Internet Browsing: Visualizing Category Map by Fisheye and Fractal Views

Christopher C. Yang†, Hsinchun Chen*, and Kay Hong**

†: Department of Systems Engineering and Engineering Management
The Chinese University of Hong Kong, Hong Kong

*: Department of Management Information Systems
The University of Arizona, USA

**: Department of Computer Science and Information Systems
The University of Hong Kong, Hong Kong

Abstract
Category map developed based on Kohonen's self-organizing map has been proven to be a promising browsing tool for solving the information overload problem in World Wide Web. The SOM algorithm automatically compresses and transforms a complex information space into a two-dimensional graphical representation. Such graphical representation provides a user-friendly interface for users to explore the automatically generated mental model. However, as the amount of information increases, it is expected to increase the size of the category map accordingly in order to accommodate the important concepts in the information space, which increases the visual load of the category map. In this paper, we propose the fisheye views and fractal views to support the visualization of category maps. Fisheye views are developed based on the distortion approach while fractal views are developed based on the information reduction approach. We have developed a prototype system and conducted a user evaluation to investigate the performance of fisheye views and fractal views. The results show that both fisheye views and fractal views significantly increase the effectiveness of visualizing category map. In addition, fractal views are significant better than fisheye views.

Introduction
Acquiring information from the huge information space of the Web is a difficult task and users are easy to become disoriented. The Kohonen self-organizing map (SOM) based category map is a browsing tool for user to explore a well-organized structure of the Internet information space where the categories are automatically generated based on the semantics of the documents. Experiments showed that it is capable to categorize a large Internet information space, Yahoo Entertainment sub-category [4].

As the information increases exponentially (information overloading problem), the typical size of category map may not be able to accommodate all the crucial concepts. However, if the size of the category increases in order to fit the increasing numbers of categories, the details of the map become too small to be seen clearly. Given a category map with over 100 nodes and a displaying window of 1000 by 1000 pixels, it is usually not viable to see the local details if the global structure of the map has to be maintained. This problem is common in information visualization and is known as visual load [1]. Users are difficult to locate their points of interest from a large pool of information if the information is packed closely together on a limited size of output window. There are several possible techniques to assist in visualizing the local details of a large category map, 1) zooming into a part of the map to observe the local details, or 2) using two or more windows, one with the global view and others with the zoomed portions of the map. Zooming is able to present the local details, however, it loses the global structure of the category map. Using multiple windows allows global structure and local details to be presented simultaneously on the same screen; however, it requires mentally integration of multiple views.

In this paper, we propose to fisheye views and fractal views to for visualizing category maps. Fisheye views increase the visualization power of the focus area and shrink the distant area by a distant function. The visualization effect is similar to the wide-angle fisheye camera [10]. The local details are magnified by the fisheye views and the global structure is maintained. Fractal views are derived based on the fractal theory [9] to abstract the displaying objects as well as controlling the amount of information displayed. Information that is less relevant to the area of interest in category map is filtered. Although the global structure is lost, users can interact with the system by adjusting the fractal value to control the amount of information to be filtered.

1. Category Map for Internet Browsing
Browsing is an exploratory and information-seeking strategy that depends upon serendipity, which is appropriate for ill-defined problems or exploring new task domains [12]. It is characterized by the absence of planning [11]. In essence, browsing explores both the organization and structures of the information space and its content, based on the pre-existing mental
models of information organization. That’s why it is often used in exploring relatively new or unexplored information spaces, such as World Wide Web.

The mental models, which are defined as the cognitive representations of information [12], are important for exploring an unknown information space. They represent the content, structure, and relationship of information in the information space [1]. Given the mental models, users may understand the organization of the information space so that they may draw better inferences and make better response during the navigation of the information space. Category maps are mental models generated by SOM [7] automatically. The important concepts in the information space are displayed by an intuitive graphical map.

SOM has an advantage of mapping complex and non-linear statistical relationship of high-dimensional input data into simple geometric relationships on a lower dimensional space. Such compression retains the original topological relation of the data items. There are only one layer of input neurons and one layer of output neurons in SOM. Each neuron of the input layer represents an input signal. Each neuron in the output layer represents a node of the final structure. After the unsupervised learning process, all input signals are organized into clusters of regions on a two-dimensional spatial representation. Neurons that are closer together in the output pattern have similar semantic relationship while neurons that are further away are less relevant to each other. Features that occur frequently in input dataset are mapped to larger clusters.

The SOM category map clusters similar documents together and automatically assigns label to the categories. The SOM algorithm is summarized as follows:

**Initialize input nodes, output nodes, and connection weights**

Use the most frequently occurring N terms from all documents as the input vector. Create a two-dimensional map of M output nodes. Initialize weights from N unit input nodes to M output nodes to small random values.

**Present each document in order**

Represent each document by a vector of N terms and present it to the system.

**Compute distances between the input node and each output node**

Distance, $D_j$, between input node i and each output node j is

$$D_j = \sum_{i=0}^{N-1} (x_i(t) - w_{ij}(t))^2$$

where $x_i(t)$ is the input to node i at time t and $w_{ij}(t)$ is the weight from input node i to output node j at time t.

**Updating weights of the winning output node and its neighbors to reduce their distance**

Select the winning output node, $j^*$, that has the minimum $D_j$.

$$\arg \min_{j \neq j^*} D_j = w_{ij}$$

Update weights for node j and its neighbors.

$$w_{ij}(t+1) = w_{ij}(t) + \eta h_j(t)(w_{ij}(t) - w_{ij}(t))$$

where $\eta$ represents the learning factor, $h_j(t)$ represents the neighborhood function, and $w_{ij}(t)$ represents the weights of the winning node.

**Label regions in map**

After the network is trained through repeated presentation of all documents in the information space, submit unit input vector of individual terms to the trained network and label the winning node the input term. Neighboring nodes, which have the same labeling term, form a *region* with the same concept. Submit each document of the information space to the trained network and assign it to a particular node in the map.

An example of a 20 by 20 category map is shown in Figure 1.

![Figure 1. Sample of SOM.](image)
space into a two-dimensional map. Information control techniques supplement the visualization of details in large category maps by incorporating with the rendering techniques for human interaction. Users interact with the visualization system by adjusting parameters of viewing and image modification operations or changing the points of interest while navigating the information space. Using fisheye views and fractal views, users may select the point of interest and adjust parameters to modify the viewing effects of category map and reduce information being displayed in category map.

3.1. Fisheye Views
Fisheye views, first developed by Furnas [5] and further enhanced by Sarkar and Brown [3,13], are known as distortion techniques in information visualization [6]. Regions of interest are enlarged and the other regions are diminished so that one or more parts of a view are emphasized to show the importance of those regions. Both local details of the regions of interest and global structure of the overall display are maintained. By moving the point of interest, users may explore different areas of the two-dimensional map. When users are navigating the category map, they may start from the concepts of interest and explore the similar concepts in the neighbourhood of the category map.

The output nodes of SOM are mapped onto a two dimensional M1 by M2 grids (M1 × M2 = M). Therefore, each output node will then be represented by a square with four vertices. Using fisheye views, the normal coordinates of the vertices, (xnorm, ynorm), are transformed to the fisheye coordinates, (xfeye, yfeye), based on the focus point, (xfocus, yfocus), and Cartesian transformation. Equation 1 presents the fisheye transformation based on Cartesian transformation.

\[
P_{\text{feye}} = (x_{\text{feye}}, y_{\text{feye}}) = \left( G\left(\frac{D_{\text{norm}}}{D_{\text{max}}}\right)D_{\text{max}} + x_{\text{norm}}, G\left(\frac{D_{\text{norm}}}{D_{\text{max}}}\right)D_{\text{max}} + y_{\text{norm}}\right)
\]

where
\[
G(z) = \frac{(d+1)z}{d \times z + 1}
\]
\[
d \text{ corresponds to the distortion factor}
\]
\[
D_{\text{norm}}, \text{ and } D_{\text{norm}} \text{ are the horizontal and vertical distance between the vertex and the focus point in normal coordination}
\]
\[
D_{\text{max}}, \text{ and } D_{\text{max}} \text{ are the maximum horizontal and vertical distance between the boundaries of the window and the focus point in normal coordination}
\]
\[
x_{\text{focus}} \text{ and } y_{\text{focus}} \text{ are the coordinates of the focus point}
\]

When the distortion factor, \(d\), equals to zero, there is not any magnification of the focus area. As \(d\) increases, the focus region will be magnified and the further regions will be diminished. Figure 2 illustrates the effect of the distortion values on 20 by 20 grids. The differences between the size of the grids near the focus point and the size of the grids further away from the focus point increases as \(d\) increases.

![Figure 2. Fisheye views of 20 by 20 grids with distortion values (a) \(d = 0.0\), (b) \(d = 3.0\), (c) \(d = 5.0\), (d) \(d = 10.0\).](image)

Fisheye views support users to navigate the category map by exploring the neighbourhood regions of the point of interest while their corresponding location in the overall structure can be determined. However, users may run into the problem of mental overload because of different distortion ratio within a view [2].

2.2 Fractal Views
Fractal views uses the information reduction approach to control the amount of information displayed by focusing on the syntactic structure of the information. Fractal is an important concept in fractal theory [9] to describe the complexity of an entity in both quantitative and mathematical terms. Fractal view [8] is an approximation mechanism to abstract complex objects and control the amount of information to be displayed with a scale (threshold) set by users. Using such mechanism, the total amount of information displayed is consistent given any focuses of attention but only details near the focus point and important landmarks further away are displayed. The amount of information displayed is also flexible to the interest of users.

Fractal dimension of a structure, \(D\), is the similarity dimension of a structure, which is controlled by a scale factor and branching factor.

\[
D = -\log_{r_s} N_s \quad \text{(Equation 2)}
\]

where \(r_s\) represents the scale factor and \(N_s\) represents the branching factor

In order to satisfy the fractal requirement, the relation between the number branches and the scale factor at each node of the structure as shown in (Equation 4) must exist.

\[
\log_{r_s} N_s = \text{constant} \quad \text{(Equation 3)}
\]

To formalize the fractal views, the focus point is taken into account and is regarded as a root. Fractal values
are propagated from the root to other nodes based on the following formulations:

Fractal value of focus point = \( F_{\text{focus}} = 1 \)

Fractal values of child of region \( x \) in category map 
\[ F_{\text{child}, x} = r_x F_x \]

where \( F_x \) is the fractal value of \( x \)
\[ r_x = C \times N_x^{-D} \]

\( C \) is a constant; \( 0 \leq C \leq 1 \)
\( D \) is the fractal dimension,
\( N_x \) is the branching factor

The regions with their fractal values less than a threshold value will become invisible. The threshold value controls the degree of abstraction by fractal views. The lower the threshold value is, the higher the degree of abstraction is.

Figure 3 illustrates the effect of fractal views. Figure 3 (a) presents a 4 by 5 category map with 7 regions. Given that region A is selected as the focus, Figure 3 (b) and Figure 3 (c) shows the effect of fractal views with fractal values equal to 0.3 and 0.1, respectively.

2.3 Comparisons and Illustrations of Fisheye Views and Fractal Views on category map

Fisheye views and fractal views support the visualization of large category maps by two different approaches. Fisheye views use the distortion approach while fractal views use the information reduction approach. Figure 4 (a) illustrates the fisheye views of the category map shown in Figure 1. Figure 4 (b) and (c) illustrates the fractal views of the same category map with thresholds equal to 0.41 and 0.19, respectively.

The global structure of category map is maintained by fisheye views but it is not maintained by fractal views. Therefore, users are not able to determine the relative location of the areas of interest in the category map by using fractal views. As shown in Figure 4 (b) and (c), the number regions being filtered increases as the threshold decreases from 0.41 to 0.19. 34 regions out of 65 regions are filtered in Figure 4 (b) but 54 regions out of 65 regions are filtered in Figure 4 (c). The relative location of the region, “child”, in Figure 4 (b) can still be determined because some regions on the boundaries of the category map are not filtered yet. However, the relative location of the same region can no longer be determined in Figure 4 (c) because only the regions within a small neighbourhood of that region are not filtered. Fractal views reduce the information being displayed to avoid the problem of overloading, however, it loses the global structure of category map.

The relative size of the regions in category map is maintained by fractal views but it is not maintained by fisheye views. The relative size of the regions represents the importance of the corresponding concepts labelled on the region. The actual size of regions is distorted by using fisheye views, and therefore, the relative importance of the concepts presented by category map is distorted simultaneously. As shown in Figure 1, the size of the regions, “child” and “property”. However, as shown in Figure 4 (a), the size of the region “property” is much smaller than the size of the region “child”. Fisheye views magnify the areas of interest in order to present the corresponding details, however, it loses the relative size to show the importance of the regions in category map.
3. Experiments
In this section, we present the experiment that we have performed to compare the effectiveness of the visualization techniques, fisheye views and fractal views, on category map.

3.1 Experiment Design
3.1.1 Participants
Twenty subjects were recruited to participate in the experiment. Three quarters of them were undergraduate students and one quarter of them were graduate students in the engineering school. No subjects have any prior experience of category map browser and visualization tools. However, a training session was given to all subjects before they conduct the test. The training session includes 1. the introduction of category map and its characteristics, 2. the introduction of fisheye views and fractal views and how their control parameters affect the view of category map, and 3. a 15 minutes session of practice to allow subject get familiar with the tools.

There were totally three setups of the category map visualization techniques, 1) fisheye views, 2) fractal views, and 3) no visualization technique. Each Subject was assigned to use the three setups in a random order.

3.1.2 Experimental Web Space and Searching Tasks
The Hong Kong government Web site is selected as the Web information space to be explored in the experiment. Over 1000 web pages are available in this information space containing information such as information from the office of chief executive, news update, government and related organizations, etc. A 20 by 20 category map is automatically generated to categorize the Web pages of this site.

Searching tasks were randomly assigned to the subjects in the experiment. First, information were extracted from the randomly selected Web pages of the Hong Kong government Web site and presented to the subjects. Subjects had three minutes to interpret the extracted information to understand what information was of interest. Then, subjects were asked to use the assigned visualization setup to retrieve the Web page that provided the requested information. The time taken to complete such task was recorded. The task being assigned to each subject for the four different setups are different.

3.1.3 System Interface for the Experiment
For each visualization setup provided to the subjects, there are three interface panels in the experiment prototype system, 1) category map interface, 2) control interface, and 3) documents interface. The category map interface panel display the automatically generated SOM of Hong Kong government Web site. Each region is assigned a color and its boundaries are blackened. The labels are presented on the regions followed by the number of available documents in the regions parenthesised.

When users click on a region in the category map interface panel, the documents of the region will be presented in the document interface panel together with the documents' keywords and titles. A window showing the complete document will be opened when users click on the button “Show document”. Users may read the document to determine if the requested information is available.

The control interface panel provides the control panel to adjust the parameters for fisheye views and fractal views. The result of using such parameters will then be displayed on the category map in the category map interface.

3.2 Results and Discussion
The average time taken for twenty subjects to complete the assigned tasks using three different setups is shown in Figure 6. We have conducted the t-test to investigate the difference among the visualization techniques. The results are shown in Table 1 and Table 2. We find that the time taken to accomplish the tasks by using fractal views is significantly less than that...
better than fisheye views' performance. Besides, fractal views' performance is significantly improvement in the visualization of category map. Fisheye views and fractal views produce significant improvement in the visualization of category map. Besides, fractal views' performance is significantly better than fisheye views' performance.

Table 1. T-test between fisheye views and no visualization technique, between fractal views and no visualization technique, and between combination of both and no visualization technique

<table>
<thead>
<tr>
<th>Comparison</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between fisheye views and no visualization techniques</td>
<td>0.057383</td>
</tr>
<tr>
<td>Between fractal views and no visualization techniques</td>
<td>0.006728</td>
</tr>
</tbody>
</table>

Table 2. T-test between fisheye views and fractal views

<table>
<thead>
<tr>
<th>Comparison</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between fisheye views and fractal views</td>
<td>0.077897</td>
</tr>
</tbody>
</table>

Based on the above observations, the information reduction approach (fractal views) is more effective that the distortion approach (fisheye views) in visualization of category map although both are useful. Maintaining the global structure is important in visualization, but reducing the visual load is more helpful to users in order to explore the areas of interest. The major drawback of fisheye views is the distortion of the original shape of regions in category map. Users have difficulty to integrate the views produced by using different focus points and distortion values. On the other hand, fractal views require less mental integration between the views produced by using different thresholds. It is because there are only additions or reductions of regions being displayed but shapes of regions are not changed.

4. Conclusion
Category map has been proven a powerful browsing tool for large information space and World Wide Web. However, visualization of category map becomes less effective as the size of the map increases to accommodate more concepts of the information space. In this paper, fisheye views and fractal views are proposed to support the visualization of large category map. Fisheye views distort the map to present the local details and maintain the global structure. Fractal views reduce the information presented on the map by keeping the most relevant regions of the region of interest. A user evaluation has been conducted to investigate the performance of the proposed techniques. The experiment results show that both fisheye views and fractal views improve the effectiveness of visualization significantly. In addition, fractal views perform significantly better than fisheye views. Most users have difficulty to integrate the change of views caused by fisheye views. In a few cases, users were even disoriented when the distortion values are large. On the other hand, most users were able to navigate the category map using fractal views by decreasing the threshold from a large value to smaller values.

Acknowledgement
This project is supported by the Earmarked Grant for Research of the Research Grants Council of Hong Kong, CUHK 7034-98E.

References