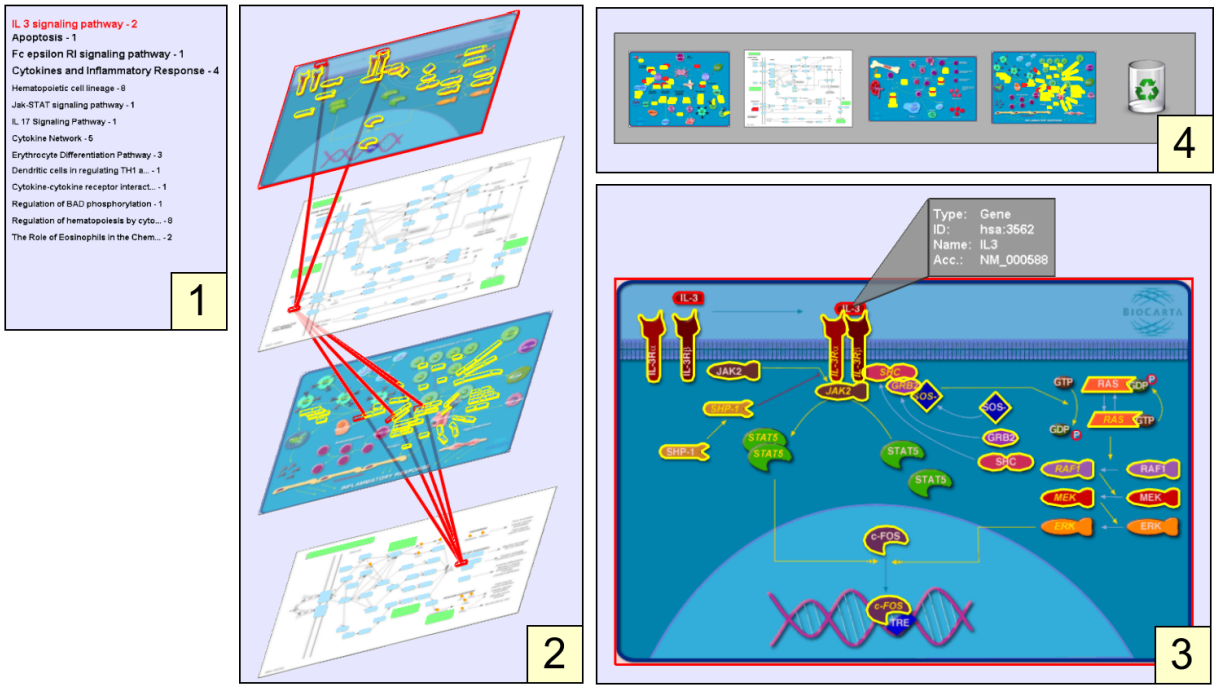

Figure 3: *The jukebox view combines a textual list menu for browsing related pathways by name (1) with an interconnected pathway stack (2), an area designated for a detailed examination of a graph (3) and a memo pad (4).*

1. Pathway pool list
2. 2.5D stacked layer view
3. Graph under interaction view
4. Memo pad

Figure 3 shows the four stages of the jukebox setup annotated in a screenshot. The setup works analogous to a jukebox where audio records can be selected from a larger collection and loaded to the turntable. A list of pathways is presented as a textual list containing the pathway names. By selecting an entry in the list the graph is loaded to the intermediate level where a predefined number of graphs is shown in a 2.5D stacked layer view. The third stage shows a large pathway for interactive inspection.

Caleydo employs a stack of graphs arranged in a 2.5D layout to densely pack pathway information in the available screen space, but also relate multiple planar graphs to one another. Individual pathways are scaled down and also compressed due to the orthographic (tilted) view. However, signature features and proportions familiar to the expert are retained, which makes the pathways still highly recognizable despite the compression. The relations between elements in different pathways are visualized using straight lines among the layers [Str07]. This approach is related to the *VisLink* approach [CC07] where the concept of inter-plane edge propagation was generalized from a graph specific towards a visualization independent solution. In contrast to *VisLink* the jukebox setup provides a solution for managing related 2D visualizations in a hierarchical way.

In the context of pathways, the connection lines among layers enable a fast identification of identical nodes in the whole network. Stage 3 in figure 3 shows the gene *IL3* that is selected by the user in the *IL 3* signaling pathway. Consequently all representations of that specific gene are highlighted and interactively connected. Selecting a particular gene allows the user to quickly determine its global relevance to various aspects of the working set. Circumstances where a gene appears in multiple pathways as part of an identical chain of reactions can also be discovered. The user can then choose a different pathway from the stack, which is exchanged with the pathway in the main interaction view. Moving a pathway up or down the hierarchy to another jukebox stage is visually supported by animated transformations [Sho85]. We found that users value the continuous transitions when the complex networks, which are scattered over multiple pathway views, demand their full attention.

There are multiple reasons why an entity can be represented multiple times in a set of pathways. One reason are layouting considerations. Another possibility is that a particular gene is catalyzing a specific reaction in a large variety of biological processes in the cell. As a consequence, genes appear several times in various pathways. Without global selection it is hard to identify such situations. We therefore

provide a mechanism that automatically searches the whole pathway pool for the selected entity. The resulting pathways are shown in the pathway list (stage 1) by displaying their names plus a score which is based on how often selected genes occur in that particular pathway. According to this score the most relevant pathways (e.g. highest score) are moved to the stacked pathway view, from where the user can continue to explore inter-pathway relationships.

For larger-scale problems, the automatic management of the stack based on a least frequently used policy can create the undesirable situation that a pathway vanished from the stack, but is later needed again and must be manually retrieved. We therefore provide the memo pad, an area of the screen where the user can place important pathways for semi-permanent safekeeping. Storing and retrieving pathways works by simple drag and drop (see stage 4 in figure 3). The memo pad does not only store the graph, but also the current selection of nodes, so that a particular working state can be completely restored instantaneously. The memo pad and the stacked pathway view are complementary: While the memo pad is designated to hold pathways persistently, the 2.5D layered view is a volatile stack that may be changed during the dynamic loading of dependent pathways.

3.3. Neighborhood Visualization

Algorithms for the calculation of adjacencies in graphs and their visualization are well researched. They are mostly applied to single, planar graphs. We implemented neighborhood visualization of arbitrary depth among the pathway boundaries. As mentioned before, nodes in pathways can be part of the same graph multiple times as well as be incorporated in other pathways. Caleydo combines the dynamic loading of dependent graphs with the highlighting of adjacencies. After the user selects a node the system presents all pathways that contain this entity. We apply Dijkstra's algorithm to all instances of the entity in the selected working set of pathways. As a result the neighborhood is propagated throughout all involved graphs. This extended adjacency visualization allows a comprehensive exploration among multiple pathways as shown in figure 4. Searching the neighborhood in all pathways simultaneously is extremely important, since it enables the user to visually reveal hidden biological dependencies and to detect reaction cascades in several pathways without stepping through all pathways manually.

The internal graph data structure facilitates the implementation of this approach. We manage a single master graph holding all logical data entities. At the first occurrence during the parsing of a pathway a data entity is created and inserted into the master graph. All visual representations are then registered with the entity in the master graph. Each further occurrence in the visual representation of another graph is only added to the corresponding list of representations.

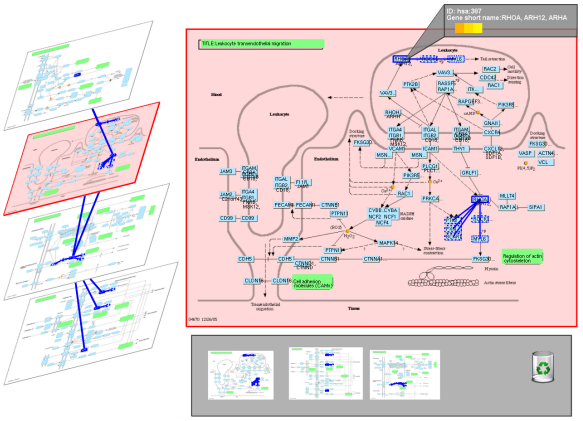

Figure 4: *The view shows the propagation of a signal in several pathways.*

3.4. Pathway Gene-Expression Mapping

The enzyme nodes in the pathways represent special proteins that perform certain tasks in the cell. These enzymes can be encoded by several genes. Vice versa, a single gene can be involved in the encoding of multiple enzymes. This *n:m* relation between enzymes and genes must be considered in the design of the graph data structure.

Moreover, it is common that gene-expressions are collected in a time-series experiment. Hereby the numeric regulation value of a gene in a tissue probe is sampled at different

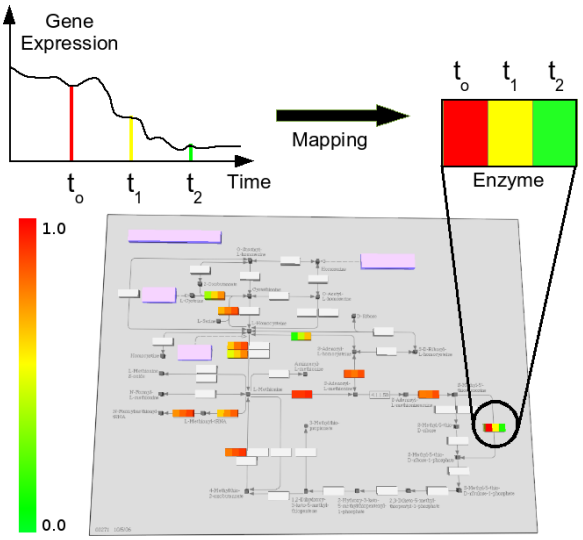

Figure 5: *The expression of gene-expression experiments, a snapshot of the regulation of genes at a certain time, are mapped onto the pathways. In this case a three part time-series experiment is color coded directly onto the nodes.*

points in time. Such a dataset can be shown as a bar of color coded icons (figure 5). If multiple genes contribute to the encoding process of that enzyme, a star layout is used. Nodes that have multiple gene mappings available are depicted in a predefined color in the pathway. When a mouse-over action above the enzyme node is triggered an animated star containing the requested information about the genes is opened.

3.5. Integrated Meta-Data Browser

The predominant part of the time when working with pathways and genetic data is consumed by the investigation of meta-data. This knowledge is essential to understand the complex networks and processes. Following a detail on demand technique all nodes in our system are linked to an information browser. The browser is connected to major database websites and shows detailed information about genes, enzymes, protein structures and other entities. In addition the system provides publications about chosen entities on demand in a designated publication browser. To achieve that we pack the selected data entity in a search query on *PubMed*, an extensive citation source that accesses *MEDLINE* and other life science journals. In contrast to the well established linked meta-information browser where the browser is a passive data output facility, we implemented active ID and name parsing of all system-wide known entities. In our medical use case the system detects pathways as well as genes in selected hyperlinks inside the integrated browser. Pathways are opened instantly. Genes or enzymes that were found are highlighted throughout the system. The system dynamically searches all pathways that contain the entity and loads them to the jukebox.

4. Use Case

In the following a use case is presented to explain the requirements of the system from the perspective of a life scientist. During the design phase of Caleydo the users were highly integrated in the requirements elicitation.

The gene pool can be divided into several subsets. A part of the genes, called disease genes, are known to be involved in one or more diseases. Others are accurately referred to as candidate genes because they are suspected to be at least one of the causing factors of a disease. Others are widely unknown in their function. This observation is reflected in our use case which assumes that the user has a predefined set of candidate genes.

The user starts by loading the gene-expression experiment data into the system. Each experiment consists of a complete snapshot of a genome. The gene data can stem from humans as well as other organisms, such as mice or rats. One tissue snapshot is called experiment. A common data base for analysis could consist of healthy human tissue, tissue from various cancer patients as well as tissue from a cultured carcinoma in an animal model.

Before the gene-expression data is mapped onto the pathways, statistical preprocessing needs to be applied to the data. Our partners are primarily interested in highly differentiated gene-expression values. In our context, differentiated refers to e.g. a high expression value for the healthy tissue and a low one for the carcinoma tissue. Therefore the overall gene pool of several thousand entities is filtered. Although statistical analysis is part of the Caleydo system, it is not in the scope of this paper. The set of suspect genes are automatically mapped to all KEGG and BioCarta pathways. By selecting the genes in the linked heat map view, the system performs the dynamic loading of pathways in the jukebox setup where the entity is involved. Usually the user knows certain pathways well, while other pathways are unfamiliar. By performing a detailed inspection of the biological context, the life scientist aims for deducing new gene functions related to the examined disease. The integration of the adjacency visualization among single pathway boundaries supports this process significantly.

In this phase, the user switches between the pathways by using the jukebox's functionality. During this process, the already filtered candidate gene set is further reduced by entities that were found to be irrelevant in the experiment's context. For example, a gene might perform house keeping functions in the cell, that are well known and therefore not interesting for the current analysis. In another biological context, i.e. in another pathway, the same gene carries out completely different tasks that could be of great interest. This analysis style is only possible with a system that supports the visual exploration like Caleydo does.

Figure 6 provides an exemplary workflow documented by a series of 4 screenshots. First the user triggers a search action for the *PTK2* gene that is, according to the statistical analysis, suspected to be relevant for a disease from which the experimental tissue samples are taken. The system loads all pathways, that contain the selected gene, to the jukebox setup (cf. part a in figure 6). The user starts to investigate the *Erb* signaling pathway in detail. Obviously the *PTK2* gene is located at the end of a signal cascade. By performing the in-depth adjacency visualization, previously unknown relationships emerge (cf. part b). By switching to the *Focal Adhesion* pathway from the top of the stack (see part c) the user can further investigate the neighboring genes. In part d, the user selects an adjacent enzyme of *PTK2* on which multiple genes are mapped. The system again performs the dynamic pathway search and loads a new set of pathways into the jukebox setup.

5. Discussion

When working in biomedicine one has to keep in mind that the underlying data is imperfect. Relations inside pathways are extracted from scientific publications. Therefore pathways are subject to a constant update process: new relations appear, existing connections are changed or invalidated, and

many entities and relations are currently still unknown. Also the data mapping of genes to enzymes is incomplete or even incorrect. The user must be aware that the data visualization is only a snapshot of the current state of knowledge.

Although the presented methods have the potential to facilitate the knowledge acquisition in huge relational networks the system is subject to restrictions. A good portion of the pathways has a size and complexity which is suitable for our design. However, some of the graphs are degenerated in size, which is problematic for the stacked visualization. The mixing of extraordinary small graphs with big graphs can be aesthetically displeasing, and even disturb the user's ability to interpret the visualization. These effects can be mitigated by adaptive scaling, but only to a limited degree.

We tested the jukebox setup in multiple configurations varying in the degree of scene customization permitted to the user. It turned out that the most restrictive setup was perceived best. While a rotation of the pathway stack can give a better perception of the line connections between the layers, most of the users' became disorientated. According to that result the stack planes are tilted in a fixed angle of 60 degrees and the camera cannot be altered by the user. During the tests with expert users we turned off the connection lines between the graph planes in the stacked visualization while only using linking+brushing for highlighting the selected nodes. A predominant part of the users complained about the missing edges between the layers. A user study could further investigate whether linking+brushing or the direct connection of selected entities in the scene leads to better results in perception. Furthermore user tests showed that the maximum stack size should not exceed five planes. Otherwise users began to feel overwhelmed by too many graphs at the same time. Nevertheless this restriction of the stack is inherently absorbed by the jukebox's hierarchical staging concept.

Another open issue is the determination of a reasonable color coding. One widely-used color scheme for visualizing gene-expression values ranges from red (high regulated gene) over yellow (mean) to green (low regulated gene). The binding of 2/3 of the RGB color space exhausts the possible uses of color for conveying additional information.

It is in the nature of information visualization that the available number of pixels is a restricting factor. We run the system on a high resolution multi-projector wall as well as on various desktop setups. The usual work environment of expert users that deal with pathways are clinical facilities and laboratories where the technical resources are limited. Only a few of them will have access to expensive multi-projector facilities. We are therefore investigating the combination of multiple monitors with a low-cost video projector for a physical focus+context representation.

6. Conclusions and Future Work

In this paper we have presented an efficient way of navigating and exploring a network of interlinked multiple graphs to discover previously unknown knowledge. The system itself is capable of visualizing all kind of networks that are split up into interconnected sub-graphs. Therefore the proposed way of interaction with these graphs lends itself to the application to other domains. In the field of software design our approach could help to understand the complexity of huge software systems. A network of UML diagrams consists of entities like classes, objects and others. Similar to pathways these software engineering graphs can contain identical entities multiple times as well as diagrams can be contained in other diagrams. These properties open promising possibilities in this field of application. Also economical processes as well as social networks fit the presented techniques.

7. Acknowledgements

The authors want to thank Kurt Zatloukal and his team from the pathology group at the Medical University of Graz for the close and fruitful collaboration. We would also like to express our gratitude to Alexander Lex for his valuable input. This research was partly funded by the FIT-IT program of the Austria research funding agency FFG and the Austrian Genome Program (GEN-AU).

References

[BDS04] BRANDES U., DWYER T., SCHREIBER F.: Visualizing related metabolic pathways in two and a half dimensions. In *Graph Drawing* (2004), pp. 111–122.

[Bou07] BOURQUI ET AL.: Metabolic network visualization eliminating node redundance and preserving metabolic pathways. *BMC Systems Biology 1*, 29 (2007).

[BR01] BECKER M. Y., ROJAS I.: A graph layout algorithm for drawing metabolic pathways. *Bioinformatics (Oxford, England) 17* (2001), 461–7.

[CC07] COLLINS C., CARPENDALE S.: Vislink: Revealing relationships amongst visualizations. *IEEE Trans. Vis. Comput. Graph. 13* (2007), 1192–1199.

[DGH03] DOLEISCH H., GASSER M., HAUSER H.: Interactive feature specification for focus+context visualization of complex simulation data. In *Data Visualization* (Grenoble, France, 2003), pp. 239–248.

[Dwy05] DWYER T.: *Two and a Half Dimensional Visualisation of Relational Networks*. PhD thesis, The University of Sydney, 2005.

[Kan06] KANEHISA ET AL.: From genomics to chemical genomics: new developments in kegg. *Nucleic acids research 34* (2006), D354–7.

[KP94] KARP P. D., PALEY S. M.: Automated drawing of metabolic pathways. In *Proceedings on Bioinformatics and Genome Research* (1994).

[KS07] KLUKAS C., SCHREIBER F.: Dynamic exploration and editing of kegg pathway diagrams. *Bioinformatics 23* (2007), 344–350.

[LA02] LINDROOS H., ANDERSSON S.: Visualizing metabolic pathways: comparative genomics and expression analysis. *Proc. of the IEEE 90* (2002), 1793–1802.

[Mic99] MICHAL G.: *Biochemical Pathways. Biochemie-Atlas*. Spektrum Akademischer Verlag, 1999.

[Mle05] MLECNIK ET AL.: Pathwayexplorer: web service for visualizing high-throughput expression data on biological pathways. *Nucl. acids res. 33* (2005), W633–7.

[NS00] NORTH C., SHNEIDERMAN B.: Snap-together visualization: a user interface for coordinating visualizations via relational schemata. In *Advanced Visual Interfaces* (Palermo, Italy, 2000), pp. 128–135.

[Pel01] PELLEGRINI ET AL.: Computational method to assign microbial genes to pathways. *Journal of Cellular Biochemistry. Supplement 37* (2001), 106–9.

[PKH04] PIRINGER H., KOSARA R., HAUSER H.: Interactive focus+context visualization with linked 2d/3d scatterplots. In *Coordinated & Multiple Views in Exploratory Visualization (CMV)* (2004), pp. 49–60.

[Roj03] ROJDESTVENSKI I.: Metabolic pathways in three dimensions. *Bioinformatics 19* (2003), 2436–2441.

[SA06] SHNEIDERMAN B., ARIS A.: Network visualization by semantic substrates. *IEEE Transactions on Visualization and Computer Graphics 12* (2006), 733–740.

[Sho85] SHOEMAKE K.: Animating rotation with quaternion curves. In *Computer Graphics and Interactive Techniques* (1985), pp. 245–254.

[SND05] SARAIYA P., NORTH C., DUCA K.: Visualizing biological pathways: requirements analysis, systems evaluation and research agenda. *Information Visualization 4* (2005), 191–205.

[SSDB95] SCHENA M., SHALON D., DAVIS R. W., BROWN P. O.: Quantitative monitoring of gene expression patterns with a complementary dna microarray. *Science 270* (1995), 467–70.

[Str07] STREIT M.: *Metabolic Pathway Visualization Using Gene-Expression Data.* Master's thesis, Graz University of Technology, 2007.

[TMK03] TOYODA T., MOCHIZUKI Y., KONAGAYA A.: Gscope: a clipped fisheye viewer effective for highly complicated biomolecular network graphs. *Bioinformatics 19* (2003), 437–438.

[YWCND06] YANG Y., WURTELE E. S., CRUZ-NEIRA C., DICKERSON J. A.: Hierarchical visualization of metabolic networks using virtual reality. In *Virtual Reality Continuum and its Applications* (2006), pp. 377–381.

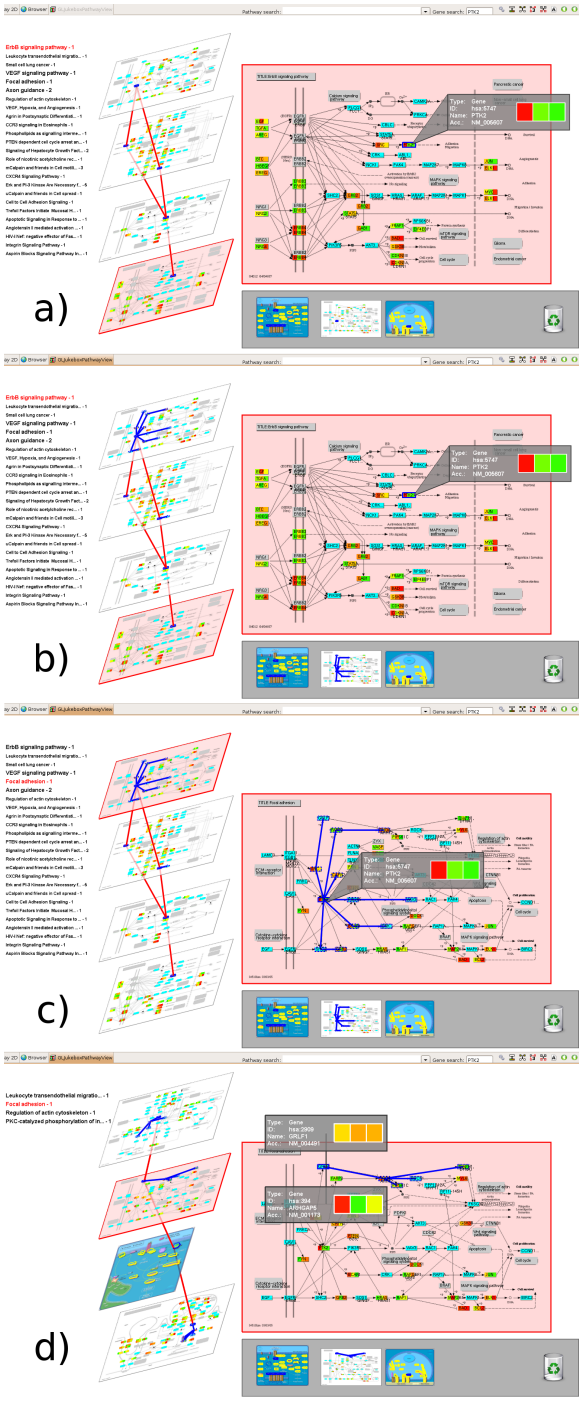

Figure 6: *The series of screenshots depicts an exemplary part of the visual exploration process in Caleydo. In (a) the system presents the pathways for the PTK2 search query. By investigation of the adjacencies in the graph stack the user can identify other genes connected to PTK2 which are not present in the local pathway context (b). In (c) the user switches to the topmost graph in the stack. (d) shows multiple genes mapped on a neighboring enzyme.*