VCSR: Video Content Summarization for Recommendation

Chi-Cheng Tsai[1], Ching-I Chung[1], Yi-Ting Huang[1], Chia-Hsing Shen[1], Yu-Chieh Wu[2] and Jie-Chi Yang[1]
[1]*Graduate Institute of Network Learning Technology,* [2]*Department of Computer Science and Information Engineering, National Central University, Taiwan*
[1]*{ken, cecilia, coral, arcadia, yang}@cl.ncu.edu.tw*
[2]*bcbb@db.csie.ncu.edu.tw*

Abstract

In this paper, the authors present a Video Content Summarization for Recommendation (called VCSR) system to auto-recommend suitable multimedia learning materials for learners. The VCSR system firstly extracts important content as summarization from input raw video data, while the generated summarization will be auto-routed to learners according to their profiles. Video captions are initially recognized using Optical Character Recognition (OCR), then a set of key passages with corresponding frame images are extracted to form a video summary. The recommendation is achieved by calculating the relevance of the video summarization for each learner. Also, this paper indicates how the VCSR system effectively plays the intermediate role in a modern digital library.

1. Introduction

Multimedia instruction has recently become a promising information source to the traditional instruction. Many studies reported that multimedia content is useful for learning and teaching, in comparison to traditional in-class, and text-based learning [1] [7]. The use of video is a more effective medium for learner motivation, attention, and satisfaction [1]. Unlike text, reading videos requires much more time since it is displayed linearly. Video summarization therefore enables the learner to skim through the video content.

With the rapid growth of video nowadays, it is a difficult task of acquiring appropriate video from huge amount of videos. Adaptive recommendation is mainly designed to help learners in filtering non-relevant information. Traditionally, these tasks, for example content annotation and recommendation in e-learning are done manually, which is very time-consuming. The situation is even worse for video content annotation since it could not be read nonlinearly. Therefore, there is a strong demand for automatic video summarization and recommendation.

Automatic text summarization aims at abstracting important sentences from source document. However, these techniques focus on generating summaries from news-like articles that are usually short and coherent. On the contrary, video-based summarization techniques offer a sketch with description of an object [2]. Nevertheless, they may not work well for text-based video and also not useful for learners due to ignore lexical information. Besides, the traditional video-based summarization does not attach to educational purpose. On the other hand, learning material recommendation via comparing the learner profiles can provide adaptive learning objects easily [3] [4]. The learning object features in these mechanisms are pre-defined where the content of features is usually just a simply text description. In modern digital library scenario, readers could select their interested films as their learning content, however, traditional text-based recommendation does adopt merely article names and titles while ignores the important contents inside the learning materials.

In this paper, the authors present a VCSR (Video Content Summarization for Recommendation) system to auto-recommend suitable multimedia learning materials for learners. The VCSR system firstly extracts video content as summarization, while the generated summarization with corresponding frames was collected. These materials are combined into the hypermedia documents and auto-recommend to learners. The system also sends the hypermedia document as email to learners in response to their profiles. Unlike traditional recommendation methods, the VCSR does not only recommend video titles, but also included extracted important content that contain summarization and corresponding image frames.

While dealing with a great deal of videos, the VCSR can extract summarization rapidly and save time. Besides, it can recommend learners video

Seventh IEEE International Conference on Advanced Learning Technologies (ICALT 2007)
0-7695-2916-X/07 $25.00 © 2007 IEEE

materials probably related to what they have learnt and taught. Thus, learners can avoid much unnecessary information and quickly take the new video information they need.

2. System Overview

An overview of the proposed VCSR system is illustrated in Figure 1. Once a new video is incoming, the *Video OCR Module* starts to recognize captions as video caption document. These documents are then passed through the *Summarization Module* and the summary document for the video is generated by extracting the key passages. Finally, the video recommendation emails are generated by the *Recommendation Module*, which estimates the relevance for each learner according to their profiles. Each of the three modules is described as follows.

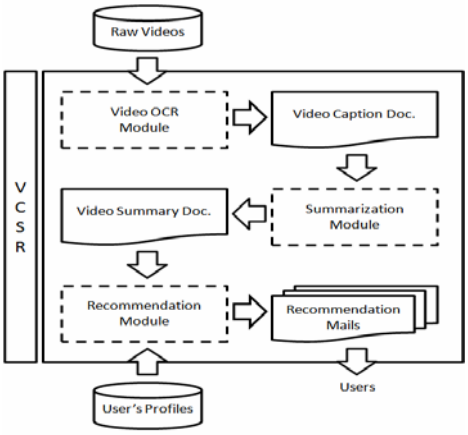

Figure 1. VCSR system overview

First, *Video OCR Module* processes the input video frame sequence and recognizes all the caption words. Video images can provide rich visual information to people, while the caption information often plays an important role for understanding. In many educational films, for example *Discovery* and *National Geography*, usually rich caption content are presented which is very close and sufficient to describe current video scenario. By employing off-the-shelf video Optical Character Recognition (video OCR) techniques [5] [6], these caption words could be automatically identified. In this paper, the authors use the OCR systems as [5] [6] for content extraction. As reported by [5], the performance of video OCR was about 70-80%. Finally, the recognized words formed the video caption document.

Second, *Summarization Module* processes the video caption document from the previous module. Over the past few years, text summarization was developed well. However, most of the traditional summary generation methods aim to extract set of key sentences from documents, which are not the case for video recommendation. Because the key sentences in the video are usually not what users concerned with and interested in. In this paper, the authors adopted the Q/A-based (Question & Answering) approach [6], which is more likely to route what users want to know, for generating summaries. In this module, the video caption document is initially segmented into five segments based on the time sequence analysis [5]. For each segment, the video Q/A system [6] extracts passage-level answers to form the video summarization. Here, the authors assume that using the answers of each segment could be more complete and comprehensive for video content understanding. The top-5 ranked video summarizations are provided for users.

Third, *Recommendation Module* processes the video summarization generated by the above two modules. In this module, it compares the video summarization with the user profiles, which record user's personal information, such as email and interest. An XML-based format is used to store user's profile description. Figure 2 illustrates a fragment of user profiles. As shown in Figure 2, the XML file depicts an example of profile. Users can modify their profile at anytime. For the recommendation, the authors focus on calculating the relevance of the video summarization for users. In order to match users more effectively, the authors integrate video name, video description, and video summarization as the sources for comparing with the profiles. The similarity measurement is estimated with cosine value. The higher relevant score the more likely this video summarization is what the user interests in. If the likelihood score exceeds a threshold, it will send the auto-generated video recommendation emails to the users.

```
<profiles>
        ...
        <user id="UCSR_00011">
                <firstname>靜怡</firstname>
                <lastname>鍾</lastname>
                <email>cecilia@cl.ncu.edu.tw</email>
                <interest>歷史</interest>
        </user>

        <user id="UCSR_00012">
                <firstname>其澈</firstname>
                <lastname>麥</lastname>
                <email>955204005@cc.ncu.edu.tw</email>
                <interest>軍事</interest>
                <interest>醫學</interest>
                <interest>天文</interest>
        </user>
        ...
</profiles>
```

Figure 2. Extract from user's profiles

By combining with these three modules, VCSR system can automatically generate and send the video recommendation emails when a new video incomes. The entire process described above is automation in the proposed VCSR system.

3. System Usage

In order to explain the system usage, the authors provided the following scenario as an example for how VCSR system helps to recommend new video.

There are often multimedia video materials ordered newly in a library, such as Discovery videos. However, if the library wants to make a video introduction, it can only refer to the simple description of the video cover. Otherwise, the library needs to spend more efforts and time viewing and annotating the video. In addition, it is another issue that how to recommend the new video to those people who need.

Using VCSR system, all of the above problems can be turned into automated process. By VCSR system, librarian could input new video as a source and then system automatically produce video summarization with corresponding image frames directly. This solution could save time and human efforts. Next, based on the profile information, system will automatically send emails to users who may need this information. The email content is composed of video cover, video description, hyperlinks to top-5 ranked video summaries, and the video summaries: summary text and key frames in corresponding to the text (see Figure 3).

4. Conclusion

In this paper, the authors present a novel automatic video summarization and recommendation system, call VCSR. The system can automatically extract video summarization from raw video data. In addition, it can automatically recommend videos based on user's profiles.

In the future, the authors plan to integrate the function of online broadcasting movie clips related to the summarization in VCSR. Based on each user's profile information, the authors also plan to customize the video summarization for each user. It means that different user can obtain different summarization even though the video source is the same. Finally, an effective algorithm to improve OCR errors is necessary for upgrading the quality of the summarization of VCSR.

5. References

[1] H.J. Choi, and S.D. Johnson, "The Effect of Context-Based Video Instruction on Learning and Motivation in Online Courses", *The American Journal of Distance Education*, 19(4), 2005, pp. 215-227.

[2] Y. Liu, and F. Li, "Semantic Extraction and Semantics-Based Annotation and Retrieval for Video Databases", *Multimedia Tools and Applications*, 2002, pp. 5-20.

[3] K.H. Tsai, T.K. Chiu, M.C. Lee, and T.I. Wang, "A Learning Objects Recommendation Model based on the Preference and Ontological Approaches", *IEEE International Conference on Advanced Learning Technologies*, 2006, pp.36-40.

[4] R. Jin, L. Si, C. Zhai, and J. Callan, "Collaborative Filtering with Decouple Models for Preferences and Ratings", *International Conference on Information and Knowledge Management*, 2003.

[5] Y.C. Wu, Y.S. Lee, J.C. Yang, and S.J. Yen, "A New Passage Ranking Algorithm for Video Question Answering", *Advances in Image and Video Technology*, 2006, pp.563-572.

[6] Y.S. Lee, Y.C. Wu, and C.H. Chang, "Integrating Web Information to Generate Chinese Video Summaries", Software Engineering and Knowledge Engineering, 2005, pp. 514-519.

[7] T.P. Mackey, and J. Ho, "Exploring the relationships between Web usability and students' perceived learning in Web-based multimedia (WBMM) tutorials", *Computers & Education*, 2006.

Figure 3. An illustration of VCSR-email content