

# A Browsing Method for Presentation Slides Based on Semantic Relations and Document Structure for e-Learning

YUANYUAN WANG<sup>1,a)</sup> KAZUTOSHI SUMIYA<sup>2,b)</sup>

Received: March 28, 2011, Accepted: September 12, 2011

**Abstract:** Currently, many universities use Web services, such as SlideShare and edubase, to store presentation files. These files provide varying levels of knowledge, and are useful and valuable to students. However, self-learners retrieving such files still lack support in identifying which slides meet their specific needs. This is because, amongst presentation files intended for different levels of expertise, it is difficult to understand the context, and thus identify relevant information, of a user query in a slide. We describe a novel browsing method for e-learning by generating snippets for target slides. For this, we consider the relevant information between slides and identify the portions of the slides that are relevant to the query. By analyzing the keyword conceptual structure on the basis of semantic relations, and the document structure on the basis of the indent levels in the slides, not only can target slides be precisely retrieved, but their relevant portions can also be brought to the attention of the user. This is done by focusing on portions from either detailed or generalized slides at the conceptual level; this gives the surrounding context to help users easily determine which slides are useful. We also present a prototype system and the results of an evaluation of its effectiveness.

**Keywords:** presentation slide retrieval, e-learning, snippet generation, semantic relation

## 1. Introduction

Presentation slides (e.g., PowerPoint, Keynote) are now one of the most frequently used tools for educational purposes. A considerable amount of slide-based lecture material, often prepared from teaching material used in actual classes at universities or other educational institutions, is freely shared on Web sites such as SlideShare [28] and edubase [6]. In particular, students can view lectures on their iPhone or iPad by using MPMeister [21], which has hosted presentation slides and recorded lecture videos from Kyoto University since 2010 [24] (see Fig. 1). Other online e-learning material archives include those of the Nara Institute of Science Technology [23], which has provided presentation content recorded from lectures for about seven years, and the Database Society of Japan (DBSJ) [5], which stores 1,200 presentations from workshops (DEWS and DEIM) for members of the society. These presentations provide varying levels of knowledge, and are useful and valuable to students. Thus, content can be reviewed and studied alone and when convenient, not only by students who missed a lecture or presentation, but also by anyone interested in the topic.

Currently, self-learners (e.g., students) must formulate a query consisting of proper keywords in order to retrieve the required

lecture slides. However, e-learning material provides varying levels of knowledge associated with the various levels of university courses or seminars, and so many presentation slides will require prior knowledge and expertise. Moreover, if the keywords in the query are common, the large number of search results returned will make it difficult for self-learners to find material appropriate to their level of understanding. This current method does not consider the relevance of the information contained in slides returned by the query, so it is impossible for students to easily determine which of the slides retrieved by the query are appropriate for study.

We present a novel slide retrieval method to meet user requirements for presentations containing different levels of knowledge. This retrieval method has a specific focus on high levels of expertise by using snippet generation. As depicted in Fig. 2, the method can be implemented by (1) extracting the relationships between all slides in a presentation in terms of a user query (shown on the left side in Fig. 2), and (2) generating snippets for target slides that present the relevant portions of the target slides satisfying the query, based on easily understandable relationships between the target and other slides (shown as the right side in Fig. 2). To achieve our goal, we analyzed the implicit semantic relations between keywords, and how the keywords at different indent levels of slides are related to a user query. We derived keyword conceptual structure focusing on ‘is-a’ and ‘part-of’ relations between keywords extracted from the slide text. However, the usage of keywords in slides varies depending on the author. Figure 3 shows the layout of a slide containing a slide title, in-

<sup>1</sup> Graduate School of Human Science and Environment, University of Hyogo, Himeji, Hyogo 670-0092, Japan

<sup>2</sup> School of Human Science and Environment, University of Hyogo, Himeji, Hyogo 670-0092, Japan

<sup>a)</sup> ne11u001@stshse.u-hyogo.ac.jp

<sup>b)</sup> sumiya@shse.u-hyogo.ac.jp



Fig. 1 Captured image of a student studying lecture material with an iPad.

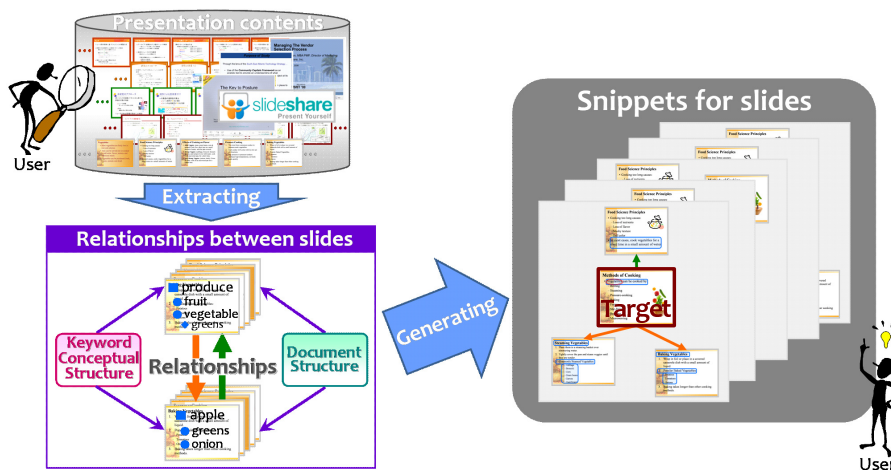


Fig. 2 Conceptual image of slide retrieval using snippet generation.

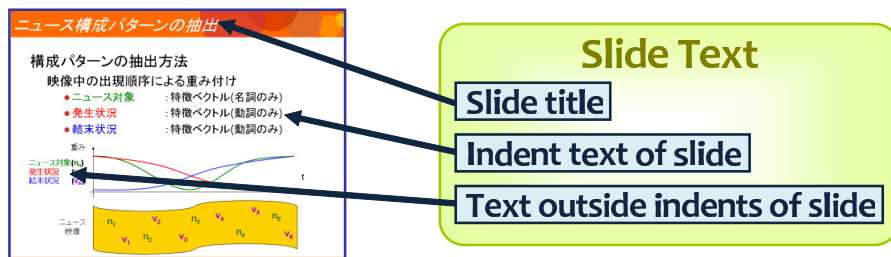


Fig. 3 Example of slide text.

dedented text, and text outside indents, which together we call the slide text. We derived document structure by focusing on certain features of the slides, such as the levels of indents in the slide text, as these are often used to help users to better understand the content in slides. It was then necessary to use the semantic relations and document structure to determine the portions of slides related to the user query; furthermore, we detected the relationships between slides in terms of the query

As an example, consider the user query “vegetable.” The snippet for the target slide, which we call slide *a*, is shown in Fig. 4. Some presentation slides may be related to other slides in terms of detailed and generalized information. Therefore, we generate snippets of the relevant portions of slides based on relationships between slides related to the query. For instance, we define slide *b* as being conceptually related to slide *a* as the specific content of the two slides is related; the keyword conceptual structure and document structure for slides *a* and *b* are shown in the

callout rectangle in Fig. 4. The explanation provided in slide *b* (“spinach is a leafy vegetable”) is likely to be more specific and detailed than the general definition provided in slide *a* (“vegetables”). Therefore, slide *b* has a *detailed* relationship with slide *a* in terms of “vegetable.” In this case, a snippet for slide *a* looks like portion  $P_a$  of slide *a* with portion  $P_b$  of slide *b* related to “vegetable.”

The remainder of this paper is organized as follows. In the next section, we describe our approach and review related work. Section 3 contains an explanation of the keyword conceptual structure and document structure, and we determine the mathematical relationships between slides. Section 4 describes the generation of snippets using the relationships between slides, and Section 5 introduces a prototype application for study support based on our method, illustrating the results of an experiment conducted using a real dataset of presentation content. Finally, Section 6 concludes this paper with suggestions for further work.

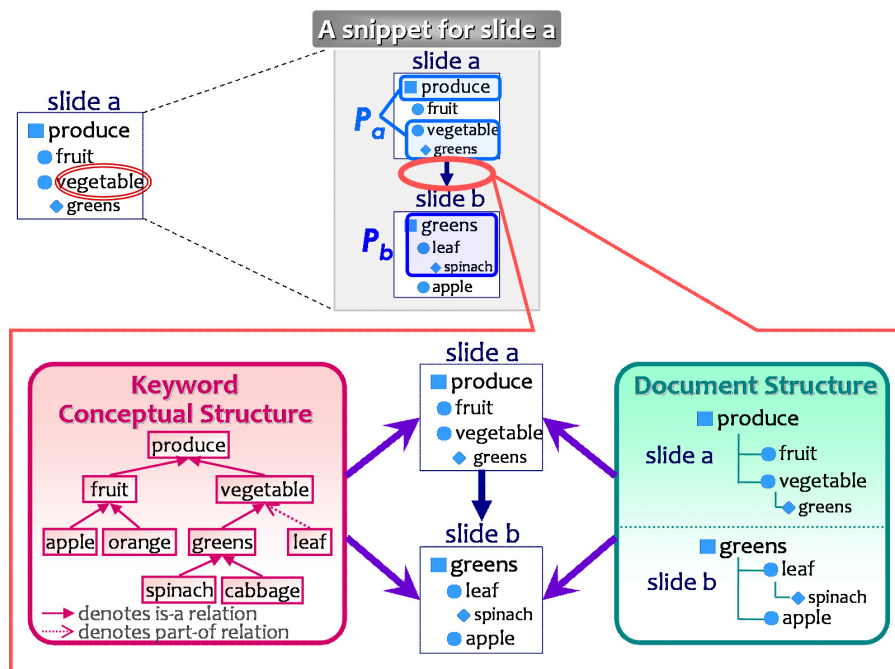


Fig. 4 Snippet for a slide using the relationships between slides.

## 2. Our Approach and Related Work

### 2.1 Our Approach

In this paper, we discuss how to help users understand the context of slides so they can select appropriate ones for self-learning purposes from retrieval results. A traditional snippet uses a keyword-in-context (KWIC) index (or permuted index), which is an index of text documents with a wide layout, to obtain a retrieval result consisting of a portion containing the user query; this is displayed with surrounding words to provide the context of particular index words [17]. This kind of index helps determine the meaning of index words of interest through their context. This context visualization technique is also useful for search engines, such as Google, to display search results. Google returns a ranked list of webpages to users, showing page titles together with a few lines of the text as a segment (called a snippet) containing the query terms (i.e., query terms are shown in the context of returned web documents). This is particularly helpful for judging the relevance of the retrieved document.

We propose a browsing method referred to as Slide KWIC to help enhance user comprehension of the context for target slides related to their query. Also, snippets are generated with Slide KWIC to show the surrounding context of the target slides. A snippet for the target slide, which we call the focused slide, is shown in Fig. 5. There are three layers: the basic layer is the focused slide, the high layer is a generalized slide of the focused slide, and the low layer is a detailed slide of the focused slide. We then generate a snippet consisting of a captured portion of the focused slide with the relevant portions of the related slides in terms of the query, helping users to understand the presentation content in the focused slide.

The semantic relations between keywords, and how the keywords in different levels of indents in slides are related to a user

query, were then analyzed. We defined the keyword conceptual structure for the semantic relations between keywords to be extracted from the slide text using the conceptual dictionary, WordNet [10], [31]. In addition, we defined the document structure based on indents in the slide text extracted from presentation files. As mentioned above, the semantic relations and document structure can then be used to identify portions of sentences for a given indent level relevant to the query in the focused slide, along with the relevant portions from other slides.

### 2.2 Related Work

Due to the changing nature of domestic and international studies, e-learning has become a new teaching model, characterized mainly by self-learning that is supported by information technology and centered on the learners. Ando et al. [1], Nozaki et al. [25], and Kishi et al. [11] analyzed the effect of using tablet PCs for making annotations during e-learning, such as underlining or note-taking, and found that students take meaningful notes freely in this environment. These studies focused on the effect of the interactions between humans, and between a human and the learning material. In contrast, Wu et al. [32] presented three levels—*theoretical, technical, and activity*—of the e-learning teaching process and system design. These three levels solve different issues and strongly promote both theoretical and practical e-learning research [2], [3]. In our study, we focused solely on the interaction between a human and the learning material, to enhance user comprehension of how to easily browse and select the appropriate information for study.

In general, context is useful for understanding, and several studies, which we now briefly explain, have exploited context in different ways. Pattanasri et al. [22] utilized information in textbooks to construct an entailment ontology, finding that two types of entailment relations were helpful for identifying context when

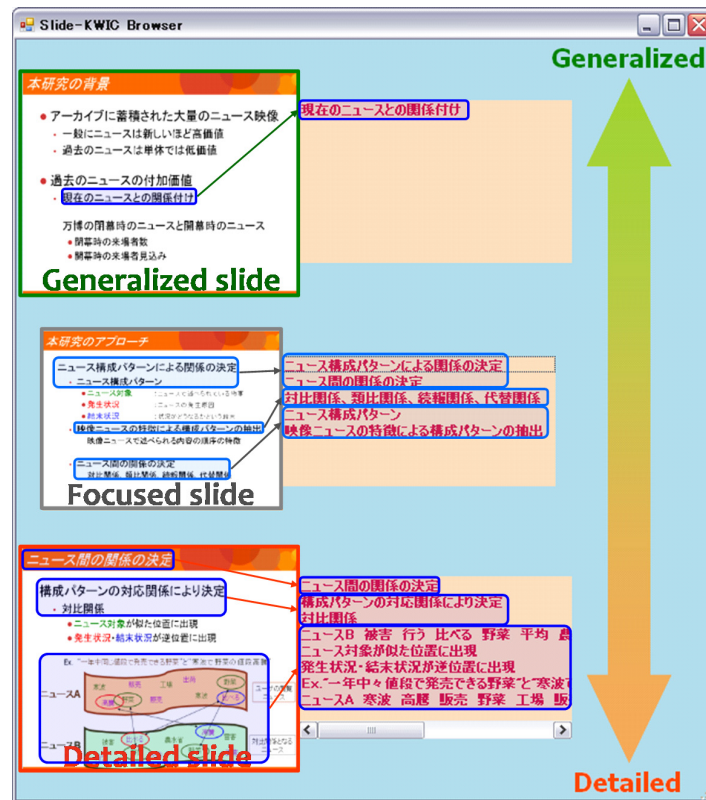


Fig. 5 Screenshot of a snippet generated by Slide KWIC for focused slide.

trying to understand search results inside e-learning material. There are also various techniques for visualizing information. In a literature review, Leung et al. [16] proposed the focus-plus-context technique, using a small display window through which the information of interest can be viewed without losing context (e.g., surrounding items). The basic idea is that focused items occupy a large portion of a display window, while a much smaller space is reserved for displaying contextual items (sometimes incorporating distortion techniques). Closely related to focus-plus-context is the preview-and-overview presentation style [7]. The context role of the preview is to provide coarse-grained details of items; for example, in a photo database, users may view photo thumbnails before downloading actual photos to consider them in full detail. The context role of the overview is to provide users with an idea of what is available in a collection of items and what is not. For instance, the sitemap of a company’s website provides a summary of information (e.g., webpages) available on that particular website. Thus, users can browse a collection of items or documents to understand them easily. Similarly, we generate snippets for the focused slides to give the surrounding context of these slides for a user query at the conceptual level. By conceptual level, we mean that the snippets are taking account of the structure of the slide, and the semantic relations between different keywords contained therein. Users can then browse a collection of slides to visualize the relevant information between slides.

Most studies related to e-learning material have focused on the retrieval of slides. Yokota et al. [33] and Okamoto et al. [27] proposed a system named Unified Presentation Slide Retrieval by Impression Search Engine (UPRISE) for retrieving a sequence of lecture slides from archives containing a combination of slides

and recorded videos. Kobayashi et al. [13] described a method of retrieving lecture slides with UPRISE based on the use of laser pointer information. Le et al. [14] proposed a method for extracting important slides by automatically generating digests from recorded presentation videos. Their method extracts important slides from unified content, on the basis of the metadata features of either a single medium or two heterogeneous media. However, we consider that retrieving only important slides can destroy the implicit relevance of information spread across a number of slides, particularly where it is difficult to understand the context of individual slides, and their method cannot be used to browse important slides containing information related to a query. This does not enhance the user’s understanding of slides. Therefore, our objectives were to effectively retrieve slides by implicitly accumulating relevant information between slides for a user query, and to generate snippets for the focused slides using the relationships between slides related to that query.

In previous works, Kitayama et al. [12] proposed a method for extracting slides with corresponding video scenes based on relationships between slides and their roles, and Wang et al. [30] described a process for automatically generating learning channels by using the semantic relationships that implicitly exist in the slides of a lecture with an accompanying recorded video. These studies were similar to our study, in that a method for retrieving the desired slides using relationships between, and relevant information about, these slides was proposed. Our method is an extension of these methods, developed by focusing on generating snippets for focused slides based on the relationships between slides.



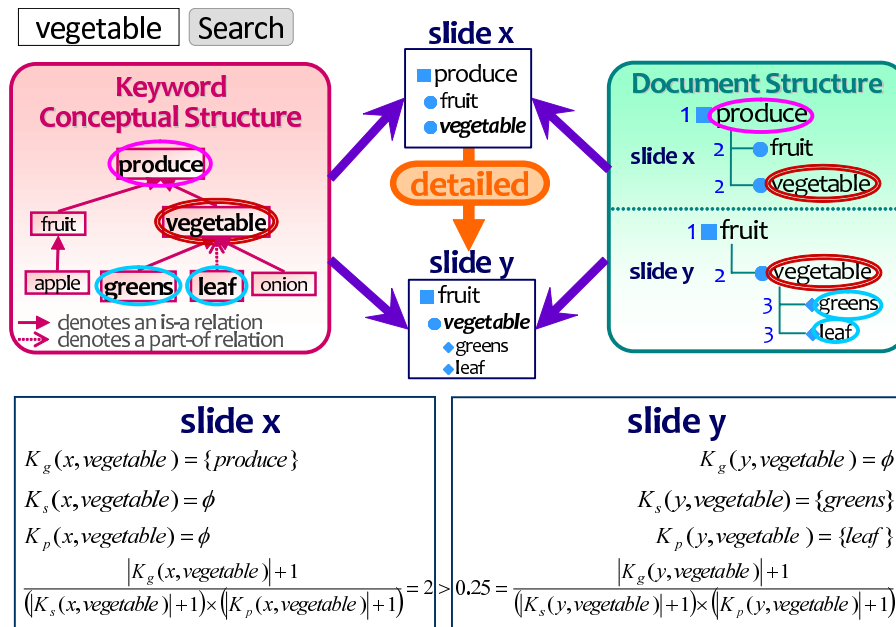


Fig. 6 Example of a detailed relationship between two slides.

### 3. Determining Relationships Between Slides Using Semantic Relations and Document Structure

#### 3.1 Keyword Conceptual Structure and Document Structure

The content of one presentation contains thumbnails (images) of slides and their text information. We consider semantic relations exist implicitly between keywords extracted from the slide text. For example, when the keyword “fruit” is included in the user query “kinds of fruit,” a semantic relationship is assumed to exist between the keyword “fruit” and other keywords in the slide text; for example, another keyword “apple” describes a specialized explanation of the keyword “fruit.” Furthermore, other keywords such as “pulp” and “peel” also give explanations of the keyword “fruit.” Therefore, various semantic relations such as is-a and part-of [19], [20] are used as a basis for the most common semantic relations between keywords. “X subsumes Y, or Y is-subsumed-by X” (Y is-a X) usually means that concept Y is a specialization of concept X and that concept X is a generalization of concept Y. Moreover, “Z is part of X, or X has Z as a part of itself” (Z is part-of X) usually means that Z is a meronym of X and that the whole X has Z as a part. For example, “fruit” is a generalization of an “apple,” “orange,” and many other fruits; in other words, an “apple” is-a “fruit.” Furthermore, “fruit” is a holonym of “pulp,” “peel,” and many other meronyms; in other words, “pulp” is a part-of “fruit.” Therefore, we define the keyword conceptual structure as consisting of an is-a or part-of relation between keywords extracted using WordNet [10], [31].

We define document structure from slides that appear in the outline pane, based on indents in the slide text. We defined the slide title (first level indent) as the upper level. The first item of text is considered to be on the second level, and the depth of the sub-items increases with the level of indentation (third level, fourth level, etc.). Objects outside of the text, such as figures or

tables, are considered to be at the same indent level as the text in which they are placed. If a given keyword appears in the title of the slide or in lines with smaller indents, we implicitly assume that the lower-level indented keywords are supplementary and that they explain the upper-level keywords.

#### 3.2 Preliminary Experiment: Usage Tendency of Document Structure

We conducted a preliminary experiment to confirm the usage tendency of this document structure, using presentation slides from the DBSJ Archives [5], HandsOut [8], and SlideShare [28] as experimental data. We extracted 50 slides from each site, giving a total of 150 slides, and analyzed the document structure and presentation category. The results are summarized as follows:

- Slides often used levels of indentation in the categories of academic, educational, and business content.
- This structure was used in 94.0%, 91.7%, and 100% of academic, educational, and business presentations, respectively.
- In 25% of the experimental presentations, no text indent levels were used, as the presentations contained only visual elements, such as pictures or videos.

In this experiment, we confirmed that document structure is used in academic, educational, and business slide presentations. Therefore, we considered semantic relationship types to adapt to the presentation content based on a method involving the document structure.

#### 3.3 Determination of Relationship Types

To determine the relationship between two slides, we define one as a focused slide, and consider the other to be conceptually related to the focused slide through one of two types of relationships: *detailed* and *generalized*. In other words, the relationship has a direction. The focused slide is the starting point, and the other slides are end points, of the direction of the relationship. In the example shown in Fig. 6, slide x is the focused slide, and the

relationship from slide  $x$  to slide  $y$  is a *detailed* relationship. If a slide has a *detailed* relationship with the focused slide, it is called a *detailed* slide. If a slide has a *generalized* relationship with the focused slide, it is called a *generalized* slide. Let  $x$  be the slide number of a focused slide and  $y$  be the slide number of the slide that we want to retrieve. The relationship types are determined for all slides containing the keyword  $q$  from a user query.

### 3.3.1 Determination of Detailed Relationships

If a slide has more information about a user query than the focused slide, its relationship with the focused slide is *detailed*. We explain the determination of *detailed* slides using the query keyword  $q$ , present in both the focused slide  $x$  and slide  $y$  (slide  $y$  as the *detailed* slide needs to be retrieved). As an example, Fig. 6 shows the determination of the *detailed* relationship between slides  $x$  and  $y$  for a query on the word “vegetable.”

When the query keyword  $q$  and other keywords in slides  $x$  and  $y$  satisfy certain conditions, slide  $y$  is determined to be the *detailed* slide of slide  $x$ . This is because  $q$  has more specific content in slide  $y$  than it does in slide  $x$ .

$$K_g(x, q) = \{k_i | k_i \in x, l(x, q) \geq l(x, k_i), q \text{ is-a } k_i\} \quad (1)$$

$$K_s(x, q) = \{k_j | k_j \in x, l(x, q) < l(x, k_j), k_j \text{ is-a } q\} \quad (2)$$

$$K_p(x, q) = \{k_m | k_m \in x, l(x, q) < l(x, k_m), k_m \text{ part-of } q\} \quad (3)$$

Here,  $K_g(x, q)$  is a set of keywords that can be considered as general information in terms of  $q$  in slide  $x$ . In Eq. (1),  $l(x, q)$  is a function that returns the level of indentation of  $q$  in slide  $x$ , and will thus return a value greater than 1. When  $q$  appears frequently in slide  $x$ ,  $l(x, q)$  will return the lowest possible value; that is, the uppermost level at which  $q$  occurs in slide  $x$ . This is because we consider that when  $q$  appears in an upper level, all of the other levels in which  $q$  appears in the body of that slide are explanatory points related to the upper level occurrence of  $q$ . Keyword  $k_i$  is included in the levels that have a hierarchical relationship with the level of  $q$ , and  $k_i$  belongs to the set of keywords  $K_g(x, q)$  in slide  $x$ .  $l(x, k_i)$  is less than or equal to  $l(x, q)$  in the document structure, and  $q$  (e.g., “vegetable”) has an is-a relation with  $k_i$  (e.g., “produce”) in the keyword conceptual structure (see Fig. 6). When  $k_i$  does not exist in slide  $x$ ,  $K_g(x, q)$  will be empty. In our method, the keyword conceptual structure is extracted as a tree-shaped structure. In general, an is-a or part-of relation between keywords is equivalent to a parent-child relation, and so our method may classify is-a or part-of as a descendant relation.  $K_s(x, q)$  is a set of keywords that can be considered as specific information in terms of  $q$  in slide  $x$ .

In Eq. (2), keyword  $k_j$  is included in the levels that have a hierarchical relationship with the level of  $q$ , and  $k_j$  belongs to the set of keywords  $K_s(x, q)$  in slide  $x$ .  $l(x, k_j)$  is greater than  $l(x, q)$  in the document structure, and  $k_j$  (e.g., “greens”) has an is-a relation with  $q$  (e.g., “vegetable”) in the keyword conceptual structure (see Fig. 6). When  $k_j$  does not exist in slide  $x$ ,  $K_s(x, q)$  will be empty.  $K_p(x, q)$  is the set of keywords that can be considered as additional information in terms of  $q$  in slide  $x$ .

In Eq. (3), keyword  $k_m$  is included in the levels that have a hierarchical relationship with the level of  $q$ , and  $k_m$  belongs to the set of keywords  $K_p(x, q)$  in slide  $x$ .  $l(x, k_m)$  is greater than  $l(x, q)$  in

the document structure, and  $k_m$  (e.g., “leaf”) has a part-of relation with  $q$  (e.g., “vegetable”) in the keyword conceptual structure (see Fig. 6). When  $k_m$  does not exist in slide  $x$ ,  $K_p(x, q)$  will be empty. For the conditions mentioned above, when  $k_i$ ,  $k_j$ , or  $k_m$  does not exist in slide  $x$ , then  $K_g(x, q)$ ,  $K_s(x, q)$ , or  $K_p(x, q)$ , respectively, will be empty. In general, detailed information means a more specific explanation of a term; a *detailed* relationship seems to be a mixture of is-a and part-of relations.

Based on the above criteria, we compute the ratio of general information to detailed information related to  $q$  for slides  $x$  and  $y$ , and compare their ratios using the following formula:

$$\frac{|K_g(x, q)| + 1}{(|K_s(x, q)| + 1) \times (|K_p(x, q)| + 1)} > \frac{|K_g(y, q)| + 1}{(|K_s(y, q)| + 1) \times (|K_p(y, q)| + 1)} \quad (4)$$

where the function  $|K_g(x, q)|$  extracts the total number of  $k_i$  in  $K_g(x, q)$ ,  $|K_s(x, q)|$  extracts the total number of  $k_j$  in  $K_s(x, q)$ , and  $|K_p(x, q)|$  extracts the total number of  $k_m$  in  $K_p(x, q)$  in slide  $x$ .  $K_g(y, q)$ ,  $K_s(y, q)$ , and  $K_p(y, q)$  are also sets of keywords in slide  $y$ , satisfying the same conditions as  $K_g(x, q)$  in Eq. (1),  $K_s(x, q)$  in Eq. (2), and  $K_p(x, q)$  in Eq. (3). Thus, Eq. (4) can be used to calculate the ratio of  $|K_g(x, q)|$  to  $|K_s(x, q)|$  and  $|K_p(x, q)|$  for slide  $x$  and the ratio of  $|K_g(y, q)|$  to  $|K_s(y, q)|$  and  $|K_p(y, q)|$  for slide  $y$ .

If the ratio calculated for slide  $x$  is higher than that calculated for slide  $y$  using Eq. (4), slide  $y$  is determined to be the *detailed* slide of slide  $x$  with regard to  $q$ .

### 3.3.2 Determination of Generalized Relationships

If a slide contains content about the query in the outline given in a *generalized* slide, it is described in relation to the focused slide. We explain the determination of *generalized* slides using the query keyword  $q$  present in the focused slide  $x$  and slide  $y$ ; this keyword needs to be retrieved.

$$\frac{|K_g(x, q)| + 1}{(|K_s(x, q)| + 1) \times (|K_p(x, q)| + 1)} < \frac{|K_g(y, q)| + 1}{(|K_s(y, q)| + 1) \times (|K_p(y, q)| + 1)} \quad (5)$$

When the query keyword  $q$  and other keywords in slides  $x$  and  $y$  satisfy Eqs. (1), (2), (3), and (5), then slide  $y$  is determined to be a *generalized* slide of slide  $x$  with regard to  $q$ . This is because slide  $y$  has more general content on  $q$  than does slide  $x$ . Equation (5) can be used to calculate the ratio of  $|K_g(x, q)|$  to  $|K_s(x, q)|$  and  $|K_p(x, q)|$  for slide  $x$  and the ratio of  $|K_g(y, q)|$  to  $|K_s(y, q)|$  and  $|K_p(y, q)|$  for slide  $y$ .

Thus, *detailed* and *generalized* slides are functionally interchangeable, whereas a focused slide is a *generalized* slide from the viewpoint of a *detailed* slide.

## 4. Snippet Generation Using the Relationships Between Slides

To generate snippets, Slide KWIC takes the portions of the focused slides relevant to a user query by using the relationships between slides. It is difficult for users to understand the relevant information between portions of slides in terms of the query. For example, a user may want to study slide 4 to further understand

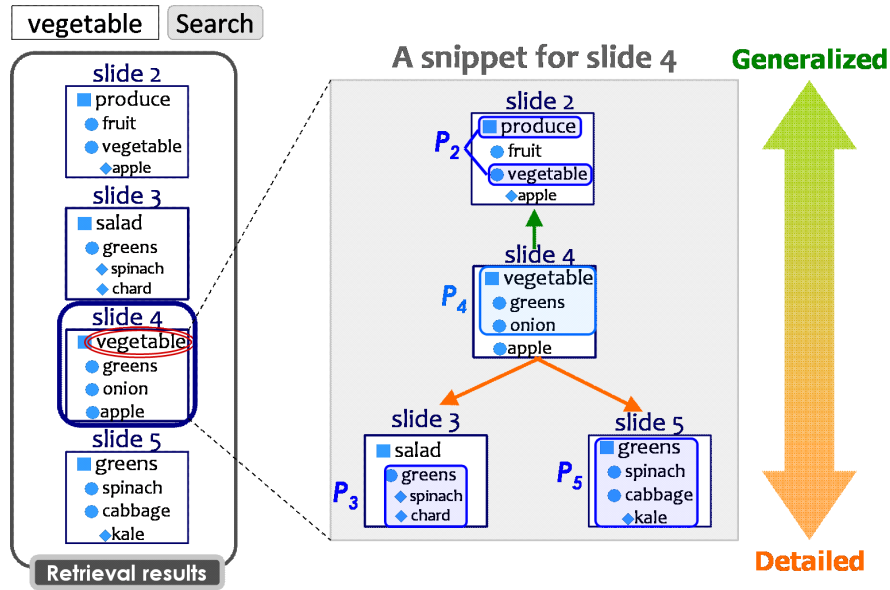


Fig. 7 Example of snippet generation.

“vegetable” in the lecture content about Vegetable as food. Our method generates a snippet for slide 4 that captures portion  $P_4$  of slide 4, along with portion  $P_2$  of slide 2 that includes text on the indent levels, explaining “produce” with regard to “vegetable.” Portions  $P_3$  and  $P_5$  include text on the indent levels, explaining “cabbage and spinach are green vegetables,” with regard to “vegetable” for slides 3 and 5 (see Fig. 7). In this case, slide 2 explains that “produce” has a *generalized* relationship with slide 4 with regard to the keyword “vegetable,” and slides 3 and 5 both explain that “cabbage and spinach are green vegetables,” implying a *detailed* relationship with slide 4 in terms of “vegetable.” When the user browses the snippet for slide 4, consisting of portion  $P_4$  from slide 4 and portions  $P_2$ ,  $P_3$ , and  $P_5$  from slides 2, 3, and 5, respectively, he or she is provided with more information on “vegetable” than just that in slide 4, and this enables the user to further his or her understanding easily. Therefore, our snippet-generation method is based on the context of slides to present snippets, which contain portions related to the user query in a detailed order to enable snippet comprehension at the conceptual level. This section describes how to generate snippets, based on the relationships between slides related to the query, through the following procedures.

#### 4.1 Identifying the Portions of Focused Slides

Although our method can retrieve slides related to a user query, the relevance of the information contained on the focused slides must be determined. Therefore, our method first identifies the portions of the focused slide related to a user query based on the keyword conceptual structure and document structure. Let  $x$  be the slide number of the focused slide. When the query keyword  $q$  and other keywords in slide  $x$  satisfy Eqs. (1), (2), (3), (6), (7), (8), and (9), portion  $P$  of slide  $x$  is determined to be related to the query keyword  $q$ .

$$K_w(x, q) = \{k_h | k_h \in x, l(x, q) \geq l(x, k_h), q \text{ part-of } k_h\} \quad (6)$$

$$L_g(x, q)$$

$$= \{s_n | l(x, k_u) \leq l(x, s_n) \leq l(x, q), k_u \in K_g(x, q) \cup K_w(x, q)\} \quad (7)$$

$$L_s(x, q) = \{s_t | l(x, q) \leq l(x, s_t) \leq l(x, k_v), k_v \in K_s(x, q) \cup K_p(x, q)\} \quad (8)$$

$$P = L_g(x, q) \cup L_s(x, q) \quad (9)$$

Here,  $K_w(x, q)$  is a set of keywords that can be considered as a whole concept in terms of  $q$  in slide  $x$ . In Eq. (6), keyword  $k_h$  is included in the levels that have a hierarchical relationship with the level of  $q$ , and  $k_h$  belongs to the set of keywords  $K_w(x, q)$  in slide  $x$ ;  $l(x, k_h)$  is less than or equal to  $l(x, q)$  in the document structure, and  $q$  (e.g., “vegetable”) has a part-of relation with  $k_h$  (e.g., “leaf”) in the keyword conceptual structure (see Fig. 6). When  $k_h$  does not exist in slide  $x$ ,  $K_w(x, q)$  will be empty. A set  $L_g(x, q)$  consists of sentences from the levels that contain general information related to  $q$  in slide  $x$ . Sentence  $s_n$  belongs to the set of sentences  $L_g(x, q)$  in slide  $x$  if the following condition is satisfied:  $s_n$  must be included in one of the indent levels ranging from the level of the sentence containing  $q$  to the level of the sentence containing keyword  $k_u$ , where  $k_u$  belongs to  $K_g(x, q)$  or  $K_w(x, q)$ , and  $q$  is-a  $k_u$  or  $q$  is part-of  $k_u$  in slide  $x$ . The selection and extraction of the  $s_n$  is performed according to Eq. (1) or Eq. (6).

In Eq. (7),  $l(x, s_n)$  is not greater than  $l(x, q)$  in the document structure, so  $L_g(x, q)$  will extract the sentences,  $s_n$ , containing  $q$  in levels ranging from  $l(x, q)$  to  $l(x, s_n)$ . In addition,  $l(x, s_n)$  is greater than or equal to  $l(x, k_u)$  in the document structure, so  $L_g(x, q)$  will also extract sentences containing  $k_u$  in levels ranging from  $l(x, s_n)$  to  $l(x, k_u)$ . A set  $L_s(x, q)$  consists of sentences from levels that contain specific information related to  $q$  in slide  $x$ . Sentence  $s_t$  belongs to the set of sentences  $L_s(x, q)$  in slide  $x$ , where  $s_t$  is included in the indent levels of sentences from the level of the sentence containing  $q$  to the level of sentence containing  $k_v$ . The keyword  $k_v$ , which has an is-a or part-of relation with  $q$ , belongs to  $K_s(x, q)$  or  $K_p(x, q)$ . This extraction is performed

using Eqs. (2) or (3). In Eq. (8),  $l(x, s_t)$  is not greater than  $l(x, k_v)$  in the document structure, so  $L_s(x, q)$  will extract sentences containing  $q$  in levels ranging from  $l(x, k_v)$  to  $l(x, s_t)$ . As  $l(x, s_t)$  is greater than or equal to  $l(x, q)$  in the document structure,  $L_s(x, q)$  will also extract sentences containing  $q$  in levels from  $l(x, s_t)$  to  $l(x, q)$ . Thus, Eq. (9) can be used to extract a portion  $P$  of slide  $x$ , and thus combine the sets of sentences from different levels,  $L_g(x, q)$  and  $L_s(x, q)$ .

## 4.2 Determining the Relevant Portions of Related Slides

When slide  $x_g$  is a *generalized* slide that has a *generalized* relationship with the focused slide  $x$ , related to query keyword  $q$ , portion  $P_g$  of slide  $x_g$  provides the general content of portion  $P$  of the focused slide  $x$  related to  $q$ . Therefore, portion  $P_g$  of the generalized slide  $x_g$  is determined using the query keyword  $q$  from the focused slide  $x$ .

### 4.2.1 Determining the Portions of Generalized Slides

When slide  $x_g$  is a *generalized* slide that has a *generalized* relationship with the focused slide  $x$ , related to query keyword  $q$ , portion  $P_g$  of slide  $x_g$  provides the general content of portion  $P$  of the focused slide  $x$  related to  $q$ . Therefore, portion  $P_g$  of the generalized slide  $x_g$  is determined using the query keyword  $q$  from the focused slide  $x$ .

$$P_g = L_g(x_g, k_u) \cup L_g(x_g, q) \quad (10)$$

When the query keyword  $q$  in slide  $x_g$  satisfies Eqs. (1), (6), (7), and (10), then portion  $P_g$  of the generalized slide  $x_g$  is determined. This is because the amount of content in slide  $x_g$  that is generic to  $q$  is greater than that in slide  $x$ . A set  $L_g(x_g, q)$  consists of sentences from levels that contain general information related to  $q$  in slide  $x_g$ , and satisfies the same conditions as the set  $L_g(x, q)$  (these conditions apply to slide  $x$  and are given by Eq. (7)). In addition, when slide  $x_g$  contains the keyword  $k_u$ , which belongs to  $K_g(x, q)$  or  $K_w(x, q)$ , then a set  $L_g(x_g, k_u)$  is used to extract a further set of sentences. These come from levels that provide general information in terms of  $k_u$ , the more generalized concept related to  $q$  in slide  $x_g$ , and satisfy the same conditions as the set  $L_g(x, q)$  (these conditions apply to slide  $x$  and are given by Eq. (7)). When slide  $x_g$  contains two or more  $k_u$ , as determined from the focused slide  $x$ , then we can extract two or more sets of sentences from  $L_g(x_g, k_u)$ . Thus, Eq. (10) can be used to determine the portion  $P_g$  of slide  $x_g$  that combines the sets of sentences from  $L_g(x_g, k_u)$  and  $L_g(x_g, q)$ .

### 4.2.2 Determining the Portions of Detailed Slides

When slide  $x_d$  is a *detailed* slide that has a *detailed* relationship with the focused slide  $x$  in respect of the query keyword  $q$ , portion  $P_d$  of slide  $x_d$  provides specific, detailed information about portion  $P$  of the focused slide  $x$  related to  $q$ . Therefore, we determine portion  $P_d$  of the *detailed* slide  $x_d$  using the query keyword  $q$  from the focused slide  $x$ .

$$P_d = L_s(x_d, q) \cup L_s(x_d, k_v) \quad (11)$$

When the query keyword  $q$  in slide  $x_d$  satisfies Eqs. (2), (3), (8), and (11), then portion  $P_d$  is determined from the *detailed* slide,  $x_d$ . This is because the amount of content in slide  $x_d$  specific to

$q$  is greater than that in slide  $x$ . A set  $L_g(x_g, q)$  consists of sentences from levels that contain specific information related to  $q$  in slide  $x_d$ , and satisfies the same conditions as the set  $L_s(x, q)$  (these conditions apply to slide  $x$  and are given by Eq. (8)). Moreover, when slide  $x_d$  contains the keyword  $k_v$ , which belongs to  $K_s(x, q)$  or  $K_p(x, q)$ , then a set  $L_s(x_d, k_v)$  is used to extract an additional set of sentences. These are extracted from levels that provide specific information in terms of  $k_v$ , the more specified concept related to  $q$  in slide  $x_d$ , and satisfy the same conditions as the set  $L_s(x, q)$  (these conditions apply to slide  $x$  and are given by Eq. (8)). When slide  $x_d$  contains two or more  $k_v$ , as determined from the focused slide  $x$ , we can extract two or more sets of sentences from  $L_s(x_d, k_v)$ . Equation (11) can then be used to determine the portion  $P_d$  of slide  $x_d$  that combines the sets of sentences from  $L_s(x_d, q)$  and  $L_s(x_d, k_v)$ .

As mentioned above, our method for generating snippets of the focused slides satisfies user demand by relating portions of the *generalized*, *focused*, and *detailed* slides to provide content varying from generalized to detailed based on a user query for specific content.

## 5. Evaluation

### 5.1 Prototype System

We built a prototype system to support slide retrieval (see Fig. 8), using Microsoft Visual Studio 2008 C#. From user queries, the system aims not only to identify and precisely retrieve target slides, but also related content based on semantic relevance and surrounding context, in order to enhance comprehension. This prototype has three stages: analysis, determination, and application. In the analysis stage, we analyze the features of the slide text according to the keyword conceptual structure, using WordNet [9], [10]<sup>\*1</sup> to extract is-a and part-of relations between keywords from presentation content. The document structure of the slide, and thus information on the indent level of keywords, is constructed by using Office Open XML files from PowerPoint in Microsoft Office 2007<sup>\*2</sup>. The terms in the slides are extracted by using the morphological analyzers *MeCab* [18] and *SlothLib* [26], [29]. In the determination stage, all types of relationships between slides are extracted based on the keyword conceptual structure and document structure. We call our application for browsing the retrieval results the Slide KWIC Browser; it is shown in the right-hand side of Fig. 8. Snippets are generated by identifying portions of the focused slides relevant to the related slides based on the relationships between them.

After a user selects the presentation content for study, enters a query in the textbox, and presses the “Search” button, the retrieved slides are presented in the retrieval results section of the browser. When the user selects a certain retrieved slide to be the focused slide, the Slide KWIC Browser presents a snippet of this slide in an adjacent window. There, a portion of the focused slide, with relevant sentences extracted, is presented in a listbox

<sup>\*1</sup> In our implementation, only the conceptual dictionary WordNet was used. The system can also use other dictionaries.

<sup>\*2</sup> In our implementation, we developed a PowerPoint parser but parsers for Keynote, Open Office Impress, and so on, can also be developed. Therefore, we can also use content made by other presentation formats.



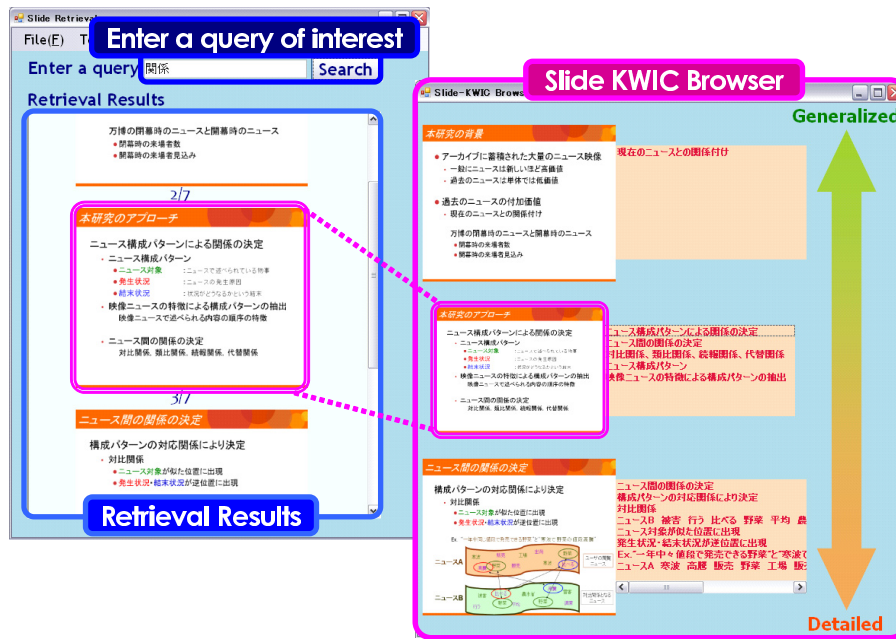


Fig. 8 Screenshot of the prototype system.

Table 1 Experimental dataset.

(1) Academic Contents		
No.	Title	Number of slides
P-W	Mining Disjunctive Tree Patterns	22
P-X	A Web Archive Search Engine Based on the Temporal Relation of Query	15
P-Y	Video Archive Contents Browsing Method based on News Structure Patterns	7
P-Z	Improvement on Processing Rules Stored in Individual Metadata for Flexible Contents Management	30
(2) Lecture Contents		
No.	Title	Average number of slides
L-W	Methods of Education Research	16.4
L-X	Introduction to Psychology	29.6
L-Y	Social Statistics	24.1
L-Z	Introduction to Statistics	27.4

in the center of the Slide KWIC Browser window as a basic layer. Other, related slides, with their relevant portions in listboxes, are displayed above and below the focused slide (shown as the right part in Fig. 8).

5.2 Experimental Dataset

In our experiments, we examined the proposed method of snippet generation for slide-browsing support based on the relationships between slides. We prepared a dataset using actual content, as shown in Table 1, consisting of (1) four actual academic presentations from a session of DEWS2006 in the DBSJ Archives [5], and (2) 36 actual lecture presentations [15] of four introductory courses from the lecture archives of the Social Informatics department at Aoyama Gakuin University. There were 5-15 students in the School of Human Science and Environment, University of Hyogo, taking the Social Informatics course and Information Media lab who participated in the following experiments. We assumed that the academic content in Informatics requires a certain level of expertise and is difficult to understand, and that introductory lectures provide a basic level of knowledge

in Informatics and are thus easily understandable for the students who participated in the following experiments. We show and discuss the experimental results in the follow sections.

5.3 Experiment 1: Validity of Determining Relationship Types

This experiment was designed to assess the generation of snippets based on relationships between slides related to a user query. Five participants freely described the relationships which existed between two slides, assessing 199 slide pairs containing keywords sampled at random from the four academic presentations in the dataset. Relationships between the slide pairs were determined if and when three or more participants described the same relationship. We calculated the coverage using the slide pairs, which were determined according to any relationship type identified by participants; we also defined the *others* relationship for those that could not be determined by our method, as shown in Fig. 9.

The results and our findings were as follows:

- Coverage reached 92.6% (63/68). 68 slide pairs were de-

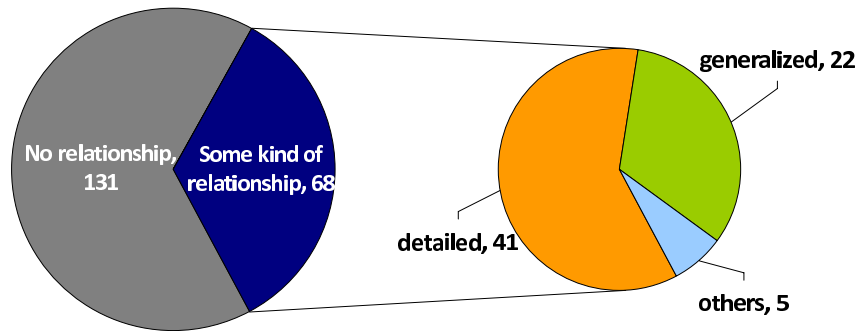


Fig. 9 Coverage of the relationship types.

Table 2 Classification of relationship types.

		Results determined using our system		
		<i>detailed</i>	<i>generalized</i>	<i>others</i>
Correct answers	<i>detailed</i>	<b>91</b>	17	58
	<i>generalized</i>	35	<b>65</b>	48
	<i>others</i>	6	6	91

terminated to have some kind of relationship; 41 slide pairs were described as *detailed*, 22 slide pairs were described as *generalized*, and 5 slide pairs were classified as *others*. No relationship was determined for 131 slide pairs. We concluded that our defined relationships can account for slide relationships.

- Coverage reached a low of 60.3% (41/68) using our system. Of the 68 slide pairs that we had determined as having a relationship (see Fig. 9), the experiment participants only agreed with our opinion of the relationship type on 41 occasions; we thus concluded that the slide relationships in presentations cannot be expressed comprehensively by using our method alone.

Table 2 lists the classification results. A correct answer was defined as a relationship between two slides where three or more participants described the same relationship. Participants did not have any particular bias<sup>\*3</sup>, and we consider that the correct answers can be defined using the answers obtained from participants. Only one type of relationship defined in our system was determined by the participants for any given slide pair, and then this answer was duplicated by more participants. For example, if three of five participants give the answer of “detailed” while each of the remaining two participants does the answers of “generalized” and “others,” respectively, the correct answer becomes “detailed” and the numbers of relationship of “detailed,” “generalized,” and “others” are accumulated by 3, 1, and 1, respectively. We found that *detailed* includes “instance” relationships, where slides show the specific examples with their explanations, and *generalized* includes “parallel” relationships, where slide pairs describe information derived from a single topic on equal terms. However, these relationships did not occur frequently, and are thus difficult to define. We should therefore improve the defini-

tions of *detailed* and *generalized* relationships. This experiment confirmed that the relationships between slides containing any keyword could be covered by using the concept of relationship types. In our method, we focused on *detailed* and *generalized* relationships at the conceptual level, but this should be expanded to determine other types of semantic relationships.

We used three representative keywords from each academic presentation to extract 678 slide pairs. We evaluated the validity of the rules for determining the two types of relationships by precision<sup>\*4</sup>, relative recall [4]<sup>\*5</sup>, and F-measure<sup>\*6</sup> using the results obtained from four methods, and a correct answer was considered to be a slide pair where three or more participants found some relationships present in their free description. The four methods are: “Frequency,” using the keyword frequency, “document,” using the document structure only, “concept,” using the keyword conceptual structure only, and “proposed,” which used our proposed method.

The results for the slide relationships found by the four methods are shown in Fig. 10, and they can be explained as follows:

- The relative recall of *detailed* or *generalized* was low, and many correct answers were detected to have no relationships with our method. We consider the limitations of WordNet [9], [10] to be one factor for the low relative recall. Although WordNet [9], [10] is a large lexical database, it does not necessarily contain all concepts related to an experimental keyword, as there may be new concepts associated with a technical term or new words used in the academic presentation. For instance, while keywords such as “disjunction,” “mining,” and “preorder” frequently appear in the main content of the academic presentation *P-W*, our method based on WordNet [9], [10] cannot extract semantic relations for them, as they are not included in WordNet [9], [10].
- For *detailed*, our method returned more than half of the correct answers. The precision of our method performed well. However, there was a little confusion between “detailed” and “instance.” “Detailed” means a more specific explanation of a term; “instance” means a specific explanation of a term through the use of cases or examples. Therefore, our method

<sup>\*3</sup> Five participants, i.e., A to E. The ratio of the same answers by A and B was 70.9% (141/199); by A and C was 55.3% (110/199); by A and D was 70.4% (140/199); by A and E was 69.8% (139/199); by B and C was 50.3% (100/199); by B and D was 61.8% (123/199); by B and E was 60.3% (120/199); by C and D was 50.8% (101/199); by C and E was 61.3% (122/199); by D and E was 72.9% (145/199).

<sup>\*4</sup> Precision =  $\frac{\text{Number of correct answers of the relationships determined by a method}}{\text{Total number of the relationships determined by a method}}$

<sup>\*5</sup> Relative recall =  $\frac{\text{Number of correct answers of the relationships determined by a method}}{\text{Total number of correct answers of the relationships determined by all four methods}}$

<sup>\*6</sup> F-measure =  $\frac{2 \times \text{Precision} \times \text{Relative recall}}{\text{Precision} + \text{Relative recall}}$

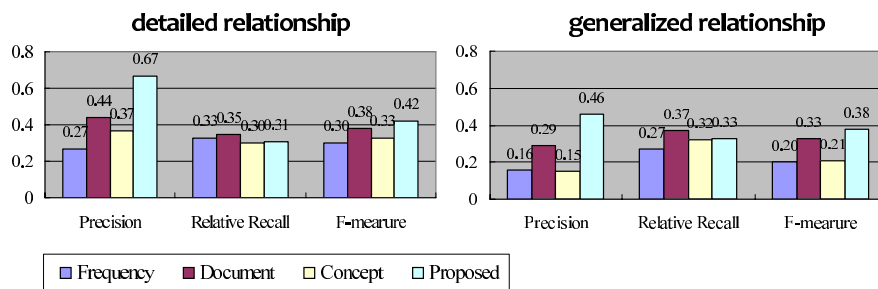


Fig. 10 Performance measure graph.

returned some results as “detailed” when participants labeled their correct answers as “instance.” We focused on whether the specific explanation of a term contained more information, but not how the specific explanation was given, such as in examples.

- For *generalized*, even if the same slide set was considered, participants’ answers differed in terms of “generalized” and “parallel.” If participants answered “generalized,” this means that they understood the content well. However, if the participants answered “parallel,” it means that they understood only that the slides had a relationship. We consider *generalized* to be effective when a user can understand that slides have at least some relationship, but cannot determine the relationship type.

The graph in Fig. 10 shows that the precision and F-measure of our system were higher than those of other methods for determining *detailed* and *generalized* relationships. This experiment confirmed that slides with academic content have some kind of relationship between each other. Our proposed relationships may provide an appropriate definition for using the semantic relations between keywords and the document structure of indents. Furthermore, we believe that a considerable number of slides in the academic presentations provide detailed explanations. However, we should enhance our method for extracting semantic relations between keywords to consider the semantic data model of keywords. In particular, for academic content containing a lot of technical terms, this method should not only involve the use of WordNet [9], [10], but also include such aspects as the use of domain-specific dictionaries, such as the Handbook of Information Processing<sup>\*7</sup> and the Medical Dictionary<sup>\*8</sup>. As mentioned above, we can improve the accuracy of our method for determining relationships between slides.

### 5.4 Experiment 2: Validity of Identifying the Portions of Slides

This experiment aimed to verify whether the proposed method is useful for identification of portions containing sentences relevant to a user query. Five participants freely captured portions containing sentences from different indent levels in the slides, and assessed three representative keywords from 40 actual presentations in the dataset to identify portions of 312 slides. A correct answer was defined as a portion where three or more participants found the sentences on the indent levels of the slides that they

<sup>\*7</sup> Information Processing Society of Japan.  
<sup>\*8</sup> <http://www.medterms.com/script/main/hp.asp>

Table 3 Results of identifying the portions of slides in academic contents.

	Academic contents by our method				
	P-W	P-X	P-Y	P-Z	Average
Precision	69.6% (298/428)	60.4% (166/275)	57.7% (142/246)	66.1% (360/545)	63.5%
Recall	67.3% (298/443)	71.2% (166/233)	64.0% (142/222)	75.8% (360/475)	69.5%
F-measure	0.68	0.66	0.64	0.71	0.67
	Academic contents by the levels contain the given keywords with their AP levels				
	P-W	P-X	P-Y	P-Z	Average
Precision	52.8% (295/559)	47.4% (180/380)	53.8% (135/251)	56.1% (415/740)	52.5%
Recall	64.6% (295/443)	77.3% (180/233)	60.8% (135/222)	87.4% (415/475)	73.0%
F-measure	0.59	0.59	0.57	0.68	0.61

Table 4 Results of identifying the portions of slides in lecture contents.

	Lecture contents by our method				
	L-W	L-X	L-Y	L-Z	Average
Precision	71.3% (196/275)	60.3% (193/320)	63.4% (716/1130)	69.6% (400/575)	66.2%
Recall	53.7% (196/365)	70.7% (193/273)	81.9% (716/874)	82.0% (400/488)	72.1%
F-measure	0.61	0.65	0.71	0.75	0.69
	Lecture contents by the levels contain the given keywords with their AP levels				
	L-W	L-X	L-Y	L-Z	Average
Precision	65.3% (261/400)	50.0% (233/466)	52.3% (792/1513)	56.8% (420/740)	56.1%
Recall	71.5% (291/365)	85.3% (233/273)	90.6% (792/874)	86.1% (420/488)	83.4%
F-measure	0.68	0.63	0.66	0.68	0.67

had captured. In this study, we evaluated the validity of the rules for identifying portions of slides in terms of the query keywords, using precision<sup>\*9</sup>, recall<sup>\*10</sup>, and F-measure<sup>\*11</sup> to compare the results obtained by our method with those obtained from participants who gave correct answers for each academic presentation and in each lecture explaining different topics. In addition, we compared the portions obtained by our method and the portions of sentences containing the given keywords on indent levels with their anteroposterior (AP) levels.

The results of the experimental identification of portions of academic and lecture presentations are listed in Tables 3 and 4,

<sup>\*9</sup> Precision =  $\frac{\text{Number of correct answers of the portions extracted by our method}}{\text{Total number of the portions extracted by our method}}$

<sup>\*10</sup> Recall =  $\frac{\text{Number of correct answers of the portions extracted by our method}}{\text{Total number of correct answers by participants}}$

<sup>\*11</sup> F-measure =  $\frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$

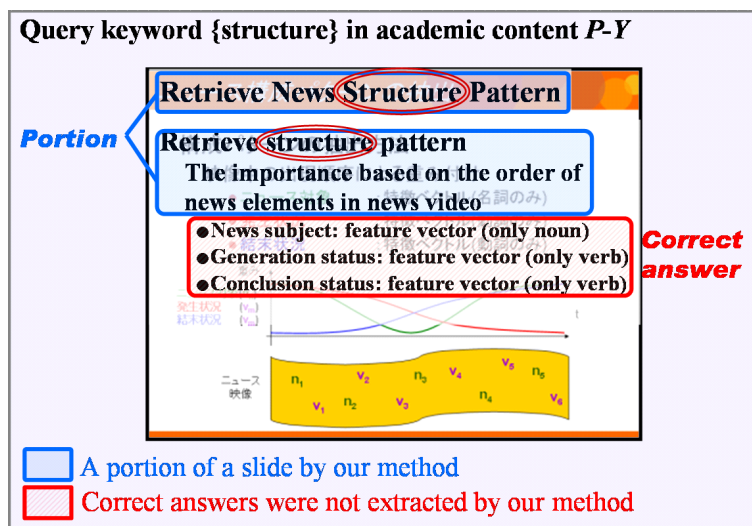


Fig. 11 Inadequate identification of a portion of a slide.

and can be explained as follows:

- The average F-measures for this experiment on academic and lecture presentations look similar. However, the average precision and recall of the lecture presentations were both higher than those for the academic presentations. We therefore concluded that it is difficult to understand the slides used in academic presentations that require some level of expertise, and we used WordNet [9], [10], which does not contain all concepts related to some general words. For example, a slide with the query keyword “structure” was used to identify portions of it in presentation *P-Y* (see Fig. 11). Sentence levels containing “news subject,” “generation status,” and “conclusion status” were correctly related to “news structure pattern” by participants. Our method, however, could not determine these keywords, as WordNet [9], [10] does not recognize “subject” or “status” as having a part-of relation with “structure.”
- The average precision of all experimental portions from academic or lecture presentations was low; our method extracted a much greater number of portions than those for which participants concurred. We believe that when determining correct answers, the participants did not consider slide titles or figures in slides in terms of the given keywords when our method was used.
- Comparing the results of the two methods, the average precision and average F-measure of our method were both higher than those of the other method. Although the results of the two methods look similar, the other method did not extract some portions containing sentences in slides that explained the given keywords, and some sentences on the AP levels were extracted which were not related to the given keyword.

This experiment confirmed that our method can extract the appropriate portions of slides, using semantic relations between keywords and the document structure of indents. However, we want to use an enhanced method for extracting mathematical formulas related to the given keywords. Furthermore, we should consider how to identify the keywords at different levels in figures or tables to improve performance in this experiment. In general,

Table 5 Results of generating snippets from academic contents.

	Academic contents by our method				Average
	<i>P-W</i>	<i>P-X</i>	<i>P-Y</i>	<i>P-Z</i>	
Precision	68.6% (175/255)	62.8% (76/125)	62.1% (59/95)	80.0% (108/135)	68.4%
Recall	66.0% (175/265)	67.0% (76/114)	57.0% (59/106)	66.7% (108/162)	64.2%
F-measure	0.67	0.64	0.60	0.73	0.67

Table 6 Results of generating snippets from lecture contents.

	Lecture contents by our method				Average
	<i>L-W</i>	<i>L-X</i>	<i>L-Y</i>	<i>L-Z</i>	
Precision	67.3% (175/260)	69.2% (229/331)	72.8% (732/1005)	74.7% (396/530)	71.0%
Recall	63.4% (175/276)	69.2% (229/331)	73.2% (732/1000)	66.6% (396/595)	68.1%
F-measure	0.65	0.69	0.73	0.70	0.70

we may also use the conceptual descriptions on the Wikipedia website<sup>\*12</sup>, an encyclopedia providing a vast amount of structured world knowledge, to build a large ontology. Therefore, we can improve the accuracy of our method for identifying portions of slides by ceasing to use WordNet [9], [10], and instead using domain-specific dictionaries for technical terms, or Wikipedia for general words.

### 5.5 Experiment 3: Validity of Generating Snippets

This experiment aimed to verify whether the proposed method is useful for generation of snippets for slides. We showed the participants 87 snippets, composed of portions of slides pertaining to the given keywords from the experimental dataset used in Experiments 1 and 2. Five participants took part in this experiment; the snippets presented a detailed explanation of the given keywords in order of the relevant portions in the slides. A correct answer was defined as three or more participants describing snippets of the focused slides with other slides as correct.

The results are shown in Tables 5 and 6; the experimental results were as follows:

- The results depended on those from Experiments 1 and 2.

\*12 <http://www.wikipedia.org/>



**Table 7** Experimental results of the efficacy of browsing snippets.

Dataset	Browsing slides only	Browsing slides with their snippets
Expertise in <i>P-Y</i>	14/24	21/24
Prior knowledge in <i>L-Z</i>	20/24	18/24
Total	34/48	39/48

However, in Experiment 1, we did not evaluate the determination of the relationship types in lecture content. For this experiment, the results for the academic and lecture content look similar. As in Experiment 2, the average precision and recall of the lecture presentations were both higher than those of the academic content; we concluded that there was no difference in the snippet generation between the slide relationships used in the academic and lecture content.

- The average recall of all experimental snippets from the academic and lecture presentations was low. When our method was used in Experiment 2, many of the correct answers were found to contain the sentences on indents in portions not extracted by our method. A snippet consists of portions of the focused slide and depends on identification of these portions, which is based on our method using WordNet [9], [10], and so did not determine that some keywords have semantic relationships between them. This was one of the reasons why the recall was low. Therefore, these portions for generating snippets also need to be considered.
- The average precision of all experimental snippets from the academic and lecture presentations was high. The results indicate that our method can generate appropriate snippets of relevant portions of slides based on the relationships between these slides, and the method can then be successfully applied to support browsing slide retrieval by generating snippets at the conceptual level.
- A few experimental snippets identified portions that did not include detailed information related to the focused slides; that is, relationships did not exist between them. In addition, many of the relevant portions were not strongly related to the portion of the focused slide, which may have reduced the precision.

This experiment showed that our method can generate snippets of relevant portions of related slides via the query, by effectively using the relationships between the slides. The results of this experiment suggest that we need to improve the determination of the snippet-generation algorithm by using the relationships between slides, and extracting the portions of slides relevant to the query. Our method used WordNet [9], [10], which will have had a bearing on the determination of the relationships between slides, and the identification of the portions of slides, due to the shortcomings already mentioned. Therefore, we plan to use domain-specific dictionaries or Wikipedia for extracting semantic relationships between keywords in the future work.

#### 5.6 Experiment 4: Efficacy of Browsing Snippets

In this experiment, we verified how the proposed method can help users to browse by introducing snippets. When users browse slides containing information, the snippets presented by our system let users easily grasp the context of the focused slides in terms

of the given keywords. We conducted this experiment with 15 participants, using four given keywords for 17 slide pages taken from two actual presentations: the academic content in *P-Y* contains seven slide pages, providing a level of expertise in Informatics that is important for the participants, and lecture material in *L-Z* containing 22 slide pages, providing basic knowledge in Informatics that is easy to understand for the participants. For evaluation purposes, we first prepared correct answers by asking three students which slide had the most detailed information related to each given keyword in each presentation from the experimental dataset. We defined a correct answer as when two or three students identified the same slide. Secondly, we provided two retrieval results for each given keyword using (a) the conventional method, where slides are retrieved by matching keywords, and (b) our method, where the corresponding snippets are generated by our system.

After providing these two retrieval results to the 12 students who did not take part in preparing the correct answers, we asked two questions in two steps as follows:

**Step 1.** Presenting the slides retrieved by method (a).

*Q1:* Which slide do you think provides the most detailed information related to the given keyword in these retrieval results? Please write your answer as the slide number and the reason for your selection.

**Step 2.** Presenting the retrieval results for method (b), including snippets.

*Q2:* When you browsed the snippets for the slides presented in *Q1*, did you change your answer to *Q1*? If so, please write the changed slide number and your reason for changing. If not, please give the reason why you did not change it.

We analyzed these answers, and the results are shown in **Table 7**. The vertical columns show how many correct answers were given when browsing the slides only, and how many correct answers were given when the snippets were also given to the participants. The horizontal rows show the breakdown of correct answers by knowledge levels required for the presentations. The experimental results are as follows:

- The total number of correct answers from browsing slides with their snippets was more than that when browsing slides only. Therefore, we believe that users browsing slides with their associated snippets can grasp the context of the focused slides, in relation to the given keywords, more easily.
- Browsing slides with their snippets provided more correct answers than browsing the slides only in presentation *P-Y*, and browsing slides with their snippets provided more correct answers in *P-Y* than in *L-Z*. *P-Y* provides expertise that is difficult for participants to understand, while *L-Z* provides knowledge that should be easily understood by participants. We confirmed that snippets are more useful when users browse slides containing a higher level of knowledge,

rather than when they browse slides containing information they are already aware of.

- In the *L-Z* dataset, there were fewer correct answers when browsing slides with their snippets than when browsing only the slides. *L-Z* provides prior knowledge that is easily understood by participants, so they are able to select many correct answers by browsing slides only. In addition, our proposed snippet-generation model only has three layers and does not consider the relevance of the related slides; thus, two participants were a little confused about the snippets for slides. However, for a majority of participants, we confirmed that our snippet-generation model is helpful for users browsing slides with their snippets.
- Our snippet-generation method is based on the relationships between slides, and works by identifying relevant portions of the focused slides. We concluded that a few generated snippets have the effect of determining the relationships between slides.

This experiment showed that our method of browsing slides with snippets is more useful than browsing slides only. In particular, our snippet-generation method is helpful for browsing slides containing higher levels of expertise alongside snippets.

In this paper, we evaluated our method by conducting Experiments 1–4, using presentation content made in PowerPoint and containing a layer structure (the levels of indentation) in the slides. We confirmed that our method is useful by satisfying certain criteria related to levels of indentation that are used to structure content in slides. Our proposed method does not only focus on presentation content made using PowerPoint, but can also be applied to a variety of other important presentation formats, such as Apple Keynote<sup>\*13</sup>, Google Docs<sup>\*14</sup>, and Prezi<sup>\*15</sup>. This is because these presentation formats contain a layer structure, in common with PowerPoint. We consider our proposed method to be applicable to a variety of presentation software, and we plan to evaluate our method with other presentation formats in the future work. Moreover, we evaluated our method by using WordNet [9], [10] to extract the semantic relations between keywords, and experimental results suggested that we can improve the accuracy of our method by using domain-specific dictionaries for technical terms in academic content, or Wikipedia for general words in presentations.

As mentioned above, we are also aware of the limitation of our method in not focusing on visual effects in slides. At the present time, authors often focus on visual effects that are easily understandable, and more attractive than slides with simple text. We do not currently use font or visual information, but it would not be difficult to improve our method by considering such data. Future developments to this method could also consider visual elements of figures, and the color distribution and animation occurrence in slides, as we can acquire this information by analyzing XML files from the various presentation formats. Furthermore, our method can be extended to consider the document structure not only in the slides, but also in associated presentation data. Finally, it is

possible to treat retrieval units for other applications to use our proposed system.

## 6. Concluding Remarks

In this paper, we proposed a snippet-generation method to support the browsing of slides based on the relationships between the slides. We described in detail how to determine the relationships between slides, using a unique platform to generate snippets for slides by analyzing the semantic relations between keywords and the document structure based on indents. In particular, by focusing on *detailed* and *generalized* slides at the conceptual level in presentations, we successfully supported users browsing slides containing higher-level expertise in our experiments, by providing snippets with the slides.

In the future work, we plan to improve the interface of the prototype system. Our method can enhance retrieval techniques if a user proposes a query including two or more keywords. Relationships between the keywords in the query need to be determined, in order to retrieve the user's desired slides by analyzing the relevance of the queried keywords. Furthermore, the results of the experiments suggested that we should use a large ontology construction, such as domain-specific dictionaries and Wikipedia, to extract the semantic relations between keywords. In addition, when presentations contain more visual elements, such as figures and videos, which do not contain the indentation levels found in text, we should consider the layout and captions of figures, the size or color information of the font, and information pertaining to the videos, by analyzing these elements of the slides. It is necessary to use these visual features to determine the relationships between slides. Moreover, we need to consider an adequate, or correct, size for snippets, so that they do not contain too much information.

We should enhance our snippet-generating algorithm to consider the relevance of the related slides to the focused slides, so as to generate snippets of different technical levels to support users of varying knowledge levels. Besides, an important question may arise—how much context is enough for understanding search results? This challenging question leads us to the quantitative study of context as another promising research direction.

## Reference

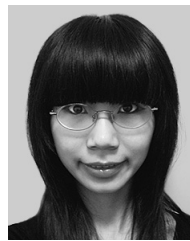
- [1] Ando, M. and Ueno, M.: Analysis of the Advantages of Using Tablet PC in e-Learning, *Proc. 10th IEEE International Conference on Advanced Learning Technologies (ICALT 2010)*, pp.122–124 (2010).
- [2] Chen, H., Wu, S., Song, C., Yu, X. and Luo, Y.: Analysis and Prospect on the E-Learning System Instructional Design, *Proc. International Symposium on Computer Science and Engineering (ISISE 2008)*, Vol.2, pp.303–306 (2008).
- [3] Chen, H., Wu, S., Song, C., Zhan, J., Chen, J. and Kang, D.: E-Learning System Model Construction Based Constructivism, *Proc. 5th International Joint Conference on INC, IMS and IDC (NCM 2009)*, pp.1165–1169 (2009).
- [4] Clarke, S.C. and Willet, P.: Estimating the Recall Performance of Web Search Engines, *Proc. Aslib*, Vol.49, No.7, pp.184–189 (1997).
- [5] DBSJ Archives, available from (<http://www.dbsj.org/Japanese/Archives/archivesIndex.html>).
- [6] edubase, available from (<http://edubase.jp/index/>).
- [7] Greene, S., Marchionini, G., Plaisant, C. and Shneiderman, B.: Previews and Overviews in Digital Libraries: Designing Surrogates to Support Visual Information Seeking, *Journal of the American Society of Information Science*, Vol.51, No.4, pp.380–393 (2000).
- [8] HandsOut, available from (<http://handsout.jp/>).

\*13 <http://www.apple.com/apps/keynote/>

\*14 <https://docs.google.com/>

\*15 <http://prezi.com/>

- [9] Isahara, H., Bond, F., Uchimoto, K., Utiyama, M. and Kanzaki, K.: Development of the Japanese WordNet, *Proc. 6th International Language Resources and Evaluation (LREC 2008)*, pp.2420–2423 (2008).
- [10] Japanese WordNet, available from (<http://nlpwww.nict.go.jp/wn-ja/>).
- [11] Kishi, T., Tsukada, H. and Nojima, E.: Relation Analysis of Presence of Note-Taking and Test Score After the Lecture [in Japanese], *Japan Society for Educational Technology*, Vol.28, pp.265–268 (2005).
- [12] Kitayama, D., Otani, A. and Sumiya, K.: An Extracting Method of Semantic Relations between Scenes for Presentation Contents [in Japanese], *IPJS Transaction on Database (TOD)*, Vol.2, No.2, pp.71–85 (2009).
- [13] Kobayashi, T., Nakano, W., Yokota, H., Shinoda, K. and Furui, S.: Presentation Scene Retrieval Exploiting Features in Videos Including Pointing and Speech Information, *Proc. International Symposium on Large-scale Knowledge Resources (LKR 2007)*, pp.95–100 (2007).
- [14] Le, H.H., Lertrusdachakul, T., Watanabe, T. and Yokoda, H.: Automatic Digest Generation by Extracting Important Scenes from the Content of Presentations, *Proc. 19th International Conference on Database and Expert Systems Application (DEXA 2008)*, pp.590–594 (2008).
- [15] Lecture Archives from Aoyama Gakuin University, available from ([http://homepage3.nifty.com/~terao/lecture/aoyama/lec\\_aoyama\\_top.html](http://homepage3.nifty.com/~terao/lecture/aoyama/lec_aoyama_top.html)).
- [16] Leung, Y.K. and Aerley, M.D.: A Review and Taxonomy of Distortion-Oriented Presentation Techniques, *ACM Trans. on Computer-Human Interaction*, Vol.1, No.2, pp.126–160 (1994).
- [17] Luhn, H.P.: Keyword-in-Context Index for Technical Literature (KWIC Index), *AMERICAN DOCUMENTATION*, Vol.11, No.4, pp.288–295 (1960).
- [18] MeCab, available from (<http://mecab.sourceforge.net/>).
- [19] Miller, G.A.: WordNet: A Lexical Database for English, *Comm. ACM (CACM)*, Vol.38, No.11, pp.39–41 (1995).
- [20] Miller, G.A., Beckwith, R., Fellbaum, C., Gross, D. and Miller, K.J.: Introduction to WordNet: An On-line Lexical Database, *International Journal of Lexicography*, Vol.3, No.4, pp.235–244 (1990).
- [21] MPMeister, available from (<http://www.rioh.co.jp/mpmeister/>).
- [22] Pattanasri, N. and Tanaka, K.: Construction of an Entailment Ontology for Enhancing Comprehension of Search Results Inside e-Learning Materials, *Proc. 6th International Conference on Creating, Connecting and Collaborating through Computing (c5 2008)*, pp.124–130 (2008).
- [23] NAIST Digital Library, available from ([http://library.naist.jp/library/archive\\_top/index.html](http://library.naist.jp/library/archive_top/index.html)).
- [24] News of Kyoto University, available from ([http://www.kyoto-u.ac.jp/ja/news\\_data/h/h1/news7/2010/100604\\_1.htm](http://www.kyoto-u.ac.jp/ja/news_data/h/h1/news7/2010/100604_1.htm)).
- [25] Nozaki, H., Yoshihashi, A., Umeda, K. and Ejima, T.: The Effects of Writing the Text on Readers' Comprehension [in Japanese], *Japan Society for Educational Technology*, Vol.29, pp.49–52 (2006).
- [26] Ohshima, H., Nakamura, S. and Tanaka, K.: SlothLib: A Programming Library for Research on Web Search [in Japanese], *The Database Society of Japan (DBSJ Letters)*, Vol.6, No.1, pp.113–116 (2007).
- [27] Okamoto, H., Kobayashi, T. and Yokota, H.: Presentation Retrieval Method Considering the Scope of Targets and Outputs, *Proc. International Workshop on Challenges in Web Information Retrieval and Integration*, pp.46–53 (2005).
- [28] SlideShare, available from (<http://www.slideshare.net/>).
- [29] SlothLib, available from (<http://www.dl.kuis.kyoto-u.ac.jp/SlothLibWiki/>).
- [30] Wang, Y., Kitayama, D., Lee, R. and Sumiya, K.: Automatic Generation of Learning Channels by Using Semantic Relations among Lecture Slides and Recorded Videos for Self-Learning Systems, *Proc. 11th IEEE International Symposium on Multimedia (ISM 2009)*, pp.275–280 (2009).
- [31] WordNet, available from (<http://wordnet.princeton.edu/>).
- [32] Wu, S., Song, C., Chen, H. and Zhan, J.: E-Learning Teaching method Research and Ideas, *Proc. International Conference on Web Information Systems and Mining (WISM 2010)*, Vol.1, pp.140–143 (2010).
- [33] Yokota, H., Kobayashi, T., Okamoto, H. and Nakano, W.: Unified Contents Retrieval from an Academic Repository, *Proc. International Symposium on Large-scale Knowledge Resources (LKR 2006)*, pp.41–46 (2006).



**Yuanyuan Wang** is a Ph.D. student in Graduate School of Human Science and Environment, University of Hyogo, Japan. She received her Master's degree from Graduate School of Human Science and Environment, University of Hyogo, Japan, in 2011. Her research interests include e-learning systems, structural analysis and multimedia database. She is a student member of Information Processing Society of Japan and the Database Society of Japan.



**Kazutoshi Sumiya** is a Professor at School of Human Science and Environment, University of Hyogo. He received his Ph.D. degree in engineering in 1998 from Graduate School of Science and Technology, Kobe University. His research interests include multimedia database and data broadcasting. He is a member of IEEE Computer Society, ACM, the Institute of Image Information and Television Engineers, IPSJ, the Database Society of Japan and IEICE.