# Caleydo: Design and Evaluation of a Visual Analysis Framework for Gene Expression Data in its Biological Context

Alexander Lex* Marc Streit† Ernst Kruijff‡ Dieter Schmalstieg§

Graz University of Technology

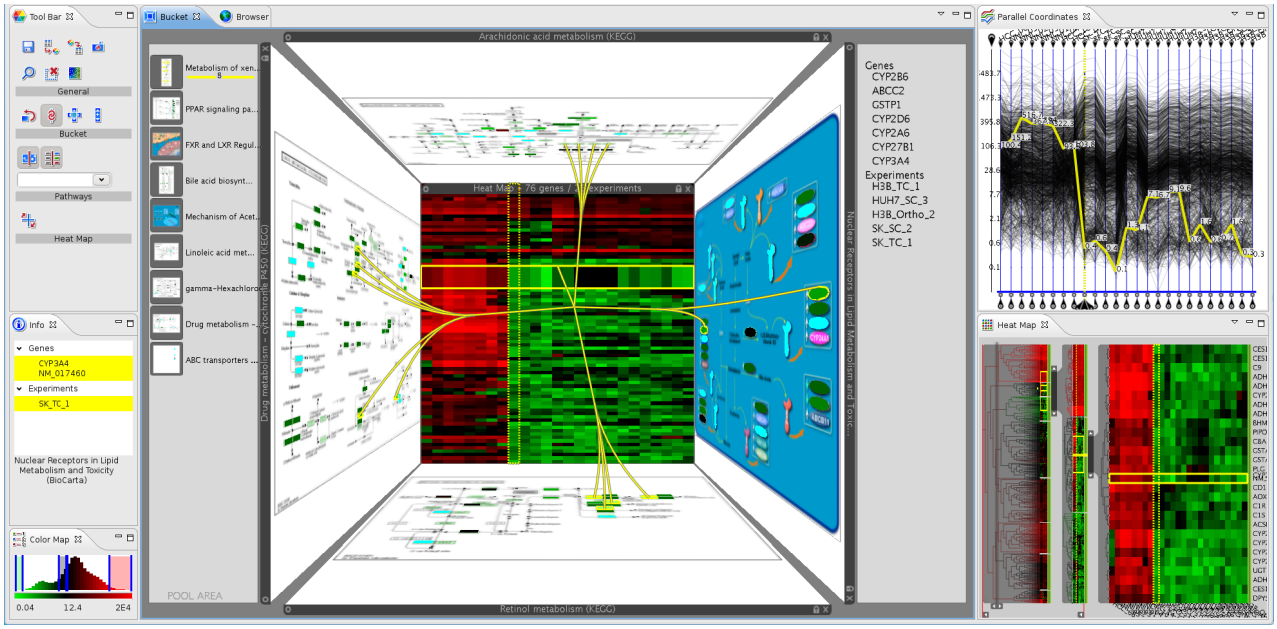



Figure 1: Screenshot of *Caleydo* with open *Bucket* view, a parallel coordinates view, a heat map and some meta-information. The *Bucket* concept is an integral part of *Caleydo* and allows us to place views for pathway and gene expression analysis in a 2.5D setup. Relations between views are shown by means of visual links.

## Abstract

The goal of our work is to support experts in the process of hypotheses generation concerning the roles of genes in diseases. For a deeper understanding of the complex interdependencies between genes, it is important to bring gene expressions (measurements) into context with pathways. Pathways, which are models of biological processes, are available in online databases. In these databases, large networks are decomposed into small sub-graphs for better manageability. This simplification results in a loss of context, as pathways are interconnected and genes can occur in multiple instances scattered over the network. Our main goal is therefore to present all relevant information, *i.e.*, gene expressions, the relations between expression and pathways and between multiple pathways in a simple, yet effective way. To achieve this we employ two different multiple-view approaches. Traditional multiple views are used for large datasets or highly interactive visualizations, while a 2.5D technique is employed to support a seamless navigation of multiple pathways which simultaneously links to the expression of the contained genes. This approach facilitates the understanding of the interconnection of pathways, and enables a non-distracting relation to gene expression data. We evaluated *Caleydo* with a group of users from the life science community. Users were asked to perform three tasks: pathway exploration, gene expression analysis and information comparison with and without visual links, which had to be conducted in four different conditions. Evaluation results show that the system can improve the process of understanding the complex network of pathways and the individual effects of gene expression regulation considerably. Especially the quality of the available contextual information and the spatial organization was rated good for the presented 2.5D setup.

**Keywords:** Bioinformatics, Visual Analysis, Pathways, Gene Expression, Multiple Views, Linking & Brushing.

**Index Terms:** H.5.2 [Information Interfaces and Presentation]: User Interfaces—and J.3 [Computer Applications]: Life and Medical Sciences—

## 1 Introduction

To understand the function of genes, it is necessary to study their biological context. In which biological processes is a gene involved? Is it involved in multiple similar processes? Pathways, representations of such processes, are consulted to provide answers to these questions. Pathway data can either be presented as a large, complex

*e-mail: lex@icg.tugraz.at
†e-mail: streit@icg.tugraz.at
‡e-mail: kruijff@icg.tugraz.at
§e-mail: schmalstieg@icg.tugraz.at

network with an automated layout, or as small functional graphs, handcrafted by experts. These small pathways frequently encode meta-knowledge, such as cell localization, in the layout. Widely used pathway databases are *KEGG* [11] and *BioCarta*[1], which together contain about 600 pathways.

There are advocates of both approaches, the large single network and the multiple small graphs. Automatic layouting of large graphs is beneficial when the goal is to modify the graph interactively, or when the subdivision of the network into predefined pathways is unwanted [6, 10]. However, the additional information implicitly encoded in the hand-drawn small graph layout turns out to be equally important for other problems of biomolecular analysis. In discussions with our biomedical focus group, we learned that the superiority of the hand-crafted layout and familiarity with the existing pathways cannot be matched by automated layouts. Nonetheless, the members of the focus group also confirmed that they had difficulties understanding inter-pathway dependencies when navigating databases of many small pathways.

Pathways are generic models which are valid for a whole species. Individual effects, such as diseases, can only be interpreted by relating gene expression data (or other kinds of biomolecular data) to pathways. Regulation of gene expression describes how actively a gene produces its functional gene products, for example RNA or proteins. In one experiment with DNA-microarrays, up to 30,000 gene expression values are acquired. Usually several experiments are studied simultaneously. Typical methods used to visualize gene expression data are heat maps [8] and parallel coordinates [6]. Regulation is relevant for understanding the general role of a gene and its particular role in a disease. A significantly down-regulated gene at the beginning of a pathway can in fact make all the following nodes irrelevant. Therefore, when the effects of diseases on pathways are studied, simultaneous consideration of gene expression information and pathways is crucial.

### 1.1 Typical workflow

After being approached by our partners from the life science domain, we analyzed the goals they were trying to achieve, and discovered two distinct workflows: The first is a pathway-centric approach, the second concerns the analysis of gene expression data with hypothesis generation and quick plausibility checks.

In the pathway-centric approach, the expert is interested in a specific biological process, like the development of colorectal cancer. The starting point for an analysis could be the *KEGG* colorectal cancer pathway. The user explores the interdependencies of this pathway with other pathways. When simultaneously exploring the pathway and gene expressions from multiple samples of cancerous tissue, the expert can detect differences in the gene expressions of groups of samples. Such a variation can indicate different sub-types of the disease or response to treatment in a time series analysis.

In a gene-expression-centric approach, the expert analyzes the expression data first. In this case, knowledge about the clinical factors that distinguish the different experimental conditions or patients is essential. For example, an expert could arrange the data in such a way that patients with short disease-free survival are grouped. He then looks for differentially expressed genes, supported by filters and analytical tools such as clustering. Such evidence may lead to a hypothesis, which can be checked for plausibility by analyzing the biological context (*i.e.*, pathways or literature) of the differentially expressed genes. Only plausible hypotheses are subjected to expensive clinical studies.

### 1.2 Contribution

We designed *Caleydo*, a visualization system that addresses both of these workflows, based on the simultaneous consideration of gene expression information and pathways. To our knowledge, it is the first approach that allows life scientists to explore relationships between multiple, handcrafted pathways, and the relationship of pathways to actual measurements of gene expression regulation directly. With this aim in mind, we developed a system that provides powerful tools for gene expression analysis, as well as the *Bucket*, a tool for pathway analysis where multiple views are presented in a 2.5D arrangement. This setup lends itself to cross-referencing views (*i.e.*, pathways and heat maps) by visual links. We report on the design and implementation of *Caleydo*, and an evaluation performed with 12 experts from the life science domain. The evaluation shows that the proposed *Bucket* arrangement is preferred over traditional list-based methods, especially in the areas of *context quality* and *spatial organization*. Participants stated that the required concentration was lower when using the Bucket. We also found that visual links significantly improve the ability to search for relevant information.

## 2 Related work

Many visualization tools for the exploration of pathways have been developed. Web tools, like *KEGG*, are largely based on lists and hyperlinks. One example of a commercial tool is Pathway Studio[2]. A more advanced approach, described in [12], starts with *KEGG* pathways and re-arranges and interconnects them, resulting in a mixture of static layouts of hand-routed pathways and automatically drawn layouts. The latter approach works well for a small number of pathways (2-3), but as the number increases, nodes become small and too many links between pathways result in visual clutter.

The problem of exploring large graphs has been subject of extensive research. Especially virtual reality systems have been investigated [3, 23], following the rationale that 3D graphs can contain more information and are easier to use with stereo vision [3]. While we consider the use of complex systems like a CAVE or multi-display environments an interesting subject, we also believe that in order for *Caleydo* to be widely adopted, we need to provide software that runs in standard office environments.

Besides pathways, gene expression data is equally relevant for our use cases. A common method of visualizing gene expressions are clustered heat maps [8]. One prominent implementation is the *Hierarchical Cluster Explorer (HCE)*, a rich framework for dynamic querying of gene expression data [19]. Parallel coordinates have been used to visualize gene expressions as well [17, 6]. These works use parallel coordinates mainly for visualization, while we employ them primarily as a selection and filter tool.

Bringing gene expressions into context with pathways has been recognized as important for several years now. Both research [13, 15, 21] and commercial tools like *GeneSpring*[3] favor an approach involving augmenting a node in the graph by color-coded rectangles, each rectangle representing an experiment. This approach works under certain conditions for a small number of experiments ([13] claims eight). Due to the tiny size of the node, however, more experiments become indistinguishable. Moreover, the text on the node itself is completely occluded, and the method is only usable for rectangular nodes, such as those used in *KEGG*, but not compatible with free-form shapes, such as those common in *BioCarta*. Finally, *KEGG* contains many nodes that are encoded by multiple genes, which is impossible to visualize by color-coding the node. Workarounds for this problem are the provision of gene lists with the expression encoding (used in GeneSpring) or tool-tips [21]. As an alternative to this approach, *Cerebral* [6] uses several small views of the visualized graph where each view corresponds to one experimental condition. While this approach does require a lot of screen space and therefore does not scale to larger numbers of experiments, it works well for a smaller (less than 20) number of small graphs that do not have multiple genes encoding one node. Com-

[1] http://www.biocarta.com

[2] http://www.ariadnegenomics.com/products/pathway-studio/

[3] www.agilent.com/chem/genespring

parability between the different experiments suffers, however, since each small view encodes only one experiment.

We believe that a combination of augmentation and using multiple views with visual linking provides a crucial benefit. Traditional linking & brushing cannot always capture the complexity of interrelations between multiple views. Therefore, recent research has introduced the general idea of visual links (i.e., drawing edges) between multiple views. The concept was used in [20, 1], and later extended to 3D and generalized for different visualization techniques in [7]. However, these examples leave the efficient use of available screen space and the convenient navigation of a focus+context hierarchy to the user. We investigated the use of visual links between pathways in [21], laying the groundwork for the solution presented in this paper.

The features of *Caleydo* are outlined in an application note in [22]. The work in [16] discusses a workflow for using biobanks, where *Caleydo* is used for one of the steps. In contrast to these works, this paper focuses on the Bucket visualization method and its analysis in a quantitative user study.

## 3 Overview of the Visualization System

*Caleydo* is a tool devised to employ visual methods where otherwise statistics is used. The different visualizations have distinctive application domains, which can be split up along the two data domains our system connects: pathways and gene expression analysis. Other views, for example an integrated linked browser that queries common websites for literature and information about genes and pathways, or a histogram also used to adapt the color coding, provide meta-information on the data.

The framework is written in Java and uses the Java OpenGL Toolkit (JOGL) for rendering 3D content. It facilitates a rapid prototyping approach for the integration of new visualization techniques. The 3D layout of 2D views is based on "style-sheets", and can accommodate various strategies for arranging them. The application can completely save and restore its state (*i.e.*, loaded data, data filters, selections, view arrangements).

### 3.1 Gene expression analysis

Gene expression data is usually delivered in tabular form. In the bioinformatics community, statistical tools are used to search for differences in the data and visualization is usually limited to static plots. To overcome this, we employ two techniques: parallel coordinates and an interactive, hierarchical heat map.

#### Parallel coordinates

Our state-of-the-art parallel coordinates implementation (*cf.* Figure 1 bottom) allows for the free arrangement of axes and various selection strategies. Not only the brushing, but also the arrangements of genes and experiments are linked. If a user decides to rearrange the axes in the parallel coordinates to group similar experiments together, this grouping is immediately reflected in all other views.

The parallel coordinates are a suitable tool to prefilter data for algorithmic methods, such as clustering. By filtering out inconspicuous genes, the dataset can be reduced to areas of magnitude that can be processed by average hardware. Once this is achieved, the user can apply one of several clustering algorithms, *i.e.*, k-means, affinity propagation [9] and the hierarchical algorithms discussed in [8], by using one of several distance measures. Due to the tendency of genes involved in the same functional process to be co-regulated, the clustering techniques assemble genes with similar functions in groups [8].

#### Heat map

A heat map [8] uses a color code for displaying gene expression regulation of several experiments. In the domain-specific convention for the color coding, green represents down-regulated genes, black a similar regulation to the reference experiment and red indicates up-regulation. Our system allows the expert to arbitrarily modify the color coding, thus providing colors suitable for red-green blind users. The elements are spatially ordered, so that all values for one gene are in one row, and all values for one experiment are in one column (or vice versa).

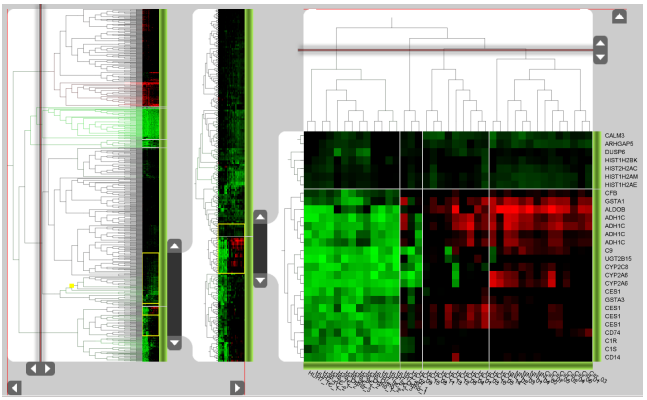


Figure 2: Hierarchical heat map showing 851 genes in three levels. Both experiments and genes have been clustered hierarchically, and dendrograms are shown in both dimensions. The desired level of granularity of groupings can be adjusted by dragging the cut-off line of the dendrogram. In the case of hierarchical clustering algorithms, the groupings are determined by the cut-off value's position whereas partitional clusterers assign them automatically.

Traditional heat map visualizations are often degenerated in size and require a lot of scrolling or panning. To overcome this limitation, we have developed a hierarchical heat map (*cf.* Figure 2), employing a focus+context approach reminiscent of [5]. Up to three levels of detail are shown simultaneously, where the first presents an overview of the whole dataset and the last visualizes an enlarged detail view of the data, where individual gene names are readable.

Additionally, if hierarchical clustering has been used, dendrograms are shown on all levels. Cluster borders are visualized with grey lines. A cluster is treated as a semantic group, and supports operations such as searching for pathways that contain several of the genes in the group. The grouping can be manually adapted.

### 3.2 Pathway analysis with the *Bucket*

Multiple linked views have proven their value for thoroughly comprehending complex data sets [4]. By interactively updating corresponding data in all views simultaneously, the investigation of interrelated aspects of a problem becomes feasible. However, the presentation of views side by side is restricted by the available screen space. High-resolution displays and multi-monitor configurations can increase the number of available pixels, but are ultimately limited by the maximum angle conveniently observed by a human. Clearly, novel compact viewing arrangements are required. We therefore developed a spatial setup of multiple 2D visualizations embedded in a 3D scene which we call the *Bucket* (see Figure 1). We use the *Bucket* to show pathways and contextual gene expression information in a heat map. The heat map in the *Bucket* contains only those genes that occur in at least one of the pathways. By clustering and sorting the genes every time a pathway is added or removed, the genes with the highest value (*i.e.*, the highest average of a cluster over all experiments) are always on top (*cf.* Figure 1). There are several ways to load pathways into the *Bucket*, for example by keyword search for a specific pathway, or by loading a pathway containing a particular gene. The new pathways are placed in the *Bucket*, where the relations can be explored.

### *Bucket* concept

The *Bucket* is a metaphor for a view arrangement where multiple related views are rendered on the inner sides and the rim of a square bucket. The users' viewport is restricted to a top view into the *Bucket*. The bottom of the *Bucket* contains the view in focus. Contextual views are rendered onto the second level, the *Bucket* walls. A third level, the rim of the *Bucket*, holds down-scaled, linked view representations that are related, but not currently in the user's focus, as well as genes, experiments or pathways that have been bookmarked previously.

To bring different views into focus, the user can move them up and down the levels. A very restrictive set of navigation operations turns out to be sufficient, providing the benefit of low cognitive load during navigation. VisLink [7] addresses the issue of navigating in a 3D multi-view arrangement by providing hotkeys for predefined camera positions, while still allowing full 3D navigation. However, our experiences indicate that a more restricted approach is beneficial – the nature of the *Bucket* layout does not require full navigational freedom. Views can simply be moved by drag and drop.

The individual views are aware of their current position, showing details accordingly. For example, text in the heat map is only shown when it is zoomed.

The *Bucket* arrangement of views in a 3D scene takes advantage of the spatial dimension by using it for multiple levels of focus+context. The visual arrangement loosely resembles the Perspective Wall [14], which also applies view stretching and shrinking as a distortion technique. However, we do not use the walls of the *Bucket* for contextual information drawn from the same visualization. Instead, we present separate, but interrelated views in a space-saving arrangement which lends itself to visual linking due to the compact hierarchical arrangement of views.

During the development of the *Bucket* concept, we experimented with different bucket shapes. We decided to use the variant with a square bottom because of its simplicity and efficient use of screen space. Unlike a hexagon or octagon, a square does not waste space in the corners assuming a rectangular shape of the views. The square allows one focus and four contextual views, which we found sufficient for most problems. The *Bucket* adapts to the available window by unfolding, if possible. This results in less perspective distortion for the side views when used with landscape-type screen resolutions, but also in unused screen space in the corners.

### Zooming

The zoom feature (see Figure 3) is restricted to two predefined z-values, which were found sufficient after some experimentation. It enables the most detailed inspection of and interaction with the visualization in focus. A zoomed visualization shows all available detail, such as labels and UI elements. A zoom action is triggered by turning the mouse wheel. It is visually supported by an animated camera flight. The contextual views from the wall are placed either to the right or on top of the focus view, depending on the window geometry, thus preserving the contextual information. The rim, containing a list of bookmarked genes (right) and the other related pathways (left), is still visible.

### Visual links in the *Bucket*

The main goal of the *Bucket* is to visualize the relations between views and the properties of a selected entity. We do this by drawing lines between the elements in multiple visualizations. In the case of genes, we have multiple occurrences in several different views, since a gene can occur in several pathways. One property of a gene is its expression regulation. In contrast to [7], our views do not contain links between similar entities, but between different representations of the same element.

The intensive use of visual links is very sensitive to optimal spatial positioning of the views relative to each other. Previous

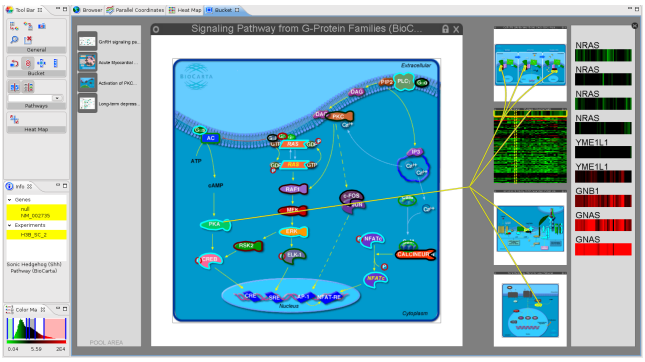



Figure 3: The *Bucket* when zoomed in, showing a 2D arrangement of focus and context views.

work [7] defers the optimal placement of views to the user. While this approach does allow flexible setups, it is not necessarily the most efficient, since the user may spend a significant amount of time simply trying to find the optimal placement in 3D space.

In previous work [21], we proposed a stacked graph setup, where different views are tilted into 3D and placed on top of each other. Views can be moved from the stack to a detailed focus view. This approach, however, suffered from two major drawbacks:

1. Relations between views that were not adjacent in the stack were not directly connected. Collins and Carpendale [7] also identified this issue as a yet unsolved problem for visually linked views.
2. There is no visual link between the view in focus and the views in the stack.

The *Bucket* avoids these pitfalls. It maximizes the opportunity of linking between the bottom and wall views with short and direct links. According to the requirements of biomolecular data analysis, the visual links connect different properties of one single gene currently under investigation.

We use multi-level edge bundling to reduce visual clutter. Edges are bundled first on a per view basis, which is important since a view often contains multiple entries. The bundled nodes from the views are then joined in a common point calculated on the fly.

### Suitability of visualization methods for the *Bucket*

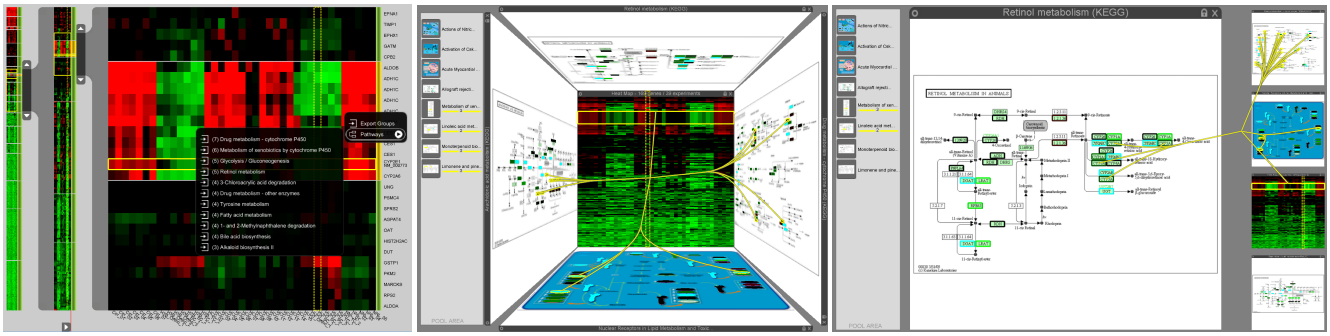
In principle, the *Bucket* can be used to show all visualization methods implemented in our system. However, not all of them are equally well represented in such a setup. We identified basic properties of visualization techniques that make them suitable for distorted analysis:

- It contains data that has many relations to other views in the setup.
- It does not suffer due to the distortion, as for example parallel coordinates do due to perspective foreshortening.
- It makes use of consistent spatial encoding, thus allowing a user to infer knowledge based purely on the location of an element.

Therefore, visualization techniques such as maps, treemaps, static graphs, heat maps, scatter plots *etc.* are well suited for use in the *Bucket*. In our framework we decided to limit the *Bucket* to show pathways and a contextual heat map to make optimal use of the available space for pathway analysis.

### Relating pathways to gene expression

By linking the contained heat map to the selected genes, the *Bucket* permits a new approach to the problem of bringing pathways into




(a) (b) (c)

Figure 4: Illustration of a gene-expression-centric analysis. (a) After filtering out inconspicuous genes and running a clustering algorithm, the pathologist finds a cluster in the heat map which has strongly diverging expression patterns for the different conditions. He checks for pathways containing the genes in the group, and in fact, several metabolic pathways contain four or more genes of the cluster. By clicking the pathways they are loaded into the *Bucket* for exploration (b). There he finds that the different pathways are heavily connected. Looking more closely at one pathway and its genes (c), he finds that the genes of the gene family *CYP* show the differential expression pattern. After checking *PubMed* and *Entrez Gene* with the integrated browser he learns that this family is a known catalyst for many reactions in the drug metabolism.

context with gene expression. The linking works for all shapes and sizes of nodes and also for 1:n relations. The number of experiments is only limited by the number of distinguishable elements in the heat map.

In pathways, nodes can be represented 0-n times, and one node can encode several genes. The linked heat map always highlights the genes which are mapped to the selected pathway node. Therefore, the gene expression values for all experiments available for this gene are shown.

In many cases, the expression values of other genes in the pathway should be considered at the same time. This allows experts to analyze the influence of the expression regulation on the pathway. If, for example, one gene in the chain is severely down-regulated, the rest of the path may be influenced. By using direct on-node color mapping for a single, selected experiment and visually linking the heat map (Figure 3), we overcome many of the problems mentioned in Section 2. We get an overview of all expression values in the pathways for this one experiment and simultaneously see information on all experiments for the currently selected gene. By selecting another experiment, the color coding on the pathway nodes is updated. This allows exploring all expression values for a pathway interactively.

For a pathway in focus, we do not color in the node thus obscuring the caption, but rather use a colored frame and thereby preserve the visibility of the text. The usage of color in this fashion is made possible by using connection lines instead of color highlighting to show identity relations between views.

Nodes that encode several genes cannot be handled by a single on-node color. We therefore render such nodes in a different (false) color, signaling that there is not only one mapping value. The concrete values can then be explored by selecting the gene node and using the linked gene expression views. On-node mapping can be turned off when gene expression is not the focus of the analysis.

## 4 Use Cases Revisited

In the course of our requirement analysis with life scientists, and during feedback sessions with prototypes, we discovered the previously mentioned distinct workflows for the analysis of biomolecular data: a pathway-centric and a gene-expression-based approach.

In the former case, a user is interested in a particular pathway. She wants to understand the pathway itself, the function of the particular genes involved and whether the genes play a similar role in other pathways. She wants to know details about a specific gene, search for publications and look it up in one of the large databases like *Entrez Gene*[4]. She may also be interested in the expression regulation values of the genes in the pathway for her experiments, in which case she is primarily interested in possible effects of the regulation on the pathway under investigation. She therefore starts her analysis by opening the *Bucket* and searching for the pathway she is interested in. Having loaded her gene expression data, she immediately sees the expression values for each selected experiment. She then notices a gene which has some interesting properties – for example, the record in the *Entrez* database, which was automatically loaded in the linked browser, tells her that the gene is involved in many forms of cancer. By right-clicking the gene, the system presents all pathways that also contain the gene – and in fact, most of them are cancer-related. However, some are not, which grasps her interest. She now moves a seemingly unrelated pathway into the focus and explores the role of the gene in this pathway.

The latter case deals with a gene-expression-centric analysis. A concrete example is illustrated in Figure 4.

## 5 Evaluation

We performed a user study to evaluate different aspects of the *Bucket* compared to traditional list-based pathway exploration methods normally used by biomedical experts. Our study specifically focused on the quality of the visualization methods to provide a useful context for finding target information in relation to the usage of both single-screen and multi-screen environments. The latter was taken into consideration, since related work [24] showed that multi-display environments can considerably advance the cognition and correlation process of information sources.

We chose not to compare *Caleydo* to other visualization frameworks, since the goal was to evaluate how much our novel visualization method can improve their previous workflows. A general comparison of visualization techniques such as the *Bucket* versus traditional multiple views should be conducted with a more general use case and average users, not life science experts. We also chose not to compare our system with other domain-specific software, since the functionality of the applications diverge in such a way that only trivial aspects could be compared.

### 5.1 Setup and procedure

The physical setup for the evaluation consisted of a desktop computer with two displays connected. Users were presented different, but comparable and complex search tasks, simulating real-life use

[4]http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene

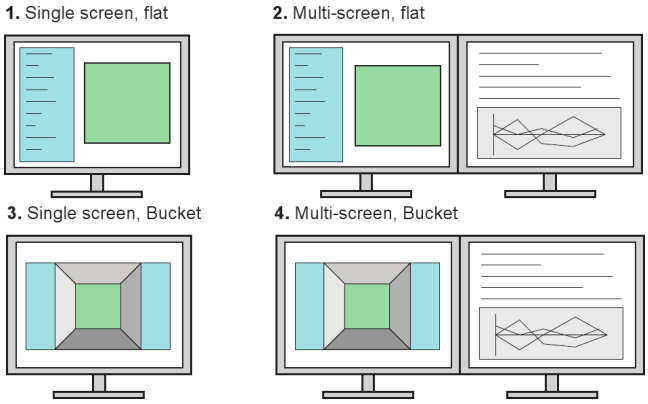



Figure 5: The four different setups for the user study.

cases. Participants performed two different tasks resembling the workflows described in section 4, under all four conditions: List-based and *Bucket*-based search tasks were performed in both single- and multi-monitor setups (see Figure 5). The first task involved pathway exploration, in which the participants were asked to detect relations between pathways, searching for a specific pathway and identify a specific gene in the pathway. As a next step, information about the gene in the *Entrez Gene* database had to be found. Finally, the participants were asked to find other pathways where the gene is also involved and determine whether there are other genes that those pathways share. The second task was based on gene expression analysis. Participants were asked to discover a specific pattern in the expression data using brushes in the parallel coordinates browser. The task required exploration of the pathways that contain these genes, and identifying a gene involved in a particular disease. During the first two conditions, the visual links were displayed in the *Bucket* view. The information on the second monitor (a web browser linking to gene databases for task one and a parallel coordinates browser for task two) was provided in a separate, tabbed window in conditions one and three. The task was subdivided into smaller units that were given step by step by the test supervisor. In order to simulate traditionally used list-based search methods (web interfaces like *KEGG*), we modified the application's user interface. It closely resembled the traditionally used list-based methods, as confirmed by our participants. It should be noted that the list condition actually had some enhancements over pure web interface methods, which would have been very hard to use for a comparative study in its original form.

We employed a 2x2 within-subjects factorial design with the factors view (*Bucket*, list) and display setup (single-monitor, multi-monitor). Analysis of main effects and interactions were performed at $\alpha = .05$ (see Table 1). Bonferroni adjustments were applied for post-hoc comparisons. To counterbalance the conditions, a Latin square distribution was used. All participants were videotaped with their consensus for later reference. The evaluation started with a ten-minute introductory session (including five minutes usage by the participant) in which the relevant functionality of the system was presented. After performing the different tests, participants answered a 7-point Likert scale questionnaire with 16 questions for both view levels and monitor-setups. Open discussions followed, where participants reflected on their experience. The total time of the user study was about 1h 15min per participant.

A third task focusing on pure observation was added with modified conditions to specifically investigate the utility of visual links: The three conditions were list-based, *Bucket* without visual links and *Bucket* with visual links, all on a single screen. Participants were asked to evaluate the quality and usefulness of the visual links under these conditions. This task was not performed in multi-screen conditions, since it is independent of the screen setup. We hypothesized the following outcomes:

**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.

For the evaluation, we recruited twelve participants with a background in life sciences. Eight participants (4 male, 4 female) were students (4 PhD, 4 master students) with beginner or intermediate experience, four participants (3 male, 1 female) were senior researchers and practitioners at a medical faculty.

### 5.2 Results

From the twelve original participants we included eleven in our analysis. One questionnaire was removed since it was highly inconsistent by itself and with respect to the interview. The results of the evaluation of the questionnaires is summarized in Figures 6, 7 and Table 1.

Table 1: *Main effects and interactions of view and display conditions. Significance: * = p<.05, ** = p<.001 (N=11)*

| **Question** | **view main** ($F_{1,10}$) | **display main** ($F_{1,10}$) | **interaction view*display** ($F_{1,10}$) |
|---|---|---|---|
| Spatial organization | 24.444** | 5.904* | 1.379 |
| Context quality | 46.414** | 6.806* | 0.313 |
| Compare information | 50.975** | 6.941* | 5.213* |
| Relate information | 30.414** | 3.978 | 2.222 |
| Detect info | 14.912** | 3.750 | 0.312 |
| Clarity of visualization | 4.290 | 6.806* | 0.132 |
| Readability | 1.000 | 2.168 | 1.000 |
| Perf. pathway explore | 4.646 | 1.957 | 1.000 |
| Perf. gene expression | 10.542* | 3.551 | 1.000 |
| Concentration | 27.121** | 0.694 | 4.808 |
| Confusion (negated) | 2.560 | 1.000 | 1.000 |

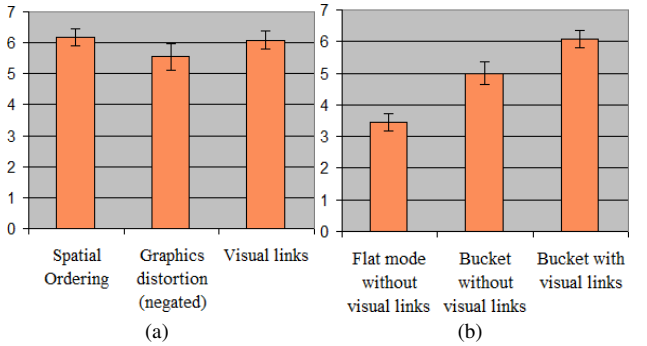



Figure 7: (a) Questionnaire results for three questions concerning the *Bucket* (N=11). (b) Comparison of perceived value of visual links compared to modes without visual links (N=11).

#### Information comparison

The performed tasks can be characterized as directed searches that aimed at accomplishing a specific, predefined goal. Participants needed to relate multiple sources of information, including different graph types and text-based sources. We found significant main

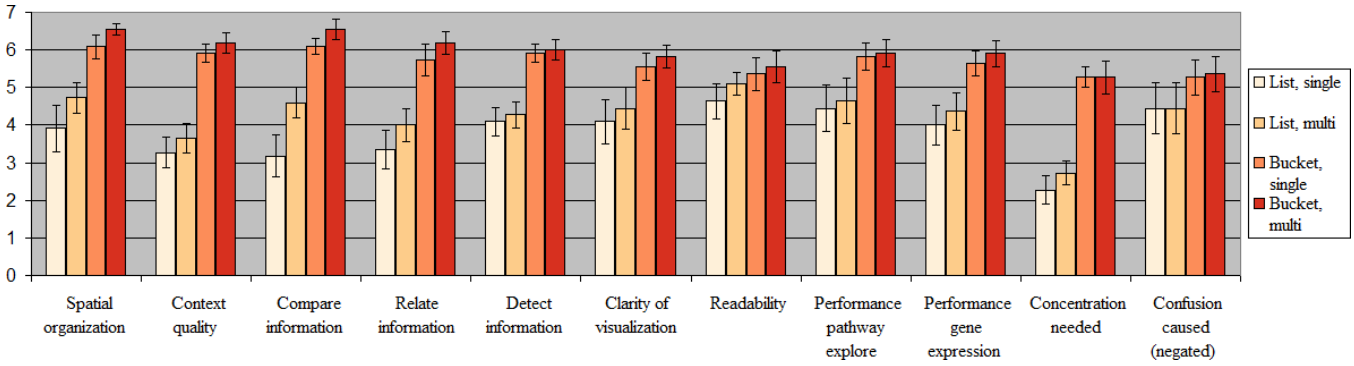



Figure 6: *Questionnaire results comparing the four different tested conditions for eleven areas of interest, comparing the four different setups of the first two tasks (N=11).*

effects of the view and display conditions on both the comparison of information and the quality of context, whereas the viewing condition also had a significant main effect on the detection of information. Additionally, an interaction between view and display was found for the information comparison. Relevant information was detected more easily in the *Bucket* conditions than in the list-based conditions, which was further improved by using the multi-monitor setup. Participants found the contextual information important for these tasks: The quality of contextual information was rated 'good' in *Bucket* conditions (in particular the multi-monitor condition), whereas the list-based multi-monitor was only rated 'mediocre'. The latter is slightly surprising, since one can clearly compare at least two different information sources in a multi-monitor setup. As we noticed during the interviews, this rating can be traced back to the participants' long experience of using just a single screen, which may be a learning problem. Overall, we found clear evidence that the detection of information is improved by the visualization aids offered in the *Bucket*, which performed significantly better than the list-based conditions.

### Visualization Method

The graphs analyzed in the tests are very dense: a large amount of information is compressed and screen space is limited. Obviously, the readability of the graphs is important to identify relevant information. The way the graphs are presented in the list and *Bucket* conditions is quite different, especially since graphics which are not in the center of the *Bucket* are distorted. When participants were asked about the readability of graphics in the different conditions, no significant difference was found: *Bucket* views even performed a little better on average. The graphics distortion was rated as negligible by most participants. This is quite surprising, since the graphs are clearly distorted at the side panels of the *Bucket*. In the interviews, some participants stated they would simply put those graphs needed for the analysis into the center of the *Bucket*. Some participants also said that distortion was not a problem since they can easily flatten the *Bucket* to a 2D view, removing perspective distortion of the side panel information (see Figure 3). In the interviews all participants stated they prefer the *Bucket* for finding interdependencies over the flat, zoomed mode.

The visual links were rated 'very useful' and also believed to speed up search tasks. During the interviews, many participants stated that the visual links aided the search for relevant information considerably. Although visual links do not affect the results of the analysis, it was easier to perceive the entire scene with its selections. In addition, some participants noted that visual links clearly helped them focus on specific parts of the graphs. We observed some participants consciously following the visual links from point to point to detect relevant information. Some also noted that the visual links are especially helpful with the pathway views, and considered them of less importance in the heat map: They argued that the gene expression views highlight the selections well by themselves, whereas the complex textures of pathways benefit from the additional visual clues.

### Effectiveness and complexity

Our main goal is to improve the workflow of users exploring pathways and gene expressions. The participants supported the hypothesis that the *Bucket* improves the workflow (speed and accuracy) significantly in comparison to the traditional list-based methods they are used to. Specifically for the gene expression task, we noted a significant main effect of the view mode (*Bucket*) on the perceived performance of the task. Participants noted that less concentration is required during the search task using the *Bucket*, which is in line with the ratings from the previous sections: The view condition had a significant effect on the level of concentration. Participants also were less confused in the *Bucket* conditions, even in comparison to the multi-monitor list condition. Likewise in single-monitor condition, the *Bucket* was rated better than the multi-monitor list-based condition in all questions.

## 5.3 Discussion

The evaluation clearly shows that the *Bucket* is a valuable improvement for pathway exploration over current practice using list-based methods. The *Bucket* performs significantly better in most conditions for most of the participants (supporting H1): in 7 out of 11 questions, we noticed a significant effect of the view condition on the outcome, and the average rating was higher for the *Bucket* conditions without exception. Three participants were even unable to fulfill the proposed task in the first list-based condition they obtained. Only a single user stated that the list-based method was preferred over the *Bucket*.

The visual links were very well appreciated, and clearly improve the search task performance in terms of (subjective) speed and lower cognitive load (supporting H3). The preference of single-monitor conditions may be related to the lack of experience our users have with multi-monitor configurations. One participant even failed to notice content on the second display entirely. These observations stand in contrast to previous evaluations like [24] that reported considerable performance boosts in multi-monitor environments. Thus, H2 turned out to be false. However, the (informally) observed performance of our participants was clearly better in the multi-monitor setup.

## 6 Conclusion and Future Work

By providing experts from the life science community with a tailored tool for the analysis of gene expression data in the context

of pathways, we aim to simplify the analysis of the large amounts of data generated. The combination of state-of-the art information visualization and visual analysis methods and the improvements in viewing arrangements, such as the *Bucket*, address their specific needs well, as confirmed by the user study.

The *Bucket* setup presented in this paper shows that a simple, restrictive approach for arranging 2D views in 3D is an effective way to visualize relations between different views. The *Bucket* can naturally accommodate focus+context as well as different levels of detail. It avoids confusion through a clear navigation concept, minimizes visual clutter with multi-level edge bundling and allows the user to manage many views conveniently. The implementation of the *Bucket* and the 2D views in OpenGL facilitate high frame rates and therefore good interactivity, even for many views and large data sets. We observed that the current implementation of the *Bucket* allows management of a sufficient but fixed number of related views. It is therefore necessary to develop a method that can handle a greater number of views for problems of a larger scale. In the context of gene expression and pathway analysis, the main focus lies on different representations of one element in several views. Visualizing several relations at the same time between views, as described in [7], will require additional methods to avoid visual clutter. The visualization of gene expression values on pathways when multiple genes encode a node can certainly be improved. We are currently considering using average values supplemented with an encoding of the standard deviation for the node.

In the future we aim to integrate biomedical data from different domains into one comprehensive framework. An example would be to integrate clinical data into the analysis, which would for example allow users to pre-filter the number of experiments shown in gene expression views based on clinical parameters. This is in line with trends toward *Biobanks* [2], large transnational databases which bring together a multitude of data. The goal of our system is to become a front-end for visual queries in such data spaces.

We also intend to conduct further user studies to emphasize our user-centered development approach. Interesting topics are to quantitatively prove the benefit of visual links, and to compare our tool with other state-of-the-art visualization techniques.

Our system has been well received by our focus group and is currently being used for real work by different departments and universities. First results acquired with the help of the tool have already been published [18].

## 7 Acknowledgments

The authors want to thank Prof. Dr. Kurt Zatloukal and his team from the Institute of Pathology at the Medical University of Graz as well as Dr. Gudrun Schmidt-Gann and Katharina Schmid from the Ludwig Boltzmann Institute for Experimental and Clinical Traumatology for their valuable input and the fruitful collaboration. We also want to thank Prof. Dr. Keith Andrews for his help when planning and Manuela Waldner when evaluating the user study. Furthermore we would like to thank the participants of the user study and acknowledge Bernhard Schlegl, Werner Puff and Christian Partl for help with the implementation.

This work was funded by the FIT-IT program (813 398), the Fonds zur Förderung der wissenschaftlichen Forschung (FWF) (L427-N15) and the Zukunftsfonds Steiermark (3007).

## References

[1] A. Aris and B. Shneiderman. Designing semantic substrates for visual network exploration. *Information Visualization*, 6(4):281–300, 2007.

[2] M. Asslaber and K. Zatloukal. Biobanks: transnational, European and global networks. *Brief Funct Genomic Prot.*, 6(3):193–201, 2007.

[3] B. Stolk et al. Mining the human genome using virtual reality. In *EGPGV '02: Proceedings of the Fourth Eurographics Workshop on Parallel Graphics and Visualization*, pages 17–21, Aire-la-Ville, Switzerland, 2002.

[4] M. Q. W. Baldonado, A. Woodruff, and A. Kuchinsky. Guidelines for using multiple views in information visualization. In *AVI '00: Proceedings on Advanced visual interfaces*, pages 110–119, New York, NY, USA, 2000. ACM Press.

[5] T. Ball and S. G. Eick. Software visualization in the large. *Computer*, 29(4):33–43, 1996.

[6] A. Barsky, T. Munzner, J. Gardy, and R. Kincaid. Cerebral: Visualizing multiple experimental conditions on a graph with biological context. *Visualization and Computer Graphics, IEEE Transactions on*, 14(6):1253–1260, Nov.-Dec. 2008.

[7] C. Collins and S. Carpendale. Vislink: Revealing relationships amongst visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1192–1199, 2007.

[8] M. B. Eisen, P. T. Spellman, P. O. Brown, and D. Botstein. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Academy of Science USA*, 95(25):14863–14868, December 1998.

[9] B. J. J. Frey and D. Dueck. Clustering by passing messages between data points. *Science*, 315(5814):972–976, January 2007.

[10] H. Chung et al. Arrayxpath II: mapping and visualizing microarray gene-expression data with biomedical ontologies and integrated biological pathway resources using scalable vector graphics. *Nucleic Acids Res*, 33(Web Server issue):W621–W626, Jul 2005.

[11] M. Kanehisa, M. Araki, S. Goto, M. Hattori, M. Hirakawa, M. Itoh, T. Katayama, S. Kawashima, S. Okuda, T. Tokimatsu, and Y. Yamanishi. Kegg for linking genomes to life and the environment. *Nucleic Acids Research*, 36(Database-Issue):480–484, 2008.

[12] C. Klukas and F. Schreiber. Dynamic exploration and editing of kegg pathway diagrams. *Bioinformatics*, 23(3):344–350, 2006.

[13] H. Lindroos and S. G. E. Andersson. Visualizing metabolic pathways: comparative genomics and expression analysis. In *Proceedings of the IEEE*, volume 90, pages 1793–1802, 2002.

[14] J. D. Mackinlay, G. G. Robertson, and S. K. Card. The perspective wall: detail and context smoothly integrated. In *CHI 1991: Proceedings on Human factors in computing systems*, pages 173–176, New York, NY, USA, 1991. ACM Press.

[15] B. Mlecnik, M. Scheideler, H. Hackl, J. Hartler, F. Sanchez-Cabo, and Z. Trajanoski. Pathwayexplorer: web service for visualizing high-throughput expression data on biological pathways. *Nucleic Acids Research*, 33(Web Server issue):633–637, July 2005.

[16] H. Mueller, R. Reihs, S. Sauer, K. Zatloukal, M. Streit, L. Alexander, B. Schlegl, and D. Schmalstieg. Connecting genes with diseases. In *Sixth International Conference BioMedical Visualization*, 2009.

[17] O. Rübel et al. Pointcloudxplore: Visual analysis of 3d gene expression data using physical views and parallel coordinates. In *EuroVis*, pages 203–210. Eurographics Association, 2006.

[18] G. Schmidt-Gann, K. Schmid, M. Uehlein, J. Struck, A. Bergmann, D. Schmalstieg, M. Streit, A. Lex, D. G. van der Nest, M. van Griensven, and H. Redl. Gene- and protein expression profiling in liver in a sepsis-baboon model. In *32nd Annual Meeting on Shock, San Antonio, Texas, June 6-9*, 2009.

[19] J. Seo and B. Shneiderman. A rank-by-feature framework for interactive exploration of multidimensional data. *Information Visualization*, 4(2):96–113, 2005.

[20] B. Shneiderman and A. Aris. Network visualization by semantic substrates. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):733–740, 2006.

[21] M. Streit, M. Kalkusch, K. Kashofer, and D. Schmalstieg. Navigation and exploration of interconnected pathways. *Computer Graphics Forum (EuroVis 2008)*, 27(3):951–958(8), May 2008.

[22] M. Streit, A. Lex, H. Müller, and D. Schmalstieg. Gaze-based interaction for information visualization. In *Proceedings of web3DW 2009 Conference, Algarve, Portugal*, 2009.

[23] Y. Yang, E. S. Wurtele, C. Cruz-Neira, and J. A. Dickerson. Hierarchical visualization of metabolic networks using virtual reality. In *VRCIA '06: Proceedings on Virtual reality continuum and its applications*, pages 377–381, New York, NY, USA, 2006. ACM Press.

[24] B. Yost, Y. Haciahmetoglu, and C. North. Beyond visual acuity: the perceptual scalability of information visualizations for large displays. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 101–110, New York, NY, USA, 2007.