〔特集論文〕 「Web インテリジェンスとインタラクションの新展開」

iPoster: Interactive Poster Generation based on Topic Structure and Slide Presentation

Yuan Yuan Wang

Kyoto Sangyo University
yuanw@cc.kyoto-su.ac.jp

Yukiko Kawai

(affiliation as previous author)
kawai@cc.kyoto-su.ac.jp

Kazutoshi Sumiya

University of Hyogo
sumiya@shse.u-hyogo.ac.jp

keywords: presentation slides, slide structure, topic structure, interactive poster, iPoster, e-learning

Summary

MOOC is a crucial platform for improving education; students are able to obtain various educational presentation contents through the Web. Recently, Prezi introduced a zoomable canvas as a substitute to the traditional presentations that allows users to zoom in and out of the presentation media. Teachers then attempt to provide presentations in a nonlinear fashion for enhancing the user interaction through these presentations; however, creation of nonlinear presentations would be time-consuming, besides posing design challenges. Therefore, we have developed a novel support system for grasping overviews of presentation slides, it generates a meaningfully structured presentation, called iPoster; this enables users to automatically navigate through the slide-based educational contents. The system places elements such as text and graphics of presentation slides in a structural layout by semantically analyzing the slide structure. The structural layout can reveal the hierarchy of elements based on topic structure by moving from the overview to a detail using automatic transitions, such as zooms and pans. Through this, the iPoster can support students to interactively browse online presentation slides for grasping an overview; it would substantially help the students navigate the presentation slides effectively for their learning purposes. In this paper, we discuss our interactive poster (iPoster) generation method and we have also included an evaluation of our method's effectiveness.

1. Introduction

Slide-based visual presentation support, such as Microsoft PowerPoint or Apple Keynote, is now one of the most frequently used tools for educational purposes currently. However, this format has been criticized repeatedly because of the limitations it imposes on authors and presenters [Tufte 03]. Furthermore, the current slideshow mode of presentation slides merely permits fluid navigation of linear structures, even while it is being presented to diverse audiences. With the proliferation of massive open online course (MOOC) in recent years, enormous amounts of slide-based educational contents of MOOC, are freely shared on Web sites such as Coursera^{*1} and SlideShare^{*2}. Then, students have to find presentation contents to learn, because each presentation content contains many slides. Thus, it is important to focus on grasping overviews of slides in presentation contents to help users (e.g., students, self-learners, etc.) find which one of the presentation con-

Table 1 Comparison of presentation slides and posters

Features	Presentation slides	Posters
Material	textbooks/ academic papers	textbooks/ academic papers/ slides
Information	Few	Less
Diagram	Yes	Yes
Flow	Yes	No
Overview	Maybe	<u>Yes</u>
Brevity	Maybe	<u>Yes</u>

tents is worth learning. In order to provide an overview, authors (e.g., teachers, presenters, etc.) often summarize the content in the limited size of a single paper as a poster.

Table 1 shows a comparison of features between presentation slides and posters. In general, presentation slides and posters are often made from textbooks or academic papers, and the posters are also often made from the presentation slides. Whereas the slides as a summary of the textbooks or the academic papers with little information, the poster has less information as a summary of the slides. Diagrams are often used both in the slides and the posters. Since the slides express presentation context, it is easy to understand the flow of the presentation. In contrast, the poster is not suited to represent the flow of the presen-

*1 <https://www.coursera.org/>

*2 <http://www.slideshare.net/>

Fig. 1 Conceptual diagram of an interactive poster by using slides

tation. However, the poster summarizes the content in a single sheet of a paper, is excellent in presenting an outline of the content, which enables users grasp an overview at a glance. Additionally, since the less information is briefly summarized in the poster, is excellent in simplicity. Therefore, we considered that significant features of the poster are an overview of the content and brevity of the content. Recently, zooming presentations are attempting to mitigate the problems posed by slideware. For instance, Prezi^{*3} provides an infinite canvas with a zooming user interface (ZUI) [Bederson 94] as an alternative to the traditional slides. This interface permits the canvas format to support the creation of expressive layouts. These layouts can be zoomed out, allowing the slide arrangement to be presented in its entirety to the audience [Lichtschlag 12a]. The canvas model was also adopted by pptPlex^{*4}. Authors and presenters will be required to create and deliver presentations in a nonlinear fashion. However, this will be time-consuming and pose challenges in designing.

As depicted in Figure 1, we aim to generate an interactive poster, called iPoster, by using presentation slides. The iPoster is a support tool for grasping overviews of online presentation contents that are offered through MOOC. Before students begin to find presentation contents to learn, they can first interactively browse our proposed iPoster for grasping an overview of slides in the presentation content. Naturally, after students learned presentation contents through MOOC, it is useful to review them using our proposed iPoster. The iPoster can be implemented by 1) generating a meaningfully structured presentation according to the significant features of the traditional poster; and 2) employing the zooming user interface for automatically navigating the presentation based on a basic idea of Prezi. To achieve our goal, we analyze textual and graphic elements in the slides and the semantic relationships between them. Furthermore, our method can organize the elements in structural layouts based on topic structure, and it can

also navigate the elements by using zooming and panning transitions between them. In topical structure analysis, we first extract elements by examining the presentation context of the particular elements in the slides. Therefore, the semantic relationships between these elements are determined using implicit hyperlinks in slides, based on a slide structure. Specifically, we derive the slide structure by focusing on the itemized sentences of bullet points present in the slide text. There are various types of structural layouts for constructing an iPoster, such as hierarchical structure, stacked Venn, and pyramid structure. In this paper, in order to provide an overview of the content, we utilize a hierarchical structure, combined with a stacked Venn for an iPoster. Finally, our iPoster is generated based on topic structure, using a ZUI for interactive browsing and automatic navigation, which can raise user interaction, besides enabling users to understand the educational presentation contents easily and efficiently.

This paper is organized as follows. Section 2 provides a brief summary of related work. Section 3 describes our topic structure analysis model to extract elements from slides and to determine the semantic relationships between elements. Section 4 explains the detailed procedure of an iPoster generation, based on the derived topic structure by employing zooming and panning transitions. Section 5 illustrates our experimental results conducted using a real dataset of online presentation contents. Finally, Section 6 concludes this paper with suggestions for further work.

2. Related Work

A variety of applications address the identified weaknesses of the current slideware tools in the presentation and authoring domains. Our approach in iPoster builds on the strength of grasping an overview of the presentation.

Many recent applications address the need to capture the complex relationships among content items, and assist in crafting compelling narratives. These applications employ both, new or unusual hardware configurations, as well as novel interfaces. Lanir et al. [Lanir 13] proposed the MultiPresenter application that leverages spatial reasoning capabilities to relate content through dual-screen projection. Although iPoster does not adopt the dual-audience-display paradigm, it addresses the need to navigate through elements dynamically during the presentation. NextSlidePlease [Spicer 12] is a novel application for authoring and delivering slideware presentations. This tool addresses issues of content integration, presentation structuring, time-management, and flexible presentation delivery. iPoster is similar to this work, as we utilize a structured layout,

*3 <http://prezi.com/>

*4 <http://www.microsoft.com/en-us/download/details.aspx?id=28558>

Fig. 2 An example of ‘fruit’ appears at different levels in slides

rather than one or more slide lists, to help users grasp the overview of the presentation.

Good and Bederson [Good 02] proposed replacing the card stack or film strip metaphor with a ZUI in their CounterPoint application, borrowing insights from the domain of mind maps [Faste 12] or concept maps [Villa 12] and visual storytelling [Pschetz 14]. Our proposed iPoster uses a ZUI to navigate to the elements from slides moving from an overview to details. The Fly application addresses graph-based presentation authoring [Lichtschlag 12b]. It provides a set of tools for authoring presentations from scratch in a two-dimensional canvas with defined paths. Our proposed iPoster generation method aims to utilize the strengths of ZUIs in providing an enhanced interactive browsing tool for learning slide-based educational contents of MOOC.

3. Topic Structure Analysis Model

In this section, we describe a topic structure analysis model for extracting elements of slides and determining the semantic relationships between them.

3.1 Element Extraction

The two most salient and dominant elements in a slide are the set of textual elements and the set of graphic elements. These are based on the itemized sentences of bullet points in the slide text. We define the slide title as the 1st level, the first item of text within the slide body as the 2nd level, and the depth of the sub-items increases with the indentation levels (3rd level, 4th level, and so on). Non-text objects such as figures or tables are considered to be at the same indentation level as the surrounding text.

§1 Textual Elements

We define textual elements as topics that focus on the nouns in slides. In general, topics can be considered as high frequency terms or slide titles. Since it may lose the context of presentation, a topic can be described as a learning point with multiple nouns that frequently appears at different levels (i.e., bullet points) in neighboring slides by considering the context in the presentation. Initially, we extract noun phrases using a language analysis toolkit MSR Splat*⁵ based on the XML files of slides.

The topics that appear at the title of a slide and the body of other slides can be considered to indicate its context in a presentation (see Figure 2). Then, we extract topics by locating the same noun phrases in different slides, at varying levels. If a noun phrase k is not always as it appears at the same indentation level in slides s_i and s_j , then k is a candidate for being one of the topics T in the presentation. In this way, if a noun phrase appears at the body of only one slide and the titles of other serial slides, it also can be a candidate of topics.

$$T = \{(k, s_i, s_j) | l(k, s_i) \neq l(k, s_j)\} \quad (1)$$

Where, T is a bag of noun phrases that can be considered as candidates for topics. $l(k, s_i)$ is a function that returns the highest level of k in the slide s_i . For instance, when the highest level is the title, i.e., the 1st level of s_i , then $l(k, s_i)$ returns 1; and when the highest level is the 3rd level of s_j , then $l(k, s_j)$ returns 3. When k appears at different levels, k is determined as a candidate for topics provided $l(k, s_i)$ is not equal to $l(k, s_j)$. Then, the weight of k in T is defined using the levels of k , and the distance between slides s_i and s_j , as follows:

$$I(k) = \frac{1}{l(k, s_i)} + \sum_{k \in T} \left(\Delta \cdot \frac{1}{dt(s_i, s_j)} \right) \quad (2)$$

$$\Delta = \left| \frac{1}{l(k, s_j)} - \frac{1}{l(k, s_i)} \right| \quad (3)$$

Where $l(k, s_i)$ indicates the weight of k in s_i , i.e., it returns the highest level of k in slide s_i by Eq. (1). Δ indicates the context of k in s_i and s_j , and it denotes variation of the weight of k both in s_i and s_j by Eq. (3). $dt(s_i, s_j)$ corresponds to the strength of the association between s_i and s_j , and it denotes the distance between s_i and s_j , that is, a number of slides between them. Thus, if the context of k appears at a high level in s_i and s_j , and the distance between s_i and s_j is short, the weight $I(k)$ of k is high. Here, k , s_i , and s_j belong to T in Eq. (1).

§2 Graphic Elements

When compared to pure textual elements, images are more attractive, appealing and informative from a psychological standpoint. Based on the study of search results presentation [Li 08], it can be noted that summaries with images assist in quicker understanding of the results, thereby helping in arriving at relevant judgments faster. Therefore, we define graphic elements as images corresponding to the topic candidates in slides, given that the surrounding text of the images is similar to the topic candidates. We considered that the images used to describe the content in slides; they often located nearby a subject of the content in the slides. Since topic candidates may appear any itemized sentence of bullet points in the slides,

*5 <http://research.microsoft.com/en-us/projects/msrsplat/>

the surrounding text of the image can be selected as a slide title or an itemized sentence of the slide and it appears at the same indentation level of the image. We then calculate the similarity of the noun phrases in the surrounding text (i.e., slide titles or itemized sentences) and the topic candidates by using the Simpson similarity coefficient. When the similarity exceeds a predefined threshold, the noun phrases in the surrounding text of the images and the topic candidates are considered similar. Then, the images are recognized as the corresponding images of the topic candidates.

3.2 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 at a glance, for a quick understanding of the content. Preliminary ideas regarding this model are given in an algebraic query model [Pradhan 06] as well.

§1 Basic Definitions and Algebra

The presentation shown in Figure 3 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, except the root, has a unique parent node. Each node n of the document tree is associated with a logical component, such as $\langle title \rangle$ or $\langle sections \rangle$, based on the bullet points on slides using an XML file in the given presentation. There is a function $words(n)$ that returns the representative noun phrases of the corresponding component in n . A partial tree of the document tree D with a given noun phrase as its root is defined as a fragment f . Thus, the fragment may consist of one node containing the given noun phrase or a bag of nodes in which the node containing the given noun phrase is the root node. It can be denoted as $f \subseteq D$. A slide is a fragment of the slide title. In Figure 3, $\langle n_1, n_2, n_3 \rangle$ is the set of nodes in slide 2 and a fragment of the sample 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. Let F_x and F_y be two sets of fragments in D of a given presentation, 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:

$$F_x \bowtie F_y = \{f_x \bowtie f_y \mid f_x \in F_x, f_y \in F_y\} \quad (4)$$

Fig. 3 An example of a presentation and its tree representation

Figure 4 illustrates an example of operation of pairwise fragment join. It refers to the sample document tree in Figure 3. For the given two topics $x = nutrition$ and $y = fruit$, where $F_x = \{ \langle n_3 \rangle, \langle n_5 \rangle, \langle n_{20} \rangle \}$, $F_y = \{ \langle n_4, n_5, n_6, n_7 \rangle, \langle n_{19} \rangle \}$, $F_x \bowtie F_y$ produces a set of fragments $\{ \langle n_3 \rangle \bowtie \langle n_4, n_5, n_6, n_7 \rangle, \langle n_5 \rangle \bowtie \langle n_4, n_5, n_6, n_7 \rangle, \langle n_{20} \rangle \bowtie \langle n_4, n_5, n_6, n_7 \rangle, \langle n_3 \rangle \bowtie \langle n_{19} \rangle, \langle n_5 \rangle \bowtie \langle n_{19} \rangle, \langle n_{20} \rangle \bowtie \langle n_{19} \rangle \}$ on applying Eq. (4).

§2 Semantic Filters

We determine semantic relationships between the given noun phrases, x and y , that are included in at least one slide of the given presentation from the extracted elements. Therefore, we analyze the set of fragments produced by taking pairwise fragment join as semantic filters, each of which combined by both a fragment of x and a fragment of y , e.g., any one of six joined fragments as shown in Figure 4. 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 of the given presentation.

Horizontal distance Logically interrelated slides of a presentation are typically close to each other. Therefore, in such presentations, the horizontal distance between nodes in different slides of a document tree is a reasonable measure of the inter-relationship between nodes. Specifically, when the horizontal distance between the nodes in slides containing x and y exceeds a certain threshold, x is irrelevant to y . Thus, if a joined fragment satisfies this filter, then x and y appear in the same slide or they appear in neighboring slides, respectively. Supposing, $hdist(t_i, t_j)$ denotes the distance between the nodes of the slide titles t_i and t_j in slides containing x and y , respectively (e.g., slide 2: $t_i = n_1, x = n_3$; slide 3: $t_j = y = n_4$, in Figure 4 (1)). Here, the distance can be measured by counting a number of

Fig. 4 An example of pairwise fragment join

slides between them. Since slides are irrelevant if the distance of them is far, in this paper, we set the empirical value α at 5 based on our experiments, that is appropriate to our target academic contents or lecture contents in which each content contains 10–30 slides. If $hdist(t_i, t_j) \leq \alpha$, then the distance between two slides containing x and y is short (i.e., relevant).

Vertical distance Logically, indentations of slides are typically close to each other. Therefore, when the distance between the slides containing nodes is long, and the nodes are at the low levels in slides, they can be less relevant in the document tree. When the vertical distance between the nodes in slides containing x and y exceeds a certain threshold, and they are at the low levels in the slides, x is irrelevant to y . Thus, if a joined fragment satisfies this filter, then x and y appear at the titles of slides or they appear at the high levels of slides. Supposing, $vdist(r, q)$ denotes the distance between the root node r of the document tree and the node containing each given noun phrase q (e.g., $r=n_0$, $q=x=n_3$ or $q=y=n_4$, in Figure 4 (1)), we set the threshold value β at $ave(depth)$, which is an average of the depth of levels in the document tree, for normalizing various presentations. If $vdist(r, q) \leq \beta$, then the levels of nodes containing x and y in

Table 2 Semantic relationships with semantic filters

Relationship 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	$f_x \subseteq f_y$

the document tree are both high (i.e., relevant).

Hierarchy For judging the semantics of x and y , we compare the levels of nodes containing x and y in the joined fragments based on the theory of hierarchical semantics. When $l(x) < l(y)$, the level of node containing x is higher than the level of node containing y ; x is a superordinate concept of y . Contrarily, $l(x) > l(y)$ denotes that the level of node containing x is lower than the level of node containing y ; x is a subordinate concept of y . When $l(x) = l(y)$, the level of node containing x is same as the level of node containing y ; they have coordinate concept with each other.

Inclusion Since fragments of x and y are partial trees of the document tree that nodes containing x and y as the roots of them, the inclusion relationships exist between the fragments of x and y in the joined fragments. When $f_x \subseteq f_y$, it denotes that the fragment of x is included in the fragment of y , i.e., f_x is a partial tree of f_y (see Figure 4 (2)). Contrarily, when $f_x \supseteq f_y$, it denotes that the fragment of x includes the fragment of y , i.e., f_y is a partial tree of f_x .

§3 Semantic Relationship Types

We determine five types of semantic relationships between the given noun phrases, x and y , which are included in at least one slide, by combining the semantic filters of Table 2. To measure 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, x shows y , x describes y , and x likewise y , by focusing on **hierarchy**. In x shows y , $l(x) < l(y)$ means x is a superordinate concept of y (y is a subordinate concept of x). In x describes y , $l(x) > l(y)$ means x is a subordinate concept of y (y is a superordinate concept of x). Then, *show* and *describe* are functionally interchangeable, when x describes y is from the viewpoint of y shows x . In x likewise y , $l(x) = l(y)$ means x and y have coordinate concept with each other.

We determine inclusion relationships, which are x *part-of* y and x *has-a* y , by focusing on **inclusion**. In x *part-of* y , $f_x \subseteq f_y$ means that the concept of x is included in the concept of y . In x *has-a* y , $f_x \supseteq f_y$ means that the concept of x includes the concept of y . Then, *part-of* and *has-a* are functionally interchangeable, when x *has-a* y is from the viewpoint of y *part-of* x . When x and y fail to match these determinations of semantic relationships, x and y are independent. Therefore, a number 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.

In this work, the semantic relationships follow a transitivity law, e.g., iff x *shows* y , y *shows* z , then it is assumed that x *shows* z .

4. iPoster: Interactive Poster Generation

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

4.1 Determination of Element Layouts

For providing an overview of elements from slides, we attempt to determine the element layouts by utilizing a hierarchical structure combined with a stacked Venn, based on the topic structure of the elements. When hierarchical relationships between two elements, i.e., either *show*, *describe*, or *likewise*, they reveal a hierarchy between those elements, as applied to a hierarchical structure. *Show* or *describe* maps a parent-child relationship in the hierarchical structure; 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 . Additionally, *likewise* maps a sibling relationship in the hierarchical structure; if x *likewise* y , then we locate x and y in the same layer. Inclusion relationships between two elements, i.e., *part-of* and *has-a*, reveal a logical relationship of inclusion and exclusion applied, as to a stacked Venn. For instance, x *part-of* y (y *has-a* x), we conceive an area of x that is included in an area of y , and that the area of y is larger than the area of x .

4.2 Determination of Transitions between Elements

To utilize a ZUI for navigating through presentations, the transitions discussed here explain the kinds of visual effects that apply to the semantic relationship types, to reflect the meaning of the elements from the slides. We

Fig. 5 System configuration diagram

animate the zooming and panning transitions for navigating through elements in the structural layout based on the topic structure; this can help users to visually understand overview and detail of the contents within a presentation.

When *show* (*describe*) between two elements is not included in an inclusion relationship, then, firstly the view must be zoomed-out from the focused element to an overview of them, following which, it must be zoomed-in to the target element. In addition, when *show* (*describe*) between two elements is included in the inclusion relationship, the transitions between them includes zooming-out from the focused element to the whole element area in the stacked Venn, and zooming-in to the target element. Therefore, the transitions include passing through the overview or the whole element area, which helps users to easily grasp the super-sub relation existing between them.

When *likewise* exists between two elements, the transitions between the two elements include zooming-out from the focused element in an area enclosing both the elements and their parent element, and then zooming-in to the target element. Therefore, the transitions indicate the presence of the parent element; thereby elucidating to the user the existence of a subservient relationship.

When *part-of* (*has-a*) 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. Therefore, these transitions help the users to easily know that they are irrelevant with respect to each other in an iPoster.

Database Systems: Lecture #1 by Prof. Jennifer Widom (Stanford University)

Fig. 6 An example of a generated iPoster based on our proposed method

5. Evaluation

5.1 Implementation

Based on the method described above, we built a prototype tool to support iPoster generation (see Figure 5), using Microsoft Visual Studio 2012 C#. The tool has three stages: analysis, determination, and generation. Firstly, in the analysis stage, we extract textual and graphic elements based on slide structure. The slide structure and information on the indent level of words are constructed by using Office Open XML files in Microsoft PowerPoint 2007. The words in the slides can be extracted using the morphological analyzers [MeCab] and [Ohshima 07]^{*6}. Secondly, in the determination stage, all types of semantic relationships between the extracted elements are determined based on the slide structure. Thirdly, in the generation stage, iPosters are generated by organizing the extracted elements in structural layouts, and attaching transitions between the extracted elements with a ZUI using Piccolo [Bederson 04]. Both the structural layouts and the transitions are determined based on the semantic relationships between the extracted elements. Thus, the generated iPoster can automatically navigate in browsing interface with a control bar.

As depicted in Figure 6, we generated an iPoster with the prototype tool using actual Lecture #1 for Database at Stanford University^{*7}. We found that this lecture emphasized the content of **DBMS** by navigating it through the iPoster. The iPoster firstly zooms-in to the area of ‘C Pro-

gram’ as shown in **ii** from the area of ‘DBMS’ as shown in **i**; this conveys to users about a whole concept ‘DBMS,’ which contains ‘C Program,’ after that, the iPoster pans in the area of ‘System Crashes’ as shown in **iii**, and zooms-out to the whole area of ‘DBMS’ as shown in **iv**. This enables users to understand that ‘DBMS’ contains ‘C Program’ and ‘System Crashes’ in this lecture.

5.2 Experimental Dataset

The purpose of this evaluation was to verify whether our proposed iPoster generation method is useful for helping users to grasp overviews of presentation contents that are offered through MOOC, and each presentation content contains many slides. As an experimental dataset, we prepared an online presentation content collection in which presentation contents are created using Microsoft PowerPoint, as shown in Table 3, consisting of (1) 20 actual lecture contents (total: 464 slides, average: 23.2 slides) of five introductory courses (four lectures of each course) about Social Informatics from the lecture archives at several universities (i.e., University of Tsukuba^{*8}, Aoyama Gakuin University^{*9}, Tokyo City University^{*10}, and Osaka University^{*11}), and (2) 20 actual academic contents (total: 327 slides, average: 16.4 slides) about Informatics from domestic workshops in the DBSJ Archives^{*12}.

There were 5–15 students in Faculty of Computer Science and Engineering, Kyoto Sangyo University, taking

^{*6} <http://www.dl.kuis.kyoto-u.ac.jp/SlothLibWiki/>

^{*7} <http://infolab.stanford.edu/widom/cs145/intro-db.ppt>

^{*8} <http://www.slis.tsukuba.ac.jp/resource/wiki/index.php>

^{*9} http://homepage3.nifty.com/terao/lecture/aoyama/lec_aoyama_top.html

^{*10} <http://www.yc.tcu.ac.jp/otsuka/es/>

^{*11} <http://www2.econ.osaka-u.ac.jp/dony/siryou2.html>

^{*12} <http://service.dbsj.org/stream/events>

Table 3 Experimental dataset

(1) Lecture contents				
No.	Course	University	Time	Slides(AVG.)
1	Media, Culture and Society	Tsukuba	90 mins	19.8
2	Human Sciences	Aoyama Gakuin	90 mins	30.5
3	Environmental Sociology	Tokyo City	90 mins	25.3
4	Marketing	Osaka	90 mins	20.5
5	Methods of Educational Research	Aoyama Gakuin	90 mins	20.0
(2) Academic contents				
No.	Title	Workshop	Time	Slides
1	Searching Earth Environmental Data with a Zooming Browser	DEWS2000	12 mins	9
2	Extraction of Spatial Relations on SVG Maps and Its Applications	DEWS2003	12 mins	24
3	Broadcasting Geographic Information with Levels-of-Details to Moving Objects	DEWS2003	12 mins	24
4	A Context-aware Map Synthesizer: Design and Implementation	DEWS2003	12 mins	17
5	Retrieval System for Digital Photographs using Observing Point of Photographs	DEWS2003	12 mins	15
6	An Analysis on Link Structure Evolution Pattern of Web Communities	DEWS2005	12 mins	19
7	A Blog Search Method using News Video Scene Order	DEWS2005	12 mins	8
8	Recommendation System for Sightseeing based on Group and User Preferences	DEWS2006	12 mins	8
9	A Web Archive Search Engine Based on the Temporal Relation of Query Keywords	DEWS2006	12 mins	15
10	Mining Disjunctive Tree Patterns	DEWS2006	12 mins	22
11	Automatic Generation of TV News Program Explicitly Expressing Emotions	DEWS2006	12 mins	20
12	A Retrieval Method of Comparative News using Content Structure Order for News Archives	DEWS2007	12 mins	15
13	An Interval Extracting Method using Semantic Relations between Scenes for Presentation Contents	DEWS2008	12 mins	16
14	Geographical Information Retrieval based on Map Operation and Category Selection	DEIM2009	12 mins	10
15	A Credibility Analysis Method based on Editor's Intention for Modified Maps	DEIM2010	12 mins	15
16	Measuring Credibility of Text using Relation between Words on the Web	DEIM2011	12 mins	22
17	Geographical Feature Extraction for Retrieval of Modified Maps	DEIM2011	12 mins	21
18	Externalization Support of Key Phrase Channel in Presentation Preparation	DEIM2011	12 mins	16
19	Generating Graphs of Tendencies and Factors from Product Reviews	DEIM2012	12 mins	15
20	The Minor Sports Search System based on Similar Major Sports	DEIM2012	12 mins	16

the Social Informatics course and Information Engineering lab., who participated in the experiments. We show and discuss the experimental results in the follow sections.

5.3 Experiment 1: Validity of Generating iPosters

This experiment was designed to assess the generation of iPosters for providing overviews of slides in presentation contents based on topic extraction. In order to evaluate the performance of topic extraction, we compared our proposed method with a conventional method [Yokota 06], which scores the words using word frequency with the indentation positions of the words in the slides. When the words frequently appear at the titles of the slides (e.g., titles of serial slides), they will be extracted as topics by the conventional method. At first, we used our proposed method to extract 607 noun phrases (lecture contents: 301, academic contents: 306), and the conventional method to extract 682 noun phrases (lecture contents: 356, academic contents: 326) from lecture and academic contents

Table 4 Results of topic extraction from lecture contents

Average	Conventional method	Proposed method
Precision	48.4%	60.1%
Relative recall	76.7%	76.6%
F-measure	0.59	0.67

Table 5 Results of topic extraction from academic contents

Average	Conventional method	Proposed method
Precision	34.9%	43.0%
Relative recall	71.3%	81.4%
F-measure	0.47	0.67

in the dataset. For each presentation content, five participants identified the extracted noun phrases are topics or not. Correct answers that the topics voted by at least three participants were considered the ground truth.

The average of precision^{*13}, relative recall [Clarke 97]^{*14}, and F-measure^{*15} of the proposed topic extraction method along with the conventional method from lecture and academic contents are shown in Table 4 and Table 5, and they can be explained as follows:

- The proposed method achieved the highest average F-measure and precision compared with the conventional method. The average relative recall of the proposed method was almost the same as that of the conventional method in lecture contents. Since our method weighted the words by using the distance between the slides in Eq. (2), it could not extract only 15 topics appear only one slide in some lecture contents (average: 12.8 slides per lecture content).
- The average precision of the proposed method in the academic contents was low; our method extracted a much greater number of noun phrases as topic candidates than those for which participants concurred. We believed that when determining the topics in the academic contents, the participants considered technical terms are more important than general words when our method weighted with high scores.

This experiment confirmed that our method can extract the appropriate topics from slides, considering the presentation context based on slide structure. Furthermore, we confirmed that our method is useful for extracting the topics from any presentation content of MOOC, which contains a certain number of slides. Conversely, the conventional method is useful for extracting the topics from any presentation content, which contains fewer slides. In addition, we want to use an enhance method for discovering the most topics for the iPoster generation. Therefore,

*13 Precision = $\frac{\text{\#correct answers by one method}}{\text{\#noun phrases were extracted by one method}}$

*14 Relative recall = $\frac{\text{\#correct answers by one method}}{\text{\#correct answers by all two methods}}$

*15 F-measure = $\frac{2 \times \text{Precision} \times \text{Relative recall}}{\text{Precision} + \text{Relative recall}}$

Table 6 Results of top-10 topics from lecture contents

Average	Conventional method	Proposed method	Proposed method (with $1/TF$)
Precision	57.5%	67.3%	64.3%
Relative recall	55.5%	64.0%	60.8%
F-measure	0.56	0.66	0.63

Table 7 Results of top-10 topics from academic contents

Average	Conventional method	Proposed method	Proposed method (with $1/TF$)
Precision	47.0%	53.3%	58.8%
Relative recall	59.5%	67.8%	73.4%
F-measure	0.53	0.60	0.65

when we evaluated the performance of the topic extraction by using the ten most important noun phrases based on the conventional and the proposed methods are presented in Table 6 and Table 7. We found that the average precision of the proposed method in lecture and academic contents were both higher than those in Table 4 and Table 5. Specially, the average relative recall of the proposed method was higher than the conventional method in the lecture contents.

Furthermore, we weighted the words from the experimental data by Eq. (2) with $1/TF$ (inverse word frequency) for discovering more technical terms, and the results of the ten most important noun phrases were listed in the right columns of Table 6 and Table 7. Although the average precision of the proposed method with $1/TF$ in the academic contents was still not high, since it was higher than the average precision of the proposed method in Table 7. In particular, the average precision of the proposed method with $1/TF$ in the lecture contents was lower than the average precision of the proposed method in Table 6. We believe that we can improve the accuracy of our proposed method for extracting topics from the academic contents by using domain-specific dictionaries for technical terms, or Wikipedia for general words.

5.4 Experiment 2: Effectiveness of Browsing iPosters

In this experiment, we verified how the generated iPosters can help users grasp overviews of slides in presentation contents. We conducted this experiment with 10 participants in groups *X* and *Y*, five participants in each group, using two academic contents (No.2 and No.15) and their iPosters are generated by our prototype tool, providing a level of expertise in Informatics that is important for the participants. No.2 contains many slides with a complex topic structure, and No.15 contains fewer slides with a simple topic structure. For evaluation purposes, we first prepared two tasks to participants, respectively. For group

Fig. 7 Results of average of browsing time and answering time **Fig. 8** Results of average of correct rate of topics

X, (1) browsing slides in No.2, and (2) browsing iPoster of No.15. For group *Y*, (1) browsing slides in No.15, and (2) browsing iPoster of No.2. For each task, participants need to finish questions in three steps as follows:

- S1** Browsing a given content (i.e., a set of slides or an iPoster) for a full understanding of it. Please write your browsing start time and finish time.
- S2** Finding ten most topics, those indicates an overview of the given content, and classifying them into two types: main topic or sub topic in the given content. Please write ten topics of the given content, and classify them into main topic or sub topic. And please record your start time and finish time.
- S3** Listing main-sub topic pairs in hierarchical structures from your written topics in **S2**, when the hyponymy relation between two topics. Please draw main-sub topics in hierarchical structures.

For each question, the participants can freely browse the slides and the iPosters without any restriction; they can browse the iPoseters with navigation many times.

For analyzing these answers, we first prepared correct answers on the topics in Experiment 1, i.e. 16 topics in No.2 and 13 topics in No.15. In general, we need to verify correct answers to authors of experimental dataset. Since the online presentation contents are created and shared by the authors (e.g., teachers, presenters, etc.); it can be considered that the online presentation contents will easily be understood to users (e.g., students, self-learners, etc.). Therefore, we defined a correct answer of a main topic or

Fig. 9 Results of average of correct rate of main-sub topics

a sub topic as when six or more participants identified the same one, then, 3 main topics and 2 sub topics in No.2, and 2 main topics and 3 sub topics in No.15. Finally, we compared the average of browsing time, correct rate of topics, correct rate of main topics, correct rate of sub topics, and their answer time between slides and their generated iPosters. The results are shown in Figure 7, Figure 8, and Figure 9; the experimental results were as follows:

- In Figure 7, the average of browsing time and answering time of the iPosters were shorter than those of slides, we confirmed that our iPosters could reduce learning time of slide-based educational contents.
- In Figure 8, the average correct rates of topics in the iPosters were also higher than those of slides. We believed that users browsing generated iPosters of slides can grasp overviews with both a complex topic structure (iPoster of No.2 is shown in Figure 10) and a simple topic structure (iPoster of No.15 is shown in Figure 11) though navigation more easily in short time.
- The average saving time of No.2 was higher than that of No.15 (see Figure 7), since No.2 contains more information than No.15 (i.e., 24 slides in No.2 and 15 slides in No.15, average 20 words per slide in No.2 and average 15 words per slide in No.15). As a result, we believed that our generated iPosters are useful and efficient for grasping overviews, when many slides of the presentation content contain a lot of information.
- In Figure 9, the average correct rate of sub topics in both iPosters and slides was very low, because they depended on correct answers of topics from Experiments 1. As a result of the average correct rate of main and sub topics in the iPoster were higher than those of the slides, respectively; we confirmed that structural layouts of the iPosters were more useful when users browse slides with a complex context of content.

Finally, main-sub topic pairs in the hierarchical structures were listed by participants while they browse slides. In order to verify the topic structure of the iPosters, we identify the participants' listed main-sub topic pairs whether

A Credibility Analysis Method based on Editor's Intention for Modified Maps

Fig. 11 Generated iPoster of No.15

they are existed in our generated iPosters or not, which allow a descendant relation between two topics in the iPosters. To identify the main-sub topic pairs, we only used the correct answers of the topics in Experiment 1. Figure 12 shows an example of the main-sub topics in the hierarchical structures were drawn by the participants while they browsed slides of No.2, and these hierarchical structures are same to those in the iPoster of No.2. In addition, Figure 13 shows the average of coverage rate of the main-sub topic pairs in the generated iPosters of No.2 and No.15, it indicates how many main-sub topic pairs of the generated iPosters are same to the participants' listed main-sub topic pairs by browsing slides. Since both the average of the coverage rate of our generated iPosters were high, there is a very small minority of the participants' listed main-sub topic pairs that are reversed in the iPosters, we then confirmed that our method can generate the iPosters based on the topic structure, this enables the participants to easily understand the sub topics of the specified topics.

This experiment showed that our method of browsing iPosters is more useful than browsing slides for grasping overviews of the slides in short time. In particular, our

Extraction of Spatial Relations on SVG Maps and Its Applications

Fig. 10 Generated iPoster of No.2

Fig. 12 An example of main-sub topic pairs were drawn by participants

iPoster generation method is helpful for learning slide-based educational contents containing higher levels of expertise. As mentioned above, we consider that it is possible to develop useful applications for MOOC based on iPosters, such as a collaborative learning platform or an exploratory search tool with ZUIs by considering users' interactions on the iPosters.

Fig. 13 Results of average of coverage rate of main-sub topic pairs

6. Conclusions and Future Work

In this paper, we have proposed a method to generate an interactive poster by using slides, called iPoster, that presents textual and graphic elements in a meaningfully structured layout with a ZUI, to promote user interaction. Especially, we introduced a topic structure analysis model for extracting elements of the slides and determining the semantic relationships between the elements. In order to generate an iPoster, we initially placed the elements in a hierarchical structure combined with a stacked Venn. We then attached the zooming and panning transitions between the elements, based on the semantic relationship types. Through our evaluation with actual academic presentations and lecture contents, we confirmed that the iPosters were successfully generated by our topic structure analysis model that extracts the topics by considering the presenta-

tion context, and we were able to confirm that the iPosters help users to easily and efficiently grasp overviews of slides in online presentation contents through our experiments.

In the future, we plan to varied structural layouts for generating iPosters to help users intuitively understand traditional presentations. Further, we need to consider the topic structure analysis on slides by considering user interaction on an iPoster. We must combine the meanings of the content with the user interaction. For instance, if the user zoomed-in from x on y on the iPoster, we can deduce that the user wanted the details of y . However, x describes y and z describes y shows that both x and z have the details of y in the slides. Then, we can suggest y zooms-in z to the user, which indicates that z has the details of y that can satisfy the user's requirement.

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 ◇

- [Bederson 94] Bederson, B. B. and Hollan, J. D.: Pad++: a Zooming Graphical Interface for Exploring Alternate Interface Physics, in *Proc. of the 7th ACM Symposium on User Interface Software and Technology (UIST 1994)*, pp. 17–26 (1994)
- [Bederson 04] Bederson, B., Grosjean, J., and Meyer, J.: Toolkit Design for Interactive Structured Graphics, in *IEEE Transactions on Software Engineering*, Vol. 30, pp. 534–546 (2004)
- [Clarke 97] Clarke, S. C. and Willet, P.: Estimating the Recall Performance of Web Search Engines, in *Aslib Proceedings*, Vol. 49, pp. 184–189 (1997)
- [Faste 12] Faste, H. and Lin, H.: The Untapped Promise of Digital Mind Maps, in *Proc. of the 30th International Conference on Human Factors in Computing Systems (CHI 2012)*, pp. 1017–1026 (2012)
- [Good 02] Good, L. and Bederson, B. B.: Zoomable User Interfaces as A Medium for Slide Show Presentations, in *Journal of Information Visualization*, Vol. 1, pp. 35–49 (2002)
- [Lanir 13] Lanir, J., Booth, K. S., and Wolfman, S.: An Observational Study of Dual Display Usage in University Classroom Lectures, in *Journal of Human-Computer Interaction*, Vol. 28, pp. 335–377 (2013)
- [Li 08] Li, Z., Shi, S., and Zhang, L.: Improving Relevance Judgment of Web Search Results with Image Excerpts, in *Proc. of the 17th International World Wide Web Conference (WWW 2008)*, pp. 21–30 (2008)
- [Lichtschlag 12a] Lichtschlag, L., Hess, T., Karrer, T., and Borchers, J.: Canvas Presentation in the Wild, in *Proc. of CHI '12 Extended Abstracts on Human Factors in Computing Systems (CHI EA '12)*, pp. 537–540 (2012)
- [Lichtschlag 12b] Lichtschlag, L., Hess, T., Karrer, T., and Borchers, J.: Fly: Studying Recall, Macrostructure Understanding, and User Experience of Canvas Presentations, in *Proc. of the 30th International Conference on Human Factors in Computing Systems (CHI 2012)*, pp. 1307–1310 (2012)
- [MeCab] MeCab, : <http://mecab.sourceforge.net/>
- [Ohshima 07] Ohshima, H., Nakamura, S., and Tanaka, K.: SlothLib: A Programming Library for Research on Web Search [in Japanese], in *The Database Society of Japan (DBSJ Letters)*, Vol. 6, pp. 113–116 (2007)
- [Pradhan 06] Pradhan, S.: 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)*, pp. 295–306 (2006)
- [Pschetz 14] Pschetz, L., Yatani, K., and Edge, D.: TurningPoint: Narrative-driven Presentation Planning, in *Proc. of the 32nd Annual ACM Conference on Human Factors in Computing Systems (CHI 2014)*, pp. 1591–1594 (2014)
- [Spicer 12] Spicer, R., Lin, Y., Lelliher, A., and Sundaram, H.: NextSlidePlease: Authoring and Delivering Agile Multimedia Presentations, in *ACM Transactions on Multimedia Computing, Communication and Applications*, Vol. 8, pp. 53:1–53:20 (2012)
- [Tufte 03] Tufte, E.: The Cognitive Style of PowerPoint, in *Graphics Press, Cheshire* (2003)
- [Villa 12] Villa, de la M., Aparicio, F., Mana, J. M., and Buenaga, de M.: A Learning Support Tool with Clinical Cases Based on Concept Maps and Medical Entity Recognition, in *Proc. of the 2012 ACM International Conference on Intelligent User Interfaces (IUI 2012)*, pp. 61–70 (2012)
- [Yokota 06] Yokota, H., Kobayashi, T., Okamoto, H., and Nakano, W.: Unified Contents Retrieval from an Academic Repository, in *Proc. of International Symposium on Large-scale Knowledge Resources (LKR 2006)*, pp. 41–46 (2006)

〔担当委員：齋藤 ひとみ〕

Received May 8, 2014.

Author's Profile

Wang, Yuanyuan (Member)

She is a researcher at Faculty of Computer Science and Engineering, Kyoto Sangyo University. She received her MS and PhD degrees in Human Science and Environment from University of Hyogo, in 2011 and 2014, respectively. She has worked at Microsoft Research Asia as a research intern in the Human-Computer Interaction Group in 2013. Her research interests include e-learning systems, structural analysis, multimedia databases and human-computer interaction.

She is a member of Information Processing Society of Japan and the Database Society of Japan.

Kawai, Yukiko

She received the MS and PhD degrees in Information Science and Technology from Nara Institute of Science and Technology, in 1999 and 2001, respectively. She has worked as a research fellow at the National Institute of Information and Communications Technology from 2001 to 2006. Since 2006 she has been a lecturer at Kyoto Sangyo University. Her research interests include data mining, information analyzing and Web information retrieval. She is a member of

Information Processing Society of Japan, the Database Society of Japan and the Institute of Electronics, Information and Communication Engineers.

Sumiya, Kazutoshi

He is a professor at School of Human Science and Environment, University of Hyogo, Japan. His research interests include multimedia databases, Web services, information retrieval, and geographical information systems. He received his PhD in information media from Kobe University and worked for Panasonic in Japan. He was an associate professor at School of Informatics, Kyoto University. He is a member of IEEE Computer Society, ACM, the Institute of

Image Information and Television Engineers, Information Processing Society of Japan, the Database Society of Japan and the Institute of Electronics, Information and Communication Engineers.