Abstract

In this paper, we present the VCSR (video content summarization for recommendation) to automatically recommend suitable multimedia learning materials for learners. The VCSR firstly extracts important contents as summaries from input raw video data, while the generated summaries auto-routes to the users 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 user. We also present usage concerning for learning and instruction, and a pilot study to evaluate VCSR was carried out. The pilot study revealed that VCSR has a strong positive effect on students’ interest of viewing video material. They also agreed that summarization in video recommendation mail is a good and convenient way of acquiring knowledge.

1. Introduction

Multimedia instruction has recently become a promising information source to the traditional instruction. Many researchers reported that multimedia content is useful for learning and teaching, in comparison to traditional in-class, and text-based learning. 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 the learner to filter out information. Traditionally, these tasks, for example content annotation and recommendation in E-learning are done manually, which is very time-consuming. Therefore, there is a strong demand for automatic video summarization and recommendation.

Research on automatic summarization is an important research topic, especially to automatic text-based and video-based summarization [4]. Text-based summarization focuses on generating summaries from news-like articles that are usually short and coherent. On the contrary, video content is quite different from news texts since it is not only long (roughly 7000 words) but also contains multiple sub-topics. Video-based summarization techniques, on the other hand, aim to offer a sketch with description of an object, i.e. color, shape, etc. Such techniques are often used in surveillance system and medical videos. However, 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.

Learning material recommendation can provide adaptive learning objects easily and efficiently for learners to improve the learning effect. Without recommendation mechanisms, learners will spend much time in selecting suitable learning objects. It had been studied that the automatic recommendation mechanisms is positive for learning object recommendation via comparing the user profiles [3] [6]. The learning object features in these mechanisms are pre-defined and 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. Even though some literatures [1] [4] provide the human-annotated objects, they were not automatic process. For larger video database, like
digital library, the annotation process might takes huge time on “watching” and “writing”.

In this paper, we present the VCSR (video content summarization for recommendation) to auto-recommend suitable multimedia learning materials for learners. It firstly extracts video contents as summaries, while the generated summaries with corresponding frames were collected. These materials are combined into the hypermedia documents and auto-recommend to users. 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 the video titles, but also included the extracted important contents that contain summaries and corresponding image frames.

While dealing with a great deal of videos, VCSR can extract summarization rapidly and save time. Besides, it can recommend user video materials probably related to what they learn and teach. Thus, users can avoid receiving a lot of unnecessary information and quickly take the new video information they need.

The following sections are organized as follows. Section 2 describes the proposed video summarization and recommendation system. Section 3 offers examples of system usage. Section 4 presents a pilot study. The conclusion and future work are given in section 5.

2. System Overview

An overview of the proposed video summarization and recommendation system (VCSR) 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 to a summarization module and the summary document for the video is generated by extracting the key passages. Finally, the video recommendation mails are generated by recommendation module, which estimates the relevance for each user according to the profiles. Each of the three modules is described as follows.

First, Video OCR module processes the input video frame sequence and recognizes all the caption words. Video images can provide rich visual information to us, while the caption information often plays an important role for understanding. In many educational films, for example Discovery, national geography, usually rich caption contents 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, these caption words could be automatically identified. In this paper, we use the OCR systems as [5] [7] for content extraction. As reported by [7], 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 previous steps. 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, we adopted the Q/A-based (Question & Answering) approach [5], 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 [7]. For each segment, the video Q/A system [5] extracts passage-level answers to form the video summarization. Here, we assume that using the answers of each segment could be more complete and comprehensive for video content understanding. We also provide top-5 ranked video summarizations for users.

Third, recommendation module processes the video summarization generated by the above two modules. In this stage, it compares the video summarization with the user’s profiles, which record user’s personal information, i.e. email, interest, and so on. We use an XML-based format to store each user’s data. Figure 2 illustrates the fragment of user profile. As shown in Figure 2, the XML file depicts an example of profile. Users can modify their profile at anytime. For the recommendation, we focus on

---

1 The online demonstration version of this video Q/A system can be found at: http://140.115.112.118/bcbb/TVQ63/index.htm
calculating the relevance of the video summarization for users. In order to match users more effectively, in this paper, we integrate video name, video description, and our video summarization as the sources for comparing with the profiles. 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 mails to the users.

Figure 2. Extract from user's profiles

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

3. System Usage

In order to explain the system usage, we proposed the following scenarios as examples for how VCSR system helps to recommend new video, and how VCSR system helps person to learn and aid instructional design.

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 annotation to 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 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 mails to users who may need this information. The mail 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).

The assistance of individual learning and education of VCSR system usage is described as follows:

Scenario1: Jerry is a junior high school teacher. He believes multimedia content is more effective than textbooks since it provides not only text but also videos, images, etc. Thus, he often makes use of the multimedia video as the supplementary material. In the past, the video database in library might not be all correlated to the courses Jerry teaches, and Jerry needed to spend extra time and attention on the library message of new incoming video. Beside, Jerry can only judge if the new video is correlated with courses according to simple description of the video. But the library has VCSR system now. Jerry filled out his personal information in the profile of VCSR, and wrote down the curriculum related keywords in the interest column of the profile. When the new video was incoming, the library would mail newsletter for users automatically. After getting the mail, Jerry could understand the outline of the new video easily, and
save much time to pick supplemental material, and then judge further whether this video is relevant information to the courses or not.

Scenario 2: Sandy is an undergraduate student. She enjoys watching TV like HBO when she has free time. In the past, she didn’t like to do her homework because she was not interested in serious textbooks. However, she becomes studious when library has begun to serve recommendation. Because the video that library recommended are related with her study, she receives content invisibly. As soon as the system mails the newest information to her, she will read the information about the summarization of newest video in the library probably appropriate for courses she studies and borrow what she wants to watch. Thus, the more video she watches the more knowledge she increases.

4. Pilot Study
4.1. Methodology
Using VCSR, a pilot study was designed to initially evaluate the following two questions:

A1. Would users think summarization in video recommendation mail is a good and convenient way of acquiring knowledge?
A2. Would VCSR mail attracts users to spend more time reading it, and further, they want to view the video that the mail recommends to users?

The Discovery video data is one of the popular video learning materials. 181 Discovery films were selected as the VCSR video sources. 20 students participated in the experimental design. These students were randomly assigned to two groups. Each student might receive at least two recommendation mails according to their profiles, the emails for the experiment group contain the video cover image, short descriptions (derived from the video cover), and the extracted summaries (roughly < 15 sentences and the image frames). Our VCSR also generated not only one summary, but also top-5 ranks, which enable users to further browsing the other ranks. As seeing in Figure 3, the hyperlinks link to the other four summary pages. On the contrary, the control group did receive the recommendation mails that contain only the video cover and short descriptions. We also monitored users and recorded the relative click through rate (CTR) and the reading time when a new recommendation mail comes. Finally, ten students were interviewed for asking their thoughts about the system.

A questionnaire was designed to evaluate VCSR system and was distributed to all subjects after they read the mail. The answers’ choices for all the questions ranged from 1 (strongly disagree) to 5 (strongly agree). The questionnaire was composed of 6 questions as follow.

Q1: I think it is very convenient that I can get the video information from the recommendation mail service.
Q2: I can understand the main outline of this recommended multimedia video after reading recommendation mail.
Q3: I want to view the recommended video after reading recommendation mail.
Q4: After reading mail, I think that the automatic recommendation service is a good and convenient way of acquiring knowledge about history, geography, science.
Q5: The information of recommendation mail let me reduce the time of collecting knowledge.
Q6: I am glad to receive recommendation mail in the future.

4.2. Results and discussion
We present the numerical result, as shown in Figure 4. Figure 4 is the questionnaire assessment result. Table 1 lists the average CTR rates and reading time of mail in two groups.

![Figure 4. The questionnaire result](image)
students in experiment group feel that they could understand more content of the recommended video (see Q2, 3.7 > 3.4). The two groups both think the automatic recommendation service is an easy way to get video information (Q1). The above description initially confirms that the answer of A1 would be positive.

Q3 is related to A2. According to the result of Q3, experiment group agree that they want to view the recommended video after reading recommendation mail (4.0). However, the rate of control group is only 3.7. We found that the video summarization is more attractive for students and also improves their motivation to view videos.

The relative click through rate (CTR) and reading time of mail are related to A2.

### Table 1. Average CTR and reading time of mail

<table>
<thead>
<tr>
<th></th>
<th>AVG relative CTR rates of mail</th>
<th>AVG reading time (sec / per mail)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment group</td>
<td>1.18</td>
<td>100.21</td>
</tr>
<tr>
<td>Control group</td>
<td>1</td>
<td>20.09</td>
</tr>
</tbody>
</table>

The experiment group spent 100.21 seconds reading VCSR mail on average (see Table 1), and the control group spent 20.09 seconds to read the recommendation mail. The result suggests that the students are glad to spend more time reading the video information when summarization content is attached to the mail. We also found that students usually click the other four extracted summaries not limited to the top-1. On the other hand, the relative click through rate (CTR) of mail of experiment group (1.18) is higher than that of control group (1). The observation suggests obviously that the video summarization of mail engages the students to read the recommendation mails more than simple description one.

Based on the interview, some interesting discussions were observed:

1. Many students in the experiment group indicated that when they found the key frame corresponding to summarization in the mail, they were attracted to click the frame image and strongly wanted to watch movie clips related to the summaries. They also clicked the hyperlinks to summaries of other rankings for further browsing. All of the above represent that the summarization encourages and motivates users to acquire more information about the video.

2. Some students in the experiment group reflected that sometimes there are lexical or semantic errors in the video caption documents. They were occasionally confused by the summarization. These errors are mainly caused by the OCR errors. In practical, the video OCR is difficult to be perfect, since lots of blurs and noise affect the performance. The result shows that our video OCR module should be improved. In the future, we plan to employ more advance video OCR techniques and speech recognition models to reduce the errors.

### 5. Conclusion

In this paper we present a novel automatic video summarization and recommendation system. The system can automatically extract video summarization from raw video data. In addition, it can automatically recommend videos based on user’s preferences. The pilot study reveals that the video summarization is a good and convenient way of acquiring knowledge. Moreover, VCSR mail attracts users to spend more time reading it, and further, they want to view the video that the mail recommends to them.

As the future work, we can integrate the function of online broadcasting movie clips related to the summarization in VCSR. Based on each user’s profile information, we also need to customize the video summarization for each user. It means that different user can obtain different summarization he is interested in even though the video source is the same. Finally, to upgrade the quality of the summarization, an effective algorithm to improve OCR errors is necessary for VCSR in the future.

### 6. References


