Ogata, H. et al. (Eds.) (2015). Proceedings of the 23rd International Conference on Computers in Education. China: Asia-Pacific Society for Computers in Education

Automatic Summarization of Lecture Slides for Enhanced Student Preview

Atsushi SHIMADA*, Fumiya OKUBO, Chengjiu YIN, Hiroaki OGATA
Faculty of Arts and Science, Kyushu University, Japan
*atsushi@artsci.kyushu-u.ac.jp

Abstract: In this paper, we propose a novel method of summarizing lecture slides to enhance preview efficiency and improve students' understanding of the content. Students are often asked to prepare for a class by reading lecture materials. However, this does not always produce good results because the attention span of students is limited. We conducted a survey involving preview of lecture materials by more than 300 students and found that they want summarized materials to preview. Therefore, we developed an automatic summarization method to reduce the original preview materials to a summarized set. Our approach is based on the use of image processing and text processing to extract important pages from lecture materials, and then optimizing the selection of pages in accordance with a specified preview time. We applied the proposed summarization method to lecture slides. In our user study involving more than 300 students, we compared the relative effectiveness of the summarized slides and the original materials in terms of quiz scores, preview achievement ratio, and time spent previewing. We found that students who previewed the summarized slides achieved better scores on pre-lecture quizzes even though they spent less time previewing the material.

Keywords: Lecture slide summarization, preview of lecture slide, enhancing preview

1. Introduction

In discussing enhancements of learning processes, it is often argued that studying in advance for a class is very important in enabling students to understand the class narrative, to become familiar with important keywords, and to discover new terms and concepts. Some studies such as Beichner (1995) report that good preparation prior to lectures leads to improved student performance. In universities, students are often asked to prepare for their next class by reading a textbook, or previewing material. Hereafter, we use the term "preview" to denote any form of studying in advance and/or reading material provided by a teacher.

We conducted a survey of student previewing in our university. In total, 326 students answered the following questions:

Q1-1: How long do you usually spend previewing?
Q1-2: How many classes ask you to preview weekly?
Q1-3: What kind of material is preferable for preview? (multiple answers allowed)

As shown in Figure 1(a), about 85% of students spend less than 30 minutes, and about half of students spend less than 20 minutes previewing. More than 90% of students attend two or more classes in which they are weekly asked to preview (see Figure 1(b)). Based on these survey results, we have to assume that students have difficulty in previewing material adequately, since the material is often extensive, necessitating considerable preview time. Meanwhile, we asked the same survey for teachers with the same questions of Q1-1 and Q1-3 (see Figure 2). The previewing time desired by teachers is much longer than the answers given in response to Q1-1. About 60% of students indicated that they wanted preview material that is summarized, rather than the entire contents, meanwhile teachers preferred all the materials for previewing. Therefore, if a teacher can prepare not only the lecture material but also a summary, this would satisfy the students' demand. However, this imposes an enormous burden on a teacher.

Our study is motivated by the background outlined above, and we propose a method by which lecturers can automatically generate a summary of their lecture material. In our study, we focus on a lecture style in which a teacher uses lecture slides (e.g., PowerPoint). These days, this lecture style is

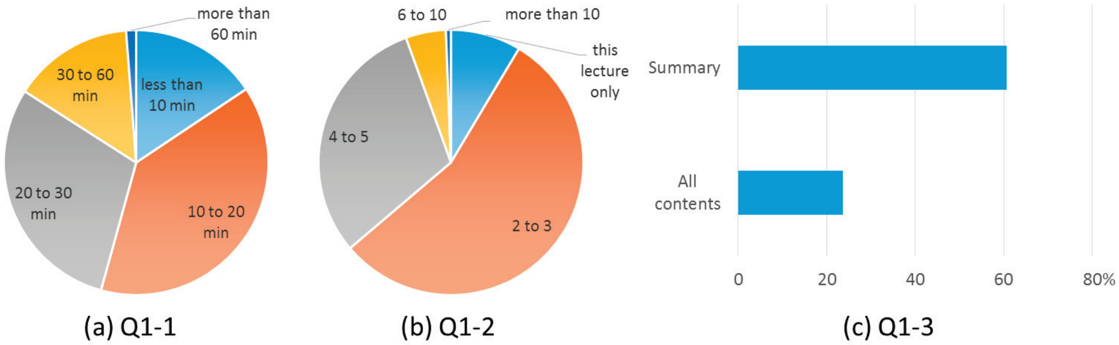

Figure 1. Student responses to survey questions 1-1, 1-2, and 1-3

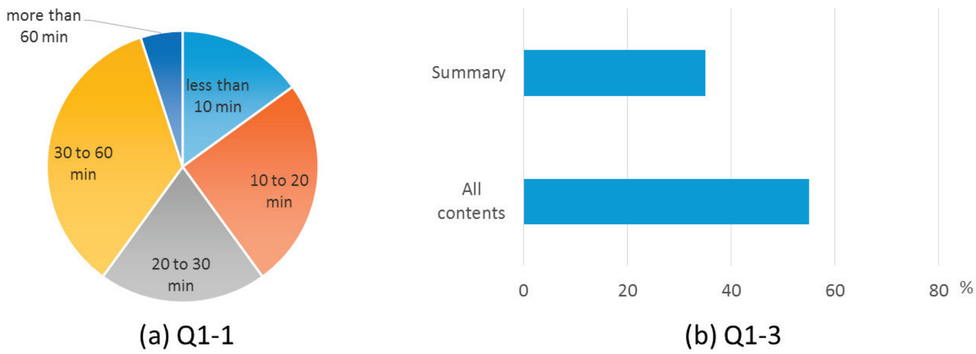

Figure 2. Teacher responses to survey questions 1-1 and 1-3

widely used, because most universities provide a projector and screen in each classroom. The proposed method enables lecturers to make a summary of the slides that they will use in their lecture, which allows students to complete their preview in a shorter time. Further, the preview time can be specified. For example, a 10-minute summary is automatically generated when a teacher specifies a 10-minute preview time. Image and text processing techniques are used to calculate an importance score for each slide page. In addition, an estimated preview time is allocated to each page in advance. Slide summarization is performed by selecting the appropriate slide pages to maximize the importance score within a given preview time.

By conducting experiments involving more than 300 students, we investigated the effectiveness of the proposed method. The students confirmed that the summarized slides were of reasonable length for preview. Further, we found that students achieved better quiz scores if they previewed the summarized slides. Details of our findings are provided in the Experimental Results section.

2. Related Work

The use of automatic summarization is often discussed in the research domains of video processing, speech processing, and document (text) processing. For example, video summarization techniques (Money 2008, Rajendra 2014) are used to make a short summary of news archives, user videos, and so on.

He (1999) applied a technique involving automatic creation of summaries to online audio/video presentations. Their technique exploits information in the audio signal, knowledge of slide transition points in the presentation, and information about access patterns of previous users. They reported that their computer-based summaries were well received by most study subjects.

Yun-Nung (2011) proposed an approach for spoken lecture summarization. In their approach, random walk is performed on a graph constructed using automatically extracted key terms and probabilistic latent semantic analysis. They applied their method to lecture documents to extract a summary of each document.

In contrast with the above summarization approaches, the target of summarization in our study is lecture slides. The proposed approach exploits slide content information such as text, images, and mathematical formulas to extract an importance score for each slide page. Consequently, a compact slide containing small number of pages are automatically generated as a summary.

In terms of automatic slide generation, Mathivanan (2009) and Sathiyamurthy (2012) proposed an approach to generate presentation slides from documents. However, our approach generates a summary set of slides from the original lecture slides themselves.

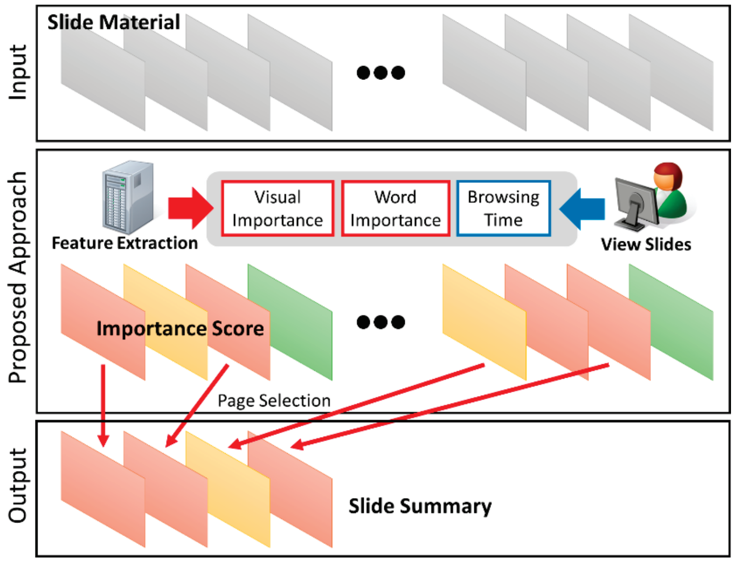

Figure 3. Overview of slide summarization

3. Slide Summarization

3.1 Overview

The purpose of slide summarization is to select a subset of pages that maximizes the importance of content under a given condition, namely, browsing time. For example, suppose that a set of slide pages requires T_{all} seconds for browsing. We would like to summarize these slide pages into a subset that can be browsed in T_{sub} ($T_{sub} < T_{all}$). To achieve this, several of the most important pages have to be selected without losing the overall narrative of the lecture. In our study, we assume that important pages display the following characteristics:

- Sufficient content to be worth browsing
- Unique content
- Keywords that appear frequently in a slide page
- Keywords that rarely appear throughout the set of slide pages.

The first assumption will select a page containing figures and/or tables which are useful to support understanding of the contents. The second one reduces redundant pages such as animations. The third and fourth characteristics locate important keywords that appear frequently in a given page, but rarely appear throughout the total set of slide pages. These characteristics are analyzed using a combination of image processing and text processing.

Figure 3 shows an overview of the proposed approach. The basic flow of the processing is inspired by Gygli (2014). First, a set of slide pages $\boldsymbol{S}$ is analyzed to extract important visual and textual features from each page. In terms of visual importance, how much text and how many figures, formulations, or other objects are contained in each slide is estimated using a background subtraction technique and an inter-frame difference technique. In addition, word importance is measured using the TF-IDF (term frequency–inverse document frequency) method (Salton (1988), Wu (2008)). Meanwhile, a teacher estimates the time that students need to spend studying each slide page. Then, these visual, textual, and temporal features are combined to predict an importance score $I(S_i)$ for each slide page, where S_i indicates the page number of the set of slide pages $\boldsymbol{S}$. Finally, an optimal subset $\widehat{\boldsymbol{S}}$ is selected whereby the importance score is maximized for a given preview time.

3.2 *Visual Features*

Two kinds of image processing techniques are applied to extract visual features from a set of pages; background subtraction and inter-frame difference. The background subtraction technique extracts the foreground mask from each page (as shown in Figure 4(b)) by subtracting a background image (see Figure 4(a)) from each page, followed by binarization processing. Then, the content volume is estimated by counting the number of foreground pixels. Let $f(S_i)$, W, and H be the number of foreground pixels in slide page S_i, the width, and the height of the slide page image, respectively. A foregroundness score, F_i, is given by the following formula:

$$F_i = \frac{f(S_i)}{W \times H}.$$

The inter-frame difference technique reveals changes between successive slide pages (as shown in Figure 4(c)). The subtracted image is also binarized to calculate a difference score D_i as follows:

$$D_i = \max\left(\frac{d(S_{i-1}, S_i)}{W \times H}, \frac{d(S_i, S_{i+1})}{W \times H}\right).$$

The term $d(S_i, S_j)$ represents the number of pixels extracted by the inter-frame difference calculation. Note that the proposed method performs the inter-frame difference calculation in both directions, between the focused frame (Slide page i) and both the previous page (Slide page $i - 1$) and the next page (Slide page $i + 1$), and then selects the larger number of extracted pixels. This process enables us to emphasize a slide page that is significantly different from neighboring pages, and also reduces the effect of small changes caused by animated motion in the page.

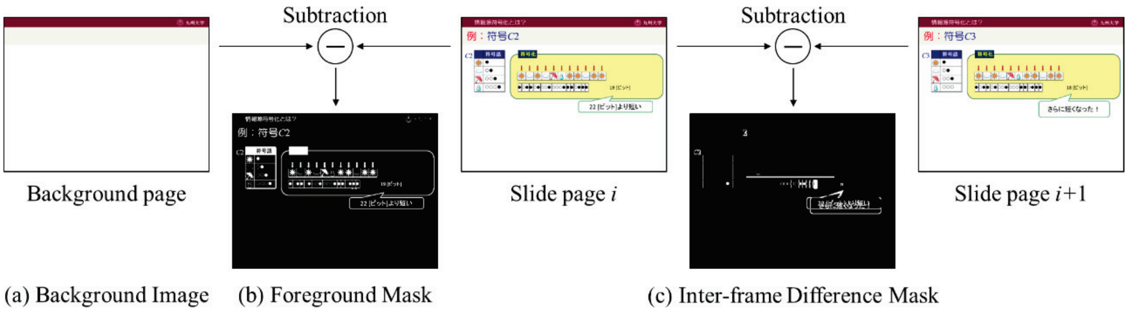

Figure 4. Example of background subtraction and inter-frame difference

3.3 *Text Features*

Text (word) importance is measured using the TF-IDF method. The basic premise of TF-IDF comes from the theory of language modeling. The terms in a given document can be divided into two categories based on whether a term is relevant to the topic of a given document. For further details on TF-IDF, see studies such as Salton (1988) and Wu (2008). In the proposed approach, all the terms (words) are extracted from each slide page S_i in $\boldsymbol{S}$. The weight $w_{p,i}$ for word p in slide page S_i is given by

$$\mathrm{w}_{p,i} = tf_{p,i} \times \log\left(\frac{N}{df_p}\right),$$

where N is the number of documents (pages in this study) in $\boldsymbol{S}$, $tf_{p,i}$ is the frequency of word p in slide page S_i, and df_p is the frequency of word p in $\boldsymbol{S}$.

Finally, a score for words W_i in slide page S_i is calculated by summing the $\mathrm{w}_{p,i}$:

$$W_i = \textstyle\sum_p \mathrm{w}_{p,i}.$$

3.4 Browsing Time

Before making a summary set of slides, a teacher is asked to specify an appropriate browsing time for each slide page S_i. In our study, a teacher who uses slide material in his lectures browsed each slide page and recorded the time T_i. In our study, the time was recorded using the rehearsal function in PowerPoint.

3.5 Combination of Features

The above features are combined to predict an importance score $I(S_i)$ for each slide page:

$$I(S_i) = (F_i + D_i + W_i) \times T_i.$$

We also tried other combination strategies, such as weighting each feature or ignoring the time term, but found that this simple combination worked best.

3.6 Optimal Slide Selection

Given a set of slide pages $\boldsymbol{S}$, the objective is to find a subset with a browsing length below a specified maximum L_S, such that the sum of the importance scores is maximized. Formally, the optimization problem is given by

$$\underset{x}{\text{maximize}} \sum_{i=1}^{P} x_i \, I(S_i), \ \text{subject to} \ \sum_{i=1}^{P} x_i \, T_i \leq L_S,$$

where $x_i \in \{0,1\}$ and $x_i = 1$ indicates that slide page S_i is selected. The maximization can be solved by a standard 0/1 knapsack problem, where $I(S_i)$ is the value of an item and its browsing time T_i is its weight (Kellerer 2004).

4. Experimental Results

4.1 Preparation

Table 1: Summarized slide sets (short, medium, and long) and the original slide set (all). The L_S is a given browsing length as defined in the previous section.

	Short		Medium		Long		All	
	No. of pages	Time L_S (s)	No. of pages	Time L_S (s)	No. of pages	Time L_S (s)	No. of pages	Time (s)
1st week	21	300	30	600	37	900	47	1600
2nd week	10	300	18	600	25	900	38	1700
3rd week	4	200	8	400	12	600	20	1100

We investigated the effectiveness of the proposed approach in a series of information science classes. In total, 372 first-year students, including both arts and science students, attended the classes, which commenced in April 2015. All of the students have their own laptops and bring them to class every week. We obtained results over a 3-week period. Each week, students were asked to preview material in preparation for the next lecture. We prepared three sets of summarized slides, short, medium, and long versions, in addition to the complete set of slides. We divided the students into four groups, and assigned one set of materials to each group. Note that we assigned the students into the groups based on a pre-test so that the average ability of the students was similar across groups. In addition, we changed the assignment of summarized slides between groups every week. Table 1 shows the details of the slide materials used in our experiments. For example, it can be seen that a teacher took about 1100 seconds to browse a set of slide pages (20 pages) in the 3rd week. Based on this browsing time, three types of summarized slide sets were generated as follows: a short version (contains 4 pages to be browsed within 200 seconds ($L_S = 200$)), a medium version (contains 8 pages to be browsed within 400 seconds ($L_S = 400$)), and a long version (contains 12 pages to be browsed within 600 seconds ($L_S = 600$)). Figure 5 depicts the full set of slide pages from the 3rd week and the summarization results.

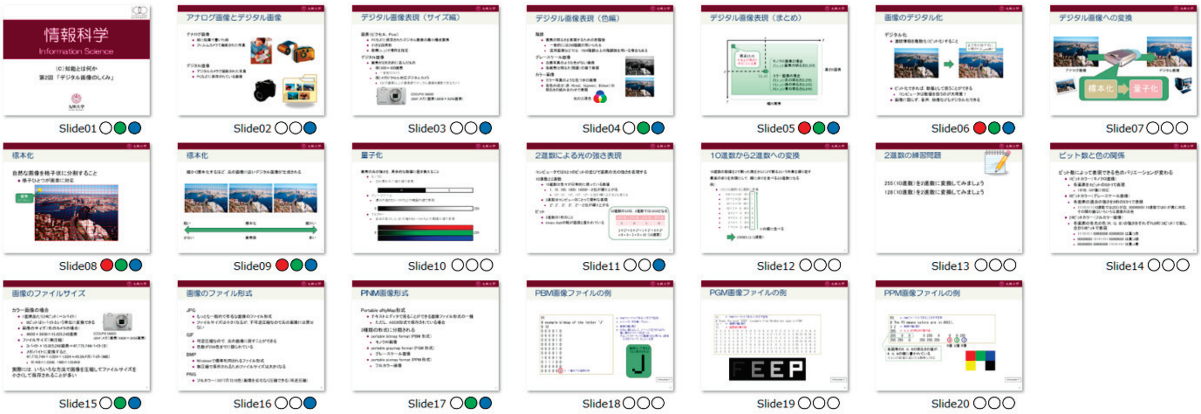

Figure 5. Original set of slide pages used in the 3rd week. The colored circles denote the selection results. Red, green, and blue correspond to short, medium, and long summarized sets, respectively.

The four sets of slides outlined above were given to four student groups using an e-book system (refer to Yin (2014) for more details). This system enables us to collect browsing logs for each set of slide pages to confirm whether a student actually previewed the material within the specified period (the week preceding the lecture). Therefore, we can classify the students into five groups, including those who did not preview the slides. Table 2 shows the number of students in each group.

Table 2: Number of students who browsed the slides.

	Short	Medium	Long	All	None
1st week	44	39	40	92	157
2nd week	40	30	33	111	124
3rd week	43	42	51	105	100

Prior to the beginning of the lecture, we conducted a short quiz to check the level of understanding. The quiz contained five to seven questions depending on the lecture of the week, and the maximum possible score was 10 points.

4.2 Summarized Slide Sets vs. Quiz Scores

We compared the quiz scores among the five groups:

- Short: students who previewed the short set of summarized slides
- Medium: students who previewed the medium set of summarized slides
- Long: students who previewed the long set of summarized slides
- All: students who previewed the original set of slides
- None: students who did not preview any slides.

We conducted t-tests comparing all possible pairings of the groups who previewed the three summarized slide sets and those who previewed all the slides, and also all the groups who previewed sets of slides with the group who did not preview any slides. Figure 6 shows the average score for each group.

The average scores for groups that previewed the various sets of slides (short, medium, long, and all) exceeded those of the group that did not preview the slides (none). With regard to the 1st week, the contents of the slides were not complex, so there was little difference among the quiz scores. The contents of the slides became more difficult as the lecture series proceeded, and therefore the differences between the scores increased in the 2nd and 3rd weeks. The most important point to note (and our major argument) is that the summarized slide sets did not have a negative effect on the quiz scores. This is supported by the fact that the null hypotheses between summarized slides and all the slides were not rejected. Instead, previewing the short and medium sets of summarized slides improved the students' quiz scores compared with previewing the full set of slides (all). From these results, it can be seen that the proposed method for summarizing the slides is very effective. In addition, students who previewed the medium set of summarized slides obtained the best scores in all the quizzes. Note that this summarized slide set was given to a different group each week. This is a very interesting result, and is closely related to the achievement ratios obtained in previewing the slides. The following section provides more detail.

4.3 Summarized Slide Sets vs. Achievement Ratios

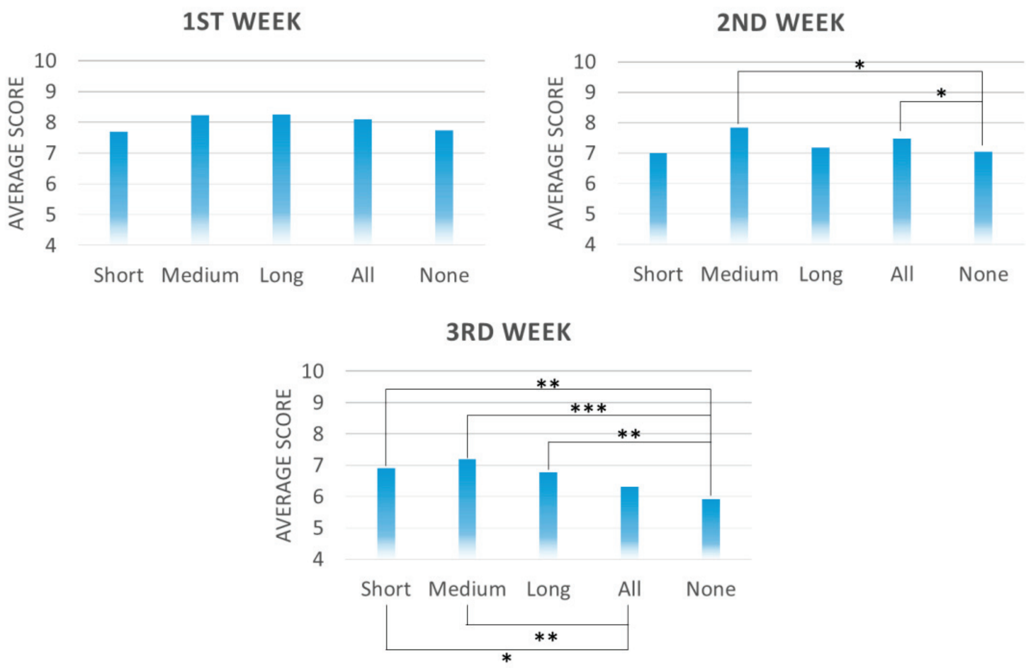

Figure 6. Average quiz scores (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$)

Table 2 shows the number of students who previewed the slide sets they were given. In this section, we analyze the preview process in more detail. By analyzing e-book action logs, we were able to ascertain the page at which a student stopped previewing the slide set. Then, we counted the number of students who previewed more than 80% of pages in the given slide set. Note that we counted the number of pages previewed by students, ignoring pages that students skipped. Finally, we calculated the achievement ratio, how many students who previewed more than 80% of slide pages in the classroom.

Figure 7 shows the achievement ratio for each slide set in each week. With regard to the 2nd week, almost 90% of students previewed their slide sets regardless of the length of the slide set. However, in the cases of the 1st and 3rd weeks, the achievement ratios clearly decreased as the number

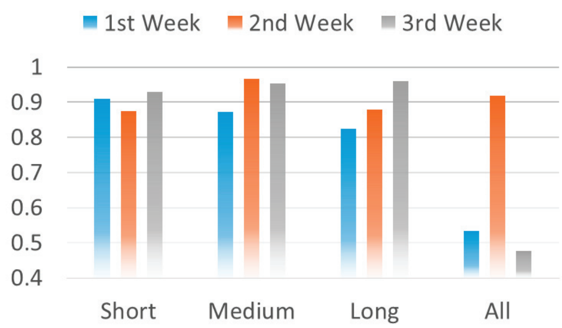

Figure 7. Achievement ratios: percentages of students who previewed more than 80% of slide pages

of slide pages and the estimated preview time increased (see Table 1). Almost one-half of all students who were given all the slides abandoned their preview early in the process. We surmise that too much material caused a decrease in the students' attention span. In other words, the attention span of students is very limited, so it is important to make a summary set of slides for preview by students.

4.4 Browsing Time Evaluation

We investigated the browsing time, i.e., how long students spent previewing the slide set they were given, using the action logs of their e-books. The average browsing time is shown in Figure 8. We found that students who previewed the summarized slide sets (short, medium, or long) all finished their preview in less than 600 seconds each week. However, students who previewed the original slide set (all) took much longer. Recall that the quiz scores for this group were worse than those of the short, medium, and long summarized slide set groups. Therefore, spending more time previewing the material is not always beneficial in terms of understanding the essential points contained in the material.

In addition, we compared the actual browsing time (shown in Figure 8) with the estimated time (shown in Table 1). Figure 9 shows the gap between these measures. In many cases (such as the short slide set group in the 2nd and 3rd weeks, the medium slide set group in the 1st and 3rd weeks, and the long slide set group in the 3rd week), the gap was less than 180 seconds. In all other cases involving summarized slide sets, the gap was much less than that for the group that was given the original slide set, except for the 1st week. Based on these results, the proposed method can be used to select the appropriate length of preview slide sets.

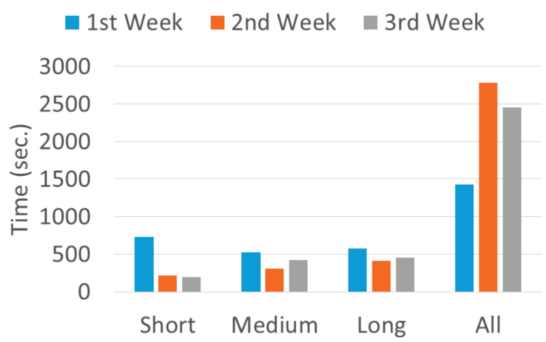

Figure 8. Average actual browsing time

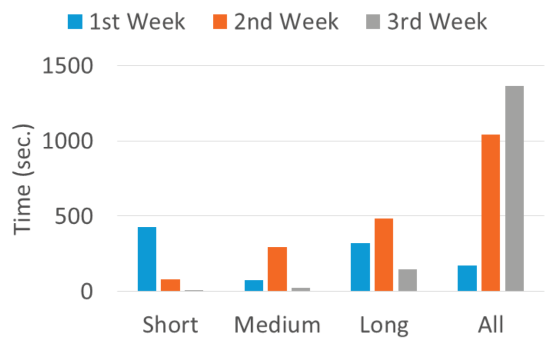

Figure 9. Gap between estimated browsing time and actual browsing time

4.5 Effectiveness of Proposed Method

We asked students the following questions:

Q2-1: What did you think about the length of the slide sets?

Q2-2: How much time did you spend previewing the slides?

Q2-3: How would you describe your level of understanding of the narrative presented by the slides?

Q2-4: How would you describe your level of understanding of the essential points, such as keywords?

In total, 325 students answered these questions. Figure 10 summarizes their answers.

More than half of the students in the long summarized slide set or all slides groups felt that the slide set was longish or long (see Figure 10(a)). As the length of the slide set shortened, the percentage of students who felt that the length was appropriate increased. From Figure 10(b), we can see that about 80% of students finished their preview within 20 min. This result roughly corresponds to that for question Q1-1, as shown in Figure 1(a) above. With regard to the survey of satisfaction (Figure 10(c) and (d)), less than 20% of students had a negative impression, and this figure was comparable across all groups. In other words, the summarized slide sets did not create a poor impression.

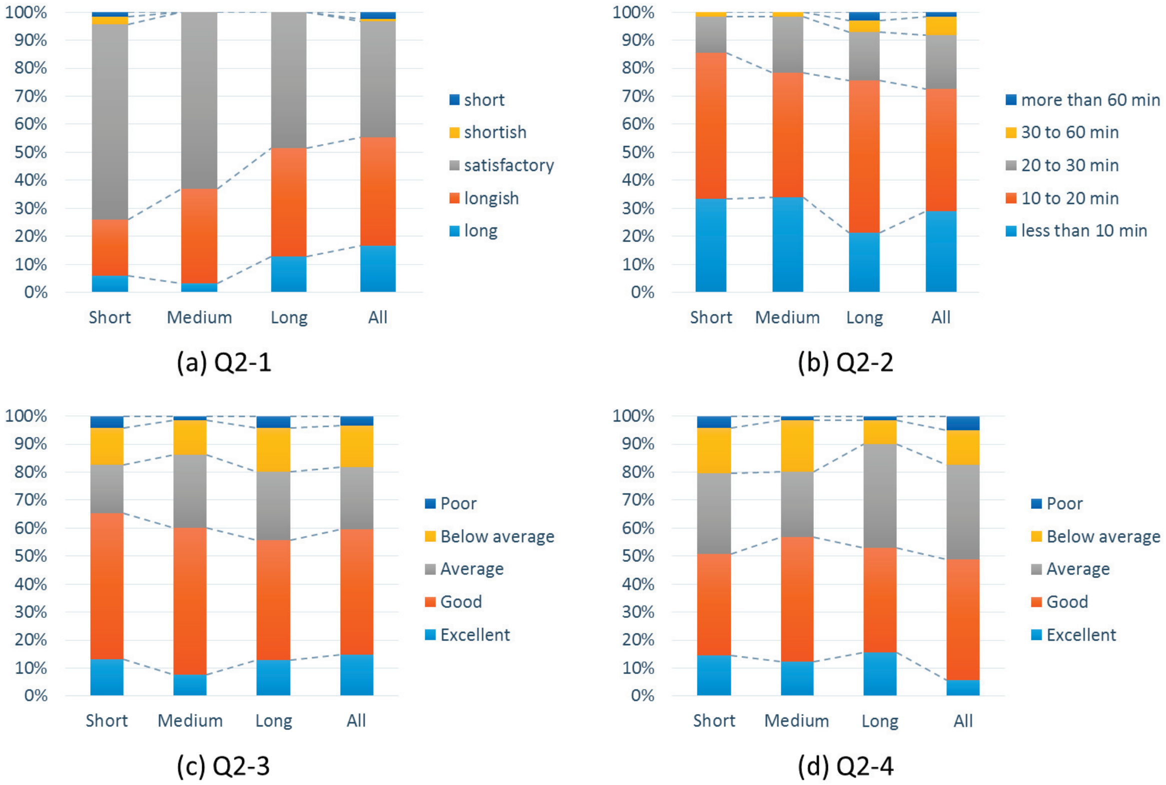

Figure 10. Summary of students' answers to questions regarding length of slide sets, time spent previewing, and level of understanding

4.6 Summarized Slide Sets vs. Teacher's Selection

We evaluated how accurately the proposed method selects the pages of slides to be used in the summary slide sets. First, a teacher who uses slides in lectures made a summarized set of slides for preview. Then, we compared the summarized set of slides made by the teacher with the set made using the proposed method. Accuracy was evaluated using the precision ratio, recall ratio, and F-measure. The F-measure quantifies the balance between precision and recall, with a larger value reflecting a better result.

Table 3 shows the results. In the cases of the medium and long summarized sets, more than 60% of the slides were correctly selected. Although the selection accuracy was not perfect, it seems clear that sufficient slides were selected to enable understanding of the lecture content, because the students in these groups obtained better scores in the quizzes.

Table 3: Precision ratio, recall ratio, and F-measure of summarized slides.

	1st week			2nd week			3rd week		
	Precision	Recall	F	Precision	Recall	F	Precision	Recall	F
Short	71.4	51.7	60.0	60.0	30.0	40.0	75.0	33.3	46.1
Medium	60.0	62.1	61.0	72.2	65.0	68.4	75.0	66.7	70.6
Long	62.2	79.3	69.7	64.0	80.0	71.1	58.3	77.8	66.7

5. Conclusion

In this paper, we proposed a method to make a summary set of lecture slides (currently the target id Power Point, but applicable to other formats such as Keynote, Open Office, etc.) for preview by students prior to lectures. First, image and text features are extracted from a set of slides to calculate importance scores. Then, the page selection problem is optimized by solving a standard 0/1 knapsack problem. In the experiments, we compared the quiz scores, preview achievement ratios, and time spent previewing of groups of students previewing summarized sets of slides of varying lengths. All the groups who previewed summarized sets of slides obtained better scores on pre-lecture quizzes than the group who previewed the entire set of slides. The summarized sets of slides were also effective in terms of preview achievement ratios. Most students browsed at least 80% of the summarized sets of slides, while students who were given the original slides gave up browsing before they reached the halfway point. Based on these findings, we argue that the summarized sets of lecture slides enhance students' preview study by not only improving the students' understanding of the material, but also maintaining the students' concentration, because they can grasp the essential points of the lecture slides in less time. According to their answers to the questionnaires, most students were impressed with the summarized sets of slides in terms of their length, and therefore the time needed to study them, and also with their ability to understand the overall narrative of the lecture and keywords.

In future work, we will continue our analysis of lecture materials with a view to obtaining further evidence supporting our conclusions. As the next step, we will create a summary set of slides taking into account varying levels of student ability. We are confident that students' level of ability can be estimated by analyzing e-book logs and e-learning system logs. Further, we suspect that the amount of time spent on previewing depends on the student's individual situation. Therefore, it is very important to create a summary set of slides that is well suited to each individual student's requirements.

Acknowledgements

This research was supported by "Research and Development on Fundamental and Utilization Technologies for Social Big Data" (178A03), the Commissioned Research of the National Institute of Information and Communications Technology (NICT), Japan.

References

Beichner, R.J. (1995), "Improving the effectiveness of large-attendance lectures with animation-rich lecture notes", AAPT Announcer 20, 917.

Money, A.G., Agius, H., (2008), "Video summarisation: A conceptual framework and survey of the state of the art", Journal of Visual Communication and Image Representation 19, 121–143.

S. P. Rajendra and K. N, (2014), "A Survey of Automatic Video Summarization Techniques", International Journal of Electronics, Electrical and Computational System (IJEECS), vol. 3.

He L.W., Sanocki E., Gupta A., Grudin J. (1999), "Auto-Summarization of Audio-Video Presentations", ACM MM'99 489-498.

Yun-Nung Chen, Yu Huang, Ching-feng Yeh, Lin-Shan Lee (2011), "Spoken Lecture Summarization by Random Walk over a Graph Constructed with Automatically Extracted Key Terms", INTERSPEECH 2011, 933-936.

Harish Mathivanan, Madan Jayaprakasam, K. Gokul Prasad, T. V. Geetha, (2009), "Document Summarization and Information Extraction for Generation of Presentation Slides", Proceedings of the 2009 International Conference on Advances in Recent Technologies in Communication and Computing, 2009. ARTCom '09, 126-128.

K. Sathiyamurthy, T. V. Geetha, (2012), "Automatic Organization and Generation of Presentation Slides for E-Learning", International Journal of Distance Education Technologies, v.10 n.3, p.35-52.

Michael Gygli, Helmut Grabner, Hayko Riemenschneider, Luc Van Gool, (2014), "Creating Summaries from User Videos", ECCV (7) 2014: 505-520.

G. Salton and C. Buckley, (1988), "Term-weighting approaches in automatic text retrieval", Information Processing & Management, 24 (5).

H. Wu and R. Luk and K. Wong and K. Kwok, (2008), "Interpreting TF-IDF term weights as making relevance decisions", ACM Transactions on Information Systems, 26 (3).

Kellerer H., Pferschy U., Pisinger D., (2004). Knapsack Problems. Springer.

Chengjiu Yin, Fumiya Okubo, Atsushi Shimada, Kojima Kentaro, Masanori Yamada, Hiroaki Ogata, Naomi Fujimura, (2014), "Smart Phone based Data Collecting System for Analyzing Learning Behaviors", 22nd International Conference on Computers in Education (ICCE 2014).