

## A Rapid Auto-Indexing Technology for Designing Readable E-Learning Content

Pao-Ta Yu, Yuan-Hsun Liao, Ming-Hsiang Su, Po-Jen Cheng and Chun-Hsuan Pai

Volume 13, Number 5, December 2012

Special Issue: Technology-Enhanced Information Retrieval for Online Learning

URI: <https://id.erudit.org/iderudit/1066985ar>

DOI: <https://doi.org/10.19173/irrodl.v13i5.1246>

[See table of contents](#)

Publisher(s)

Athabasca University Press (AU Press)

ISSN

1492-3831 (digital)

[Explore this journal](#)

Cite this article

Yu, P.-T., Liao, Y.-H., Su, M.-H., Cheng, P.-J. & Pai, C.-H. (2012). A Rapid Auto-Indexing Technology for Designing Readable E-Learning Content. *International Review of Research in Open and Distributed Learning*, 13(5), 20–38. <https://doi.org/10.19173/irrodl.v13i5.1246>

Article abstract

A rapid scene indexing method is proposed to improve retrieval performance for students accessing instructional videos. This indexing method is applied to anchor suitable indices to the instructional video so that students can obtain several small lesson units to gain learning mastery. The method also regulates online course progress. These anchored points not only provide students with fast access to specific material but also can link to certain quizzes or problems to show the interactive e-learning content that course developers deposited in the learning management system, which enhances the learning process. This allows students to click on the anchored point to repeat their lesson, or work through the quizzes or problems until they reach formative assessment. Hence, their learning can be guided by the formative assessment results. In order to quickly find the scene to index, some specific description of it was needed. Actually, most of the instructional videos were recorded by teachers, and were part of their PowerPoint presentations. Based on the features of the PowerPoint slides in the instructional videos, such as the title or page number, the specified scene can be found. Since we used specific scene descriptions, it was easy to employ the rapid scene detection method using an image filter and Sobel mask. Finally, we applied an experimental design to check the precision of the scene detection and evaluate user satisfaction. The results showed that rapid scene indexing can definitely assist learners in their online learning; that is, it gives them better learning mastery and provides regulation for the online learning environment.

Copyright (c) Pao-Ta Yu, Yuan-Hsun Liao, Ming-Hsiang Su, Po-Jen Cheng and Chun-Hsuan Pai, 2012



This document is protected by copyright law. Use of the services of Érudit (including reproduction) is subject to its terms and conditions, which can be viewed online.

<https://apropos.erudit.org/en/users/policy-on-use/>

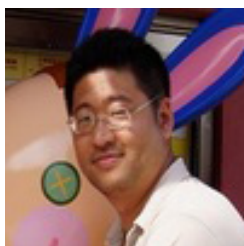
**érudit**

This article is disseminated and preserved by Érudit.

Érudit is a non-profit inter-university consortium of the Université de Montréal, Université Laval, and the Université du Québec à Montréal. Its mission is to promote and disseminate research.

<https://www.erudit.org/en/>

## A Rapid Auto-Indexing Technology for Designing Readable E-Learning Content



Pao-Ta Yu, Yuan-Hsun Liao, Ming-Hsiang Su, Po-Jen Cheng, and Chun-Hsuan Pai  
National Chung Cheng University, Taiwan, Province of China

### Abstract

A rapid scene indexing method is proposed to improve retrieval performance for students accessing instructional videos. This indexing method is applied to anchor suitable indices to the instructional video so that students can obtain several small lesson units to gain learning mastery. The method also regulates online course progress. These anchored points not only provide students with fast access to specific material but also can link to certain quizzes or problems to show the interactive e-learning content that course developers deposited in the learning management system, which enhances the learning process. This allows students to click on the anchored point to repeat their lesson, or work through the quizzes or problems until they reach formative assessment. Hence, their learning can be guided by the formative assessment results.

In order to quickly find the scene to index, some specific description of it was needed. Actually, most of the instructional videos were recorded by teachers and were part of their PowerPoint presentations. Based on the features of the PowerPoint slides in the instructional videos, such as the title or page number, the specified scene can be found. Since we used specific scene descriptions, it was easy to employ the rapid scene detection method using an image filter and Sobel mask. Finally, we applied an experimental design to check the precision of the scene detection and evaluate user satisfaction. The results showed that rapid scene indexing can definitely assist learners in their online learning; that is, it gives them better learning mastery and provides regulation for the online learning environment.

**Keywords:** Scene detection; scene indexing; instructional video; anchored or access point; mastery learning; regulated learning; e-learning; PowerPoint

## Introduction

As a result of the rapid advance of asynchronous learning, there are not enough interactive learning objects available to meet the numerous demands of learners (Girasoli & Hannafin, 2008). The fastest and easiest way to provide an adequate amount of e-learning content is to record teachers' presentations in a classroom or studio and then directly put those recordings into a learning management system (LMS) (Mittal, Pagalthivarathi, & Altman, 2006). Using a high speed Internet connection and streaming technology, learners can repeatedly view the recorded instructional videos from anywhere, at any time (Blanchette, 2012). However, this kind of streaming data lacks flexibility and interactive capability. Therefore, a user-friendly interface is required to let students easily capture any segment of the recorded instructional videos (Liu & Kender, 2004).

Most users who spend a long period of time on a course Web site often lose their concentration and start surfing elsewhere. A regulating mechanism is required to pull students' attention back to their online learning. However, a mechanism to improve the interactivity of the recorded instructional videos is also needed. This motivated us to apply the technique of image indexing to the recorded instructional videos in order to divide them into segments with several entry points. Those entry points and the technique of image retrieval easily provide interactive features for the lessons (Zhang, Wang, Shi, & Zhang, 2010).

Most of the streaming data can run on a web browser in its corresponding embedded media player (Bouvin & Schade, 1999). This embedded player can interpret a given XML or HTML document and then follow the instructions to display the correct video content (Sun, Kim, & Kuo, 2005). In this experiment, we gave the media player a series of instructions about time segments, where each segment corresponded to a short clip of the recorded instructional video (Feng, Lu, & Ma, 2002). As we knew, mastery learning requires teachers to partition content into several small logically sequenced units. Then, teachers guide students on the learning path through one of these small units until students master that particular content (Block & Burns, 1976). Therefore, we also needed a pattern recognition method to determine the logical viewing order combined with the image indexing technique to find the access point of each unit of the recorded instructional video (Antani, Kasturi, & Jain, 2002). We will give a detailed description of this process in the Method section of this paper. After we found a set of access points, we assigned this information to an XML document using appropriate tags (Keyvnpour & Izadpanah, 2011). The media player can then follow the instructions of this XML document to provide some interactive options, including giving a student access to online supervision.

During the self-learning component of a distance education lecture, most learners cannot understand the complicated concepts or formulas in a short period by simply reading the e-learning content (Chiu, Chow, & Chang, 2007). This inability to grasp the lesson causes their learning progress to fall far behind the lecture schedule. Although the asynchronous e-learning environment permits learners to arbitrarily control the course process so they can repeat their lessons as needed, learners cannot successfully discipline themselves to follow the learning schedule just by reading monotonous content (Bergamin, Werlen, Siegenthal-

er, & Ziska, 2012). Therefore, we designed a mechanism of regular testing which requires learners to answer questions corresponding to pop-up information triggered when they click on an access point found by the indexing mechanism. If learners select the wrong answer, they have to play the video from the beginning access point to repeat the lesson until they reach the expected learning level. The behavior of repeat learning is also described in the XML document, and the video can be based on the document to play at a specific time.

This paper presents an automatic method for detecting the changes of the PowerPoint feature and embedding this instructional video in an online course as an interactive component. This is a simple and useful way to tailor the learning context and help learners to enhance meaningful learning. The rest of this paper is organized as follows: First, we present a literature review and the auto-indexing technology. Next, we discuss its application and the experiment's results. Finally, we present our concluding remarks.

## Literature Review

### Mastery Learning

Mastery learning is an effective way to make learners reach higher learning levels under the appropriate conditions (Bloom, 1982). Its proponents argue that all students can reach high levels of mastery of instructional material with the right support. Bloom (1968) stated that well organized teaching materials and effective management of a student's learning process are two factors that help individuals achieve successful mastery. According to Bloom (1976) and others (Block & Burns, 1976; Fuchs, Fuchs, & Tindal, 1986), mastery learning can be accomplished by following specific procedures. The first step is to divide the concepts and materials into relatively small and sequential learning units. Each unit is associated with concrete learning objectives, and the structure is organized by partitioning difficult content into several smaller units that are easier to grasp. After teaching each unit, the instructor conducts a formative assessment to determine whether the learners have reached the desired level or not; the assessment also provides students with feedback on their learning (Yang & Liu, 2006). The learners who have not mastered the unit begin a process of remedial activities or corrections to assist them in achieving this goal. This learning process is shown in Figure 1. Mastery learning is a suitable approach to use with students due to their weak discipline in self-directed learning settings.

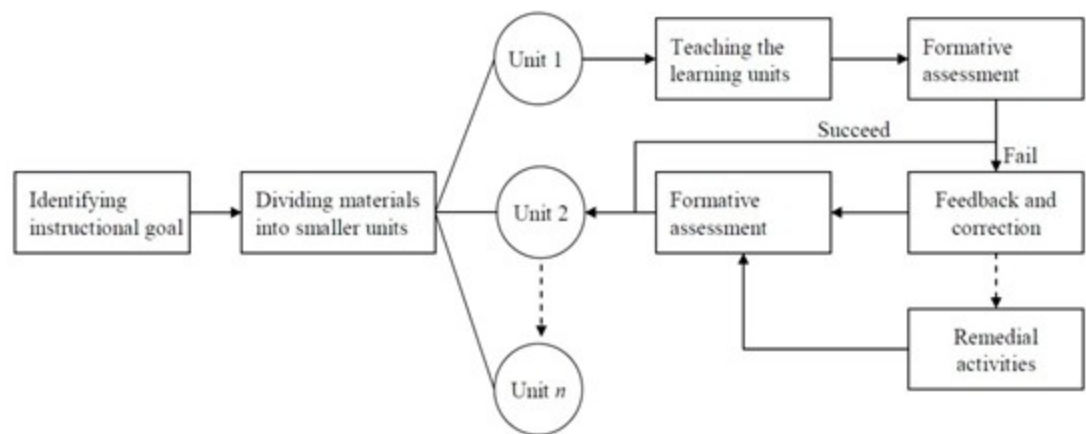


Figure 1. The strategy for mastery learning.

## Detection Technology

Multimedia indexing and retrieval has become a challenge in light of the huge amount of data that must be organized. This is not a trivial task for large visual databases. Hence, dividing videos into brief segments might improve the likelihood of completing this task (Sakarya, Telatar, & Alatan, 2012). Enabling users to search all digital multimedia data in a library and access the relevant information requires efficient analysis, indexing, and re-organizing for suitable browsing (Naphade, Mehrotra, Ferman, Warnick, Huang, & Tekalp, 1998; Koprinska & Carrato, 2001; Gargi, Kasturi, & Strayer, 2000; Song & Ra, 2001; Hanjalic, 2002; Yuan, Wang, Xiao, Zheng, Li, Lin, & Zhang, 2007; Asan & Alatan, 2009).

With the development of the low-cost digital recorders, it is easy to obtain various kinds of recorded data, such as traveling or instructional videos. Since the instructional videos that we deal with usually include different lesson segments, the video files must be pre-processed first (Ahanger & Little, 1996). In this study, the proposed approach addresses the problem of instructional video segmentation by using scene detection. Each frame in the instructional video is a snapshot of a PowerPoint slide. The feature area selector is designed to choose features such as the title or page number of the slide. In this way, the automatic detection technology only examines the feature area of the image. The technology performs better compared to other methods; indeed, it is faster and simpler.

This detection technology is based on the idea of finding the dominant scene in the video according to the selected feature. The approach employs the Sobel mask operator, which is commonly used to find the edge of grayscale two-dimensional images. This has the ability to find the gradient of neighboring points. If we construct a line“image”using the z-coordinates of the scanned points as the gray scale, then a Sobel mask may detect any sharp changes in the gray values (Gonzalez & Woods, 1992).

## Method

### Auto-Indexing Technology

Our proposed method provides an automatic scene detection process to find the scenes of an instructional video and provide the corresponding access points. After all the data to be streamed is fixed with suitable anchor points, the rapid indexing technology can translate it to e-learning content in HTML format using those anchored points, as shown in Figure 2. For convenience, we call this process an RIT system.

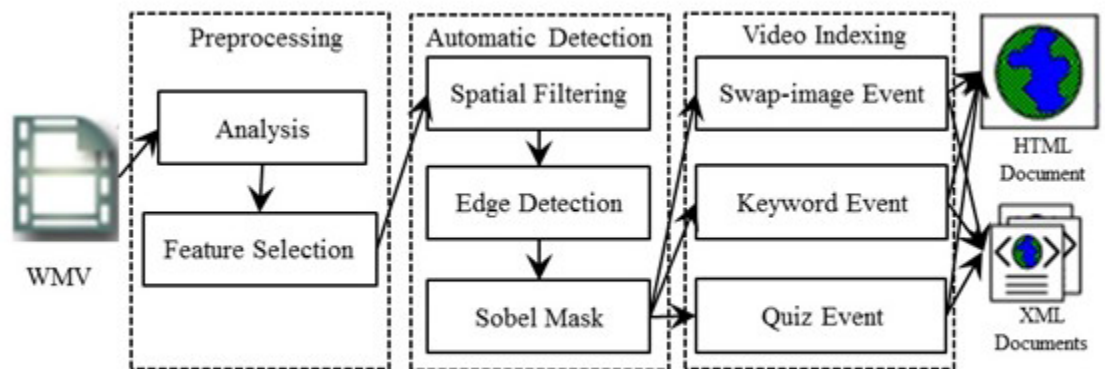


Figure 2. The auto-indexing method.

An illustrated example is shown in Figure 3. Based on snapshots obtained using automatic detection, the lecture is partitioned into four scenes and the access point of each scene can be easily found in the streaming data.

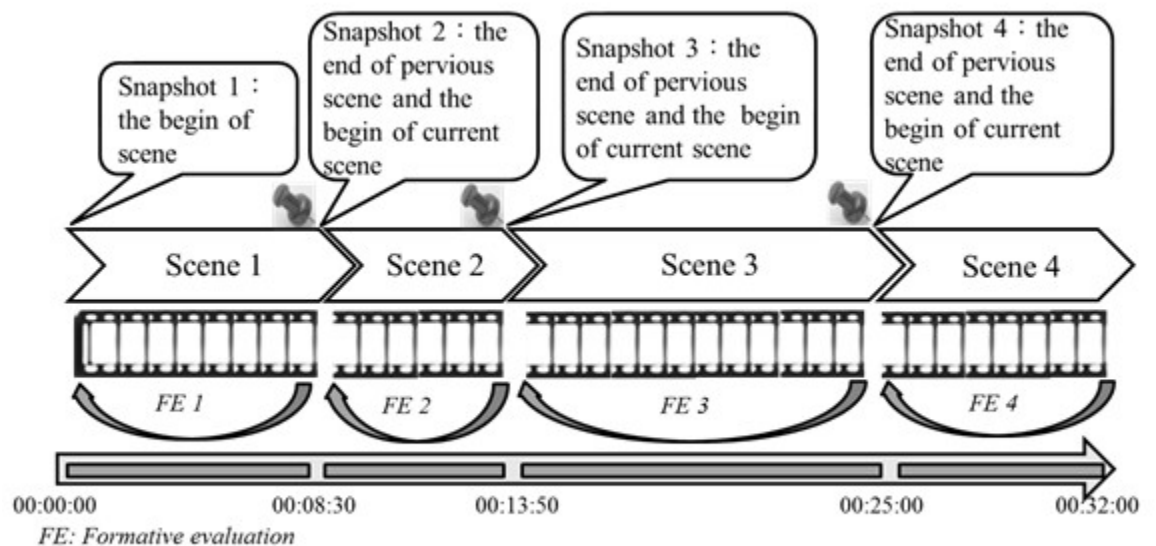


Figure 3. Diagram of slides and their divisions in the timeline.



As shown in Figure 4, we applied spatial filtering, edge detection, and a Sobel mask in our method to process the images from the instructional video, using a binary searching algorithm to find the anchored points. The basic rule is that if the comparison result of the first image and the last image in an interval is the same, the process will determine that the interval has no target image. Otherwise, the target image exists in the interval. This function is designed to provide a user with the feature selection area and can automatically detect the proper feature area.

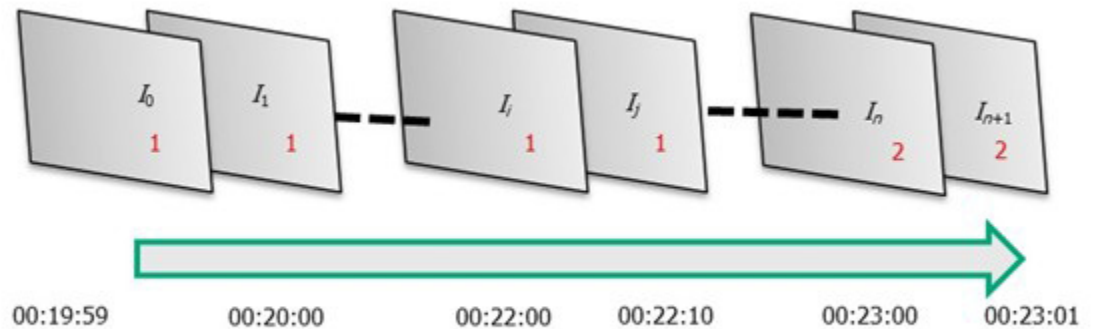


Figure 4. Automatic scene detection approach.

## Video Indexing

Early video indexing methods were based on detecting shot boundaries and extracting key frames from which visual features like color, texture, shape, or edge were extracted to be used as indices (Smoliar & Zhang, 1994; Deng, Manjunath, Kenney, Moore, & Shin, 2001; Manjunath & Ma, 1996; Park, Jeon, & Won, 2000).

Whereas initial research in this area was directed at image databases, it did not take long for researchers to start using these methods to address the similar issues of retrieval and classification for video data. These tasks are guided by suitable indexing methods based on the content of the video itself or semantic descriptors that could be extracted from the data (Yi, Rajan, & Chia, 2005). Our video indexing technology created a script XML file within Windows Media Video (WMV) (Liao, Tsai, Su, Li, & Yu, 2011). The event information is defined by the time the event occurred and is embedded in the WMV files. Three types of event information are defined: the swap-image event, the keyword event, and the quiz event. The swap-image event consists of a shot image and access point for each slide. The shot image is the final snapshot of each slide, and the access point of each slide is its ending time. Included in the keyword event is both text content and the access point. The text content is made up of the keyword information, and the access point for the time of each keyword is used. The quiz event is made up of text content and an access point; the former consists of the questions and answers and the latter the time of the question period.

When the user moves forward through the lesson, the system will automatically initiate a swap-image event to change the current slide to the previous page, as shown in Figure 8. Then the keyword event will be triggered to display pop-up information, which always

provides keywords (or hints) and is edited by teachers. In the instructional video, the quiz event will be triggered to display questions in order to examine the learner's formative status.

In this study, the RIT system used XML tags to define the information elements. The information elements contain the type of event, the command of execution, the time the event occurred, and the playback function. The instructional video was played in accordance with the information elements, as shown in Figure 5.

```

- <Topic>
  <TopicTitle> 基本理論</TopicTitle>
  <Keyword/>
  <TopicTime>00:00:50</TopicTime>
  <TopicContent> 基本理論 2 寫學會 什麼是認知理論及其操作流程? 什麼是多媒體學習理論? 那一些認知原則? 說說看, 你會學到什麼? </TopicContent>
  <TopicMemo/>
  <Thum> Thum000094000.jpg</Thum>
  <TopicThum> 2</TopicThum>
  <TopicPage> 2</TopicPage>
</Topic>
- <Topic>
  <TopicTitle/>
  <Keyword/>
  <TopicTime>00:01:34</TopicTime>
  <TopicContent> 3 認知理論(Cognitive Theory) 雙軌或雙頻道(Dual channels)是人類慣用的接收模式。主要有 Auditory/Verbal channel
    Visual/Pictorial channel 有限的記憶能力 Working memory 主動的處理能力 Selecting Organizing integrating</TopicContent>
  <TopicMemo/>
  <Thum> Thum000376000.jpg</Thum>
  <TopicThum> 3</TopicThum>
  <TopicPage> 3</TopicPage>
</Topic>
- <Topic>
  <TopicTitle/>
  <Keyword/>
  <TopicTime>00:06:16</TopicTime>
  <TopicContent> 4 認知處理流程理論(Cognitive Processing Flow) </TopicContent>
  <TopicMemo/>
  <Thum> Thum000709000.jpg</Thum>
  <TopicThum> 4</TopicThum>
  <TopicPage> 4</TopicPage>
</Topic>

```

Figure 5. The content of an XML file.

## Detection Process

Teachers select a video as a source and use the RIT system to retrieve the images of the slides from the video, as shown in Figure 6. The feature areas determined by the feature area selector serve as the parameters for retrieving images. After the images are retrieved with the RIT system, anchored points are created in the video. In addition, the system provides teachers with the ability to arbitrarily retrieve images from the video.





Figure 6. The interface of the detection process.

## Embedding Interactive Information

The RIT system analyzes the access points at which it retrieved the slides and records the related information in a temporary XML document, which stores three types of embedded information: image, quiz, and keyword. Teachers can add to the content of a learning object or edit it by filling in the interface's fields; the RIT system not only saves this new content in XML file format but also records the embedded information. Finally, after teachers click the button to embed related information into a video, the system generates a postprocessed WMV file and a few XML documents. The system's outputs are integrated into a learning object by the predesigned HTML document.

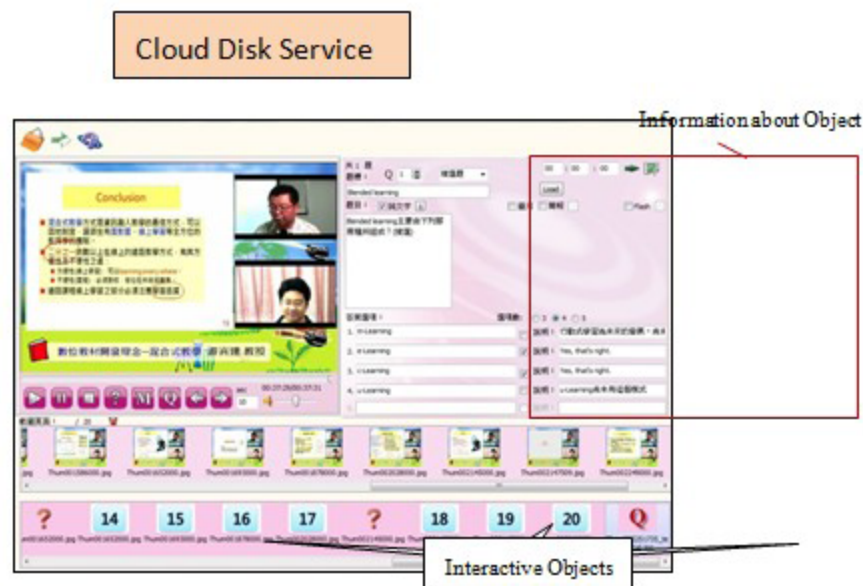


Figure 7. The interface of embedded interactivity information.

## Learner Interface

Through the predesigned XML document, teachers can integrate the postprocessed video and several XML documents into a learning object. The learner's interface for e-learning courses is shown in Figure 8.



Figure 8. Learner interface of an e-learning course.

The instructional video component is the postprocessed video generated by the teacher, and the interactive index area shows the complete course content. The previous page component shows the previous slide or video image to assist students by reminding them of the content that came before. When the instructional video component sends the image message, the previous page component changes to the specified image. When the instructional video component sends the quiz message, the interactive object component will open a new window and the playback of the instructional video component will be suspended. If students answer the question correctly, the video will continue to play; otherwise, the media player will replay the related video clips. Figure 8 shows the quiz window, one of the interactive object components.

## Learning Activity

When students' learning activities are ongoing, they can easily focus on the video content rather than skip it (as shown in Figure 9). When the quiz window appears, the video will be suspended. Students need to answer questions correctly in order to continue the learning activity. If they select a wrong answer, the video will be played from the start of the anchor point. The quiz window not only gives students feedback but also lets them review the content in stages. The keyword information was designed to allow teachers to place emphasis on the keywords of relevant hints or concepts. When the keyword window is open, students can continue to watch the video without interruption. The third interactive element shows the change of slide images. Students can refer to the previous slide to organize the discrete concepts.



*Figure 9. The timeline of the learning activity.*

As we have already mentioned, students spend their energy to concentrate on learning. Employing animation and interactivity principles (Mayer, 2005), we have provided the RIT system to instruct students using an interactive learning object. The interactive learning object is composed of an HTML document, a video, and several XML documents. Teachers can embed instructions into the video. When students use the learning objects, the video sends instructions according to the predesigned course and manner of progress. If students can continue to learn without frustration or distraction, they will be more likely to concentrate on their learning activities.

## The Experiment

### Detection Analysis

There are many scenes in the instructional videos, and each one represents a different unit of the course, arranged in logical order. In order to distinguish between different scenes, teachers can select a specified region of snapshot images from the instructional video to retrieve the desired features. The features retrieved from consecutive snapshot images in the same scene possess the same features. The detailed specifications are shown in Figure 10.



Figure 10. Different selection regions in an instructional video.

In order to get better detection results, we defined several selected regions for this study. Then, we tried the experiment in those regions to measure the accuracy of the retrieved features. We considered different sizes of the selected region: custom pixels, 80 x 60 pixels, 160 x 120, 320 x 240, 640 x 480, and full screen. The custom pixel size is appropriate for covering the feature region and is dynamic, in contrast to the fixed cases. We analyze the advantages of dynamic and fixed sizes under different considerations in the following section.

In the experiment, the performance was evaluated objectively by measures of precision, recall, f-measure, and misdetection rates (Shivakumara, Phan, & Tan, 2010). The scenes were classified into the following categories by our detection method.

- Truly detected key frame (TDK): The detected frames contain the corresponding key frame from instructional video. For a page from the course slide, a key frame is the frame in which the page content first appeared in the video.
- Falsely detected key frame (FDK): The detected frames contain the redundant frames besides the key frame.
- Key frame with missing data (MDK): The detected frames lose the key frames (MDK is a subset of TDK).

For each key frame in the instructional video, we also manually counted the number of Ac-

tual key frames (AKS), that is, the true key frames. The performance measures were defined as follows:

- $\text{Recall (R)} = \text{TDK} / \text{AKS}$
- $\text{Precision (P)} = \text{TDK} / (\text{TDK} + \text{FDK})$
- $\text{F-measure (F)} = 2 \times P \times R / (P + R)$
- $\text{Misdetecion Rate (MDK)} = \text{MDK} / \text{TDK}$

The experimental samples were obtained from four lectures: Digital Learning Design: Multimedia Learning Theory, Blended Learning, Mastery Learning Theory, and Design Experiences, respectively.

If the precision of the key frame detection was too low, it meant that a lot of unnecessary frames were recognized and search efficiency was poor. Moreover, if the recall rate was too low, it meant that many key frames were not detected.

Based on the difference of feature region, the precision, recall, and f-measure analysis were compared in Table 1. The precision in the custom feature region was higher than the other feature regions, with a recall rate of more than 94%.

The precision was related to the feature region. The precision and recall calculated by the f-measure proves that the custom region was better than all the other sizes. The experiment's results showed that the feature region was most effective, with better detection results, as shown in Table 1.

Table 1

*Feature Region Results for Lectures*

| Feature size(pixels) | Lecture 1 |      |      |      | Lecture 2 |      |      |      | Lecture 3 |      |      |      | Lecture 4 |      |      |      |
|----------------------|-----------|------|------|------|-----------|------|------|------|-----------|------|------|------|-----------|------|------|------|
|                      | R         | P    | F    | M    | R         | P    | F    | M    | R         | P    | F    | M    | R         | P    | F    | M    |
| 80 x 60              | 0.94      | 0.65 | 0.77 | 0.06 | 0.95      | 0.66 | 0.78 | 0.05 | 0.89      | 0.70 | 0.78 | 0.13 | 1.00      | 0.40 | 0.57 | 0.00 |
| 160 x 120            | 0.94      | 0.35 | 0.52 | 0.06 | 0.90      | 0.24 | 0.38 | 0.11 | 0.89      | 0.40 | 0.55 | 0.13 | 0.94      | 0.18 | 0.31 | 0.07 |
| 320 x 240            | 1.00      | 0.21 | 0.35 | 0.00 | 0.90      | 0.11 | 0.19 | 0.11 | 0.89      | 0.24 | 0.38 | 0.13 | 1.00      | 0.15 | 0.27 | 0.00 |
| 640 x 480            | 1.00      | 0.17 | 0.30 | 0.00 | 0.95      | 0.07 | 0.12 | 0.05 | 1.00      | 0.16 | 0.27 | 0.00 | 0.94      | 0.10 | 0.18 | 0.07 |
| Full screen          | 1.00      | 0.06 | 0.11 | 0.00 | 1.00      | 0.06 | 0.11 | 0.00 | 1.00      | 0.05 | 0.10 | 0.00 | 1.00      | 0.07 | 0.13 | 0.00 |
| Proposed             | 0.94      | 0.94 | 0.94 | 0.06 | 0.95      | 1.00 | 0.97 | 0.05 | 0.94      | 1.00 | 0.97 | 0.06 | 1.00      | 0.80 | 0.89 | 0.00 |

*Note:* R = Recall (Detection Rate); P = Precision (1 False Positive Rate); F = F-measure; M = Misdetecion Rate



## Student Satisfaction with the System

### Participants

We conducted the experiment to investigate how satisfied students were with the RIT system in the online learning course. In total, 46 college students in central Taiwan (20 men and 26 women, with a mean age of 33.1 years) who enrolled in the Digital Learning Design course participated in the experiment. The experiment ran for 12 weeks. Students watched one or two instructional videos online each week.

### Instruments

Although information technologies can theoretically be useful to support the learning process, a careful evaluation of their benefits is needed (Yengin, Karahoca, & Karahoca, 2011; Wu, Tennyson, & Hsia, 2010). In this experiment, we evaluated the RIT system from the functionality and usability perspective. We created a questionnaire based on a modified and edited version of the questionnaire DeLone and McLean used in their study (1992, 2003) and the teachers' own experiences (Hwang, Tsai, Tsai, & Tseng, 2008). We collected data from the 18-item questionnaire that the students completed at the end of the experiment. The questionnaire includes questions about system quality, information quality, service quality, usage, user satisfaction, and the net benefits of the RIT system. Each item was assessed on a 5-point Likert scale, where 1 was *strongly disagree* and 5 was *strongly agree*. The Cronbach's alpha reliability coefficient for the questionnaire was 0.89, with a criterion-related validity of 0.78.

### Procedure

One teacher and 46 students in a single class participated in a trial of the system. The class agreed to use the RIT system to support their lectures in the winter 2011 semester. The reading materials were uploaded to the learning management system (LMS). Students were able to watch the audiovisual materials produced by the RIT system. The procedure of the experiment, as shown in Figure 11, was as follows.

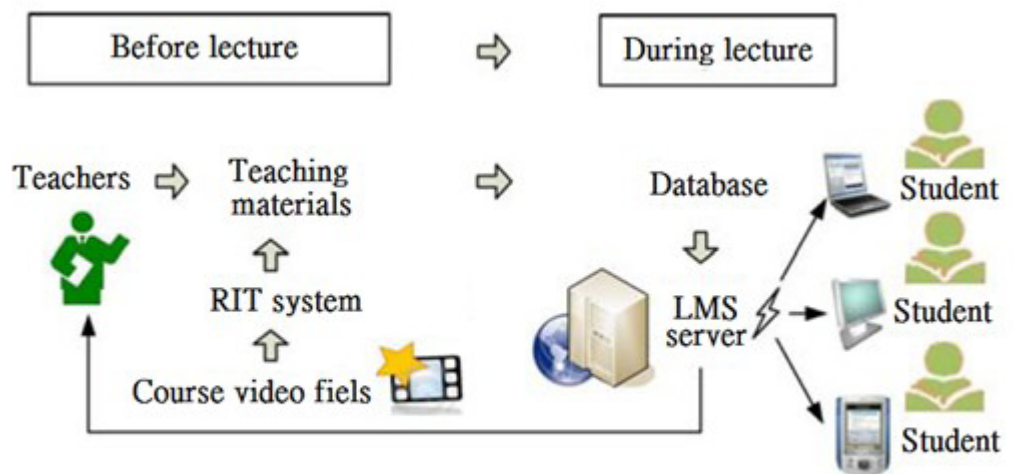


Figure 11. The RIT experiment procedure.



1. Before the lecture: Teachers prepare and produce the instructional videos, turning them into the teaching materials by using the RIT system. While preparing the course materials, teachers can subdivide lessons into multiple concepts according to the difficulty of the material. Then they can post questions for different concepts to test whether students understand them. Finally, teachers upload the materials to the LMS.
2. During the lecture: Students log in to the LMS system to watch the online materials. Once online, they can learn one concept step by step or review other lessons they have learned. In addition, students can discover which concepts they still don't understand by working through tests. Teachers can always amend course materials based on the situation of the students who are using them.
3. After the class: Students were asked to fill out the questionnaire.

## User Satisfaction

Table 2 shows that the RIT is able to effectively construct a learning environment. Most participants (95.6%) found the application to be satisfactory; the mean satisfaction score was 3.81. Of the students, 4.3% were dissatisfied with this regulated learning, possibly because they didn't like the test conditions that determined their learning status. Most students, 73.9%, were satisfied that the system's interface was user friendly. Some of them suggested the quiz window should be locked onscreen to ensure that students will answer the questions.

Of the students, 71.7% were satisfied that the application could be used and accessed with ease. Furthermore, 69.6% were satisfied that their learning progress was integrated into the learning management system so that they could search for scores through one single system. The majority of the students, 76.1%, were satisfied that the application could be operated rapidly with the interface, and 84.8% of them were satisfied with the quality of the quiz questions, which made the content clearer. Of the students, 80.5% were satisfied with the services provided by the application, such as FAQ and security. Most, 69.6%, said they liked to use the application, and 78.3% stated they would often use it in the future.

Moreover, 82.6% stated that they found the application helpful and benefited from the learning process. Respondents who chose 4 or 5 (satisfied or very satisfied) said that they used the learning application to help them practice and remember important content, not only to watch the video. Thus we can conclude that the RIT system was satisfying for most students.

Table 2

*Internal Consistency Results*

|                     | Construct               | Non-existent | Poor    | Fair      | Good      | Excellent | Alpha |
|---------------------|-------------------------|--------------|---------|-----------|-----------|-----------|-------|
| System quality      | Ease of use             | 0(0%)        | 2(4.3%) | 11(23.9%) | 29(63%)   | 4(8.7%)   | 0.871 |
|                     | Integration             | 0(0%)        | 0(0%)   | 14(30.4%) | 28(60.9%) | 4(8.7%)   | 0.875 |
|                     | Ease of access          | 0(0%)        | 1(2.1%) | 10(21.7%) | 29(63%)   | 6(13%)    | 0.875 |
|                     | Interface quality       | 0(0%)        | 1(2.1%) | 11(23.9%) | 30(65.2%) | 4(8.7%)   | 0.868 |
|                     | Rapidity                | 0(0%)        | 1(2.1%) | 10(21.7%) | 31(67.4%) | 4(8.7%)   | 0.867 |
| Information quality | Information quality     | 0(0%)        | 0(0%)   | 7(15.2%)  | 31(67.4%) | 8(17.4%)  | 0.868 |
| Service quality     | Service quality         | 0(0%)        | 0(0%)   | 9(19.6%)  | 32(69.6%) | 5(10.9%)  | 0.868 |
| Usage               | Intensity               | 0(0%)        | 2(4.3%) | 12(26.1%) | 28(60.9%) | 4(8.7%)   | 0.875 |
|                     | Frequency               | 0(0%)        | 1(2.1%) | 9(19.6%)  | 28(60.9%) | 8(17.4%)  | 0.881 |
| User satisfaction   | User satisfaction       | 0(0%)        | 0(0%)   | 2(4.3%)   | 41(89.1%) | 3(6.5%)   | 0.875 |
| Net benefits        | Helpfulness and benefit | 0(0%)        | 0(0%)   | 8(17.4%)  | 35(76.1%) | 3(6.5%)   | 0.872 |
| Mean evaluation     |                         | 3.831        |         |           |           |           |       |

## Conclusions

In this paper, we used an auto-indexing approach to create interactive learning objects. The RIT system provides automatic detection technology to assist teachers in producing learning materials. The automatic detection technology can retrieve slide images from a video and record access points. Then, the system transfers the record into information embedded in the video. Students can organize the concepts in meaningful ways, viewing the current slide with the previous one.

Students can take advantage of elements in this online course which allow them to repeat lessons until they reach a specific level. In addition, the anchored point of formative evaluation provides an appropriate way to test students and determine whether they have mas-

tered the content or not. The course data, including streaming data, text and pictures with anchored points can be integrated with the interactive learning objects in HTML format. Teachers can easily use the RIT system to give their instructional videos added value.

Finally, the results of our research show that students who use the RIT system have a more positive attitude toward the learning process. The RIT system benefits students, allowing them to review content asynchronously in the cyber classroom.

## Acknowledgment

We would like to thank the National Council of Taiwan, R.O.C., for partially supporting this research under Contract No. NSC 100-2511-S-194-001-MY2 and No. NSC 99-2511-S-194-003-MY3.

## References

- Ahanger, G., & Little, T. D. C. (1996). A survey of technologies for parsing and indexing digital video. *Visual Communication and Image Representation*, 7(1), 28–43.
- Antani, S., Kasturi, R., & Jain, R. (2002). A survey on the use of pattern recognition methods for abstraction, indexing and retrieval of images and video. *Pattern Recognition*, 35, 945–965.
- Asan, E., & Alatan, A. A. (2009, September). Video shot boundary detection by graph-theoretic dominant sets approach. *Proceedings of the 24th International Symposium on Computer and Information Sciences* (pp. 7–11). Washington, DC: IEEE Computer Society.
- Bergamin, P. B., Ziska, S., Werlen, E., & Siegenthaler, E. (2012). The relationship between flexible and self-regulated learning in open and distance universities. *The International Review of Research in Open and Distance Learning*, 13(2), 101–123.
- Blanchette, J. (2012). Participant interaction in asynchronous learning environments: Evaluating interaction analysis methods. *Linguistics and Education*, 23(1), 77–87.
- Block, J. H., & Burns, R. B. (1976). Mastery learning. In L. S. Shulman (Ed.), *Review of research in education* (Vol. 4, pp. 3–49). Itasca, IL: Peacock.
- Bloom, B. S. (1968). Learning for mastery. *Evaluation Comment*, 1(2), 1–5.
- Bloom, B. S. (1976). *Human characteristics and school learning*. New York, NY: McGraw-Hill.
- Bloom, B. S. (1982). *All our children learning*. New York, NY: McGraw-Hill.
- Bouvin, N. O., & Schade, R. (1999). Integrating temporal media and open hypermedia on the World Wide Web. *Computer Networks*, 31, 1453–1465.
- Chiu, M. M., Chow, B. W.-Y., & Chang, C. M. (2007). Universals and specifics in learning strategies: Explaining adolescent mathematics, science, and reading achievement across 34 countries. *Learning and Individual Differences*, 17(4), 344–365.
- Deng, Y., Manjunath, B. S., Kenney, C., Moore, M. S., & Shin, H. (2001). An efficient color representation for image retrieval. *IEEE Transactions on Image Processing*, 10(1), 140–147.
- Feng, S. F., Lu, H., & Ma, S. (2001). Mosaic representations of video sequences based on slice image analysis. *Pattern Recognition Letters*, 23, 513–521.
- Fuchs, L. S., Fuchs, D., & Tindal, G. (1986). Effects of mastery learning procedures on student achievement. *Journal of Educational Research*, 79(5), 286–291.

- Gargi, U., Kasturi, R., & Strayer, S. H. (2000). Performance characterization of video-shot-change detection methods. *IEEE Transactions on Circuits and Systems for Video Technology*, 10(1), 1–13.
- Girasoli, A. J., & Hannafin, R. D. (2008). Using asynchronous AV communication tools to increase academic self-efficacy. *Computers & Education*, 51(4), 1676–1682.
- Gonzalez, R. C., & Woods, R. E. (1992). *Digital image processing*. Reading, MA: Addison-Wesley.
- Hanjalic, A. (2002). Shot-boundary detection: Unraveled and resolved? *IEEE Transactions on Circuits and Systems for Video Technology*, 12(2), 90–105.
- Liao, Y.-H., Tsai, C.-Y., Su, M.-H., Li, H.-H., & Yu, P.-T. (2011, August). Digital learning video indexing using scene detection. In R. Kwan, J. Fong, L.-F. Kwok, & J. Lam (Eds.), *4th International Conference on Hybrid Learning 2011* (pp. 336–344). New York, NY: Springer.
- Liu, T., & Kender, J. R. (2004). Lecture videos for e-Learning: Current research and challenges. *Proceedings of IEEE Sixth International Symposium on Multimedia Software Engineering* (pp. 574–578). Washington, DC: IEEE Computer Society.
- Keyvnpour, M., & Izadpanah, N. (2011). A functional measure-based framework for evaluation of multi-dimensional point access methods. *Procedia Environmental Sciences*, 10, 784–789.
- Koprinska, I., & Carrato, S. (2001). Temporal video segmentation: A survey. *Signal Processing: Image Communication*, 16, 477–500.
- Mayer, R. E. (2005). *The Cambridge handbook of multimedia learning*. Cambridge, UK: Cambridge University press.
- Manjunath, B. S., & Ma, W.Y. (1996). Texture features for browsing and retrieval of image data. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 18(8), 837–842.
- Mittal, A., Pagalthivarthi, K. V., & Altman, E. (2006). Content classification and context based retrieval system for e-Learning. *Educational Technology & Society*, 9(1), 349–358.
- Naphade, M., Mehrotra, R., Ferman, A. M., Warnick, J., Huang, T. S., & Tekalp, A. M. (1998). A high-performance shot boundary detection algorithm using multiple cues. *Proceedings of the IEEE International Conference on Image Processing* (Vol. 2, 884–887). Washington, DC: IEEE Computer Society.
- Park, D. K., Jeon, Y. S., & Won, C. S. (2000). Efficient use of local edge histogram descrip-

- tor. *Multimedia 2000: Proceedings of the 2000 ACM workshops on Multimedia* (pp. 51–54). New York, NY: ACM.
- Sakarya, U., Telatar, Z., & Alatan, A. A. (2012). Dominant sets based movie scene detection. *Signal Processing*, 92, 107–119.
- Shivakumara, Phan, & Tan, (2010.)
- Skinner, B. F. (1999). The science of learning and the art of teaching. *Cumulative record, definitive edition* (pp. 179–191). Cambridge, MA: B. F. Skinner Foundation (Original work published 1954).
- Smoliar, S. W., & Zhang, H. (1994). Content-based video indexing and retrieval. *IEEE Transaction on Multimedia*, 1(2), 62–72.
- Song, B. C., & Ra, J. B. (2001). Automatic shot change detection algorithm using multi-stage clustering for MPEG-compressed videos. *Journal of Visual Communication and Image Representation*, 12, 364–385.
- Sun, X. S., Kim, C.-S., & Kuo, C.-C. (2005). MPEG video markup language and its applications to robust video transmission. *Journal of Visual Communication and Image Representation*, 16, 589–620.
- Wu, J.-H., Tennyson, R. D., & Hsia, T.-L. (2010). A study of satisfaction in a blended e-Learning system environment. *Computer & Education*, 55(1), 155–164.
- Yengin, I., Karahoca, A., & Karahoca, D. (2011). E-Learning success model for instructors' satisfactions in perspective of interaction and usability outcomes. *Procedia Computer Science*, 3, 1396–1403.
- Yuan, J., Wang, H., Xiao, L., Zheng, W., Li, J., Lin, F., & Zhang, B. (2007). A formal study of shot boundary detection. *IEEE Transactions On Circuits and Systems for Video Technology*, 17, 168–186.
- Zabih, R., Miller, J., & Mai, K. (1999). A feature-based algorithm for detecting and classifying production effects. *Multimedia Systems*, 7(2), 119–128.
- Zhang, D., Wang, F., Shi, Z., & Zhang, C. (2010). Interactive localized content based image retrieval with multiple-instance active learning. *Pattern Recognition*, 43(2), 478–484.

Athabasca University 

