ABSTRACT

In this paper, the authors illustrate their motivation and method in the automated generation of tree-structure outline for lecture videos with supplementary synchronized slides, and then propose a further application, lecture video segmentation by slide-group-change event, based on the outline previously generated. Starting with OCR (Optical Character Recognition) result, with an approximate accuracy of 90%, the authors attempt to reconstruct the text system of each slide into an up-to-3-level content tree, and then explore logical relations between slides in order to set them hierarchical. A final up-to-6-level outline will be achieved after removing all the redundancy. And the hierarchy of the slides, which is saved in the outline, will largely simplify the additional segmentation process. Evaluation result shows that, the final outline generated based on the test dataset retains about only 1/4 of the original texts from all slides and is organized well, with a high accuracy of 85% at slide title level. And the majority of the segments the authors’ get are logically reasonable, while the average length of them is about 5–15 minutes.

Keywords: e-Lecture, Optical Character Recognition (OCR), Slide Analysis, Tree-Structure Outline, Video Segmentation

E-learning is nothing new today. The huge amount of online courses can be easily found online in all kinds of e-learning web portal to meet people’s exploding demand for that. These numerous e-lectures enable people to learn wherever, whenever and somehow even whatever they want, but also bring them the difficulty to choose. No one wants to waste time by taking the wrong lecture. Since the video is the fundamental and helpful resource in most e-learning systems (McGreal & Elliott, 2008; Wieling & Hofman, 2010; Ronchetti, 2011), it’s important to help the portal users to have a clear idea on what this lecture video is talking about before they open it. However, a glance on the title is far from enough. More detailed but still accurate information should be provided.

DOI: 10.4018/ijtem.2014010103
Tagging is the most favorable choice to do so, which has been researched in various ways over decades. But no matter the tags come from user feedback (Bateman, Brooks, McCalla, & Brusilovsky, 2007; Torniai, Jovanovic, Gasevic, Bateman, & Hatala, 2008; Rahimi, Ban den Berg, & Veen, 2013), or derive by automated semantic analysis (Sack & Waitelonis, 2006; Kamabuthula & Iyer, 2011; Imran, Rahadianti, Cheikh, & Yayilgan, 2012), the only problem tagging can solve is “what it is”, not “how it goes”. The information tags can offer is too limited to give the learner an adequate lecture outlook. Besides, Lecture description is another option, which can clearly tell the e-learning portal users how this lecture goes. But only the professions are capable to write professional descriptions, and it’s not realistic to hire a group of professions or to force the lecturers to do this additional task by themselves. However, some investigations show that a lecture outline has a positive effect on the students’ overall performance who taking online courses (Grabe & Christopherson, 2008; Lonn & Teasley, 2009). But the problem is how to achieve the lecture outline and obviously, this work should not be done manually because of the huge potential consumption of time and money.

Due to the extensive usage of the slides by the lecturers when giving their lectures in recent years, which is gradually replacing the blackboard in the classroom, extracting the content of the slides can be an excellent choice. In many cases, the slides are exactly the detailed and accurate outline of the lecture. Some attempts have already be done in generating outline from lecture slides (Yang, Grünewald, & Meinel, 2012), but we want to make the outline more logical and friendly to the users, rather than simply listing the slide titles.

Another problem is the length of quite a lot of online lectures. It is not comfortable for people to stare at a computer screen for a long time, especially when they just have interest in some parts of the lecture rather than the whole. As a result, to segment the lecture video becomes necessary. The traditional segmentation methods for natural videos (Koprinska & Carrato, 2001; Sidiropoulos et al., 2011) are not suitable for lecture videos, which lack scene changes. Instead, some research on how to segment lecture videos focus on blackboard writing (Onishi, Izumi, & Fukunaga, 2000; Friedland & Rojas, 2008), speech recognition (Lin, Chau, Cao, & Nunamaker, 2005; Repp & Meinel, 2008) or additional transcript (Imran et al., 2012).

But in our case, we previously intend to generate a logical outline by extracting content from the synchronized slides of the lecture videos. Outline consists of texts, and there is a lot of research works on how to segment the text (Hearst, 1997; Cardoso, Taboada, & Pardo, 2013). But the outline extracted from slides is special, which can be easily further split into subtopics consisted by the slide group, and the items in the outline can undoubtedly contain some time information. Then the subtopics with time tags in the outline can be perfectly transformed into the segments of the lecture video.

In this paper, we propose a solution to generate outline for lecture videos based on OCR result. OCR technology enables us to get the textual data from the slides automatically, with a fairly good accuracy, approximately 92% about the characters and 85% about the words (Yang, Siebert, Luhne, Sack, & Meinel, 2011). Then we can reconstruct each slide into an up-to-3-level content tree by analyzing the size, location and all other possible attributes of the texts achieved by OCR process. Next we will remove repeated or useless slides and explore any logical connection between the remaining. Then slides also have 3 different levels. Along with the up-to-3-level intra-slide content tree, the complete outline generated by our solution will be stored in an up-to-6-level tree structure. Finally, an outline-based lecture video segmentation method will be employed as an additional application to the solution.

SOLUTION FRAMEWORK

Figure 1 depicts the framework of our solution. Preprocessing is the first part, which contains logo removing and text modification. Outline
Analyzing Causality Among the Service Quality, Customer Satisfaction and Behavioral Intention Variables with Respect to E-Shopping: An Empirical Take

Lighting the Fires of Entrepreneurialism?: Constructions of Meaning in an English Inner City Academy
www.igi-global.com/chapter/lighting-fires-entrepreneurialism/75720?camid=4v1a