一般社団法人 電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
IEICE Technical Report

A Word Cloud Visualization for Comparing Presentation Contents

Yuanyuan WANG† and Kazutoshi SUMIYA‡

† Graduate School of Human Science and Environment, University of Hyogo

1-1-12 Shinzaike-honcho, Himeji, Hyogo, 670-0092 Japan

E-mail: † {ne11u001@st, sumiya}shse.u-hyogo.ac.jp

Abstract iTunes U and SlideShare are a crucial platform for improving education; students are able to search various presentations through Web. However, self-learners still need support to decide which files are worth learning, because most of search results are similar; it can be difficult to identify differences in them. We consider presentation flow is very important for comparing the differences in the results. We developed a quick browsing tool to help users effectively compare presentations for their specific learning needs. Our tool provides a word cloud visualization that summarizes information to help the users visually understand the context of each presentation. Words important to the "presentation context," that is, the relevant information on the slides, is first extracted based on components of the presentation (i.e., intra-slide and inter-slide structures). Our word cloud visualization shows the words are interactively presented with visual effects in presentations.

Keyword Presentation Slides, Quick Browsing, Word Cloud Visualization, E-learning

1. Introduction

1.1. Motivation

Presentations (e.g., PowerPoint, Keynote) are now one of modern tools for educational purposes. A huge amount of slide-based lecture material, often used in actual classes at universities or other educational institutions, is freely shared on Web such as iTunes U[1] and SlideShare[2]. Thus, not only students who missed a lecture or conference, but also anyone interested in the topic can study the presentation on their own. Therefore, techniques are in demand that will efficiently find one or more appropriate presentations with content worth learning.

Although many techniques for searching and recommending presentations have been proposed, some problems remain from the viewpoint of understandability for users browsing search engine results. One problem is a search engine doesn't consider context when matching user query words within presentation, leading to a large number of search results are similar. Another problem is the difficulty of general quick browsing for visualizing results of presentations, that is, when browsing slide titles only, users cannot grasp specifics of the content. In addition, important words of slides that summarize information simply based on TF-IDF scores can destroy the relevant information between slides and decrease the relevance of words in slides to the overall context. It difficult to understand the context of each presentation.

As depicted in Fig. 1, we present a quick browsing tool that considered the context of presentation for supporting users to decide presentations effectively. It can be implemented by 1) generates word clouds for each slide by considering the words within the context of the presentation (i.e., the intra-slide and inter-slide structures); and 2) determines transitions between the generated word clouds based on relationships between the words in serial slides to lead a word cloud visualization for identifying differences in presentations. In order to achieve our goal, we derive the intra-slide structure that slide structure by focusing on bullet points in the slide text, and determine the inter-slide structure that links between slides by considering words that appear at different bullet points in other slides. For example, 'Keys' appears in the body of text in a slide entitled 'Relational model', which related to the slide entitled 'Relational database' and 'Relational model' appears in the body of text in it. 'Keys' is a title of other slide that is related to the slide entitled 'Relational model'.

1.2. Our Method

In this paper, we define presentation context to mean the context in one presentation, represented by relevant information on the slide and allowing for the relevant information from the rest of the presentation that is not included on the slide. We define two types of presentation context for a slide: link context and structural context, based on the links between slides and slide structure,

[1] http://www.apple.com/jp/education/itunes-u/
[2] http://www.slideshare.net/

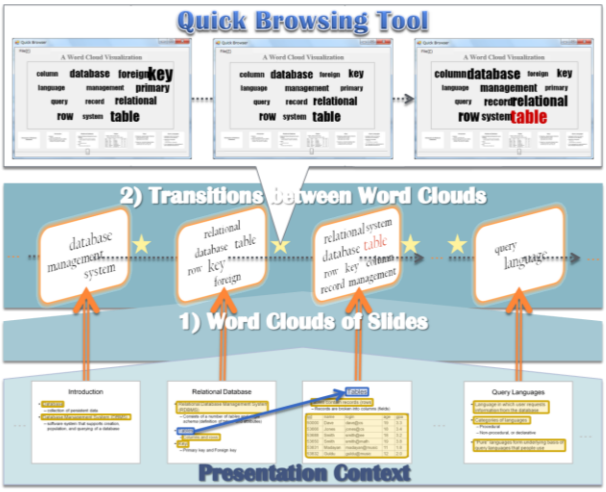

Fig. 1 Conceptual diagram of quick browsing tool

respectively. Using presentation context, we can generate word clouds of slides by weighting words. There are two concepts that are particularly helpful when quick browsing presentations for finding differences in them:

- Presentation flow: link or break [1]. Often, presentations are formed of a chain of slides such one slide links to the next one. Sometimes, the slide will move from the point in a previous slide to a different point. In this case, there is a break between them.
- Highlight points: semantics. This occurs when one slide describes a point from a previous slide in detail.

In this case, our approach presents words interactively from one slide to another as a streaming word cloud reflecting the flow of points in the slides, helping users to compare presentations easily and effectively.

In this remainder of this paper, Section 2 reviews related work, and Section 3 describes how to determine presentation contexts. Section 4 presents word cloud generation and transitions determination between the word clouds, and Section 5 introduces a prototype application for comparing presentations. Finally, Section 6 concludes this paper with a brief description of future work.

2. Related Work

Our work is directly related to the research efforts in two areas: text analytics and information visualization. We also review the effects of word cloud visualization on browsing tasks. In the area of text analytics, there are two main techniques: sentence-based and word-based text summarization. Sentenced-based approaches identify the most salient sentences in a document [2]. For example, Murai and Ushiama [3] proposed a review-based recommendation of attractive sentences in a novel. However, it may be time consuming for users to read several sentences per document especially when handling a large number of documents. Alternatively, word-based methods summarize documents by topics, each of which is characterized by a set of words [4]. Our quick browsing tool is built on the latter method, but it focuses on enhancing the summarization results through word and visualization. Moreover, we provide users a word cloud visualization for comparing presentations.

In the area of information visualization, researchers have developed various visualization approaches to text analysis: metadata-based and content-based text visualization. Metadata-based text visualization focuses on visualizing the metadata of text documents. In email or news articles analysis for instance, metadata-based text visualization can use a time-based visualization to explain text summarization results [5], and a visual topic analysis system to help users explore and understand topic evolutions for news articles [6]. For content-based text visualization, Viegas et al. [7] used Themail to visualize words based on TF-IDF scores in an email collection. In this toolkit, the content evolution is visually encoded by a set of word lists at different time points. In addition, several approaches have been suggested for representing content changes using tag frequencies [8]. Similarly, Strobelt et al. [9] used a mixture of images and TF-IDF-based keywords to create a compact visualization of a document. In contrast, others have concentrated on representing text content at the word or phrase level, including TextArc (www.textarc.org), WordTree, and Phrase Net. Our work focuses on visualizing presentations by generating word clouds of slides, and presenting transitions between the word clouds.

3. Determination of Presentation Contexts

We determine two types of presentation context: link context and structural context, based on the links between slides and slide structure, respectively. We define the slide whose presentation context is discussed to be the target slide. We construct the slide structure based on the bullet points in the slide. The slide title is 1st level; the 1st item of text within the slide body is 2nd level, and the depth of the sub-items increases with level (3rd, 4th level, etc.).

3.1. Determination of Link Context

The link context for a target slide consists of links and anchors (as hyperlinks) related to the bodies with titles of other slides. They refer to words in the title of the target slide and titles of other slides that contain words in the body of the target slide. They also indicate from what type

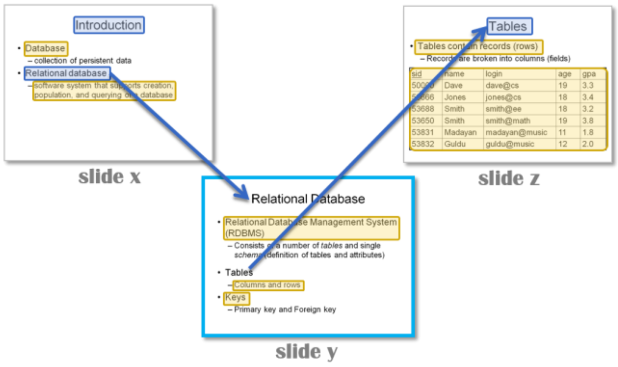

Fig. 2 Presentation context for slide *y*

of content the target slide is referred. We extract the link context of the target slide by finding the same words at different levels in the target and other slides.

For a given bag of words *M* in the title and a given bag of words *N* in a level in the body of the target slide, words in the titles and levels in the body of other slides are extracted: …, T_2, B_2, T_1, B_1. Here, T_i is the title of slide *i* and B_i is a level of the body of slide *i*. If B_i corresponds to *M*, B_i can be considered as a link anchor. Then, B_i links to the target slide that the words in B_i and its slide title T_i belong to the link context for the target slide, while the words in B_i are similar to that in *M*. This is calculated using the Simpson similarity coefficient, as $Sim(B_i,M)=|B_i \cap M|/min(|B_i|,|M|)$. When $Sim(B_i,M)$ exceeds a predefined threshold, the words in B_i and *M* are similar. Meanwhile, if *N* corresponds to T_i, *N* can be considered as a link anchor. Then, *N* links to the slide titled T_i in that the words in T_i belong to the link context for the target slide, while the words in *N* are similar to that in T_i.

In Fig. 2, the link context for slide *y* shows that slide *y* explains about "Relational Database," which is referred to on slide *x* as a subheading of 'Introduction,' and the subheading 'Tables' in slide *y* is described in slide *z*.

3.2. Determination of Structural Context

The structural context for the target slide consists of lower, current, and upper levels of the target levels corresponding to the link context, and lower, current, and upper levels of the link context in other slides based on slide structure. When the target slide doesn't have a link context, we take the title of the target slide as the target level, and then we extract the structural context for the target slide that consists of the levels below the title in it.

For a given bag of words *N* is a level in the body of the target slide, words in the lower, current, and upper level of *N* are extracted: …, l_3, l_2, l_1. Here, l_j represents a bag of words at a particular level *j*. When l_j and *N* are the current level in the target slide, the words at level l_{j+1} are at the lower level of *N* and l_{j-1} are at the upper level of *N*. Therefore, the lower, current, and upper levels of the link context in other slides are extracted in the same way.

When the target level in the target slide corresponds to the link context in more than one slide, we just extract the link context of the slide nearest to the target slide. The link context and structural context are extracted within a minimal range of surrounding information, containing enough words to characterize the presentation context. Therefore, the presentation context expresses presentation flow and highlight points well.

Fig. 2 illustrates the structural context for slide *y*, where 'Tables' is related to 'RDBMS,' 'Keys,' and 'Columns and rows' at the upper, current, and lower levels in slide *y* and 'Tables' includes a link at a lower level, 'Tables contain records (rows),' in slide *z*; 'Relational Database' includes a link to a lower level at 'RDBMS' in slide *y*, and 'Relational Database' is related to current and lower levels at 'Database' and 'Software system' in slide *x*.

4. A Word Cloud Visualization

4.1. Generation of Word Clouds of Slides

Word clouds of slide are generated by weighting the words the presentation context to determine font size. For each type of presentation context, we calculate the degree of the words that 1) appear close to the target slide and 2) appear frequently near the target slide.

Let us consider each word of target slide *S* as a relevant object, denoted by *o*. The degree of *o* for the presentation context *P*(*S*) is defined as follow:

$$W(o,P(S)) = \frac{density(o,P(S))}{dist(o,S)} \quad (1)$$

Here, $density(o,P(S))$ is the density of *o* for *P*(*S*), and $dist(o,S)$ is the distance between *o* and *S*. Intuitively, $density(o,P(S))$ means how densely the same word as *o* appears in *P*(*S*). If the same words as *o* appear frequently in *P*(*S*) but less frequently in other presentation contexts, $density(o,P(S))$ becomes large. Suppose that *S* is the *k*-th slide among all slides (the target slide). The density of *o* in $P(S_k)$ is calculated as follows:

$$density(o,P(S_k)) = \frac{N_{A(o,P(S_k))}}{N_{A(o,U)}} \quad (2)$$

where $A(o,P(S_k))$ is a set of relevant objects representing the same word as the object *o* in the presentation context $P(S_k)$, and $A(o,U)$ is the set of relevant objects in the presentation context of all slides: $U=P(S_1) \cap P(S_2) \cap \ldots$. $N_{A(o,P(S_k))}$ and $N_{A(o,U)}$ represent the number of objects in $A(o,P(S_k))$ and $A(o,U)$, respectively. Because it is difficult

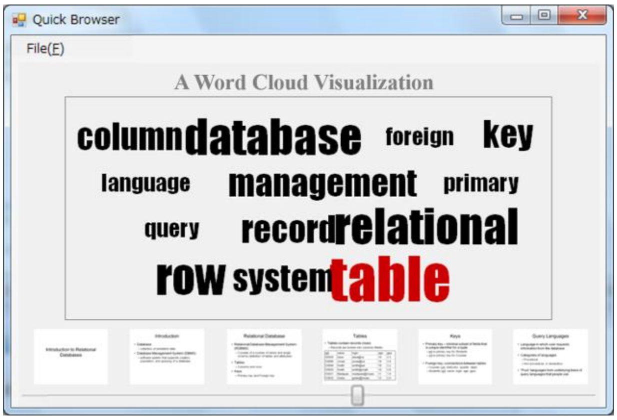

Fig. 3 Screenshot of quick browsing interface

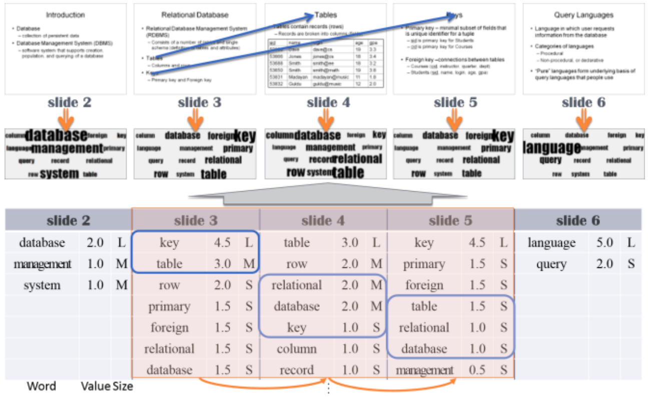

Fig. 4 Example of extracted words for word clouds of slides

to identify the set U due to mutual dependencies between the presentation contexts, we approximate U as the set of relevant objects of all slides.

$dist(o,S)$ indicates the strength of the associations between the relevant object o and the target slide S, and is defined for each type of presentation context as follows:

Distance in link context: The number of link relationships from S to o.

Distance in structural context: The number of parent, brother, and child nodes to be followed from the target levels in S to o.

We generate word clouds of slides based on the ratio of the degree of each word and the highest degree of the word in each slide. We also sort the words into three different font sizes as follows:

$$STag(c,P(S)) = \left\{ c \,\middle|\, \frac{W(c,P(S))}{W_{\max}(c,P(S))} \geq \theta_1, \frac{W(c,P(S))}{W_{\max}(c,P(S))} < \theta_2 \right\} \quad (3)$$

$$MTag(c,P(S)) = \left\{ c \,\middle|\, \frac{W(c,P(S))}{W_{\max}(c,P(S))} \geq \theta_2, \frac{W(c,P(S))}{W_{\max}(c,P(S))} < \theta_3 \right\} \quad (4)$$

$$LTag(c,P(S)) = \left\{ c \,\middle|\, \frac{W(c,P(S))}{W_{\max}(c,P(S))} \geq \theta_3 \right\} \quad (5)$$

In Eqs. (3), (4), and (5), $W(c, P(S))$ is the degree of c and $W_{\max}(P(S))$ is the highest degree of the word in S using Eqs. (1) and (2). Including too many words in each word cloud does not help users to browse them effectively, so we extract c in S such that the ratio of $W(c, P(S))$ and $W_{\max}(P(S))$ is greater than a threshold (i.e., *0.25*). $STag(c, P(S))$, $MTag(c, P(S))$, and $LTag(c, P(S))$ are the groups of weighted words to be displayed in small, medium, and large font size such that the ratios satisfy Eqs. (3), (4), and (5), respectively. In this paper, we empirically set the values of the thresholds to be $\theta_1=0.25$, $\theta_2=0.50$, and $\theta_3=0.75$. Although, in general, the word position is important for a word cloud, our word cloud instead places the word randomly so that the user is not biased to any specific terms based on their placement position.

4.2. Determination of Transitions between Word Clouds

Based upon the presentation contexts for slides, we present word clouds with visual effects that reflect relationships between words interactively. For this purpose, we use the relationships between words in the word clouds similar to the relationships between slides defined in our previous work [12], which fall into four types based on the presentation contexts for slides:

- ***Detailed* relationship**: titles of other slides belonging to the link context for the target slide. The other slides have more information about the link context than the target slide.
- ***Generalized* relationship**: bodies of other slides belonging to the link context for the target slide. The target slide contains the words about the link context in the outline given in the other slides.
- ***Parallel* relationship**: titles of other slides belonging to the link context for the current levels in the target slide, other slides are parallel with each other.
- ***Independent* relationship**: slides do not have a link context for each other.

To present a word cloud visualization, the transitions discussed here explain the kinds of visual effects, reflecting presentation flow or highlights. Presentation flow consists of many chains of serial slides such that each chain and each transfer switch between chains must be presented. For one chain of serial slides, *detailed*, *generalized* or *parallel* relationships exist between them. For a transfer switch between different chains of slides, *independent* relationship exists between them. Highlights are the words belonging to the link context of one chain in detail. The three types of transitions are as follows:

- ***Font size changes***: a *shrinking* or *expanding* effect is set between serial slides presenting one chain in the presentation flow. When the font sizes of the words

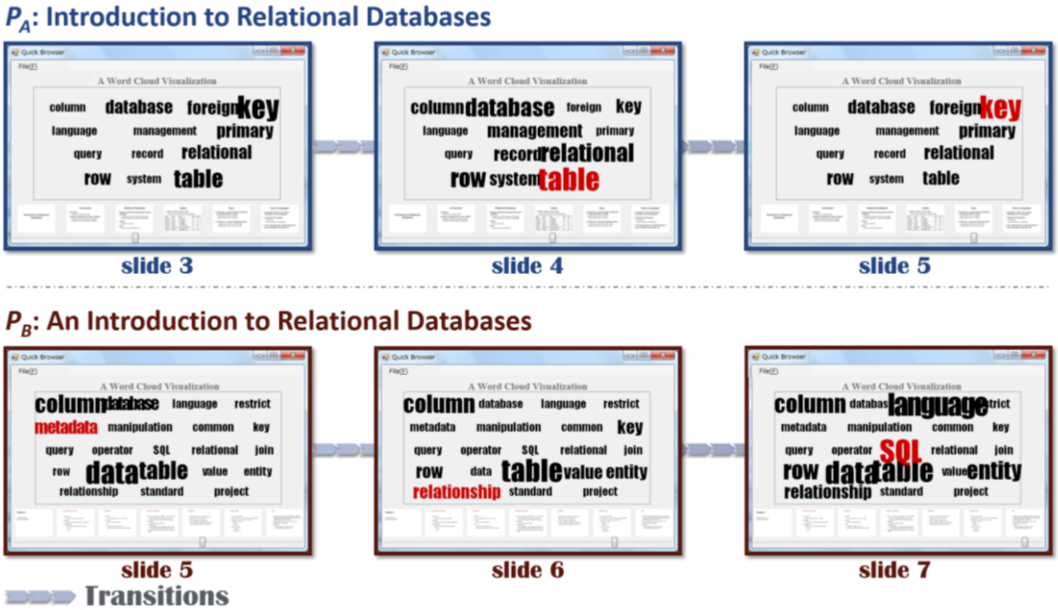

Fig. 5 Examples of transitions between word clouds of P_A and P_B

in the current word cloud are smaller than those in the previous one, a *shrinking* effect is implemented. When the font sizes of the words in the current word cloud are larger than those in the previous one, an *expanding* effect is applied after a *shrinking* effect. Users can easily understand that they are following a chain of slides where the words are mentioned.

- ***Color changes***: a *coloring* effect is set between serial slides to highlight detailed points in one chain of the presentation flow. For a current word cloud, the words belonging to the next word cloud, which are described in detail on the next slide, are drawn in catchy color. Users easily see that these words are highlighted in one chain. When the highlighted words in the current word cloud are not detailed in the next word cloud, these words are drawn the default color.
- ***Switching***: a *dissolve* effect is applied to a transfer switch between different chains in the presentation flow. The current word cloud disappears and the next word cloud appears gradually in its place. Users easily grasp that a transfer switch has occurred.

5. Application

5.1. Prototype System

In this paper, we built a novel word cloud visualization to support users to quickly identify differences in presentations by gaining a broad understanding of presentations (see Fig. 3). Users can specify any presentation from search results, and the browser presents all words from all slides with an initial font size (i.e., 20pt). When a user moves a seekbar to turn over slides, and the weighted words belonging the word clouds of slides with their font sizes (i.e., 30pt, 40pt, 50pt) are interactively presented with visual transitions. We plan to attempt to build other kinds of interfaces to express the presentation well.

5.2. Validity of Generated Word Clouds

We confirmed our word cloud generation method by using four presentations from our dataset, P_A, P_B, P_C, and P_D. In here, P_A[3] and P_B[4] are online lectures related to database; P_C and P_D are academic presentations from DEWS workshops for members of the society. We show an example of extracted weighted words with their values and determined sizes for generating word clouds of P_A named “Introduction to Relational Databases” (see Fig. 4).

In this example, slide *3* entitled “Relational Database” that ‘relational’ and ‘database’ are important in general. However, we considered the context of slide *3* that ‘key’ and ‘table’ have high value in slide *3*. In addition, for slides *4* and *5*, we can extract weighted words such as ‘relational’ and ‘database’ that are not included in slides *4* and *5*, but these words are related to them. Therefore, we considered that users can grasp the flow of slides *3* to *5* about ‘tables and keys in relational databases’ well.

5.3. Application Examples

When a user wants a presentation about ‘relational

[3] Portland State University: http://web.cecs.pdx.edu/~howe/cs410/lectures/Relational_Intro_1.ppt

[4] Atilim University: http://www.atilim.edu.tr/~mrehan/Chapter\%203.ppt

database overview: tables and keys,' he can specify any presentation for quick browsing from the results. An example of browsing the word cloud visualization of two presentations from our dataset, P_A and P_B, is shown in Fig. 5. In this case, the user browses the flow of slide *3* to slide *5* in P_A named "Introduction to Relational Databases," and the flow of slide *5* to slide *7* in P_B named "An Introduction to Relational Databases." When the user moves the seekbar to go from slide *3* to *5* in P_A, the font sizes of 'table,' 'database,' 'relational,' and 'row' are increased in slide *4*, and the font sizes of 'key,' 'foreign,' and 'primary' are increased in slide *5*. In particular, 'table' and 'key' are drawn in red in slides *4* and *5*, respectively. On the other hand, when the user moves the seekbar to go from slide *5* to *7* in P_B, the font sizes of 'table,' 'relationship,' and 'entity' are increased in slide *6*, and the font sizes of the words such 'SQL,' 'language,' and 'data' are increased in slide *7*. In particular, 'metadata,' 'relationship,' and 'SQL' are also drawn in red in slides *5*, *6*, and *7*, respectively.

In the case of P_A, we found 'table' and 'key' are core points in the flow of slides *3* to *5*, explains 'tables and keys in relational databases,' detailed relationships exist between 'table' in slides *3* and *4*, and 'key' in slides *3* and *5*. There is also a parallel relationship between 'table' and 'key' in slides *4* and *5*. For P_B, 'metadata,' 'relationship,' and 'SQL' are core points in the flow of slides *5* to *7*, explains 'characteristics of data in relational database,' and a parallel relationship exists among 'metadata,' 'relationship,' and 'SQL' in slides *5*, *6*, and *7*. Therefore, P_A is worth learning in that it better meets the user's needs. Although we confirmed that our proposed method enables a user effectively and easily select presentations with contents that meet his needs, we encountered difficulties when presentations (i.e., P_A and P_B) had a similar title.

We need to consider how best to present the differences in similar presentations (e.g., different topics or same topics with different context information, etc.). Additionally, we must also provide a word cloud visualization that compares presentations simultaneously with the differences clearly marked.

6. Concluding Remarks

In this paper, we proposed a word cloud visualization for comparing presentations that presents words interactively with visual effects to help a user visually understand the context of presentations. We described how presentation context can be determined from slide structure and the links between slides. In order to generate word clouds of slides, we extracted weighted words from presentation context, and then presented transitions that highlighted the relationships between slides. Finally, we confirmed our method with some application examples.

In the future, we plan to develop various other visual effects for presenting transitions to help users intuitively compare presentations. We have to evaluate the usability of our word cloud visualization to confirm that it can enable users to gain a broad understanding of presentations meet their needs from our collected 47 database lectures effectively and easily.

References

[1] Presentation Flow: Link Or Break, http://philpresents.wordpress.com/2011/11/23/presentation-flow-link-or-break/.

[2] D. Wang, T. Li, S. Zhu, and C. Ding, "Multi-document Summarization via Sentence-level Semantic Analysis and Symmetric Matrix Factorization," Proc. of the 31st Annual International ACM SIGIR Conference, pp.307-314, Toronto, Canada, Jul. 2008.

[3] S. Murai and T. Ushiama, "Review-Based Recommendation of Attractive Sentences in a Novel for Effective Browsing," International Journal of Knowledge and Web Intelligence, vol.3, no.1, pp.58-69, 2012.

[4] J. Kaye, A. Lillie, D. Jagdish, J. Walkup, R. Parada, and K. Mori, "Nokia Internet Pulse: A Long Term Deployment and Iteration of a Twitter Visualization," Proc. of International Conference on Human-Computer Interaction, pp.829-844, Austin, Texas, USA, May 2012.

[5] S. Liu, M. X. Zhou, S. Pan, Y. Song, W. Qian, W. Cai, and X. Lian, "TIARA: Interactive, Topic-Based Visual Text Summarization and Analysis," ACM Transactions on Intelligent Systems and Technology, vol.3, no.25, pp.1-28, 2012.

[6] W. Cui, S. Liu, L. Tan, C. Shi, Y. Song, Z. Gao, H. Qu, and X. Tong, "TextFlow: Towards Better Understanding of Evolving Topics in Text," IEEE Transactions on Visualization and Computer Graphics, vol.17, no.12, pp.2412-2421, Dec. 2011

[7] F. Viegas, S. Golder, and J. Donath, "Visualizing Email Content: Portraying Relationships from Conversational Histories," Proc. of the Conference on Human Factors in Computing Systems, pp.979-988, New York, NY, USA, Apr. 2006.

[8] W. Cui, Y. Wu, S. Liu, F. Wei, M. X. Zhou, and H. Qu, "Context Preserving, Dynamic Word Cloud Visualization," IEEE Computer Graphics and Applications, vol.30, no.6, pp.42-53, 2010.

[9] H. Strobelt, D. Oelke, C. Rohrdantz, A. Stoffel, A. D. Keim, and O. Deussen, "Document Cards: A Top Trumps Visualization for Documents," IEEE Transactions on Visualization and Computer Graphics, vol.15, no.6, pp.1145-1152, Oct. 2009.

[10] D. Kitayama, A. Otani and K. Sumiya, "An Extracting Method of Semantic Relations between Scenes for Presentation Contents," IPSJ Transaction on Database (TOD), vol.2, no.2, pp.71-85, Jun. 2009 (in Japanese).