Chapter I
Automatic Digital Content Generation System for Real–Time Distance Lectures

Hiroshi Takeda
Hosei University, Japan

Hisashi Yaginuma
Hosei University, Japan

Hajime Kiyohara
Hosei University, Japan

Akira Tokuyasu
Hosei University, Japan

Masami Iwatsuki
Hosei University, Japan

Norio Takeuchi
Hosei University, Japan

Hisato Kobayashi
Hosei University, Japan

Kazuhiro Yana
Hosei University, Japan

ABSTRACT

This article describes a new automatic digital content generation system we have developed. Recently some universities, including Hosei University, have been offering students opportunities to take distance interactive classes over the Internet from overseas. When such distance lectures are delivered in English to Japanese students, there is a pressing need to provide materials for review after class, such as video content on a CD-ROM or on a Web site. To meet this need, we have developed a new automatic content generation system, which enables the complete archiving of lectures including video/audio content, synchronized presentation materials, and handwritten traces on virtual whiteboards. The content is generated in real time and is immediately available at the end of the class. In addition, this system incorporates a unique video search algorithm which adopts a phonetic-based search technology. This enables quick review of the video content by typed-in keywords. The system can automatically create a vast amount of digital content and provide students with an efficient learning tool.

Copyright © 2009, IGI Global, distributing in print or electronic forms without written permission of IGI Global is prohibited.
INTRODUCTION

Hosei University has been providing numerous classes and lectures in the form of distance lectures through Hosei University Research Institute, California (HURIC). Since such distance lectures from abroad are delivered in English, we have strongly required an e-learning system for postproducing a lecture, including video/audio and other presentation materials, quickly into digital content, and making it available for review after class to Japanese students lacking sufficient language ability. It would be possible, were it merely a question of delivering distance lectures in real time, to suitably combine commercially available systems to display handwritten data in addition to video/audio and slide materials. However, in order to turn information presented in the lecture, including handwritten data, into video-synchronized digital content automatically, a data storage server is required in addition to a PC for the lecturer (Kaneko, Sugino, Suzuki, & Ishijima, 2000; Li, 2000; Panasonic, 2005). There was hitherto no system that was very portable and could be used with only a PC and a video camera. Further, since videotapes of the many lectures held every week create an enormous amount of stored content, students would be forced to waste time and energy if they had to play back a video in order to search for sections of a lecture they wanted to review. In order to conduct such a video search with precision, a search function using the lecture audio as keywords is required in addition to the conventional video search function using slide title indexes. To achieve such an audio-based keyword search function, the audio portion needs to be converted to text after a lecture and time information for each word or phrase added/edited manually. This would take too much time and be so costly that it would be difficult to achieve.

This article proposes a new automatic digital content generation system for lectures, developed and improved based on practical experience accumulated through distance classes at Hosei University (Hayashi, 2003), which allows handwritten data to be incorporated and audio-based keyword searching to be performed. Since the system can digitally reproduce and distribute videotaped information of a lecture, including handwritten data, immediately after class without any editing, students can play back sections that they did not fully understand as much as they like and continue their studies in detail after the lecture for a thorough understanding. Also, although still at a prototype stage, a function allowing applicable sections to be searched using audio data as keywords is automatically added so that any section requiring review may be located and played back immediately.

REAL-TIME DISTANCE LECTURE SYSTEM

Since April 2002, Hosei University has been offering a pre-MBA course toward an MBA to be acquired by further studying abroad for a minimum of one year. Figure 1 shows photos of the pre-MBA class. In this course, a service is provided in which students can attend some MBA-accredited courses in advance at Kudan Hall in Ichigaya Campus through distance classes from the Hosei University Research Institute California. Furthermore, an international distance class on welfare engineering given by lecturers in the US, Korea, and Hosei University in Japan is being offered as an interdisciplinary as well as a crossover subject to Ichigaya, Tama, and Koganei Campuses from April 2003. Figure 2 shows a set of scenes of the classroom.

These e-learning classes employ the system illustrated in Figure 3 for the linkage between the HURIC and Ichigaya Campus, and they use the multicampus LAN and teleconferencing system shown in Figure 4 for linkages among the campuses. In the classroom, the teleconferencing system presents the lecturer’s motion and voice to students, as shown in Figure 5, and the
Automatic Digital Content Generation System for Real-Time Distance Lectures

Figure 1. Photos of Pre-MBA class

Lecturer site at HURIC in California  Student site at Kudan Hall in Ichigaya Campus

Figure 2. Photos of international distance class on “Welfare Engineering transmitted from South Korea

Student site in Tama Campus  Student site in Koganei Campus

Figure 3. Real-time e-learning system

System allows lecturing with participants freely writing on PC screen by means of StarBoard.

Connection between HURIC and Kudan Hall at Ichigaya Campus in Tokyo
StarBoard function supplied by Hitachi Software Engineering displays Microsoft’s PowerPoint slides bilaterally between lecturers and students. Handwritten data drawn on the StarBoard can be performed in sync and in real time via the Internet. Meanwhile, the EZ Presenter (EZP), supplied by Hitachi Advanced Digital and installed in Ichigaya Campus, digitizes and stores class work data so that students can retrieve any part of it as Web content immediately after the end of class work (Hitachi, 2005).

In the past, however, data on slides and handwritten messages on the StarBoard could not be directly captured and stored in digital form in sync with classroom images for automatic generation of classroom content for later use. We were
unable, therefore, to make available to students the classroom information in complete duplicate form. Hence, we first developed the software by working with Hitachi for linking the digital board with the EZP so as to store all class video images, including handwritten data.

In the meantime, if about 15 rounds of class work per semester are turned into digital content, it would be almost impossible to search and retrieve any portion of the recorded audio/video data for review or restudy, for one has to go through vast quantities of data, amounting to more than 20 hours worth per class subject. Unless this problem is overcome, students would be forced to expend many hours of wasted effort in data search, to the detriment of their motivation for learning. To embed a useable video search function in digital contents, it was usually necessary to add time stamps and Meta-data to the contents manually after the end of class work. We then found that adoption of a search technique based on phoneme analysis of voice data (NEXIDIA, 2005; Clements, Robertson, & Miller, 2002) would permit instantaneous search and retrieval by means of keywords without the intervening editing process. Thus, we started joint development with Network Solutions for incorporating such a search technique in our automatic digital-content production system.

**AUTOMATIC DIGITAL CONTENT GENERATION SYSTEM**

This section describes the automatic digital content generation system for lectures developed and improved based on practical experience accumulated through distance lectures offered hitherto, which allows incorporation of handwritten data and audio-based keyword searching.

**Real-Time Storage Function of Handwritten Data**

Hosei University has been using the EZP, which can be built on a minimal set of PCs and video cameras. The EZP system cannot, however, capture handwritten data. To fill this gap and improve the system, we have added a linking function between the EZP and the StarBoard Software (SBS), so that
handwritten messages entered on a digital board or a PowerPoint slide may be captured and stored as vector data in sync with motion video. As shown in Figure 6, the SBS first extracts vector data out of handwritten data. Then, the linking software converts the extracts into a format readable by the EZP, which then automatically produces digital content in sync with motion video. Figure 5 shows a browsing image displaying class content from our pre-MBA class. The right-hand side frame in Figure 6 clearly shows the lecturer’s handwritten notation on a PowerPoint screen. The linking software developed in this project is expected to be commercially available as a part of the automatic digital-content production system from four companies, including Hitachi.

Search Function by Means of Voice-Data Keywords

The Fast-Talk Phonetic-Based Searching technique developed by Fast-Talk Communications (now NEXIDIA) is designed to build search databases from phonemes, the smallest elements contained in voiced speech, and to allow high-speed searches into voice data obtained from unspecified speakers (Clements et al., 2002). Usually, when searches are to be made into video and audio files by means of keywords, the target voice data, for example, will have to be converted by preprocessing into text form, which is then subjected to searches. Accurate conversion of voices of unspecified speakers into text form is often difficult, and searches of such voice-based texts cannot be performed with high precision. The Fast-Talk Phonetic-Based Searching does not search for words but rather for phonemes, or basic elements of pronounced English words. The technique thus allows searches on novel words, slang, and inaccurate spellings, as well as proper nouns, phrases, and initials. Since phoneme patterns in voice data are indexed for comparison and searching in this method, recognition rates of as high as 98% can be obtained on unspecified speakers. The specifications published by Fast-Talk call for a recorded voice quality of “studio-quality S/N ratios,” but recognition rates sufficiently high for practical purposes are attained when voice communications, encoded by G.728 at a bandwidth of 16Kbps, are transmitted across the Internet and recorded. For index searches on phoneme patterns, one has to build an index called the PAT (Phonetic Search Track) files. This preprocessing on classroom data can be performed on regular PCs in about one third of the time spent in class. Thus, high-performance servers would allow the preprocessing to be done in real time, making it possible to deliver classroom data immediately after the end of the class in the form of Web content that is searchable by voice-data keywords. We have prototyped a new digital-content production system hitherto unavailable, by adding the new search capability by means of voice-data keywords to the conventional content produced by the EZP. In the uppermost frame of the screen shown in Figure 6 (Class Content), a text box is provided on the left for entry of the voice-data keyword and a list box on the right for display of search results. The list box shows the search hits in the order of relative scores and the reference times of their occurrence in the class. By clicking on a hit entry, students can instantly jump to the spot of interest in the content.

Since our system conclusively generates about 600Mbyte content for a 3-hour lecture, the total contents of one-semester classes consumes about 9Gbyte disk space. The students can access the Web contents through the Internet with bandwidth connections of about 384Kbps.

EVALUATION BY STUDENTS

At the end of the last class on welfare engineering introduced in section 2, the students were surveyed on distance lectures and the use of lecture content. Figure 7 shows the results relating to distance lectures. As the Figure shows, approximately 60% of
the students were favorably inclined to the video and audio. It can be seen that the results were good, with close to 80% feeling they might take similar distance lectures in the future. However, only about 40% felt favorably towards interactivity. In a lecture involving a small number of students and delivered point-to-point, such as one of the pre-MBA classes, interactivity was viewed favorably, but in a distance lecture such as a welfare engineering lecture, which is distributed to many locations and to a large audience, the problem of insufficient interactivity arises probably because it is difficult for the participants to communicate with one another. On this point, the instructor may need to proactively call upon the students to check on questions and reactions.

A series of lectures was post produced into digital content with the EZP and made available on the Web server after the lecture. Since the audio-based keyword search function was still at a prototype stage, however, it was not incorporated in the Web content. Approximately 90% referred to the Web content, as shown in Figure 8, indicating that an unexpectedly large number of students used the service to review materials.

Figure 9 shows the ways in which the service was used. From the figure, it can be seen that a majority of the students referred to the content to prepare assignment papers. In Welfare Engineering, the lectures focus on the latest topics, and in preparing a paper, students need to base their considerations on knowledge acquired from the lecture itself rather than delving into reference books. This is believed to be why the ratio of students using the content to prepare papers was so high.
CONCLUSION

This article described an automatic digital content generation system for lectures, developed and improved based on practical experience accumulated through distance classes at Hosei University. By adopting this system, it is possible after a class not only to reproduce presentation material, including handwritten data, without any editing work, but also to provide a Web content that allows instant retrieval through an audio-based keyword search. Because the audio-based keyword search interface is still at a prototype stage, content generation is not yet fully automatic, and after extracting audio data manually, a PAT file is created for the content generated automatically by the EZP. We hope to complete a system that can automatically generate digital content that includes a phonetic-based keyword search function without a need for post processing.

We also plan to develop a system in which data written on a whiteboard in a normal classroom can be extracted and stored as vector data, and to adapt the audio-based keyword search to accommodate the Japanese language.

REFERENCES


