Published as: Che, Xiaoyin, Haojin Yang, and Christoph Meinel. "Adaptive E-Lecture Video Outline Extraction Based on Slides Analysis." In Advances in Web-Based Learning–ICWL 2015, pp. 59-68. Springer International Publishing, 2015.

Adaptive E-Lecture Video Outline Extraction Based on Slides Analysis

Xiaoyin Che, Haojin Yang, and Christoph Meinel

Hasso Plattner Institute, University of Potsdam
Prof.-Dr.-Helmert-Str. 2-3, 14482 Potsdam, Germany
{xiaoyin.che,haojin.yang,christoph.meinel}@hpi.de

Abstract. In this paper, we propose an automated adaptive solution to generate logical, accurate and detailed tree-structure outline for video-based online lectures, by extracting the attached slides and reconstructing their content. The proposed solution begins with slide-transition detection and optical character recognition, and then proceeds by a static method of analyzing the layout of single slide and the logical relations within the slides series. Some features about the under-processing slides series, such as a fixed title position, will be figured out and applied in the adaptive rounds to improve the outline quality. The result of our experiments shows that the general accuracy of the final lecture outline reaches 85%, which is about 13% higher than the static method.

Keywords: E-Learning, Lecture Outline, Adaptive Slides Analysis

1 Introduction

E-lectures are very close to our daily lives today, and video is the core material for most online courses, no matter in traditional tele-teaching or MOOC (*Massive Open Online Course*). But how can learners find the lecture videos or segments exactly what they want among such huge amount of choices? Metadata is so far the best answer. Currently tags and manual descriptions are the main sources of such metadata, but we believe a lecture outline would be a better option.

Some studies suggest that students benefit from the lecture outline when taking online courses [1, 2]. And a survey offered with a MOOC course [3] shows that 91% of the respondents (*90 of 99*) believe an accurate outline could be a positive factor in their learning process. A proper outline contains much more information than tags and is much better structured than descriptions, which enables multiple functions such as preview, navigation, segmentation and retrieval.

It would be an extra burden for the lecturers if we ask them to provide outlines for their courses, and hiring others to create outline manually would cost lots of time or/and money. These facts prompt us to search for automated outline generation possibilities. However, concluding outline from lecture speech transcript [4] is too challenging, while capturing the teacher's writing on the blackboard [5] becomes less practical, because slides have occupied the front of the classroom nowadays[6, 7].

Published as: Che, Xiaoyin, Haojin Yang, and Christoph Meinel. "Adaptive E-Lecture Video Outline Extraction Based on Slides Analysis." In Advances in Web-Based Learning–ICWL 2015, pp. 59-68. Springer International Publishing, 2015.

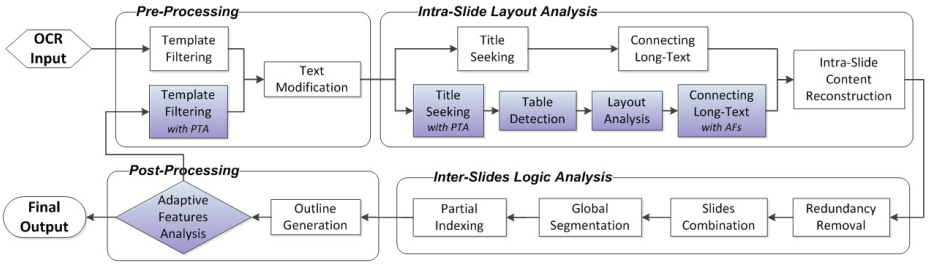


Fig. 1. The Framework of Proposed Solution

Actually most of the lecturers use the slides exactly as the outline of their speeches. Parsing the digital slide files seems to be a good option. But in the purpose of navigation or segmentation, timing information is crucial, which the digital files cannot provide. Instead we will take the slide images extracted from real-time lecture video as the input. Slide images can be located and restored from the projection screen in traditional single-stream lecture video [8, 9]. And with double-stream recording systems, such as Tele-TASK, a desktop stream recodes directly from the lecturer's computer. The accuracy of applying OCR (*Optical Character Recognition*) on the high quality slide images extracted from desktop stream reaches 85% [10], which is available for further use.

Some early efforts have been made with the slides, by exploring the hierarchical semantic concepts [11] or locating the slide title [12]. The first slide-based lecture outline generator has been proposed in 2013 [13]. It explores different components within the slides according to a pre-defined "template" and achieves decent result. But the slide layout could be very diverse due to numerous different slide templates developed by agencies all over the world, the performance of [13] drops drastically when the slides do not fit the pre-defined template.

Therefore we intend to develop a "smarter" outline generator, with the ability to detect the characteristics of the template used by the under-processing slides series and adjust the analyzer adaptively. The initial round of the proposed solution is based on the static method described in [13]. Then four Adaptive Features (AF) describing the differences between slide templates will be analyzed from the output, including Potential Title Area (PTA), General Hierarchical Gap (GHG), Low Case Start (LCS) and Item Bullet (IB). PTA indicates the default title area and the others focus on template characteristics of text-blocks.

The AF-involved steps, along with a few steps inplemented in adaptive rounds only, are marked with light blue background in Fig. 1. All AFs will be updated after each round and any change of them will trigger a new round. But the maximum of adaptive rounds has been set to 3, in order to avoid potential "dead loop". The result of slide transition analysis and OCR is taken as system input. Each round can be separated into pre-processing, intra-slide layout analysis, inter-slides logic analysis and post-processing, which will be further illustrated by Section 2, 3, 4 and 5 respectively. The evaluation and conclusion come afterwards.

Published as: Che, Xiaoyin, Haojin Yang, and Christoph Meinel. "Adaptive E-Lecture Video Outline Extraction Based on Slides Analysis." In Advances in Web-Based Learning–ICWL 2015, pp. 59-68. Springer International Publishing, 2015.

2 Pre-Processing

With OCR result as input, each slide can be simplified as a blank background and a group of text-lines. But not all the text-lines should be included in the outline. Since most of the lecturers or presenters will create their slides with affiliation-related templates, logos, names or the presentation titles may appear repeatedly throughout the whole slides series. Thus, a searching scheme is applied to traverse all the slides and mark those repeatedly appearing text-lines with same content and position. If the number of a text-line's accumulated appearances is beyond the threshold, it will be removed from all slides. Sometimes a real slide title, or part of it, might also be shared in multiple continuous slides under the risk of being marked as redundant. So we introduce one adaptive feature, the potential title area, to the pre-processing in adaptive rounds. Any text-lines in the PTA will bypass the pre-processing.

Another task in pre-processing is to modify ill-recognized text-lines caused by OCR errors. Nobody wants to see meaningless strings in the lecture outline, so they need to be deleted. The standards to define meaninglessness include the average word length smaller than 2 characters, extreme text-lines sizes (*either too large or too small*), containing too many same letters or symbols, etc. Moreover, the extra space in the beginning of the text-lines will also be removed here.

3 Intra-Slide Layout Analysis

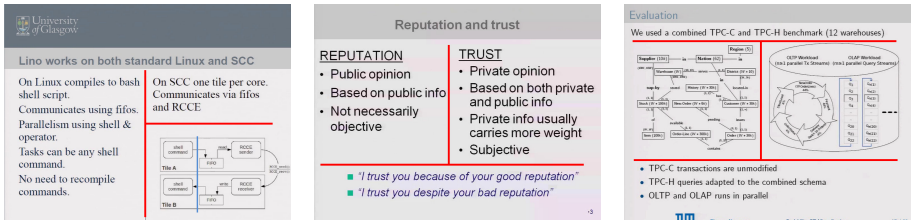
3.1 Title Seeking

Slide title is the most important component in building a lecture outline. We search title candidates in the top 1/3 of the slide. A text-line must have an above-average height and locate not too close to the slide edges. Since the title may occupy multiple rows, including potential subtitle, we accept up to 3 candidates as long as they locate closely. In adaptive rounds, we search prior within the PTA zone. PTA has a strict limitation on vertical position but is quite open horizontally. If a text-line located in this bar-shaped PTA zone, it will be accepted as title candidate and exclude all text-lines out of the zone. With this effort some text components in the slides, which are occasionally near the title, will no longer be mistakenly recognized as the subtitle.

3.2 Table Detection

Table is a type of frequently used data structure in slides. When existing, it contains a lot of detailed information, which we do not need in the outline. We develop a table detection algorithm specialized for slide images. It begins with detecting rows and columns from the text-lines. And then their intersections will be found. An evaluation process comes afterwards to examine whether these intersections belong to a table by analyzing their content, structure and location. When confirmed, these intersections will be considered as a table prototype and further expanded alongside the rows and columns. A final rectangle table area will form and all text-lines inside it will be removed.

Published as: Che, Xiaoyin, Haojin Yang, and Christoph Meinel. "Adaptive E-Lecture Video Outline Extraction Based on Slides Analysis." In Advances in Web-Based Learning–ICWL 2015, pp. 59-68. Springer International Publishing, 2015.



(a) Left-Right-Bottom (b) Left-Top-Bottom (c) Top-Left-Right-Bottom

Fig. 2. The Examples of Page Layout Analysis

3.3 Page Layout Analysis

The slide layout can be very diverse. Besides the tables, diagrams or multi-column layout may also lead to complex text-lines distribution, which cause a lot of problems in the static method. Detecting them will definitely improve the accuracy of slide content reconstruction. Ignoring the title area, we attempt to split text-lines into groups with proper vertical and horizontal axis, as the red lines shown in Fig. 2[1]. Each group can be represented as a block, either a text block to be further processed, or a diagram block to be deleted. The specific procedures are introduced as follow:

1. Attempt to find a middle line which horizontally divides all text-lines, except for the title, into left and right blocks. If there are more than one option, apply the one closest to the absolute middle. (*Fig. 2-a*)
2. For every left-block or right-block, if there is a huge line space inside the block, split it again vertically. (*Right block in Fig. 2-a*)
3. If step 1 failed, attempt top-left-right or left-right-bottom layout. In this case no further vertically splitting is applied. (*Fig. 2-b*)
4. If step 3 still failed, attempt top-left-right-bottom layout, treat the middle parts as diagram and remove them. (*Fig. 2-c*)
5. Any blocks which can be further horizontally split will be removed as diagram. (*Right-bottom sub block in Fig. 2-a, with light blue line*)
6. Analyze all remaining blocks by their content. A block contains many digits, single words or not well-aligned will also be considered as a part of diagram or chart and gets deleted.

After removing the blocks supposed to be diagrams, every block left is acknowledged as text block and will be treated as an independent text system in following procedures. For those slides which cannot be split into blocks, all their text-lines is considered as a whole, in other words, an entire block.

[1] The copyright of the example slides belongs to original authors: Mr. Paul Cockshott, Prof. Audun Jøsang & Prof. Thomas Neumann

Published as: Che, Xiaoyin, Haojin Yang, and Christoph Meinel. "Adaptive E-Lecture Video Outline Extraction Based on Slides Analysis." In Advances in Web-Based Learning–ICWL 2015, pp. 59-68. Springer International Publishing, 2015.

3.4 Continuous Text-Line Combination

When several text-lines locating in different rows belong to a long statement, they need to be reconnected. We take the factors like line spaces, horizontal positions and text-line initials under consideration, in addition with some adaptive features. Here we take t_{n-1} and t_n to represent the text-lines supposed to be combined and explain the decisive factors as follow:

- ◇ When IB is positive, a combination will be suggested if t_{n-1} has a bullet but t_n does not. if t_n has a bullet, the combination will be strongly opposed.
- ◇ The text-line initial can be upper-case, no-case (*digit e.g.*) or lower-case, with descending values. A combination will be suggested when the value of t_{n-1} is larger than t_n. If LCS is positive, the weight of this factor decreases.
- ◇ If the line space of t_{n-1} and t_n is way larger than their heights, or obviously larger than the line space of t_n and t_{n+1}, a combination is opposed.
- ◇ The left-ends of t_{n-1} and t_n should be horizontally close if they belong to same sentence. Please note if the difference of their horizontal starting points fits the GHG, the combination will be vetoed.
- ◇ All text-lines sharing same horizontal starting point with t_{n-1} will be traversed and the widest one will be taken as reference. Only if the difference between the width of t_{n-1} and the reference is smaller than the width of the first word in t_n, the combination could be suggested.

All above factors will be quantified, with "suggested" into positive values while "opposed" or "vetoed" into negative. Finally if the sum is above 0, a combination will be applied. Please note the combination of multi-rows text-lines is implemented within text blocks, but if a combination is necessary between same-row text-lines, it would be done before the page layout analysis.

3.5 Tree-Structure Outline Reconstruction

The default content reconstruction method, which is the only option in initial round, begins with searching a large enough text-line (*above average height at least*) whose left-end locates in the left-top quarter of the slide. It will be taken as the datum. Then by checking the horizontally aligned and vertically adjacent text-lines of this datum, up to 3 hierarchies will be marked and the tree-structure outline can be generated.

In adaptive rounds more possibilities can be provided. Taking text blocks from the layout analysis as input, a second method specialized for center-aligned situation is applied on such text system which contains less than 5 text-lines. The first text-line is directly set to level-1, and all other text-lines center-aligned to the first will be set to level-2. If there are more than 5 text-lines in the block, a third method is introduced. It traverses all the text-lines to find out the most frequently used left-end as the datum, then locates the 3 potential hierarchies.

By comparing the result of the default method and the alternative, the method that makes more text-lines involved in the system will be adopted. If a slide is divided into several blocks, their content will be combined together according to the order of top-left-right-bottom.

Published as: Che, Xiaoyin, Haojin Yang, and Christoph Meinel. "Adaptive E-Lecture Video Outline Extraction Based on Slides Analysis." In Advances in Web-Based Learning–ICWL 2015, pp. 59-68. Springer International Publishing, 2015.

(a) Tag-Page (b) Split-Page (c) Section-Page

Fig. 3. The Examples of "Border" Slides

4 Inter-Slides Logic Analysis

4.1 Redundancy Removal and Slides Combination

Redundancy is something that every lecturer attempts to avoid when creating slides. But in real-time presentation, it is quite natural to roll back to a previous slide for further explanation or just misoperate, either of which causes extra slide transitions and results in repeated slides, such as making an original slide sequence 'A-B-C-D' into 'A_1-B_1-C_1-B_2-A_2-B_3-C_2-D'. We use Levenshtein Distance [14] (L.D.) to evaluate how similar two slides are, along with counting the number of same words (S.W.) existing in compared slides. For repeated slides just like A_1 and A_2, L.D. is supposed to be small while S.W. is large. Then the redundancy can be removed.

Slide combination takes place when two adjacent slides share the same title, sometimes with additional numbers, like "(1/3)", "<b>" or "III". There are two possibilities: either a progressive displayed slide is detected as several independent slides, or a key point needs to be discussed in several continuous slides. In such cases the slides will be merged and the text-lines involved in intra-slide content-trees keep their hierarchies unchanged.

4.2 Global Segmentation and Partial Indexing

In many cases a lecture consists of several segments, each of which focuses on a subtopic. Some lecturers mark those segments directly by using special "border" slides, such as a tag-page which is a rough outline of the lecture with a certain highlighted text-line, a split-page which contains only one text-line indicating the subtopic, or a section-page which has special phrases like "Part 1" or "Section two" in its title. Examples can be found in Fig. 3[2]. By confirming these border slides, the whole lecture can be easily segmented globally.

For the lectures without available border slides, or when the globally segmented subtopics are still too general, we propose partial indexing to explore

[2] The copyright of the example slides belongs to original authors: Prof. Rudi Studer, Mr. Paul Cockshott & Prof. Gil Rosenman

Published as: Che, Xiaoyin, Haojin Yang, and Christoph Meinel. "Adaptive E-Lecture Video Outline Extraction Based on Slides Analysis." In Advances in Web-Based Learning–ICWL 2015, pp. 59-68. Springer International Publishing, 2015.

the connections within neighboring slides. Sometimes a slide may act as a preview of several following slides by listing their titles. We address it as index-page, mark it the root node of a partial segment and link previewed slides as its leaves. The range of a partial segment led by an index-page cannot beyond the border of global segments. In extreme instance, a leaf node linked to an index-page inside a global segment can be in level-3.

Besides, we also search for continuous slides with same keyword or prefix in their titles, by which these slides are supposed to illustrate different aspects of a certain topic. If a noun or a phrase repeatedly appears in the floating search interval, a virtual index-page will be created, taking the keyword or prefix as the title. This virtual index-page will be inserted before the interval and act as a real index-page. And virtual index-page searching is only available for slides not included in any global segments.

5 Post-Processing

By setting the slide title as level-0 root element in the intra-slide content tree, every text-line has an intra-slide hierarchy ranged 0∼3. In addition with an inter-slides hierarchy ranged 1∼3, a final hierarchy of each text-line can be calculated by a simple addition, which is ranged 1∼6 in tree-structure outline.

Meanwhile, the adaptive features are also achieved or updated. The title position of each slide is recorded and the repeatedly used positions (*1/4 of all slides*) are stored as PTAs. The gap between level-1 and level-2 text-lines within each slide is also recorded. If more than 1/4 slides have similar hierarchical gaps, the value will be saved in GHG. The boolean attributes LCS and IB derive from the statistics of the final outline. When more than 30% of outline items begin with lower-case letters, LCS will be set to "true". And threshold for IB, which means a subtopic begins with a bullet recognized as single character, just like 'o' or 'u', is 20%. Any change in AFs will trigger a new round, unless this is already the 3rd adaptive round.

6 Evaluation

In our evaluation session, we select 12 complete e-lectures or academic presentations of 12 different lecturers from Tele-TASK platform to build the test dataset. A total number of 354 pages of original slides are supposed to be extracted from the desktop stream of 437 minutes of lecture videos, by which the diversity of the dataset could be assured. A lecture ID is used to identify certain lecture, and all these lectures are publicly available[3], in addition with the manually created ground-truth of the lecture outlines[4]. We would like to compare the performances of proposed adaptive solution with the static outline generation method introduced in [13].

[3] The lecture (ID = 'id') is in http://www.tele-task.de/archive/lecture/overview/id/

[4] https://drive.google.com/folderview?id=0B13Cc1a7ebTufmV6WFRCbmxPYllxR3hYNE1SRUtWN3hxZl9tdHBPaHU0THZwOXVpM29sZEE&usp=sharing

Published as: Che, Xiaoyin, Haojin Yang, and Christoph Meinel. "Adaptive E-Lecture Video Outline Extraction Based on Slides Analysis." In Advances in Web-Based Learning–ICWL 2015, pp. 59-68. Springer International Publishing, 2015.

Table 1. Intra-Slide Accuracy Report

	ID	Character Aspect			Item Aspect					
		Length	L.D.	Precision	G.T.	Hit	Recall	All	Correct	Precision
Static Solution	5626	2305	393	83.0%	66	50	75.8%	82	36	43.9%
	5759	5596	1625	71.0%	208	162	77.9%	215	128	59.5%
	6011	6888	596	91.3%	123	115	93.5%	132	101.5	76.9%
	6031	6103	961	84.3%	158	129	81.6%	139	114	82.0%
	6102	5637	1223	78.3%	162	147	90.7%	184	137	74.5%
	6106	4417	2800	36.6%	107	87	81.3%	169	68.5	40.5%
	6196	7742	2585	66.6%	268	152	56.7%	194	132.5	68.3%
	6201	3381	786	76.8%	118	104	88.1%	133	100.5	75.6%
	6261	4268	1273	70.2%	132	98	74.2%	132	93.5	70.8%
	6266	2569	272	89.4%	65	63	96.9%	81	61	75.3%
	6663	4014	245	93.9%	98	94	95.9%	113	89.5	79.2%
	7314	2820	1352	52.1%	83	63	75.9%	109	60	55.0%
	All	55740	14111	**74.7%**	1588	1264	**79.6%**	1683	1122	**66.7%**
Adaptive Solution	5626	2305	205	91.1%	66	64	97.0%	68	58.5	86.0%
	5759	5596	755	86.5%	208	185	88.9%	205	156	76.1%
	6011	6888	612	91.1%	123	110	89.4%	124	98	79.0%
	6031	6103	515	91.6%	158	154	97.5%	166	146.5	88.3%
	6102	5637	1263	77.6%	162	153	94.4%	184	141	76.6%
	6106	4417	1074	75.7%	107	102	95.3%	139	89	64.0%
	6196	7742	1767	77.2%	268	230	85.8%	257	220	85.6%
	6201	3381	328	90.3%	118	112	94.9%	117	111	94.9%
	6261	4268	506	88.1%	132	115	87.1%	136	105.5	76.1%
	6266	2569	302	88.2%	65	63	96.9%	76	60.5	79.6%
	6663	4014	205	94.9%	98	91	92.9%	94	87	92.6%
	7314	2820	407	66.6%	83	68	81.9%	78	59.5	76.3%
	All	55740	7939	**85.8%**	1588	1447	**91.1%**	1644	1330.5	**80.9%**

First we focus on the intra-slide phase with two aspects: characters and items. In character aspect, the whole content-tree of a slide is connected together as a string, and a Levenshtein distance (L.D.) will be calculated against the ground-truth (G.T.). The smaller the L.D. is, the higher the precision reaches. Then in item aspect, whether a content-tree is hierarchically accurate will be tested. The content and the hierarchy of an outline item value 0.5 respectively. And by comparing with the G.T., both recall and precision can be obtained. Please note minor differences in characters leading no misunderstanding will be ignored here, because they have already affected in character aspect. Statistics can be found in Table 1.

The second phase of evaluation focuses on the inter-slides logic. Slide title will represent the whole slide, with a string and a hierarchy ranged 1∼3. Similar to the item-aspect intra-slide evaluation, the string and the hierarchy weight 50% each. Table 2 shows the result. Total Slides (T.S.) indicates how many slides have been extracted from the video originally, which in many cases differs from the G.T., due to the logical inter-slides processing.

Published as: Che, Xiaoyin, Haojin Yang, and Christoph Meinel. "Adaptive E-Lecture Video Outline Extraction Based on Slides Analysis." In Advances in Web-Based Learning–ICWL 2015, pp. 59-68. Springer International Publishing, 2015.

Table 2. Inter-Slides Accuracy Report

ID	T.S.	G.T.	Static		Adaptive	
			Correct	Accuracy	Correct	Accuracy
5626	17	13	10	76.9%	10	76.9%
5759	45	20	5.5	27.5%	19.5	97.5%
6011	18	10	5	50.0%	8	80.0%
6031	22	20	7	35.0%	19	95.0%
6102	36	30	26.5	88.3%	28	93.3%
6106	28	28	25	89.3%	25	89.3%
6196	81	31	19	61.3%	21.5	69.4%
6201	25	25	22	88.0%	23	92.0%
6261	27	20	12	60.0%	14	70.0%
6266	18	18	12.5	69.4%	14	77.8%
6663	20	20	19	95.0%	19	95.0%
7314	17	18	12	66.7%	10.5	58.3%
All	354	253	175.5	**69.4%**	211.5	**83.6%**

Table 3. General Accuracies

	Intra-Slide (A_1)			Inter-Slides (A_2)	A_{final}
	Character (P_C)	Item			
		R_I	P_I		
Static	74.7%	79.6%	66.7%	69.4%	**71.5%**
Adaptive	85.8%	91.1%	80.9%	83.6%	**84.7%**

A final accuracy is calculated generalizing all aspects, as shown in Formation (1). The general intra-slide item-level accuracy is achieved by applying F-measure (*harmonic mean*) on recall (R_I) and precision (P_I). Then the G-measure (*geometric mean*) of both item-level accuracy and character-level precision (P_C) represents the general intra-slide accuracy. The final accuracy derives from the G-measure of both intra-slide and inter-slides accuracies (A_1 *and* A_2). From the statistics illustrated in Table 3, we can easily figure out that the general accuracy of proposed solution reaches approximately 85%.

$$A_{final} = \sqrt{A_2 \cdot \sqrt{P_C \cdot \frac{R_I \cdot P_I}{R_I + P_I}}} \tag{1}$$

7 Conclusion and Future Work

In this paper we introduce our effort in generating outline adaptively for online lectures by analyzing the slides extracted from videos. The proposed solution goes by analyzing and further utilizing the specified features of the certain under-processing slides series. The evaluation shows that the accuracy of final output reached 85%. In the future we tend to adjust our system for the compromised slide images from single-stream lecture video, enable the digital slide file parsing

Published as: Che, Xiaoyin, Haojin Yang, and Christoph Meinel. "Adaptive E-Lecture Video Outline Extraction Based on Slides Analysis." In Advances in Web-Based Learning–ICWL 2015, pp. 59-68. Springer International Publishing, 2015.

for accuracy improvement and apply further applications, such as lecture video retrieval, based on the outline generated.

References

1. Grabe, M., Christopherson, K.: Optional student use of online lecture resources: resource preferences, performance and lecture attendance. Journal of Computer Assisted Learning 24(1), 1–10 (2008)
2. Lonn, S., Teasley, S.D.: Saving time or innovating practice: Investigating perceptions and uses of learning management systems. Computers & Education 53(3), 686–694 (2009)
3. Che, X., Luo, S., Wang, C., Meinel, C.: An attempt at mooc localization for chinese-speaking users. International Journal of Information and Education Technology 6(2), 90 (2016)
4. Zhang, J., Chan, R.H.Y., Fung, P., Cao, L.: A comparative study on speech summarization of broadcast news and lecture speech. In: Interspeech. pp. 2781–2784. Citeseer (2007)
5. Onishi, M., Izumi, M., Fukunaga, K.: Blackboard segmentation using video image of lecture and its applications. In: Pattern Recognition, 2000. Proceedings. 15th International Conference on. vol. 4, pp. 615–618. IEEE (2000)
6. Hill, A., Arford, T., Lubitow, A., Smollin, L.M.: im ambivalent about it the dilemmas of powerpoint. Teaching Sociology 40(3), 242–256 (2012)
7. Levasseur, D.G., Kanan Sawyer, J.: Pedagogy meets powerpoint: A research review of the effects of computer-generated slides in the classroom. The Review of Communication 6(1-2), 101–123 (2006)
8. Li, K., Wang, J., Wang, H., Dai, Q.: Structuring lecture videos by automatic projection screen localization and analysis. IEEE Transactions on Pattern Analysis & Machine Intelligence 37(6), 1233–1246 (2015)
9. Schroth, G., Cheung, N.M., Steinbach, E., Girod, B.: Synchronization of presentation slides and lecture videos using bit rate sequences. In: Image Processing (ICIP), 2011 18th IEEE International Conference on. pp. 925–928. IEEE (2011)
10. Yang, H., Siebert, M., Luhne, P., Sack, H., Meinel, C.: Lecture video indexing and analysis using video ocr technology. In: Signal-Image Technology and Internet-Based Systems (SITIS), 2011 Seventh International Conference on. pp. 54–61. IEEE (2011)
11. Atapattu, T., Falkner, K., Falkner, N.: Automated extraction of semantic concepts from semi-structured data: Supporting computer-based education through the analysis of lecture notes. In: Database and Expert Systems Applications. pp. 161–175. Springer (2012)
12. Yang, H., Gruenewald, F., Meinel, C.: Automated extraction of lecture outlines from lecture videos-a hybrid solution for lecture video indexing. In: Proceedings of 4th International Conference on Computer Supported Education (CSEDU). pp. 13–22 (2012)
13. Che, X., Yang, H., Meinel, C.: Tree-structure outline generation for lecture videos with synchronized slides. In: e-Learning and e-Technologies in Education (ICEEE), 2013 Second International Conference on. pp. 87–92. IEEE (2013)
14. Levenshtein, V.I.: Binary codes capable of correcting deletions, insertions and reversals. In: Soviet physics doklady. vol. 10, p. 707 (1966)