MPVR: A Multi-Perspective Visual Retrieval Toolkit for Multi-dimensional Data

Yi Peng,* Li Chen,† Quan Li, Zhu Zhu, Jie Dong, Jun-Hai Yong
Tsinghua University
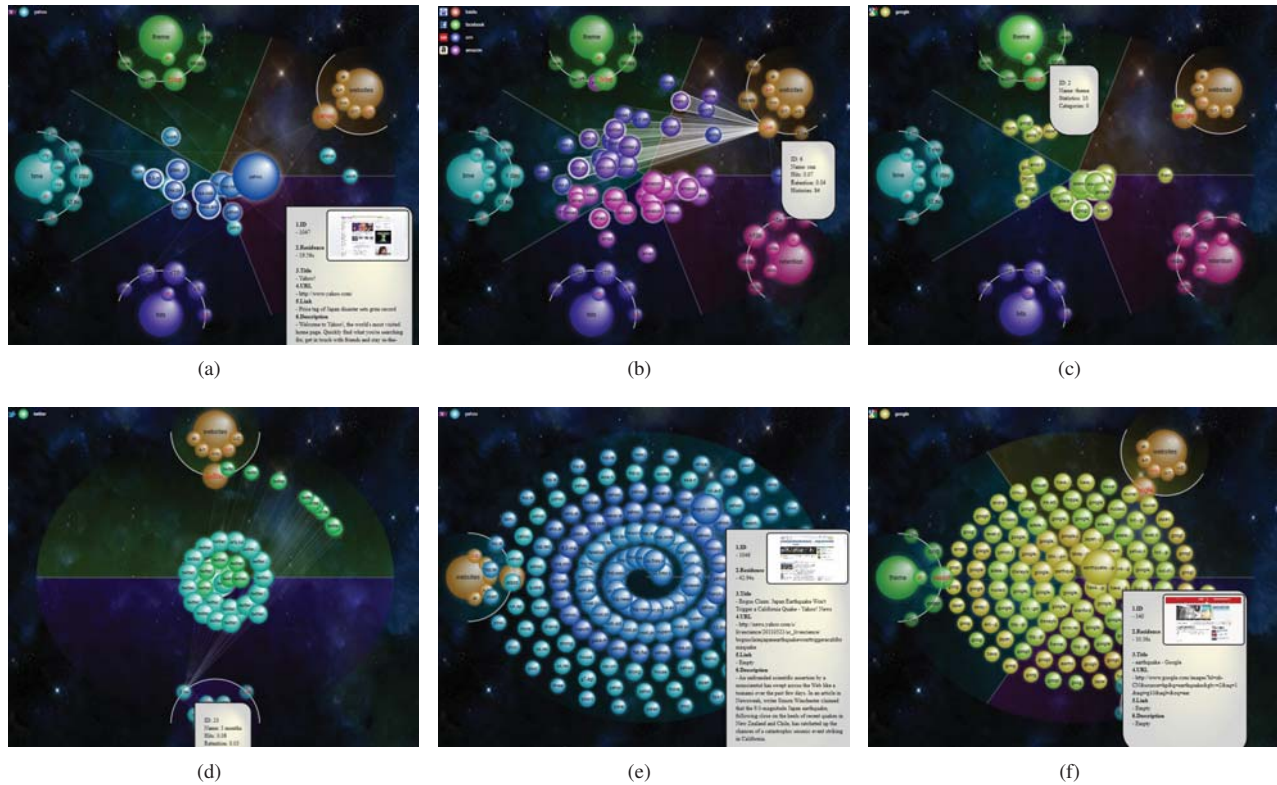
Figure 1: *MPVR: a multi-perspective visual retrieval toolkit*

Abstract

Rapid retrieval of specified items in a large amount of multi-dimensional data has become an urgent and challenging topic for improving the effectiveness of data searching and analysis approaches. In order to help users quickly re-find the desired items, we present a novel multi-perspective visual retrieval method which is totally different from traditional visualizations for searching results, such as list-based techniques. In our prototype, customized dimensions are generated adaptively which enables users to fast approach to the desired targets from multiple perspectives. We also design several display modes to layout the items by introducing degree-of-interest(DOI) ranking, which is implicitly defined by the geometric relationships of different dimensions. The display modes includes the *Cluster Mode* using importance-driven dimensionality reduction techniques, the *Radial Mode* diffusing the items into separate directions, the *Uniform Mode* evolved by Archimedean Spiral which keeps items non-overlapping and the *Custom Mode* for deep analysis. Moreover, we combine *Focus + Context*, *Visual Cues* which presents item properties, *Collaborative Filtering* and animations to recommend results according to the features users are interested in. Finally, we evaluate our method in different scenarios: text message retrospection, desktop search and web browsing history retrieval. The experiments show that our method is flexible, extensible, and efficient.

CR Categories: I.3.3 [Computer Graphics]: Three-Dimensional Graphics and Realism—Display Algorithms

Keywords: information retrieval, multi-perspective, visual navigation mode, focus+context

---

*E-mail: 15pengyi@gmail.com, Department of Computer Science and Technology.

†E-mail: chenlee@tsinghua.edu.cn, School of Software, Tsinghua University, Beijing 100084.

1 INTRODUCTION

Data Navigation for a large collection of information items is a difficult problem. It the information is well-defined, the process becomes a simple retrieval task. However, we may know vaguely

about what we are searching for, as is the usual case. If the rapid retrieval of specified items involves complex logical thinking and memory recall, improving the effectiveness of data searching and analysis becomes urgent and challenging.

Existing traditional retrieval methods, such as list-based techniques which aim to rearrange items based on only one or two dimensions are not adequate for complex retrieval task. Moreover, the process of exploration of desired items is vital to analysis, which greatly helps users to understand the data in different applications.

Several existing techniques are proposed to enhance information retrieval, such as list-based item view, thumbnails of items to provide an overview, or re-sorting based on a certain attribute, etc. Meanwhile, researchers have found other ways to make full use of visualization to aid data retrieval, such as tree-based view, content-dependent bookmark [Nakajima et al. 2002], timeline-based toolkit [Hoeber and Gorner 2009], semantic map [Šimko et al. 2010], tree map [Clarkson et al. 2009] and so on. However, they often focus on the hierarchical structure of data while ignoring other characteristics, such as multi-dimensional, entry date, etc., which may lead to ineffectiveness and impractical manipulation.

In order to solve these problems, we present a novel Multi-Perspective Visual Retrieval method (MPVR) which is different from traditional methods and enables users to re-find the desired items in high efficiency. Compared with other techniques, our method has four primary advantages.

- Rapid Retrieval

  Customized dimensions are adaptively generated in a multi-perspective view, making re-finding easy and efficient.

- Progressive Searching

  Collaborative Filtering is provided to help user fast approach to the target progressively and self-feedback technology enables users to see details in fine scale.

- Visual Navigation Modes

  In view of different favorites, we provide four visual navigation modes with adjustable parameters to analyze the data and understand their underlying relationships.

- Coordinated Highlighting

  For better analysis, we integrate several visualization techniques, such as *Focus + Context*, *Visual Cues* , hierarchical view and animations, with our display modes to recommend results according to user interests.

Primitive evaluations in three scenarios: text message retrospection, desktop search and web browsing history retrieval are carried out, showing that our prototype is flexible, extensible, and efficient for re-finding and data analysis.

The remainder of this paper is organized as follows. We discuss the background and motivation behind MPVR before illustrating the functional capabilities of MPVR in Section 2. In Section 3 we make a comprehensive description of our work. The results of evaluation in several scenarios and different applications are given in Section 4. Section 5 summarizes the characteristics of our method and present the planned future work.

2 RELATED WORK

Research on visualization of information items has been conducted for several years, and related techniques that influence our design can be divided into several groups as follows.

2.1 Layout Approaches

Researchers have designed a variety of intelligent techniques to improve the effectiveness of information retrieval tasks. Crnovrsanin et al. propose a network visualization and exploration system which recommends relevant nodes to users based on different metrics.[Crnovrsanin et al. 2011] We extend this idea by using different metaphor and automatically the satisfied results will appear within an elaborate designed layout. Yi et al. present a novel multivariate visualization techniques that are easier to understand and more accessible by using a magnet metaphor and various interactive techniques, which enables the public to deal with multivariate data effectively [Yi et al. 2005].

In regard of retrieving previously revisited information, several techniques are proposed. Simko, Tvarozek and Bielikova present a novel approach to exploit web search and navigation history by collecting streams of user actions, identifying user goals and constructing visual trees to represent visiting history[Šimko et al. 2010]. The result is based on semantic trees. However, this work requires a lot of efforts on semantic interpretation, which is not the usual case. Some works are based on tree layouts, such as Gandhi's work that implements a Domain Tree Browser to enable users to find their way back to the previously visited pages by building a tree structured visual navigation using domain names and thumbnails. [Gandhi et al. 2000] Also, Mase and Yamada propose a structure-based hierarchical clustering method to show the topics and the relationships between them by extracting topic maps from histories. [Mase and Yamada 2006] To break the limit of two-dimensional screen, Frecon and Smith gives a virtual reality (VR) based application to be used alongside traditional web browsers, which provides users with a flexibly tailorable real-time visualization of their history. [Vartiainen et al. 2008]

2.2 Focus + Context

In PaperVis Chou and Yang modify the existing Radial Space Filling (RSF) and Bullseye View techniques to arrange involved papers that better depict the relationships [Chou and Yang 2011]. Leung and Apperley give a survey on distortion-oriented visualization techniques including Polyfocal Display, Bifocal Display, Fisheye View and Perspective Wall [Leung and Apperley 1994]. We use the same idea to highlight the focused targets while maintaining context information.

2.3 Treemap

Jarke J. Van and Wijk Huub Van de Wetering conceive the concept of Cushion Treemaps for hierarchical information visualization. They use a simple shading model to render the subdivision ridges providing insight in the hierarchical structures [Van Wijk and Van de Wetering 1999]. ClarkSon, Desai and James apply Result Maps based on treemap techniques on online search interfaces to enhance digital library search engine result pages, which leads to ancillary benefits of data understanding [Van Wijk and Van de Wetering 1999]. However, the hierarchy property doesn't necessary lies in data itself.

2.4 Multi-Resolution

Hoeber and Gorner employ a two-dimensional timeline based visual interface to visualize browsing histories [Hoeber and Gorner 2009], which allows users to identify temporal patterns visually and re-find previously viewed pages effectively. However, their method is mainly based on time dimension and is not efficient enough to deal with large amount of data. Hightower et al. have implemented

a browser companion which builds a graphical history-map of visited web pages dynamically using minimal screen space to enhance the usability of web browsers [Hightower et al. 1998].

2.5 Other Related Visual Techniques

Several other related techniques also provide important references for our design.

Roberts surveys the techniques on coordinated and multiple views [Roberts 2007], which we also integrate into our prototype. In order to help users analyze the histories from different aspects, we design visual navigation modes which are highly configurable and easy to operate.

Liang and Huang summarize the existing highlighting techniques that can be applied in information visualization and propose a new three-layer model of highlighting [Liang and Huang 2010]. In our method, we make a comprehensive use of highlighting techniques. The distortions of opacity, color, size and shape are used for view control, tooltips including lighting effects, directed edges and thumbnails are used as interaction control, and an extended fisheye lens is used to expose the most relevant data in the view center for visual recommendation.

3 THE PROPOSED VISUAL WEB HISTORY RETRIEVAL TOOLKIT

Our toolkit is an integrated system with three basic components (like abstract class): a data collection engine, a data viewer and an interactive visual interface. In different scenarios, the three components are application-specific derived from the basic one (like derived class).

3.1 Framework of MPVR

- Data Collection Engine

  We design an abstract data collection engine by connecting to different data-sources. For example, in the application of web browsing history analysis, we develop a plugin using the kernel of Internet Explorer which provides users basic functionalities for surfing the Internet and connect it to our data collection engine (Figure2(a)). In another application of text message retrospection, this engine is connected to an export interface of smart phones.

- Data Viewer

  After data collection, users can get an overview of the raw data and its statistics shown by four interdependent relational tables. The statistical data(Figure2(b) and 2(c)) is carried out automatically by a derived member function for data preprocessing in different scenarios. We can also sort them based on one attribute.

- Interactive Visual Interface

  With the statistics, the general visual interface can directly show their relationships(Figure2(d)) with four features: rapid retrieval, progressive searching, visual navigation modes and coordinated highlighting techniques.
  The interface consists of three parts as seen in Figure2(d). The left area is for primary visualization. The top right region contains the configuration options for data dimensions and navigation modes, while the bottom right is used for adjustment of layout parameters.
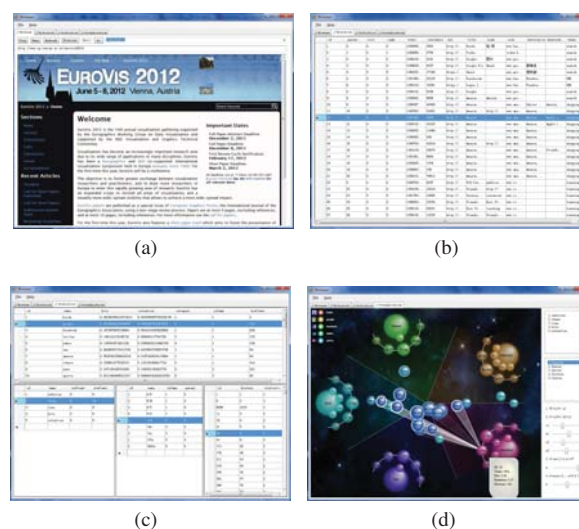

Figure 2: *The framework of MPVR.*

3.2 Multi-Perspective View

The primary idea of our design is to make information item retrieval easy and efficient. So we design a multi-perspective view based on two important observed facts during the searching process.

Fact 1 Target items are multi-dimensional and can be represented by some major dimensions.

Fact 2 For users the general impression of an information item is always a description of several keywords.

With customized perspectives, retrieval and analysis of information items becomes easy due to the following three reasons.

Reason 1 With tags in different dimensions, data filtering can be much faster than before.

Reason 2 Searching process becomes progressive when different dimensions cooperate with each other as seen in Figure4.

Reason 3 By introducing related links, a hierarchical structure on each dimension is clearly presented.

Figure1 shows the customization and inner-relationship display features of multi-perspective view under different conditions. Figure1(a), 1(b) and 1(c) show that the items are projected into five dimensions and a hierarchical structure is presented by interaction. Figure1(d), 1(e) and 1(f) show three other conditions restricted by three dimensions, two dimensions and only one dimension, respectively.

In web browsing history analysis application we elaborately choose five primary dimensions listed in Table 1 for history retrieval, which means each of the histories is projected into five dimensions with a value indicating its importance. We set up a local dictionary to filter out and extract those keywords related to semantic aspect, such as theme and the name of website, rather than manually labeling those history items.

3.3 Progressive Searching

Progressive searching provides users searching benefits by adding searching conditions implicitly through adding incremental con-

Table 1: *Dimensions of Browsing History*

Dimension	Symbol	Description
Website	w	URL of a web page
Theme	m	Topic to which a web page belongs
Time	t	Visit time of a web page
Hits	n	Number of visits on this page
Residence	r	Time spent on a web page
Thumbnail		Detail information of a web page

straints, thus reducing the number of result items and helps users quickly find what they want. This mechanism is composed of two steps below.

3.3.1 Collaborative Filtering

By selecting more dimensions and tags, more items will be filtered out since there exists more constraints, which means by using these filterings methods alternately, the number of result items may decrease significantly, therefore, users can approach to the targets rapidly.

(a)   (b)
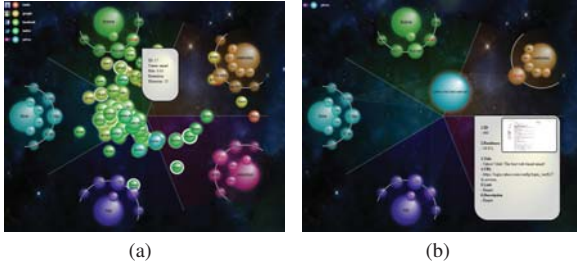
**Figure 3:** *Collaborative Filtering for progressive searching.*

Figure3 shows an example of *Collaborative Filtering* for progressive searching. For instance, the retrieval of the web page "yahoo mail" consists of only two steps. By selecting the "email" element around the "theme", the number of clusters greatly reduced as seen in Figure3(a). Further progress is made by specifying a "U-Z" tag to filter the results as shown in Figure3(b). Now the clusters are even fewer than before and the web pages "yahoo.mail.best.web" are already at the view center.

3.3.2 Self-feedback

Although *Collaborative Filtering* contributes a lot to the searching process, the number of filtered items can still be very large, which leads to more overlapping and chaotic layout. Here we use self-feedback technology to solve this problem.

This idea utilizes focus + context and takes hierarchical clustering into consideration. By interacting with overlapping items, we recalculate the position of each item in fine scale, keeping them separate from each other.

Figure4 shows an searching example using self-feedback feature. The original clusters are shown in Figure4(a). By right-clicking on the cluster "us.fires.wave.of", the overlapping items diffuse into different positions along a spiral path, as seen in Figure4(b). The location of each scattered item is calculated by its importance using fine scale statistics.

(a)   (b)
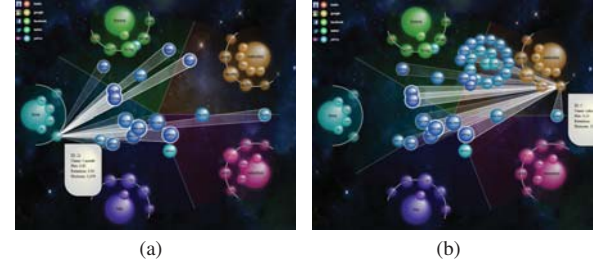
Figure 4: *Self-feedback for progressive searching.*

3.4 Visual Navigation Modes

Visual navigation modes are designed to meet the needs of different scenarios for searching. First, we will clarify some basic concepts.

Three display modes are designed to meet various needs of different stages in searching process. Before the introduction, some basic concepts need to be clarified.

- Item, Attribute and Statistic

  An item indicates a record in the raw dataset. Each dimension is named an attribute. A statistic s is a set of items with the same attributes. For instance, a history item e may have at least two attributes: access frequency A_f and residence (how long the user stays on this page).

  These three concepts have corresponding visual elements in our prototype: Item Visual Element (IVE), Attribute Visual Element(AVE) and Statistic Visual Element(SVE).

- Statistic Importance

  For each item, we use several attributes to define its importance. For example, giving a weight vector w, the importance of a history item $I(e)$ is defined by its frequency A_f and residence A_r. Then the statistic importance $I(s)$ can be derived based on the item importance as shown in Formula 1.

$$I(s) = \sum_{e \in s} I(e) = \sum_{e \in s} w_f \cdot A_f(e) + w_r \cdot A_r(e), w_f + w_r = 1$$

(1)

- Element Location

  The location of a visual element is defined by $(\rho(e), \theta(e))$ in polar coordinates system. The reason why we choose polar coordinates system is that the more central place an element locates, the more important it is. In this way, we can directly use the radius $\rho(e)$ to measure its importance.

- AVE Layout Algorithm

  All the AVEs are distributed evenly along the circumference of an ellipse, which means the included angles of two adjacent AVEs are the same. The location of an AVE is calculated by Formula 2.

$$\begin{cases} \theta_i = (\frac{i+0.5}{n_a}) \cdot 2\pi \\ \rho_i = \sqrt{a^2 \cos^2 \theta_i + b^2 \sin^2 \theta_i} \end{cases} i = 0, 1, \cdots, n_a - 1$$

(2)

Figure 5: *The layout of customized dimensions.*

Figure5 shows the layout of AVEs which are evenly distributed on the boundary of an ellipse. Since the included angles of adjacent AVEs are the same, the areas of all the sectors are also the same(according to the geometrical characteristic of ellipse). The layout can be changed if the attribute importance is defined. Which means, the area of a sector will be weighted by the importance.

- SVE Layout Algorithm

We place a SVE around the AVE it belongs to. The relative polar coordinates to the corresponding AVE of a SVE is defined by Formula 3, where c_s is a constant value, n_s indicates the number of visible SVEs. If $n_s < 3$, the field angle defined by all the SVEs is set to π, otherwise it is equal to the inner angle of a regular polygon with n_s vertices as shown in Figure6.

$$
\begin{cases}
\theta_i = \begin{cases} (\frac{i+0.5}{n_s})\pi & n_a < 3 \\ (\frac{i+0.5}{n_s})\pi \cdot (1 - \frac{2}{n_a}) & n_a \geq 3 \end{cases} \quad i = 0, 1, \cdots, n_s-1 \\
\rho_i = c_s
\end{cases}
\tag{3}
$$

Figure 6: *Different layouts of SVE.*

Figure6 shows different layouts of SVEs. If the number of SVEs is more than a default value(we choose 5 in our toolkit), users need to specify a sliding window, which means only 5 SVEs are visible. The visible SVEs are evenly distributed on a circle around the AVE they belong to.

- IVE layout algorithm

The same principle applies when we layout the IVEs. In order to place IVEs with greater importance closer to the view center, we compute the radius $\rho(e)$ carefully. The $\rho(e)$ should

take the statistic importance into consideration. For each SVE an IVE belongs to, we first compute a candidate location (ρ_c, θ_c), which is calculated by Formula 4:

$$
\begin{cases}
\theta_c = \theta_s + \frac{\pi}{2} \\
\rho_c = \left\| (\mathbf{x}_v - \mathbf{x}_s) \cdot \begin{pmatrix} \cos\theta_s \\ \sin\theta_s \end{pmatrix} \right. \\
\qquad \left. \cdot \begin{pmatrix} \cos\theta_s \\ \sin\theta_s \end{pmatrix}^T - (\mathbf{x}_v - \mathbf{x}_s) \right\|
\end{cases}
\tag{4}
$$

where $\mathbf{x}_v$ denotes the location of the view center, $\mathbf{x}_s$ denotes the location of a SVE, θ_s denotes the angle of the SVE. In specific, the candidate location is the projection of the view center on the line passing through the SVE in its direction.

$$
w_c = 1 - \frac{\left(1 - a^{I_s}\right)^2}{1 + a^{I_s}}
\tag{5}
$$

Formula 5 defines the importance of the SVE correspond to a IVE, which means SVE of less importance will have a weight-value close to 1, while greater important SVE may have a weight close to 0. In our paper we choose $a = 0.05$.

3.4.1 Cluster Mode

The Cluster Mode is designed to provide users an overview of the data, where IVEs belong to the same SVEs are clustered. In this mode, the distribution of information items is clearer. The position of an IVE is defined by Formula 6.

$$
\begin{cases}
\rho(e) = \sum_{\{s|e \in s\}} w_c(s) \cdot \rho_c(s) \\
\theta(e) = \sum_{\{s|e \in s\}} w_c(s) \cdot \alpha_c(s)
\end{cases}
\tag{6}
$$

Formula 6 indicates that the distance $\rho(e)$ between IVE and view center is mainly determined by the least important SVE, which means if and only if all the corresponding SVEs are relatively important, the location of IVE will be central as shown in Figure7.

Figure 7: *The Cluster Mode for IVE layout.*

Figure7 shows that the IVEs belonging to the same SVEs form a cluster.

3.4.2 Radial Mode

The radial mode diffuses the IVEs which belong to the same cluster into separate locations, which help users see detailed information.

$$
\begin{cases}
\rho(e) = \sum_{\{s|e \in s\}} w_c(s) \cdot \rho_c(s) \\
\theta(e) = (\frac{i+0.5}{n_i}) \cdot 2\pi
\end{cases}
\quad i = 0, 1, \cdots, n_i - 1 \quad (7)
$$

This mode follows the computation of $\rho(e)$ in the Cluster Mode, while calculating $\theta(e)$ to make the IVEs distributed evenly along the boundary of a sector as seen in Figure8.

We sort the IVEs based on their importance to determine their positions on the sector (Formula 7).
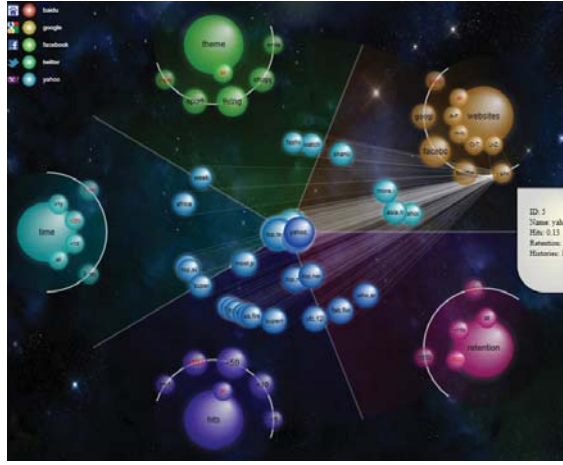
Figure 8: *The Radial Mode for IVE layout.*

Figure8 shows that the IVEs in a same cluster are diffused into different positions along the boundary of an sector.

3.4.3 Uniform Mode

The Uniform Mode is based on the *Archimedean Spiral*, which uses the same criteria with the Cluster Mode to compute $\rho(e)$ and sort them in an ascending order. Finally, we use formula 8 to compute the index of an IVE in the ordered list.

$$
t(e) = \frac{i+0.5}{n_i}, i = 0, 1, \cdots, n_i - 1
$$
$$
\begin{cases}
\rho(e) = c_2 \cdot t^{e_2} + c_3 \\
\theta(e) = c_1 \cdot t^{e_1}
\end{cases}
\quad (8)
$$

In this mode we set $e_1 = e_2 = 0.5$, $c_1 = \frac{r}{r_i}$, $c_2 = r - r_i$ and $c_3 = r_i$, where r_i denotes the radius of IVEs and r is the size of the window.

Figure9 shows the Uniform Mode with the same data shown in Figure7. In this mode the histories with the name "quan.li" are diffused along a spiral path, which means these histories lay on an ellipse boundary approximately.

3.4.4 Custom Mode

The custom mode is an extension of Uniform Mode. It allows advanced users to adjust layout parameters, such as c_1, c_2, c_3, e_1 and e_2.
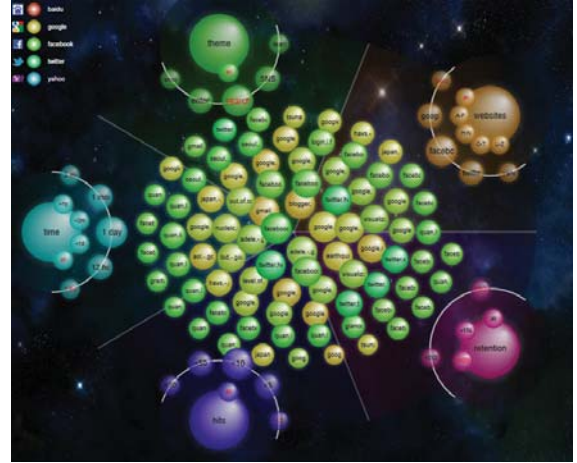

Figure 9: *The Uniform Mode for IVE layout.*

3.5 Coordinated Highlighting

For better user experience, we integrate several visualization techniques, such as Focus + Context, dynamic tooltips to present detailed information while retaining the context, therefore, users can concentrate on the specific items he/she is interested in.

For all the IVEs, the colors are defined by a user specified dimension. And for each visual element, its size and opacity are determined by its importance and when focused both of them will increase to show detailed information.

We also design dynamic tooltips to give a thumbnail and other important details. Moreover, when a user moves the mouse over a visual element, several linked light beams will be generated automatically to show the relationship between the focused one and other related elements. A color bar is also provided at the top-left corner, helping users understand the color mapping schema.

Figure 10: *The coordinated highlighting techniques.*

Figure10 shows that we use a tooltip to show its detailed information for each visual element. Especially for information item, a thumbnail is presented to help users recapture memories on it. Meanwhile, a color bar is provided to show the correspondence of items and their colors.

4 EVALUATION METHODOLOGY

In this section we will demonstrate the effectiveness of MPVR through some potential scenarios in three familiar applications: we-

b browsing history analysis, mobile text message retrospect, and desktop searching. Our case studies are based on the following three datasets:

Table 2: *Three datasets for method evaluation*

ID	Name	Items	Dimensions
1	Web browsing history	1218	12
2	Mobile text message	1886	14
3	Local Computer Files	1857	9

4.1 Scenario 1: Browsing history retrieval

Web browsing histories can be used to help users retrieve important web pages which are used frequently. It also plays a significant role in analyzing user behaviors. It has been studied that over 80% of web pages have been previously visited [COCKBURN and MCKENZIE 2001]. Several techniques are integrated into web browsers to enhance the history utilization, such as list-based history view, thumbnails of most visited sites, URL auto-complete based on browsing histories, management of favorite links and so on [Ayers and Stasko 1995] [Hightower et al. 1998] [Cockburn et al. 1999] [Gandhi et al. 2000] [Mase and Yamada 2006]. However, they often focus on the hierarchical structure of data while ignoring their multi-dimensional properties. Therefore, in this scenario, we concern about how to effectively retrieve those web pages from a large amount of data by introducing multi-dimensional information. We show two cases that can be easily encountered in browsing experience to evaluate our method.

- Case Study 1: Finding webpages with high access frequency, these webpages are visited a certain time ago.

  When serving the internet, users tend to re-visit a web page on a specific topic, such as shopping, online learning and so no. In this case, we carry out an experiment on dataset 1(Figure13(a)). Suppose the goal is to re-find the shopping webpages of amazon, which are visited three month ago.

  Since the target is clear defined, we first click the "shopping" element around "theme". Unfortunately, there are still no IVEs corresponding to "shopping" (as seen in Figure13(b)) due to the current websites having no relations with shopping. We then select the "A-F" tag and click the "amazon" element, and we can see that more relevant items appear (see Figure13(c) and 13(d)). We then choose the tag "¡3m" and click the "3 months" element to see how many webpages have been visited three month ago(see Figure13(e)).

  Although the majority of the relevant items are presented, they overlap heavily. So we choose the Radial Mode and click the "¡50" element to see the most frequently visited pages(Figure13(f)). Then, we focus on the pages with the longest time of residence by choosing the "¡600" tag and clicking the "¡600" element. After that, there is only one page which meets our conditions (Figure11g). We can move the mouse over it to see the details(Figure11h). So we get the final page which is about "guitar", and maybe the user is a guitar fun. Actually what we guess is right.

### 4.2 Scenario 2: Mobile text message retrospect

SMS (Short Message Service) has already been a popular communication technology for mobile users. However, the visualization technology for text message retrieval and retrospect is not highly developed. There are many works that explore the social and temporal interactions in instant messaging which may involve different people, such as CrystalChat. This tool visualizes personal chat history and reveals the communication patterns [Tat and Carpendale 2006]. In this scenario, we focus on how to aid users to retrospect their messages from the visualization aspect.

- Case Study 2: Finding important text messages contained key information from a certain person.

  The scenarios is that a user collects recent mobile text messages and hope to check some key information, however he can't remember clearly. Suppose we know the name and location of the contact(a person). In general, the fastest way to find these messages is to find out all candidate messages from this person and check them one by one. Here, we use our toolkit to have a try.

  The first step is to reveal the related IVEs based on the first condition: contact name and his/her location. These results are shown in Figure11(a) and 11(b). Then we find that there are only two messages left from the specific person. So we checks the two messages and find the information is that the contact talks about that the report materials, including the PPTs, videos and demos, which are all in disk D. The contents of the other messages are closely related to this topic: project briefing. Also, we can recall our memory on the communications with a certain person through text messages (Figure11(c)) by exploring the message clusters ordered by file size (Figure11(d)).

(a)   (b)

(c)   (d)
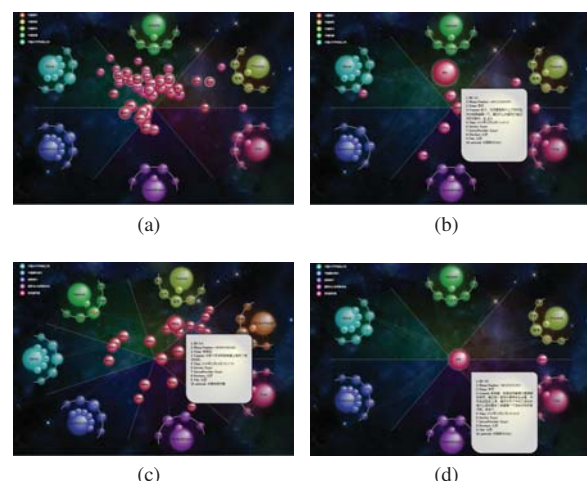
Figure 11: *Finding important text messages contained key information from a certain person.*

4.3 Scenario 3: Desktop searching

We have witnessed the huge development of online tools to search the web, however, desktop search have been slowly developed and they are still at their infancy stage. Some researchers devote themselves to developing new methods to enhance desktop search engine. *Multiple Views* are coordinated and used to carry out query refinements and analyze data. These view includes list view, tree view, map view, bubble view, tile view and cloud view [Foo and Hendry 2007]. We are interested in applying our prototype to desktop search and observe whether it has the potential to improve the usefulness of such tools.

- Case Study 3: Searching for a document modified long time ago with specific extension.

Usually users takes a long time to do documents management, especially for the searching and maintenance of important documents such as resumes, schedules etc. In this case, we use dataset 3(Figure12(a)) to carry out an experiment on important document searching. Assume that the target document with "docx" file extension contains characters of "p"-"s" and is modified a long time ago but accessed recently.

We notice the particular file extension. So we choose the "office" tag and click the "document" element(as seen in Figure12(b)). we then specify its creation time, write time and access time by clicking the "1 year" element around "creation time" and "write time", "1 month" element around "access time", and choose the "O-T" tag. After that many old documents appear as seen in Figure12(c). Since there are too many documents, we choose the Uniform Mode to diffuse them as seen in Figure12(d). Because the desired document is created a long time ago, we can make a further assumption that it may be placed in a deep directory. So we click the "4" element around "file path". Fortunately, the overlapping is eliminated(see Figure12(e)). Then we can check these 4 documents with the extension of "docx" one by one to see whether they are what we want(Figure12(f)).

4.4 Evaluation Results

The evaluation shows that our toolkit based on multi-perspective visual retrieval outweighs traditional retrieval methods. Meanwhile, we apply our method to three domains, web browsing history analysis, text message retrospect and desktop searching. And the three case studies turn out that our toolkit performs quite well in high dimensional data retrieval and the participants involved in the tasks think that our approach is attractive, efficient and quite interactive, while some of the participants maintain that they still need training before carrying out retrieval tasks.

5 CONCLUSION

In this paper, we present a novel approach named Multi-Perspective Visual Retrieval(MPVR) to help users quickly re-find the desired items of their interests. By providing four key features: Multi-Perspective View, Progressive searching, Visual Navigation Modes and Coordinated Highlighting, users can fast approach to the targets. Although the preliminary investigations give us positive feedback, there still remain a number of unsolved problems such as overlapping in Uniform Mode, unsatisfied layout with much more dimensions, less functionalities of coordinated views and so on.
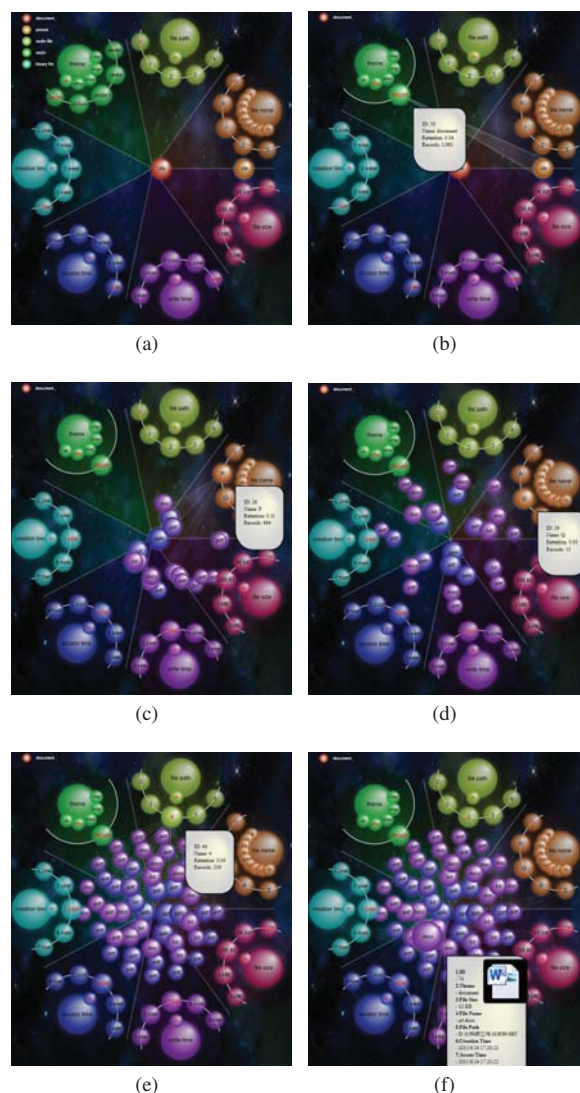
Other future work includes improving user interface, applying our method to other domains and analyzing user's behaviors.

Acknowledgements

References

AYERS, E. Z., AND STASKO, J. T. 1995. Using graphic history in browsing the world wide web.

CHOU, J. K., AND YANG, C. K. 2011. Papervis: Literature review made easy. *Computer Graphics Forum 30*, 3, 721–730.

Figure 12: *Searching for a document modified long time ago with specific extension.*

CLARKSON, E., DESAI, K., AND FOLEY, J. 2009. Resultmaps: Visualization for search interfaces. *Visualization and Computer Graphics, IEEE Transactions on 15*, 6 (nov.-dec.), 1057 –1064.

COCKBURN, A., AND MCKENZIE, B. 2001. What do web users do? an empirical analysis of web use. *Int. J. Hum.-Comput. Stud. 54*, 6 (June), 903–922.

COCKBURN, A., GREENBERG, S., MCKENZIE, B., JASONSMITH, M., AND KAASTEN, S. 1999. Webview: A graphical aid for revisiting web pages. 15–22.

CRNOVRSANIN, T., LIAO, I., WU, Y., AND MA, K.-L. 2011. Visual recommendations for network navigation. *Computer Graphics Forum 30*, 3, 1081–1090.

FOO, S., AND HENDRY, D. 2007. Desktop search engine visualisation and evaluation. In *Proceedings of the 10th international conference on Asian digital libraries: looking back 10 years*
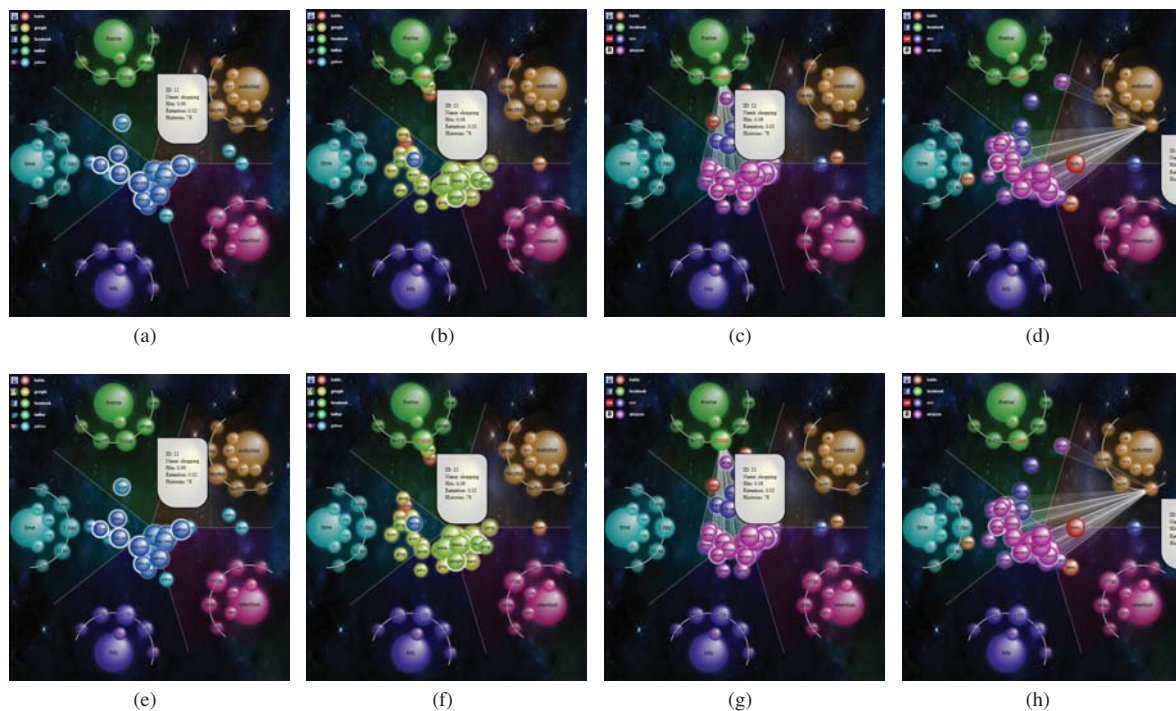
Figure 13: *Finding specific webpages of high access frequency visited a certain time ago.*

and forging new frontiers, Springer-Verlag, Berlin, Heidelberg, ICADL'07, 372–382.

GANDHI, R., KUMAR, G., BEDERSON, B., AND SHNEIDERMAN, B. 2000. Domain name based visualization of web histories in a zoomable user interface. In *Database and Expert Systems Applications, 2000. Proceedings. 11th International Workshop on*, 591 –598.

HIGHTOWER, R. R., RING, L. T., HELFMAN, J. I., BEDERSON, B. B., AND HOLLAN, J. D. 1998. Graphical multiscale web histories: a study of padprints. In *Proceedings of the ninth ACM conference on Hypertext and hypermedia : links, objects, time and space—structure in hypermedia systems: links, objects, time and space—structure in hypermedia systems*, ACM, New York, NY, USA, HYPERTEXT '98, 58–65.

HOEBER, O., AND GORNER, J. 2009. Browseline: 2d timeline visualization of web browsing histories. In *Information Visualisation, 2009 13th International Conference*, 156–161.

LEUNG, Y. K., AND APPERLEY, M. D. 1994. A review and taxonomy of distortion-oriented presentation techniques. *ACM Trans. Comput.-Hum. Interact. 1*, 2 (June), 126–160.

LIANG, J., AND HUANG, M. L. 2010. Highlighting in information visualization: A survey. In *Information Visualisation (IV), 2010 14th International Conference*, 79 –85.

MASE, M., AND YAMADA, S. 2006. Extracting topic maps from web histories by clustering with web structure and contents. In *Proceedings of the 2006 IEEE/WIC/ACM international conference on Web Intelligence and Intelligent Agent Technology*, IEEE Computer Society, Washington, DC, USA, WI-IATW '06, 405–408.

NAKAJIMA, S., OYAMA, S., SUMIYA, K., AND TANAKA, K. 2002. Context-dependent web bookmarks and their usage as queries. In *Proceedings of the 3rd International Conference on Web Information Systems Engineering*, IEEE Computer Society, Washington, DC, USA, WISE '02, 333–344.

ROBERTS, J. 2007. State of the art: Coordinated multiple views in exploratory visualization. In *Coordinated and Multiple Views in Exploratory Visualization, 2007. CMV '07. Fifth International Conference on*, 61 –71.

ŠIMKO, J., TVAROŽEK, M., AND BIELIKOVÁ, M. 2010. Semantic history map: Graphs aiding web revisitation support. In *Database and Expert Systems Applications (DEXA), 2010 Workshop on*, 206 –210.

TAT, A., AND CARPENDALE, S. 2006. Crystalchat: Visualizing personal chat history. In *System Sciences, 2006. HICSS '06. Proceedings of the 39th Annual Hawaii International Conference on*, vol. 3, 58c.

VAN WIJK, J., AND VAN DE WETERING, H. 1999. Cushion treemaps: visualization of hierarchical information. In *Information Visualization, 1999. (Info Vis '99) Proceedings. 1999 IEEE Symposium on*, 73–78, 147.

VARTIAINEN, E., ROTO, V., AND KAASALAINEN, J. 2008. Graphical history list with multi-window support on a mobile web browser. In *Internet and Web Applications and Services, 2008. ICIW '08. Third International Conference on*, 121 –129.

YI, J. S., MELTON, R., STASKO, J., AND JACKO, J. A. 2005. Dust & magnet: multivariate information visualization using a magnet metaphor. *Information Visualization 4*, 4 (Oct.), 239–256.