ELSEVIER

# Visualization of large category map for Internet browsing

Christopher C. Yang [a],[*], Hsinchun Chen [b], Kay Hong [a]

[a]*Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Hong Kong, China*
[b]*Department of Management Information Systems, The University of Arizona, Tucson, AZ 75721, USA*

## Abstract

Information overload is a critical problem in World Wide Web. Category map developed based on Kohonen's self-organizing map (SOM) has been proven to be a promising browsing tool for the Web. The SOM algorithm automatically categorizes a large Internet information space into manageable sub-spaces. It 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. It results in increasing of visual load of the category map. Large pool of information is packed closely together on a limited size of displaying window, where local details are difficult to be clearly seen. In this paper, we propose the fisheye views and fractal views to support the visualization of category map. Fisheye views are developed based on the distortion approach while fractal views are developed based on the information reduction approach. The purpose of fisheye views are to enlarge the regions of interest and diminish the regions that are further away while maintaining the global structure. On the other hand, fractal views are an approximation mechanism to abstract complex objects and control the amount of information to be displayed. 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 significantly better than fisheye views but the combination of fractal views and fisheye views do not increase the performance compared to each individual technique.
© 2002 Elsevier Science B.V. All rights reserved.

## 1. Introduction

The Kohonen self-organizing map (SOM)-based *category map* is a promising two-dimensional graphical tool for Internet browsing. Experiments showed that it is capable to categorize a large Internet information space, Yahoo Entertainment sub-category [8].

Given a large electronic information space, the SOM automatically categorize it into manageable sub-spaces so that users can navigate the map to locate the documents of interest.

There are two major objectives of individual information seeking: (1) *explore* the information space, gain familiarity and locate something of interest, and 2) *search* the information space to retrieve relevant information on a given topic [8]. A user may start with a homepage and surf through the links to explore something that may be of interest. Some other users may have something in mind and use a search engine

[*] Corresponding author. Tel.: +852-2609-8239; fax: +852-2603-5505.
*E-mail address:* yang@se.cuhk.edu.hk (C.C. Yang).

to retrieve the documents that includes the relevant information. Prior knowledge of the organization or categorization of the information is helpful to achieve the objective of exploring [5]; while the knowledge of the subject of interest is important to achieve the objective of searching [12]. For most of the web surfers, they may not have a specific subject of interest in mind yet when they begin their browsing or they may not have the knowledge of specific vocabulary to represent their interest to be submitted to the search engine. Acquiring information from the huge information space of the Web is a difficult task. It is easy for users to become disoriented. The category map developed based on Kohonen SOM 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.

As the information increases exponentially (information overloading problem [6,27]), 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. Visual load makes the display of information less intuitive and useful. 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 [26]. Zooming is able to present the local details, however, it loses the global structure of the category map. Using multiple windows allow 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 employ the fisheye views and fractal views to develop a visualization tool for visualizing category map.

The purpose of visualization is mapping information onto graphical representation to gain insight for users. Information visualization assists users to view and locate the information of interest in a limited space by interacting with the visualization system. Bertin [3] claimed that visualization technique using geometry to display quantitative data (such as scale and size) is more effective than other visualization techniques such as color. In our work, fisheye views and fractal views are utilized to solve the problem of visual load in category map. 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 [20]. The local details are magnified by the fisheye views and the global structure is maintained. Fractal views are derived based on the fractal theory [19] 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.

## 2. Browsing by category map

Browsing is an exploratory and information-seeking strategy that depends upon serendipity, which is appropriate for ill-defined problems or exploring new task domains [22]. It is characterized by the absence of planning [21]. 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 is 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 [22], are important for exploring an unknown information space. They represent the content, structure, and relationship of information in the information space [10]. 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. Hierarchical categories are the most typical mental mod-

els. Yahoo! is a prominent example providing searching services on the Web using such mental model. However, the categories are limited in their granularity and timeliness. The creating process of the categories is manual and cumbersome. On the other hand, category map generated by SOM is automatic. The important concepts in the information space are displayed by an intuitive graphical map.

The semantic network is first suggested by Quillian [23] to encode and associate word meanings to visualize and construct mental models of information space. The SOM is identified to be the most promising algorithm to organize large scale of information by creating an intuitive, graphical display of the important concepts contained in textual information [7,18]. The objective of Kohonen's self-organizing map [16] is to abstract and organize various features of input data or signals, and quantize them into categorized representation while retaining the topological relation.

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 as shown in Fig. 1. 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 cluster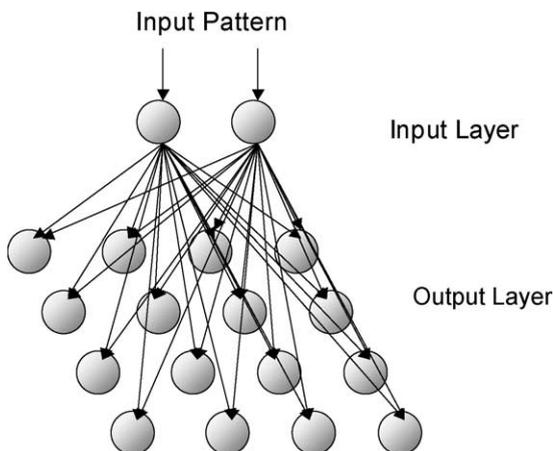s of regions on a two-dimensional spatial representation. Neurons that are closer together in the output pattern have similar semantic relationship while neurons that are further away are less relevant to each other. Features that occur frequently in input dataset are mapped to larger clusters.

The SOM category map clusters similar documents together and automatically assigns label to the categories. Similar documents, which are represented by similar noun phrase terms, are grouped together in a neighborhood. 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 $D_j$.

$$\arg \min_{j \in M} \| x_i - w_{ij}(t) \|$$

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



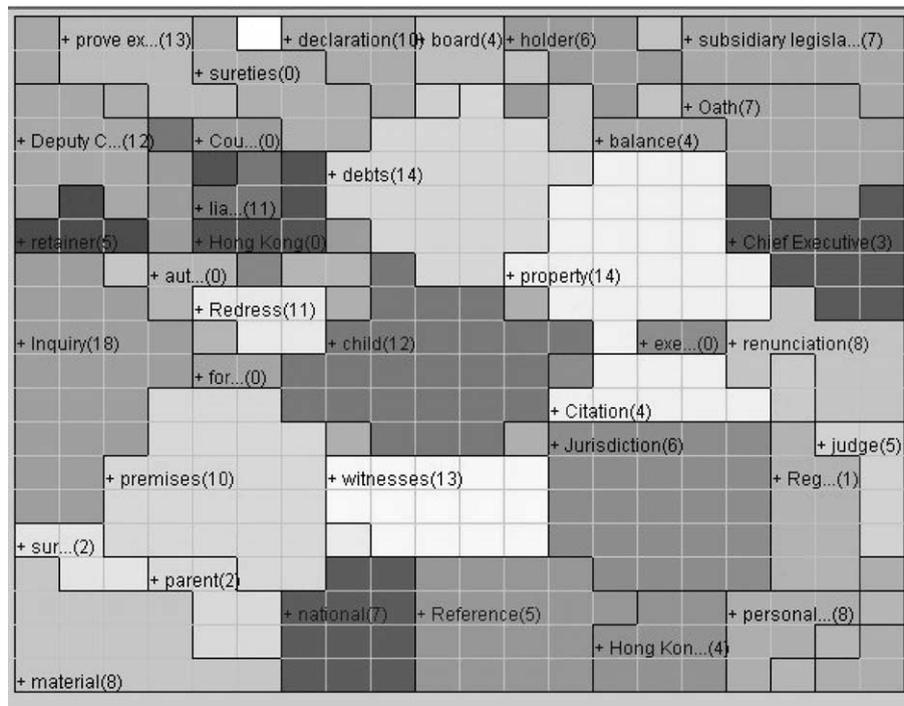Fig. 1. Structure of two-dimensional self-organizing map.

Fig. 2. Sample of SOM.

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 Fig. 2.

## 3. Information control in information visualization

Information visualization is a human computer interaction process that can be generally divided into two stages: (1) information interpretation and mapping, and (2) information display and control. In information interpretation and mapping, data are transformed from multi-dimensional quantitative and qualitative features into of two or three-dimensional graphical representations for subsequent visualization operations. The transformed representations are then rendered and displayed. The category map achieves such purpose by compressing a complex information space into a two-dimensional map. However, as the number of nodes in the two-dimensional category map increases, it may not be feasible to present the labels in some of the smaller regions. As illustrated in Fig. 2,

there are totally 65 regions in the 20 by 20 category map; however, only 22 of them are able to present the complete labels within the region. Therefore, information control techniques are required to support the visualization of the details of large category maps. The information control techniques incorporate with the rendering techniques by controlling the viewing and image modification operations, such as zooming, rotation, hiding, etc., for human interaction. Users interact with the visualization system by adjusting parameters or changing the points of interest while navigating the information space. In this paper, we develop the fisheye views and fractal views for visualization of category map and explore their effectiveness. Users may select the point of interest and adjust parameters to modify the viewing effects of category map and reduce information being displayed on category map.

### 3.1. Fisheye views

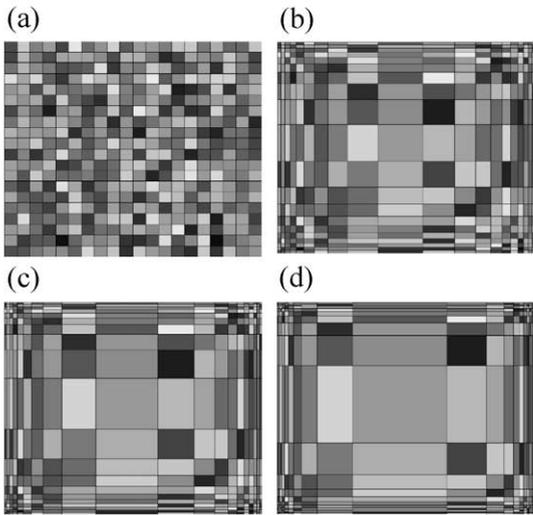Fisheye views, first developed by Furnas [11] and further enhanced by Sarkar and Brown [4,20,24], are

Fig. 3. Fisheye views of 20 by 20 grids with distortion values (a) $d = 0.0$, (b) $d = 3.0$, (c) $d = 5.0$, (d) $d = 10.0$.

known as distortion techniques in information visualization [13]. 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 neighborhood of the category map. Fisheye views enable users to explore the region of interest and visualize the relationship of the further diminished regions.

The output nodes of SOM are mapped onto a two-dimensional $M_1$ by $M_2$ grids ($M_1 \times M_2 = M$). Therefore, each output node will then be represented by a square with four vertices. Using fisheye views, the normal coordinates of the vertices, ($x_{norm}$, $y_{norm}$), are
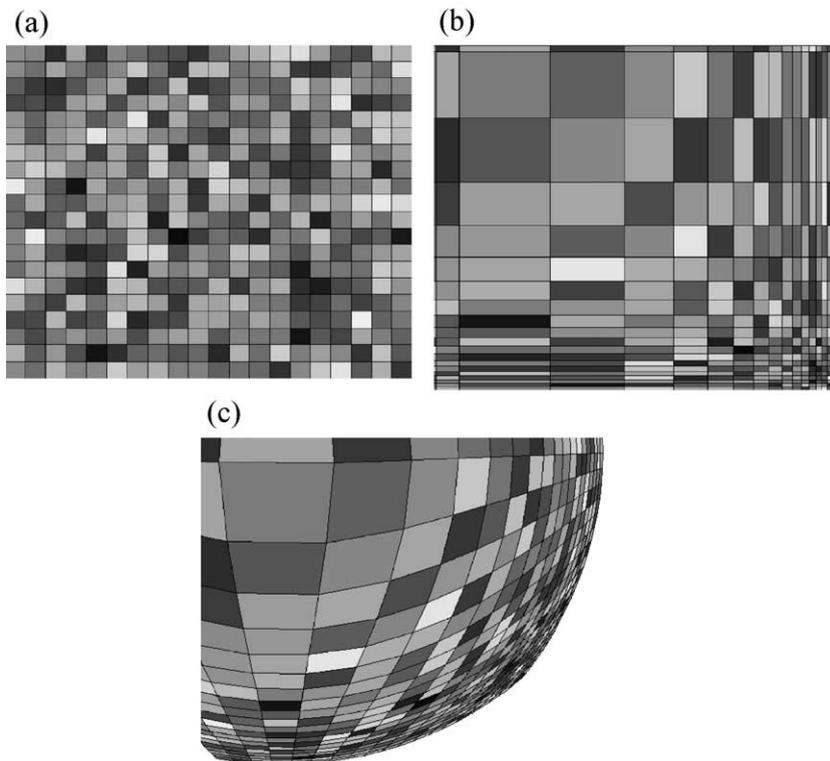


Fig. 4. 20 by 20 grids (a) with no distortion, (b) using Cartesian transformation, (c) using polar transformation.

transformed to the fisheye coordinates, $(x_{feye}, y_{feye})$, based on the focus point, $(x_{focus}, y_{focus})$, and Cartesian or polar transformation. Eq. (1) presents the fisheye transformation based on Cartesian transformation.

$$
\begin{aligned}
P_{feye} &= (x_{feye}, y_{feye}) \\
&= \left( G\left(\frac{D_{norm_x}}{D_{max_x}}\right) D_{max_x} + x_{focus}, G\left(\frac{D_{norm_y}}{D_{max_y}}\right) \right. \\
&\quad \left. \times D_{max_y} + y_{focus} \right)
\end{aligned}
\tag{1}
$$

where $G(z)=(d+1)z/(d \times z+1)$. $d$ corresponds to the distortion factor, $D_{norm_x}$ and $D_{norm_y}$ are the horizontal and vertical distance between the vertex and the focus point in normal coordination, $D_{max_x}$ and $D_{max_y}$ are the maximum horizontal and vertical distance between the boundaries of the window and the focus point in normal coordination, $x_{focus}$ and $y_{focus}$ 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. Fig. 3 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.

The Cartesian transformation produces a fisheye view that fully utilizes the displaying window. However, the width and height of the grids are distorted separately according their horizontal and vertical distances. Therefore, grids that are of the same distance from the focus point may not have the same size. It may appear as a rectangle with the width much longer than the height if its horizontal distance for the focus point is small. On the other, it may appear as a rectangle with the height much longer than the width if its vertical distance from the point is small. If the horizontal and vertical distances from the focus point are approximately the same, it appears as a square. The polar transformation avoids this problem. It distorts the distant regions based on its distance from the focus point. The polar transformation are computed as follows:

$$
P_{feye} = (x_{feye}, y_{feye}) = (r_{feye}\cos\theta, r_{feye}\sin\theta)
\tag{2}
$$

where $r_{feye}=((d+1)r_{norm})/((d \times r_{norm})/(b_{min})+1)$. $d$ corresponds to the distortion factor

$$
r_{norm} = \sqrt{(x_{norm} - x_{focus})^2 + (y_{norm} - y_{focus})^2}
$$

$$
\theta = \tan^{-1}\left(\frac{y_{norm} - y_{focus}}{x_{norm} - x_{focus}}\right)
$$

$b_{min}$ corresponds to the minimum boundary size of the displaying window, $x_{focus}$ and $y_{focus}$ are the coordinates of the focus point, $x_{norm}$ and $y_{norm}$ are the coordinates of the point before transformation.

Fig. 4 illustrates the difference between the effects of the Cartesian transformation and polar transformation. Although polar transformation does not produce different horizontal and vertical distortion, it does not fully utilize the displaying window.

Fisheye views support users to navigate the category map by exploring the neighborhood 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].
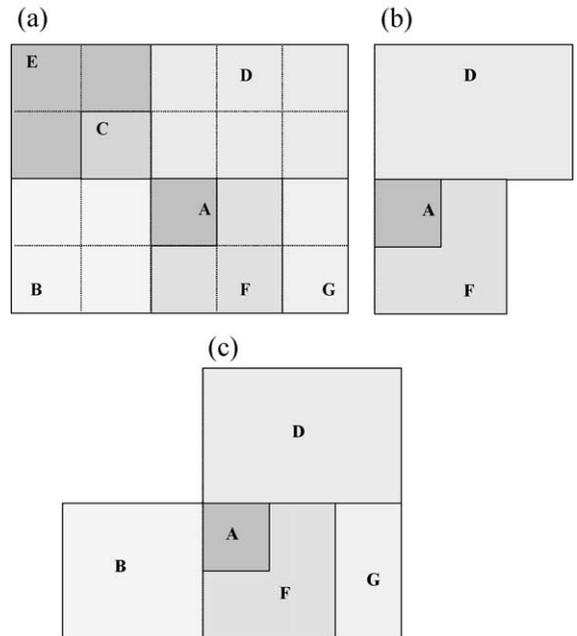


Fig. 5. (a) A sample of 4 by 5 SOM with seven regions, (b) fractal view of SOM with threshold equals to 0.3, (c) fractal view of SOM with threshold equals to 0.1.

## 3.2. Fractal views

Fractal views are information control methods for information display. It uses the information reduction approach to control the amount of information displayed by focusing on the syntactic structure of the information. Fractal [9] is an important concept in fractal theory [19] to describe the complexity of an entity in both quantitative terms and mathematical terms. Fractal view [17] 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.

$$D = -\log_{r_x} N_x \qquad (3)$$

where $r_x$ represents the scale factor and $N_x$ represents the branching factor.

In order to satisfy the fractal requirement, the relation between the number branches and the scale



Fig. 6. (a) Fisheye view of category map in Fig. 2 using Cartesian transformation, (b) fisheye view of category map in Fig. 2 using polar transformation, (c) fractal view of category map in Fig. 2 with threshold equals to 0.41, (d) fractal view of category map in Fig. 2 with threshold equals to 0.19.

factor at each node of the structure as shown in Eq. (4) must exist.

$$\log_{r_x} N_x = \text{constant} \qquad (4)$$

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 $= F_{\text{child\_of\_}x} = r_x F_x$

where $F_x$ is the fractal value of $x$, $r_x = C \times N_x^{-1/D}$, $C$ 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.

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

### 3.3. 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. Fig. 6(a) and (b) illustrates the fisheye views of the category map shown in Fig. 2 using Cartesian transformation and polar transformation, respectively. Fig. 6(c) and (d) 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 Fig. 6(c) and (d), 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 Fig. 6(c) but 54 regions out of 65 regions are filtered in Fig. 6(d). The relative location of the region, "child", in Fig. 6(c) 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 because only the regions within a small neighborhood of that region are not filtered. The location of the whole neighborhood is impossible to be identified after such large amount of information reduction. Fractal views reduce the information being displayed to avoid the problem of overloading, however, it loses the global structure of category map.
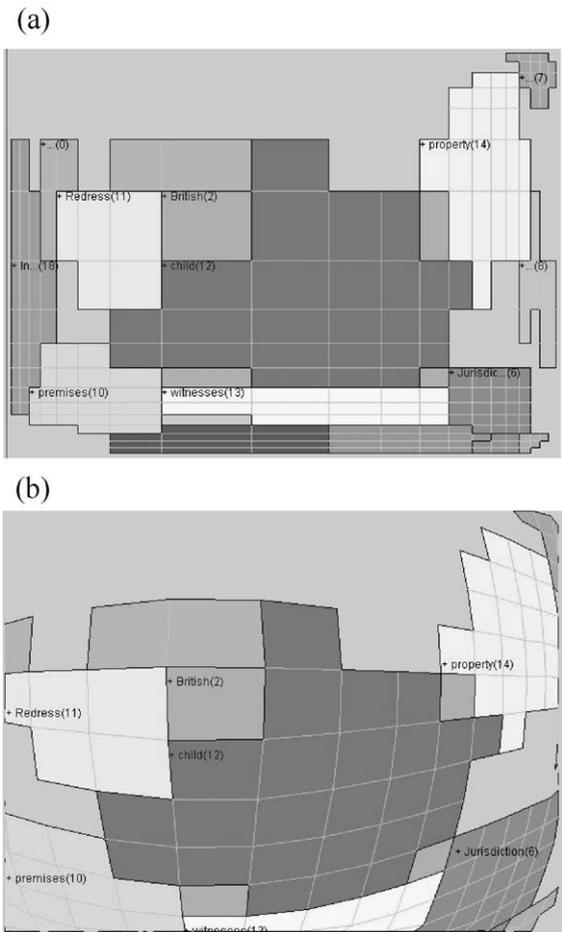


Fig. 7. Combination of fisheye views and fractal views. (a) Cartesian transformation of fisheye view and fractal view, and (b) polar transformation of fisheye view and fractal view.

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. Since category map uses the neighborhood function in training, the regions representing the important concept in the information space grow in size as more and more documents are submitted as input to the network during the iterations. 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 Fig. 2, the size of the regions, "child" and "property", are approximately the same, where the size of "child" is 25 units and the size of "property" is 21 units. However, as shown in Fig. 6(a) and (b), 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.

Fisheye views and fractal views have their own advantages and disadvantages. We combine both techniques to develop a visualization tool for category map such that users may adjust the parameters of both views to take their advantages while navigating the SOM. Fig. 7 illustrates the effect of combining fisheye views and fractal views.

## 4. Experiments

An experiment have been conducted to compare the effectiveness of the visualization techniques, fisheye views and fractal views, on category map. This study analysed how different visualization techniques affect the performance of information searching tasks on category map. The research questions concerned are:

1. Do fisheye views and fractal views improve the performance of information searching tasks on category map? Which technique performs better?
2. Does the combination of fisheye views and fractal views perform better than applying individual technique?

Visualization techniques are evaluated with respect to their suitability for certain task characteristics [14,15]. Task-based experiments are commonly used in comparing the performances of different visualization techniques [25]. Subjects are asked to use the proposed visualization techniques to accomplish some tasks depending on the applications of the visualization techniques. Time taken to accomplish the assigned tasks is measured. Statistical analysis will then be conducted to investigate if there is any significant time difference. In our experiments, we use the task-based approach where different combination of fisheye views and fractal views will be provided to the subjects and the time taken to accomplish the randomly assigned search tasks will be measured.

### 4.1. Experiment design

#### 4.1.1. 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-min session of practice to allow subject get familiar with the tools.

There were totally four setups of the category map visualization techniques, (1) fisheye views, (2) fractal views, (3) combination of fisheye views and fractal views, and (4) no visualization technique. Each subject was assigned to use the four setups in a random order.

#### 4.1.2. 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. Since all Web pages were used to generate the category map for experiments, the randomly extracted information viewed by the subjects was included in the documents available on the category map. Subjects had 3 min to interpret the extracted information to understand what 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.

All subjects are local students in Hong Kong with similar educational background. The domain is the general information about Hong Kong and Hong Kong SAR government. Therefore, none of the subjects has more advance understanding of the concepts being used in the experiment.

### 4.1.3. 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 displays 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.

For users using the setup other than the one without any visualization techniques, the control interface panel provides the control panel to adjust the param-
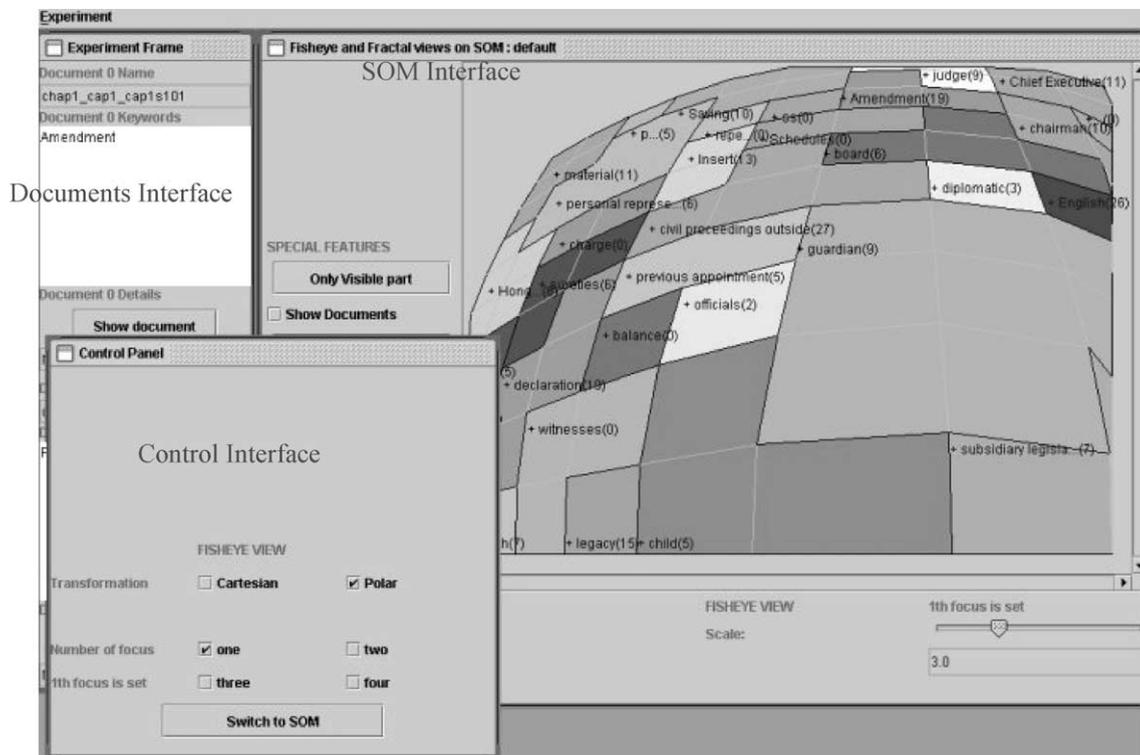


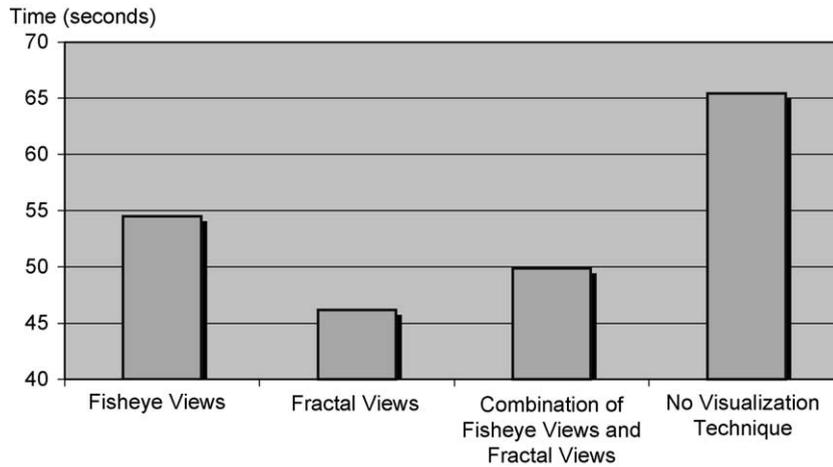Fig. 8. System interface of the experiment.

Fig. 9. Experimental results.

eters 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.

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.

Given the searching task, subjects may select their own focus point based on the labels of the category map, where the label are automatically generated by the self-organization map algorithm as described in Section 1. The labels on the category map are the most significant concept in the information space. The larger the region of the label is, the more significant the corresponding concept in the information space is (Fig. 8).

### 4.2. Results and discussion

The average time taken for 20 subjects to complete the assigned tasks using four different setups is shown in Fig. 9. We have conducted the one-way analysis of variance (ANOVA) to analyse if there is any significant difference between the performances of different visualization techniques. The result is shown in Table 1. The $F$-ratio is 3.584 and the $p$-value is 0.018. That means the difference between the time taken to complete tasks by using fisheye views, fractal views, combination of both and no visualization technique is significant at the level of 0.018.

Since there is significant difference between using different visualization techniques, we conduct the $t$-test to investigate the difference among the visualization techniques. The results are shown in Tables 2 and 3. We find that the time taken to accomplish the tasks by using fractal views is significantly less than that using no visualization technique ($p = 0.006728$).

Table 1
ANOVA result of the time taken between different visualization techniques

|  | Sum of squares (SS) | Degree of freedom ($df$) | Mean sum of square (MSS) | $F$-ratio ($F$) | $p$-value ($p$) |
|---|---|---|---|---|---|
| Between different visualization techniques | 2550.508 | 3 | 850.1693 | 3.583764 | 0.017567 |
| Within different visualization techniques | 18,029.33 | 76 | 237.228 | | |

However, the difference between using fisheye views and no visualization techniques is comparatively less significant ($p > 0.05$). We also find that the difference between using fisheye views and using fractal views is significant at the level of 0.0787, but there are no significant difference between using fisheye views and combination of both and between fractal views and combination of both. Therefore, we conclude that both fisheye views and fractal views produce significant improvement in the visualization of category map. Besides, fractal views' performance is significantly better that fisheye views' performance at the level of 0.0787. However, combination of fisheye views and fractal views does not produce any significant changes in the performance of visualizing category map.

Based on the above observations, the information reduction approach (fractal views) is more effective than 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.

During the experiment, users were free to select Cartesian transformation or polar transformation for fisheye views. However, Cartesian transformation was

Table 2
*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

|  | *p*-value |
|---|---|
| Between fisheye views and no visualization techniques | 0.057383 |
| Between fractal views and no visualization techniques | 0.006728 |
| Between combination of fisheye views and fractal views and no visualization techniques | 0.009373 |

Table 3
*T*-test between fisheye views and fractal views, between fisheye views and combination of both, and between fractal views and combination of both

|  | *p*-value |
|---|---|
| Between fisheye views and fractal views | 0.077897 |
| Between fisheye views and combination of fisheye views and fractal views | 0.202655 |
| Between fractal views and combination of fisheye views and fractal views | 0.254825 |

preferred. Most users commented that polar transformation produced more fancy interface, but it was more difficult to control. It was mainly because the shape of regions was significantly distorted. This problem becomes worse, when the distortion value is large. Since users had the flexibility to choose between Cartesian and polar transformation and most users had chosen Cartesian transformation, there was not enough statistical data to analyse if there were significant difference between using Cartesian transformation and polar transformation. We conducted another experiment where 10 subjects were assigned searching tasks using Cartesian transformation or polar transformation in random order. A *t*-test is then conducted to analyse the result. It is found that $p = 0.352931$. There is no significant time difference although polar transformation is preferred by users.

## 5. 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 per-

formance of the proposed techniques. Twenty subjects participated in the experiment. Each of them was asked to use different techniques in random order to complete tasks and the time taken was recorded. The experiment results show that both fisheye views and fractal views improve the effectiveness of visualization significantly. However, combining both techniques do not produce any significant difference compared with the individual techniques. In addition, fractal views perform significantly better than fisheye views. Most users have difficulty integrating the change of views when they move the focus point around to explore other regions. In a few cases, users were even disoriented when the distortion values are large. On the other hand, fractal views are much user-friendly. Most users were able to navigate the category map by decreasing the threshold from a large value to smaller values to locate their areas of interest.

## References

[1] J. Assa, D. Cohen-Or, T. Milo, Displaying data in multidimensional relevance space with 2D visualization maps, Proceedings of IEEE Conference on Visualization, 1997.

[2] L. Bartram, A. Ho, J. Dill, F. Henigman, The continuous zoom: a constrained fisheye technique for viewing and navigating large information space, Proceedings of the User Interface Software and Technology, Pittsburgh, PA, 1995, pp. 207–215.

[3] J. Bertin, Graphics and Graphic Information Processing Berlin: de Gruyter, translated by William J. Berg and Paul Scott, 1981.

[4] M.H. Brown, J.R. Meehan, M. Sarkar, Browsing graphs using a fisheye view, Proceeding of ACM on Human Factors in Computing Systems, Amsterdam, Netherlands, 1993.

[5] H. Chen, V. Dhar, User misconceptions of online information retrieval systems, International Journal of Man–Machine Studies 32 (6) (1990) 673–692.

[6] H. Chen, K.J. Lynch, Automatic construction of networks of concepts characterizing document databases, IEEE Transactions on Systems, Man, and Cybernetics 22 (1992) 885–902.

[7] H. Chen, C. Shuffels, R. Orwig, Internet categorization and search: a machine learning approach, Journal of Visual Communications and Image Representation 7 (1) (1996) 88–102.

[8] H. Chen, A.L. Houston, R.R. Sewell, B.R. Schatz, Internet browsing and searching: user evaluations of category map and concept space techniques, Journal of the American Society for Information Science 49 (7) (1998) 582–603.

[9] J. Feder, Fractals, Plenum, New York, 1988.

[10] C.L. Foss, Tools for reading and browsing hypertext, Information Processing and Management 25 (4) (1989) 407–418.

[11] G.W. Furnas, Generalized fisheye views, Proceedings of the SIGCHI Conference on Human Factors in Computing System, 1986.

[12] G.W. Furnas, T.K. Landauer, L.M. Gomez, S.T. Dumais, The vocabulary problem in human-system communication, Communications of the ACM 30 (11) (1987) 964–971.

[13] M. Heo, S.C. Hirtle, An empirical comparison of visualization tools to assist information retrieval on the Web, Journal of the American Society for Information Science and Technology 52 (80) (2001) 666–675.

[14] D.A. Keim, An Introduction to Information Visualization Techniques for Exploring Very Large Databases, Tutorial Notes, Information Visualization '00, Salt Lake City, UT, USA, October (2000).

[15] D.A. Keim, Visual exploration of large date sets, Communications of the ACM 44 (8) (August 2001) 39–44.

[16] T. Kohonen, The self-organizing map, Proceedings of the IEEE 78 (9) (1990) 1464–1480.

[17] H. Koike, Fractal views: a fractal-based method for controlling information display, ACM Transactions on Information Systems 13 (3) (1995) 305–323.

[18] X. Lin, D. Doergel, G. Marchionini, A self-organizing semantic map for information retrieval, Proceedings of the Fourteenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Chicago, 1991, pp. 262–269.

[19] B.B. Mandelbrot, The Fractal Geometry of Nature, Freeman, New York, 1977.

[20] S. Manojit, M.H. Brown, Graphical fisheye views, Communications of the ACM 37 (12) (1994) 73–83.

[21] G. Marchionini, An invitation to browse: designing full tex system for novice users, Canadian Journal of Information Science 12 (3) (1987) 69–79.

[22] G. Marchionini, B. Shneiderman, Finding facts vs. browsing knowledge in hypertext systems, IEEE Computer 21 (3) (1988) 70–79.

[23] M.R. Quillian, Semantic memory, Semantic Information Processing, The MIT Press, Cambridge, MA, 1968, pp. 227–270.

[24] M. Sarkar, M.H. Brown, Graphical fisheye views of graphs, Proceedings of ACM on Human Factors in Computing Systems, Monterey, CA, 1992.

[25] D. Turo, B. Johnson, Improving the visualization of hierarchies with treemaps: design issues and experimentation, Proceedings of the IEEE International Conference on Visualization, Los Alamitos, CA, USA, 1992, pp. 124–131.

[26] C.C. Yang, H. Chen, K.K. Hong, Visualization tools for self-organizing maps, Proceedings of the Fourth ACM Conference on Digital Libraries, Berkeley, 1999, pp. 258–259.

[27] C.C. Yang, J. Yen, H. Chen, Intelligent Internet searching agent based on hybrid simulated annealing, Decision Support Systems 28 (3) (May 2000) 269–277.

Christopher C. Yang is currently an associate professor in the Department of Systems Engineering and Engineering Management at the Chinese University of Hong Kong. From 1997 to 1999, he was an assistant professor in the Department of Computer Science and Information Systems and associate director of the Authorized Academic Java[SM] Campus[SM] at the University of Hong Kong. He received his BS, MS, and PhD in Electrical Engineering from the University of Arizona, Tucson, in 1990, 1992, and 1997, respectively. From 1995 to 1997, he was a research scientist in the Artificial Intelligence Laboratory in the Department of Management Information Systems, where he was an active researcher in the Illinois Digital Library project. From 1992 to 1997, he was also a research associate in the Intelligent Systems Laboratory in the Department of Electrical and Computer Engineering. His current research interests are digital library, information visualization, Internet agent, and color image retrieval. He has published over 60 research articles. He was the chairman of the Association for Computing Machinery Hong Kong in 2000 and the program co-chair of the First Asia Digital Library Workshop, 1998.

Dr. Hsinchun Chen is McClelland Endowed Professor of MIS at The University of Arizona. He was also named Andersen Consulting Professor of the Year (1999). He received his PhD in Information Systems from New York University in 1989, MBA in Finance from SUNY-Buffalo in 1985, and BS in Management Science from the National Chiao-Tung University in Taiwan. He is author of more than 90 articles covering medical informatics, semantic retrieval, search algorithms, knowledge discovery, and collaborative computing in leading information technology publications. He is an expert in medical informatics, digital library, knowledge management research, whose work has been featured in various scientific and information technologies publications including *Science*, *New York Times*, *Business Week*, *NCSA Access Magazine*, *WEBster*, and *HPCWire*. Since 1990, Dr. Chen has received more than US$10 M in research funding from various government agencies and major corporations including NSF, DARPA, NASA, NIH, NIJ, NLM, NCI, NCSA, HP, SAP, 3COM, and AT and T. Dr. Chen founded The University of Arizona Artificial Intelligence Lab in 1990, which employs over 40 full-time staff, research scientists, research assistants, and programmers. Dr. Chen is also the Founding Director of The University of Arizona Mark and Susan Hoffman eCommerce Lab (October 2000). The lab has received over US$2 M in infrastructure funding from Mark Hoffman (founder of Commerce One) and HP, and more than US$10 M of software donation from major IT software companies including SAP, Oracle, Microsoft, and IBM.

Kay Hong was a research assistant at the Department of Computer Science and Information Systems in the University of Hong Kong. He received his BS and MPhil in Computer Science and Information Systems from the University of Hong Kong in 1998 and 2001, respectively.