An Exploratory Search for Presentation Contents based on Slide Semantic Structure

Yuanyuan Wang
Kyoto Sangyo University, Japan
circl.wang@gmail.com

Yukiko Kawai
Kyoto Sangyo University, Japan
kawai@cc.kyoto-su.ac.jp

Kazutoshi Sumiya
University of Hyogo, Japan
sumiya@shse.u-hyogo.ac.jp

***Abstract*—MOOC is a crucial platform for improving education; students are able to browse various presentation contents through the Web. Any single presentation content can only cover a small fraction of knowledge in a specific domain by a given query, and thus offers a limited depth of information. Students then have to go through series of presentation contents, but this would be time-consuming and difficult to explore relevant information from various presentation contents. Therefore, we aim to build a novel exploratory search tool based on a meaningfully structured presentation, called "iPoster." The system places elements such as text and graphics of slides in a structural layout with a zooming user interface (ZUI) by semantically analyzing the slide structure. Through this, iPoster can support students interactively browsing slides, for retrieving and navigating information from other presentation contents by considering the students' browsing behavior. In this paper, we discuss two types of exploratory search, (1) focused searching based on well-matched browsing behavior that enables users obtain details of specific topics; and (2) exploratory browsing based on partially-matched browsing behavior that enables the users find various relevant information on topics of interest.**

***Keywords*-exploratory search; presentation contents; iPoster;**

I. Introduction

Slide-based presentation tools, such as Microsoft PowerPoint or Apple Keynote is now one of the most frequently used tools for educational purposes. A huge amount of slide-based educational materials for MOOC, are freely shared on Web sites such as Coursera[1] and SlideShare[2]. Thus, not only students who missed a lecture or presentation, but also anyone interested in a topic can study the presentation on their own. Therefore, techniques are in demand that will efficiently find appropriate information worth learning from the vast numbers of presentations available. Although many techniques for searching and recommending presentation slides have been proposed, some problems remain from the viewpoint of exploratory search. One problem is current slideshow mode of presentations does not allow users operate freely on the presentations for stimulating the users' interests. Recently, Prezi[3] utilizes a ZUI as an alternative

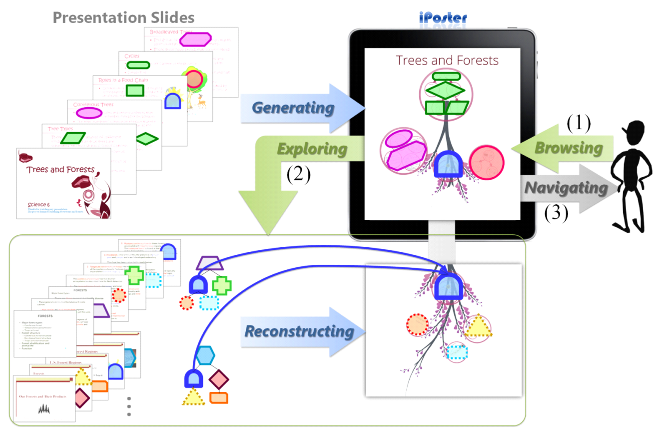

Figure 1. Conceptual diagram of an exploratory search tool by iPoster

to the traditional slides, allowing users easily operate the presentations. Another problem is any single educational material only cover a small fraction of knowledge in a specific domain by a given search query, and thus offers a limited depth of information. The users then have to go through various presentation contents, but this will be time-consuming and difficult to find relevant information from multiple presentation contents. Therefore, users will be required to browse them in structural layouts with ZUIs, and easily obtain information meets the users' specific needs by considering user browsing behavior.

As depicted in Figure 1, we present an exploratory search tool that generates a meaningfully structured presentation by using the slides, which is called an iPoster. With our exploratory search tool, (1) users can interactively browse an iPoster, therefore, (2) the iPoster can explore information from other presentations by considering user browsing behavior; and (3) represent and navigate information meets the users' specific needs. To achieve our goal, the iPoster can be implemented by 1) extracting textual and graphic elements in slides and semantic relationships between them; and 2) organizing elements in structural layouts with zooming and panning transitions based on a idea of Prezi. In semantic structure analysis, we first extract elements by examining the presentation context of the particular elements in the

[1]https://www.coursera.org/
[2]http://www.slideshare.net/
[3]http://prezi.com/

slides. The semantic relationships between these elements are determined using implicit hyperlinks in slides, based on slide structures. Specifically, we derive the slide structure by focusing on the itemized sentences in the slide text. For providing an overview of the content, we utilize a hierarchical structure, combined with a stacked Venn. Finally, our iPoster is generated in a structural layout based on semantic relationships, using a ZUI, which can enable users to explore the presentations easily and efficiently.

The next section reviews related work. Section 3 describes our semantic structure analysis model. Section 4 explains iPoster generation. Our exploratory search on presentations and conclusions are given in Sections 5 and 6, respectively.

II. Related Work

A variety of applications address the weaknesses of the current slideware tools in the presentation and authoring domains. Our approach in an iPoster builds on the strength of exploratory search. Lanir et al. [1] proposed a MultiPresenter application that leverages spatial reasoning capabilities to relate content through dual-screen projection. Although the iPoster does not adopt the dual-audience-display paradigm, it addresses the need to navigate through elements dynamically during the presentation. NextSlidePlease [2] creates and delivers presentations in a nonlinear fashion. The iPoster is similar to this work, as we utilize a structural layout with the ZUI, rather than one or more slide lists, to allow users interactively browsing and automatically navigating.

Exploratory is constantly being changed and shaped by a range of related research. White et al. [3] suggests that browsing is a cognitive and behavioral expression of exploratory behavior and she claims that it has four elements: (1) glimpse a scene; (2) target an element of a scene visually and/or physically; (3) examine items of interest; and (4) physically or conceptually acquire or abandon examined items. Therefore, our method according to this, offer an overview (glimpses), the ability to operate the content through various presentations (exploratory browsing). Detlor et al. [4] developed a model of information seeking that combines both browsing and searching. It suggests that much of Ellis's model [5] is already implemented by components currently available in Web browsers. We then applied this model to search presentations by considering user browsing behavior.

III. Semantic Structure Analysis

A. Element Extraction

There are two elements, i.e., textual elements and graphic elements, from presentation slides based on itemized sentences based on the XML files of slides. We define the slide title is the 1st level, the first item of text within the slide body is the 2nd level, and the depth of the sub-items increases with indentation level (3rd level, 4th level, etc.).

1) Textual Elements: We define textual elements as topics that focus on noun phrases in slides. Based on the presentation context, a topic that frequently appears at the higher levels (i.e., slide title) in neighboring slides. The topics that appear in the title of a slide and the body of other slides can be considered to indicate its context in a presentation. Then, we extract topics by locating the same noun phrases in different slides, at varied levels. If a noun phrase k appears at different levels in slides s_i and s_j, then k is a candidate for being one of the topics T in the presentation.

$$T = \{(k, s_i, s_j) | l_{max}(k, s_i) \neq l_{max}(k, s_j)\} \qquad (1)$$

where, T is a bag of noun phrases that can be considered as candidates for topics. $l_{max}(k, s_i)$ returns the highest level of k in s_i. For instance, when the highest level is the title, i.e., the 1st level, then $l_{max}(k, s_i)$ is 1; and when the highest level is the 3rd level, then $l_{max}(k, s_i)$ is 3. When k appears at different levels, k is determined as a candidate for topics by Eq. (1). Then, the weight of k in T is defined using the levels of k, and the distance between s_i and s_j, as follows:

$$I(k) = \frac{1}{l_{max}(k, s_i)} + \sum_{k, s_i, s_j \in T} \left(\frac{1}{l_{max}(k, s_j)} \cdot \frac{1}{dist(s_i, s_j)} \right) \qquad (2)$$

where $dist(s_i, s_j)$ corresponds to the strength of the association between s_i and s_j, and it denotes the distance between s_i and s_j. If k appears at a high level in s_i and s_j, and the distance between s_i and s_j is short, $I(k)$ of k is high.

2) Graphic Elements: When compared to pure textual elements, images are more attractive, appealing and informative from a psychological standpoint. Therefore, we define graphic elements as images corresponding to the topic candidates in slides, given that the surrounding text of the images are similar to the topic candidates. We considered that the images used to describe the content in slides, and a slide title can be a subject of the content. When the similarity exceeds a predefined threshold by calculating the Simpson similarity coefficient, the images are recognized as the corresponding images of the topic candidates.

B. Determination of Semantic Relationships

Semantic relationships between elements are determined from a document tree of a presentation to enable users obtain relevant information between the key elements. Preliminary ideas are given in an algebraic query model [6] as well.

1) Basic Definitions and Algebra: The presentation content shown in Figure 2 is represented as a rooted ordered tree $D = (N, E)$ with a set of nodes N and a set of edges $E \subseteq N \times N$. There exists a distinguished root node from which the rest of the nodes can be reached by traversing the edges in E. Each node n, except the root, has a unique parent node, it of the document tree is associated with a logical component, i.e., $< title >$ or $< sections >$, based on an XML file in the given presentation. Function $words(n)$

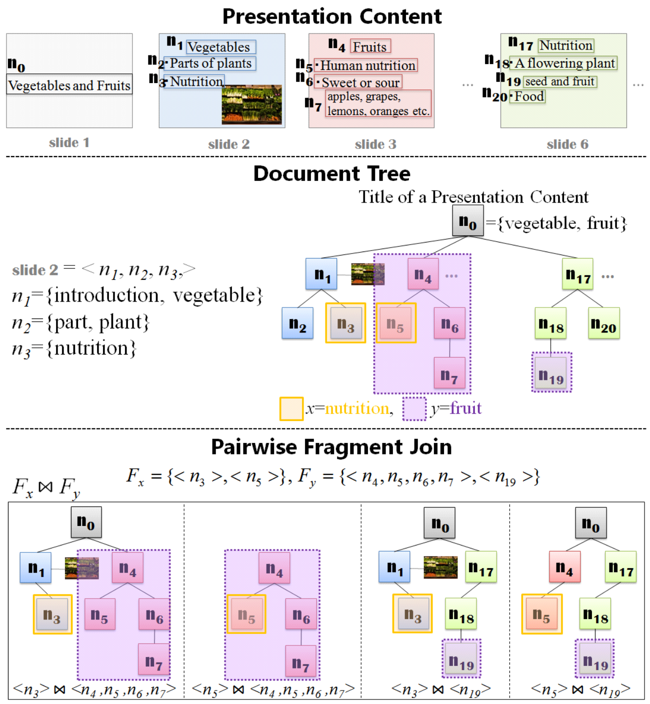

Figure 2. An example of pairwise fragment join of a document tree

returns the noun phrases of the corresponding component in n. A partial tree of D with a given noun phrase as its root is defined as a fragment f. It can be denoted as $f \subseteq D$. A slide is a fragment by the slide title. $< n_1, n_2, n_3 >$ is the set of nodes in slide 2 and a fragment of the document tree.

To formally define the semantic relationships between the extracted elements, we first define operations on fragments, and sets of fragments using a pairwise fragment join [6]. Let F_x and F_y be two sets of fragments in D, then, the pairwise fragment join of F_x and F_y, denoted as $F_x \bowtie F_y$, is defined to extract a set of fragments. This set is yielded by computing the fragment join of every combination of an element in F_x and an element in F_y, in pairs, as follows:

$$F_x \bowtie F_y = \{f_x \bowtie f_y \mid f_x \in F_x, f_y \in F_y\} \quad (3)$$

An example of operation for pairwise fragment join. It refers the sample document tree in Figure 2. For the given two topics $x = nutrition$ and $y = fruit$, where $F_x = \{< n_3 >, < n_5 >\}$, $F_y = \{< n_4, n_5, n_6, n_7 >, < n_{19} >\}$, $F_x \bowtie F_y$ produces a set of fragments $\{< n_3 > \bowtie < n_4, n_5, n_6, n_7 >, < n_5 > \bowtie < n_4, n_5, n_6, n_7 >, < n_3 > \bowtie < n_{19} >, < n_5 > \bowtie < n_{19} >$ by Eq. (3).

2) Semantic Filters: We determine semantic relationships between the given topics, x and y, using the set of fragments produced by taking pairwise fragment join as semantic filters. For this, we define four types of semantic filters by considering the horizontal and vertical relevance, as well as the structural semantics from the document tree.

Table I
SEMANTIC RELATIONSHIPS WITH SEMANTIC FILTERS

Types	Horizontal	Vertical	Hierarchy	Inclusion
x *shows* y	$< \alpha$	*either*	$l(x) < l(y)$	*either*
x *shows* y	$\geq \alpha$	$< \beta$	$l(x) < l(y)$	*either*
x *describes* y	$< \alpha$	*either*	$l(x) > l(y)$	*either*
x *describes* y	$\geq \alpha$	$< \beta$	$l(x) > l(y)$	*either*
x *likewise* y	$< \alpha$	*either*	$l(x) = l(y)$	*either*
x *likewise* y	$\geq \alpha$	$< \beta$	$l(x) = l(y)$	*either*
x *has-a* y	$< \alpha$	*either*	*either*	$f_x \supseteq f_y$
x *has-a* y	$\geq \alpha$	$< \beta$	*either*	$f_x \supseteq f_y$
x *part-of* y	$< \alpha$	*either*	*either*	$f_x \subseteq f_y$
x *part-of* y	$\geq \alpha$	$< \beta$	*either*	$f_x \subseteq f_y$

- **Horizontal distance:** supposing, $hdist(t_i, t_j)$ denotes distance between the nodes of the slide titles t_i and t_j in slides containing x and y, we set the threshold value α at $|N|/2$, i.e., half the total number of nodes in the document tree. If $hdist(t_i, t_j) \leq \alpha$, the distance between slides containing x and y is short (i.e., relevant).
- **Vertical distance:** when distance between the slides containing x and y is long, and they are at the low levels in slides, they can be less relevant in the document tree. Supposing, $vdist(r, q)$ denotes the distance between the root node r and the node containing x or y, we set the threshold value β at $ave(depth)$, which is an average of the depth of levels. If $vdist(r, q) \leq \beta$, the level of the node containing x or y is high (i.e., relevant).
- **Hierarchy:** we compare the levels of x and y in the fragments based on the theory of hierarchical semantics. When $l(x) < l(y)$, the level of x is higher than the level of y; x is a superordinate concept of y (y is a subordinate concept of x). Contrarily, $l(x) > l(y)$ denotes that the level of x is lower than the level of y; x is a subordinate concept of y (y is a superordinate concept of x). When $l(x) = l(y)$, the level of x is same as the level of y; they have coordinate concept.
- **Inclusion:** when $f_x \subseteq f_y$, the fragment of x is included in the fragment of y, i.e., f_x is a partial tree of f_y.

3) Semantic Relationship Types: We determine five types of semantic relationships between the given noun phrases, x and y, by combining the semantic filters of Table I. For measuring the relevance between x and y, we focus on the **horizontal distance** and the **vertical distance**. Here, when the **horizontal distance** between them is long, the **vertical distance** should be short. We determine hierarchical relationships, $show$, $describe$, and $likewise$, by focusing on **hierarchy**. In x $shows$ y, $l(x) < l(y)$ means x is a superordinate concept of y. In x $describes$ y, $l(x) > l(y)$ means x is a subordinate concept of y. In x $likewise$ y, $l(x) = l(y)$ means x and y have coordinate concept with each other. We determine inclusion relationships, which are $has\text{-}a$ and $part\text{-}of$, by focusing on **inclusion**. In x $has\text{-}a$ y, $f_x \supseteq f_y$ means that the concept of x includes the concept of y. In x $part\text{-}of$ y, $f_x \subseteq f_y$ means that the concept of x is included in the concept of y. When x and y fail to match these determinations of semantic relationships, x and y are independent. Therefore, a numbers of semantic relationships between x and y are formed from a set of fragments produced by taking the pairwise fragment join; a semantic relationship is determined by majority.

We conduct multiple presentations based on this semantic structure analysis, the semantic relationships follow a transitivity law, e.g., iff $x\ shows\ y$ in presentation A, $y\ shows\ z$ in presentation B, then it is assumed that $x\ shows\ z$.

IV. iPoster Generation

We generate an iPoster possessing two features: (1) an overview of elements from the slides, retaining this feature of traditional posters; and (2) a ZUI, promoting user interaction and reflecting the semantics of the elements.

A. Determination of Element Layouts

When hierarchical relationships between two elements, reveal a hierarchy applied as to a hierarchical structure. $Show$ or $describe$ maps a parent-child relationship, if $x\ shows\ y$ ($y\ describes\ x$), then we mark x in a parent area and y in a child area, suggesting that the layer of x is higher than the layer of y. Moreover, $likewise$ maps a sibling relationship, if $x\ likewise\ y$, then we locate x and y in the same layer. Inclusion relationships between two elements, reveal a logical relationship of inclusion and exclusion applied as to a stacked Venn. If x *has-a* y (y *part-of* x), we conceive an area of y that is included in an area of x, and that the area of x is larger than the area of y.

B. Determination of Element Transitions

To utilize a ZUI, (1) users can browse the iPosters with their operations, such as zoom-in, zoom-out, and pan; (2) users can browse the iPosters without their operations by automatically navigations with transitions between elements. The transitions discussed here explain the kinds of visual effects that are applied to the semantic relationship types.

When $show$ ($describe$) between two elements. Then, firstly the view must be zoomed-out from the focused element to an area of both, following which; it must be zoomed-in to the target element. Therefore, the transitions include passing through the area of both, which helps users to easily grasp the super-sub relation existing between them.

When $likewise$ between two elements, the transitions between them include zooming-out from the focused element to an area enclosing both the elements and their parent element, and then zooming-in to the target element. Then, the transitions provide their parent element helps users to easily know they are subservient to the same concept.

When *has-a* (*part-of*) exists between two elements, the transition between the two elements pans from the focused element to the target element. Therefore, this simple and direct transition between the two elements helps users to easily understand that they are dependent on each other, and that there exists an inclusion relationship between them.

The transitions between two independent elements include zooming-out from the focused element to all elements, and then zooming-in to the target element. These transitions help the user to easily know that they are irrelevant.

As depicted in Figure 3, we generated iPosters using actual Lecture ♯1 [4] for Database at Portland State University by Prof. Laura Bright. We can easily find that this lecture emphasized the content of **Relational Database**.

V. Exploratory Search on Presentations

We build an exploratory search tool that aids users to search multiple presentations in search results by a given query: (1) focused searching and (2) exploratory browsing. Then, we measure dependence of the structure of the iPoster based on user browsing behavior, as follows:

$$D(H) = \frac{1}{|H|-1} \sum_{n=1}^{|H|-1} \frac{1}{dist(e_n, e_{n+1})}, \quad e_n \in H \tag{4}$$

Here, H is a browsing history based on user browsing behavior. e_n is a browsed element in H. We define the browsed element, focusing on zoom-in operations of elements by the users, that the elements can be considered as the users are interested in. Then, we calculate a degree of $D(H)$ by using average of relevance between the browsed elements. Function $|H|$ returns the number of the browsed elements, $|H|-1$ then denotes the number of edges between them. $dist(e_n, e_{n+1})$ is a shortest distance between e_n and e_{n+1} in an order, which is calculated by the number of edges between the browsed elements on the structure of the iPoster, then, $dist(e_n, e_{n+1}) \geq 1$. When $dist(e_n, e_{n+1})$ returns 1, the relevance between e_n and e_{n+1} is closest. In this case, we set the threshold value γ at $|H|/|E_p|$, $|E_p|$ denotes the number of nodes in a partial tree included all browsed elements of the structure of the iPoster. If $D(H) \geq \gamma$, the browsing behavior can be considered well-matched on the structure of the iPoster; contrarily, if $D(H) < \gamma$, the browsing behavior can be considered partially-matched on the structure of the iPoster.

A. Focused Searching of Presentation Contents

When a user browses along the structure of the iPoster focused on a topic and its subtopics with zoom-in operations, we assume that it is focused searching based on well-matched browsing behavior, he wants to get details of the focused topics. Algorithm 1 describes a procedure for focused searching in a sub-structural layout $FS = (E_d, R, P)$. E_d is a set of elements related to x by users' operations as an input. r is a type of semantic relationships R defined in Table I. P is a set of presentations in search results by a given query. This procedure represents e' related to

[4] http://web.cecs.pdx.edu/~howe/cs410/lectures/Relational_Intro_1.ppt

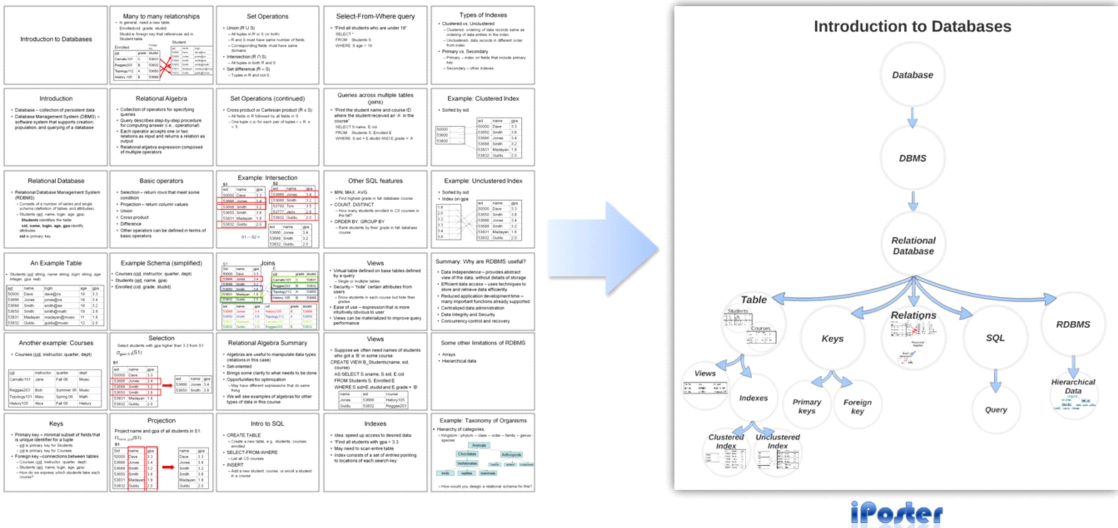

Figure 3. An Example of a generated iPoster based on our proposed method

Algorithm 1 Explore $FS = (E_d, R, P)$

Require: x is an element in a given presentation p, last browsed by a user with a zoom-in operation.

Ensure: $R = \{(e, e', r) | e, e' \in E, e, e' \in p'\}$

```
R ⇐ φ
for all presentation p′ in a given domain do
  if r is show relationship then
    e ⇐ x
    R ⇐ (x, e′, r)
    P ⇐ p′
  end if
end for
```

x according to (x, e', r), in which r is $show$ for finding details of x.

B. Exploratory Browsing of Presentation Contents

When a user browses topics in apart on the iPoster, we assume that it is exploratory browsing based on partially-matched browsing behavior, he wants to get much relevant information of the browsed topics. Algorithm 2 describes a procedure for exploratory browsing in a sub-structural layout $EB = (E_w, R, P)$. E_w is a set of elements related to x and y by users' operations as an input. This procedure represents e' related to x and y according to (x, y, r), in which r is $likewise$ for finding relevant information of x and y.

Figure 4 illustrates an example of exploratory browsing, a user firstly zooms-in to the area of 'Forest Ecosystem,' after that zooms-out it and zooms-in to the area of 'Food Chain,' and zooms-out it and zooms-in to the area of 'Products.' We considered that he wants to get a lot of information about 'Food Chain' and 'Products' along the content related to

Algorithm 2 Explore $EB = (E_w, R, P)$

Require: x, y are elements in a given presentation p, browsed by a user with a zoom-in operation.

Ensure: $R = \{(e, e', r) | e, e' \in E, e, e' \in p'\}$

```
R ⇐ φ
for x, y such that (x, y, r) ∈ R in p do
  for all presentations p′ in a given domain do
    if r is likewise relationship then
      e ⇐ x, y
      r ⇐ describe
      R ⇐ (x, e′, r) = (y, e′, r)
      P ⇐ p′
      if r is descibe relationship then
        e′ ⇐ z
        R ⇐ (e, z, r)
        P ⇐ p′
      end if
    end if
  end for
end for
```

'Forests and Humans.' Due to 'Products' $likewise$ 'Food Chain,' and they $describe$ 'Forests and Humans,' 'Forests and Humans' has its details (i.e., 'Products' and 'Food Chain') only in P_A[5]. In this work, we can extract 'Nitrogen Cycle' and 'Rainforest Animals', which $describe$ 'Forests and Humans' in P_B[6], and represent a whole of 'Forests and Humans' with its details (i.e., 'Products,' 'Food Chain,' 'Nitrogen Cycle,' and 'Rainforest Animals'). In addition,

[5]http://teacherweb.com/AB/GilbertPatersonMiddleSchool/MsDavid/Tree-Types-2b-Posting-version.ppt

[6]http://www.marinepolicy.net/cparsons/Ecology/12-Forests.PPT

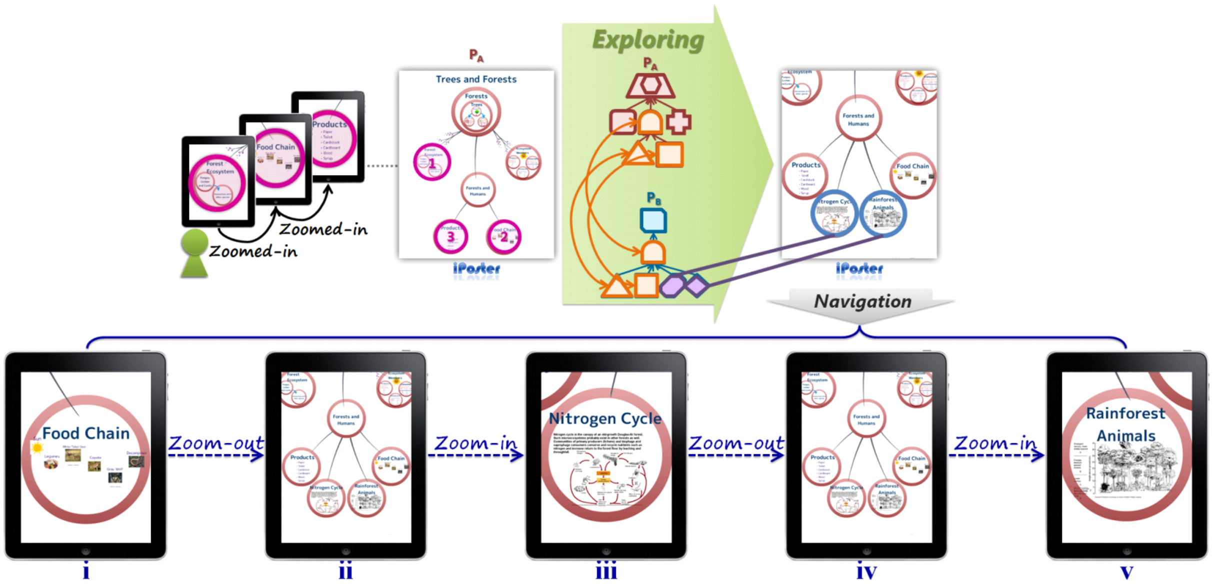

Figure 4. An example of exploratory browsing based on partially-matched browsing behavior

the iPoster can automatically navigate the whole of 'Forests and Humans' to 'Nitrogen Cycle' and 'Rainforest Animals.' The iPoster firstly zooms-out from the area of 'Food Chain' (**i**) to the whole area of 'Forests and Humans' (**ii**) that shows an overview of 'Forests and Humans' with 'Nitrogen Cycle' and 'Rainforest Animals.' Next, the iPoster zooms-in to 'Nitrogen Cycle' (**iii**) and 'Rainforest Animals' (**v**), respectively. It helps the user to understand details 'Nitrogen Cycle' and 'Rainforest Animals' of 'Forests and Humans.'

In this way, when many elements are extracted, we need to consider how to select candidate elements from presentation contents such as weights of the elements.

VI. Conclusions and Future Work

In this paper, we built an exploratory search tool for presentation contents based on iPoster generation, which represents textual and graphic elements in a structural layout with ZUIs, to promote user interaction. In order to generate an iPoster, we introduced a semantic structure analysis model for extracting elements and determining semantic relationships between them from slides. The iPoster enables users to browse and explore easily and efficiently through various presentations.

In the future, we plan to consider a collaborative exploratory searching tool, which will provide a way to summarize already encountered information. The tool could tailor these summaries to the respective skill levels of collaborators.

Acknowledgments

This research was supported in part by Strategic Information and Communications R&D Promotion Programme (SCOPE), the Ministry of Internal Affairs and Communications of Japan, and a Grant-in-Aid for Scientific Research (B)(2) 26280042 from the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

References

[1] J. Lanir, K. S. Booth, and S. Wolfman, "An Observational Study of Dual Display Usage in University Classroom Lectures," in *Journal of Human-Computer Interaction*, vol. 28, no. 4, 2013, pp. 335–377.

[2] R. Spicer, Y. Lin, A. Kelliher, and H. Sundaram, "NextSlidePlease: Authoring and Delivering Agile Multimedia Presentations," in *ACM Transactions on Multimedia Computing, Communication and Applications*, vol. 8, no. 4, 2012, pp. 53:1–53:20.

[3] R. W. White and R. Roth, "Exploratory Search: Beyond the Query Response Paradigm," in *Synthesis Lectures on Information Concepts, Retrieval, and Services*, vol. 1, no. 1, 2009, pp. 1–98.

[4] B. Detlor, C. W. Choo, M. MacKenzie, and D. Tunbull, "Information Seeking and Use in Diverse Organizational Contexts," in *Proc. of the American Society for Information Science and Technology*, vol. 46, no. 1, 2009, pp. 1–4.

[5] D. Ellis and M. Haugan, "Modelling the Information Seeking Patterns of Engineers and Research Scientists in an Industrial Environment," in *Journal of Documentation*, vol. 53, no. 4, 1997, pp. 384–403.

[6] S. Pradhan, "An algebraic query model for effective and efficient retrieval of xml fragments," in *Proc. of the 32nd international conference on Very large data bases (VLDB 2006)*, 2006, pp. 295–306.