

# Development and Evaluation of Indexed Captioned Searchable Videos for STEM Coursework

Tayfun Tuna<sup>1</sup>  
ttuna@uh.edu

Jaspal Subhlok<sup>1</sup>  
jaspal@uh.edu

Lecia Barker<sup>2</sup>  
lecia@ischool.utexas.edu

Varun Varghese<sup>1</sup>  
vkvarghese@uh.edu

Olin Johnson<sup>1</sup>  
johnson@cs.uh.edu

Shishir Shah<sup>1</sup>  
shah@cs.uh.edu

## ABSTRACT

Videos of classroom lectures have proven to be a popular and versatile learning resource. This paper reports on videos featuring Indexing, Captioning, and Search capability (ICS Videos). The goal is to allow a user to rapidly search and access a topic of interest, a key shortcoming of the standard video format. A lecture is automatically divided into logical indexed video segments by analyzing video frames. Text is automatically identified with OCR technology enhanced with image transformations to drive keyword search. Captions can be added to videos. The ICS video player integrates indexing, search, and captioning in video playback and has been used by dozens of courses and 1000s of students. This paper reports on the development and evaluation of ICS videos framework and assessment of its value as an academic learning resource.

## Categories and Subject Descriptors

H.5.2 [Information Interfaces And Presentation]: User Interfaces - *Evaluation/methodology* H.3.1 [Information Storage and Retrieval]: Content Analysis and Indexing - *Indexing method* I.4. [Image Processing And Computer Vision]: Segmentation-*Edge and feature detection* I.7.5 [Document and Text Processing]: Document Capture- *Optical character recognition (OCR)*

## General Terms

Design, Experimentation, Measurement, Performance

## Keywords

Assessment, Educational Technologies, Lecture Videos, OCR, Text Segmentation, Video Indexing, Video Search

## 1. INTRODUCTION

Video of classroom lectures is a versatile learning resource. It is often made available as additional material for a conventional course, as the core of distance learning coursework, or posted publicly for community learning or as reference material. Evidence of the popularity of lecture videos includes thousands of complete courses posted on portals such as MIT OpenCourseware and Apple's iTunes University. At the University of Houston,

video lectures have been widely used for over a decade to enhance STEM coursework. Typically Tablet PCs, which allow free mixing of prepared (PowerPoint) viewgraphs with hand annotations and illustrations, are employed for teaching and simultaneous recording of lectures. Advantages include excellent resolution, as the video consists of PC screen shots, and low video production cost, as no camera or operator is needed. The videos typically show whatever the professor is projecting on the screen (e.g., PowerPoint slides, animations, annotations, formulas, algorithms, or drawings) and the instructor's voice. Prior research has established that the recorded Tablet PC lectures are a powerful resource on par with a textbook and the classroom.

A major weakness of these recorded lecture videos is the inability to quickly access the content of interest in a video lecture. This is intuitively clear and also established by our surveys and student interviews. This paper reports on **ICS videos**: videos enhanced with *Indexing*, *Captioning*, and *Search* capability that are designed for quick access to video content. **Indexing** adds logical index points, each in the form of a snapshot representing a video segment that can be accessed directly; **Captioning** adds the text of the video lecture in a separate panel; **Search** enables identification of video segments that match a keyword provided by the user.

The central technical goal of this project is automatic conversion of standard videos to ICS Videos. Index points are identified automatically with analysis of video frames. Search framework is based on the adoption of OCR technologies enhanced with a suite of image transformations to identify text on video frames. The accuracy and efficiency of indexing and search procedures are presented in the paper. Currently, captioning is done manually. All aspects are integrated in a unified ICS Video player.

The ICS video framework has been used by dozens of STEM courses and 1000s of students at the University of Houston, a large public university system. Captioned videos have been tested and evaluated with project partners at the Texas School for the Deaf. Results are reported from extensive assessment of ICS video usage carried out with student surveys and instructor interviews. The research presented clearly demonstrates that i) Tablet PC based video lectures are an extremely important student resource, ii) The framework developed to enhance videos with indexing and search features is efficient, effective, and a significant improvement over the state of the art, and iii) Indexing and search capability significantly enhance the value of lecture

<sup>1</sup> Department of Computer Science, University of Houston, Houston, TX 77204

<sup>2</sup> School of Information, University of Texas, Austin, TX 78701

videos. The video framework developed is freely available to academic institutions.

## 2. RELATED WORK

Because of limited space, we only list representatives of the main directions of related work. The idea of capturing videos of classroom lectures and other presentations, and making them available electronically, has been in widespread use for several years. These videos are generally recorded by camera(s) operated by professionals or installed in the lecture/presentation rooms [3]. Navigating between video chunks, using annotations and hyperlinks is discussed in [9,11]. An index structure-based method to efficiently search short video clips in large video collections is presented in [17]. OCR engines and text localization has been used for indexing videos [8].

The key features of the work presented in this paper that differentiates it from most previous work are i) automation of segmentation of videos, ii) deployment of image transformations with OCR for enhanced video search, iii) an approach based on video frame analysis that is independent of any hardware or presentation technology, v) focus on recordings from Tablet PC screens, and v) large-scale evaluation of student outcomes.

## 3. ICS VIDEOS

This section discusses the technology and evaluation of the key components of ICS videos; indexing, search, captioning and the video player. The technical aspects of this work are reported in more detail in related work [17].

### 3.1 Indexing

In the context of ICS videos, indexing is the task of dividing a lecture video into segments that contain different topics. This task is accomplished by first identifying all *transition points* where the scene in the video changes significantly. A subset of these transition points is selected as *index points*, which represent the beginning of the video segments as presented to the user.

Detection of transition points is based on a comparison of successive frames in the video. Corresponding pixels in successive frames are considered to be different if they differ by a minimum *RGB threshold* when the RGB values of the pixels are compared. Successive frames constitute a transition point if the fraction of pixels that are different based on the RGB criteria exceeds a minimum threshold, which we refer to as the *transition point threshold*. The reason for using thresholds to identify transition points is that frames corresponding to the same scene in practice (e.g., exactly the same viewgraph) also have minor differences in the RGB spectrum that must be ignored to avoid false transition points. The threshold values are chosen empirically after evaluation of a large number of diverse lectures. Figure 1 illustrates a transition point in a sequence of video frames.



Figure 1. Transition point in a video: third frame is a new transition point

Comparing pairs of all successive frames in a video is rather inefficient and time consuming. We apply two methods to speed up this process: *Employing Jumping Interval* and *Binary Search*.

*Employing Jumping Interval*: In a video lecture, scene transitions are relatively infrequent. With this optimization, instead of comparing successive frames, the current frame is compared to a frame that is one *Jumping Interval* ahead. If the two frames are identical, then in all likelihood, the frames between them are also identical. Hence the search marker is moved up by a *Jumping Interval* and the frames between them are not processed. If the current frame and the frame one *Jumping Interval* ahead are different, then the frames between them are compared sequentially starting from the current frame to identify all transition points. Essentially, in a long sequence of identical frames, most comparisons between successive frames are skipped.

*Binary Search*: This procedure starts by splitting the video into two segments – from the first frame to the middle frame and from the middle frame to the last frame. For each of these segments, the first and the last frame are compared. If they are identical, then it is assumed that the segment contains no transition points and if they are different, the segment is again subdivided into two segments and the procedure is applied recursively. When a segment size of one is reached, a transition point is identified.

Employment of a jumping interval and binary search substantially improves performance but there is a chance that a legitimate transition point may be missed. This happens when the first and last frame in a segment are identical masking transition points inside the segment. A common reason for this to happen is when the instructor goes back and forth between viewgraphs while explaining a topic in a lecture. For evaluation we selected 20 different videos prepared by 15 different instructors. These videos were selected for their diversity in templates and color styles. The courses being taught were from the Computer Science, Biology, and Earth and Atmospheric sciences departments. Figure 2 and Figure 3 present results of experiments with this video suite.

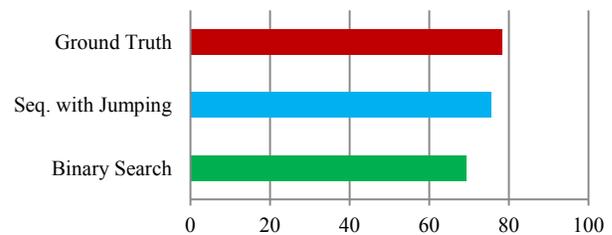


Figure 2. Average number of transition points detected with different methods in a lecture video

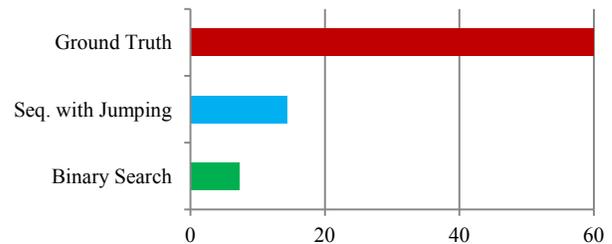


Figure 3. Average time (mins) for indexing an hour of video

Figure 2 shows the number of transition points detected on average in a video and Figure 3 shows the average processing time for indexing an hour of video. We note that a few transition points are indeed missed due to the optimizations. However, the cost of losing a few transition points is outweighed by the significant benefit of improved execution time.

The algorithms discussed above typically find all slide transitions in a lecture video with very few false positives. However, a video may have a large number of such transition points; over 100 is not unusual when an instructor writes frequently on the Tablet PC screen. To make indexing more meaningful and easier to use, we need to identify a smaller number of index points from the detected transition points. For this purpose a filtering algorithm is applied to eliminate unwanted transition points. The algorithm takes the number of index points desired and the minimum time gap between successive index points as parameters, and then applies iterative elimination. The basic idea is to repeatedly find the closest pair of remaining transition points on the timeline and eliminate one from the pair. When there is a cluster of transition points on the timeline, the algorithm ensures that a representative is included as an index point. The transition points are eliminated until their number has reached the desired number of index points, and the time difference between the remaining transition points has exceeded the predefined minimum time gap. The remaining transition points are retained as index points.

### 3.2 Keyword Search

*Keyword search* is the process of identifying all video segments where a particular keyword exists. The framework to support keyword search is as follows. The indexer creates the video segments as well as the transition point frames. Text on these frames is detected and stored in a database. The ICS Video Player loads the keywords from this database along with the corresponding video. When a user searches for a keyword, the player presents a series of index points that allow the user to navigate to the corresponding video segments. This process requires that all the text contained in the images corresponding to the transition points be identified. This can be accomplished by the use of Optical Character Recognition (OCR) tools, an approach investigated in [10]. After a comprehensive analysis of available OCR tools that could be integrated for use with the ICS Video Project, we opted to use the MODI (Microsoft Office Document Imaging) toolset. We found that OCR tools generally have limited effectiveness at recognizing text in the presence of 1) certain text and background color and shade combinations, 2) text mingled with colorful shapes, and 3) small and exotic fonts. To increase the detection efficiency of text on video frames, we used simple image processing techniques for image enhancement (IE) prior to the application of OCR tools. IE operations employed include segmentation of text, enlargement with interpolation, and color inversion.

Segmentation of text involves defining and extracting the text regions in an image. For text segmentation, we first convert the color image to a binary black and white image by using Simple Image Statistics based thresholding[14]. Then by using a dilation operation, which allows for growth of separate objects and joining of objects, we build words from characters. After that we employ Sobel edge detection. Lastly, we use blob extraction to extract standalone objects in the image using connected components labeling algorithm. Enlargement with interpolation is implemented for the segmented blocks. By this operation, small size text is enlarged so that some of it becomes visible to OCR engines. Segmentation of text and enlargement is illustrated in

Figure 4. Color inversion is done separately by altering the RGB values of images, aimed at increasing the contrast between the text and background. Image enhancement procedures often lead to new text being recognized, but can also prevent the recognition of other text. Hence OCR engines are applied to the original images as well as the enhanced images and the union of the results is taken.

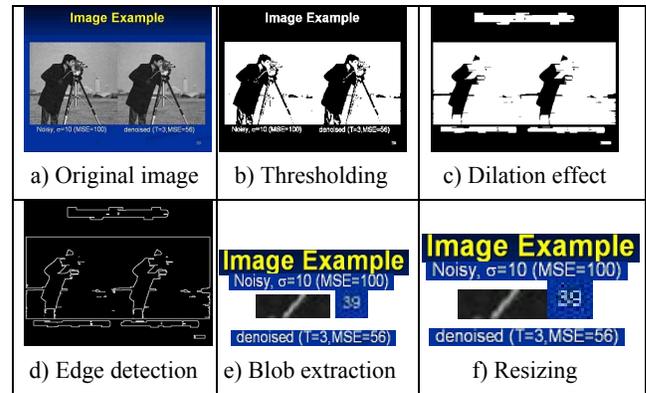


Figure 4. Segmentation and enlargement of text

To test the OCR Tools and the impact of IE procedures, we evaluated 1387 different images that are selected by the indexer from 20 diverse videos discussed in the previous section. Images in these videos contain 20,007 unique words, 27,201 total words (of more than 1 character length) for a total of 144,613 characters.

Experimental results, presented in Figure 5, show that the search accuracy of three distinct OCR engines, Tesseract, GOCR and MODI, improved, with an increase of 9% on average with IE transformations. The maximum accuracy obtained by applying all OCR engines with image enhancements was 97.1%. Alternately stated, the miss rate was 8.9% for the best single OCR engine, 5.2% for all OCR tools combined, and 2.9% for all OCR engines combined with image enhancement.

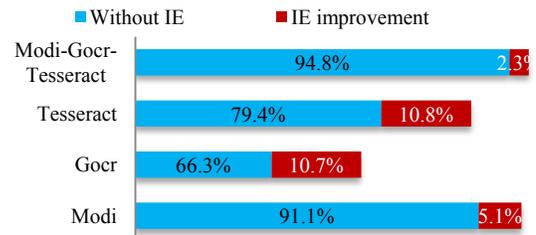


Figure 5. Search accuracy rate of OCR tools

Image enhancement provided this accuracy improvement, but increased the processing time significantly, partly because OCR engines have to be applied on the original and the enhanced images. Nonetheless, the processing time remains modest for a typical video. On average it is in the range of 2-3 minutes for an hour long video on a typical desktop. Image enhancement also doubled the false positives detected by OCR engines, i.e., more words were detected that were not actually present in the video. This often happens when an OCR engine misses a character in a word, leading to false identification of a different word. Since the main aim of the text recognition is to let the user find words of interest, the “extra words” resulting from false positives are unlikely to diminish the functionality in a meaningful way.

### 3.3 Captioning

Captioning consists of adding synchronized text blocks matching the audio stream, which typically consists of the instructor's speech as well as student interactions. Captioning was motivated by the need to have video lectures available to the deaf, although it also has significant potential for improving the experience of hearing and foreign language students. Ideally, captioning should be automatic based on speech recognition technology, which has been known to work well in many scenarios. However, our preliminary investigation revealed that available speech recognition tools were not satisfactory for technical lectures. The key reasons were the preponderance of technical jargon and abbreviations. We believe that a number of techniques can be used to make speech recognition practical, in particular training speech recognition software with instructors, dictionaries of technical terms, and a GUI for correction of speech errors. However, this has not yet been investigated systematically so we will not discuss further. The ICS video player does support manually transcribed captions. We have prepared lecture material for testing and evaluation for the Texas School for the Deaf and are in the process of preparing diverse course material with captions. We expect to report on these in the future.

### 3.4 ICS Video Player

ICS Video Player is a Flash-based player capable of playing standard H.264 encoded streaming media. The player is made up of video playback component, Index Panel, Timeline with selectable index, Search box, Search results display, and Caption display, as illustrated in Figure 6.



Figure 6. A snapshot of ICS Video Player

*Index Display:* An index panel is situated on the left side of the player, as shown in Figure 6. Index points are listed vertically in the index panel and each has an index title, a screenshot of the video at that point of time, and index time in *hh:mm:ss*. Users can select any index point from the list and the video starts playing from the selected index point. The bottom part of the player has a panel comprised of the timeline, indicating the progress of the video playback, along with index points arranged by their time locations. Users can skip to any part of the video by selecting the desired position in the timeline. The image screenshot corresponding to the index point is displayed as a pop-up when the mouse cursor is placed over an index point bar, providing a visual display of the segment represented by the index point.

*Search Display:* A search box is located on the top right corner of the video player. In order to execute a search, a user types the

keywords in the search box. All index segments that contain the search term are displayed as results. The matching keyword, number of matches in the segment, the index title, and the index image are also displayed. The search feature is illustrated in Figure 6; two video segments represented by index points are presented as matches for the search term "granite." Clicking on the search images displayed will start the video from that point.

*Caption Display:* Captions are displayed as an overlay at the bottom of the video and in a separate caption window. The caption window displays the start time of the caption together with the caption text. The most recent caption is added to the bottom of the list. Users can scroll and read the previously displayed captions. The captions can be turned on or off by clicking the appropriate button on the player. A visual display is omitted due to lack of space.

## 4. EXPERIENCE AND ASSESSMENT

The usage of ICS videos framework is widespread at the University of Houston campus. For example, in the Spring 2011 semester, raw statistics show that 34 courses employed ICS videos for instruction. 314 new video lectures were uploaded, (and many of the existing video lectures from earlier semesters were also accessed). There were a total of over 23,000 video accesses by the students throughout the semester. Students in many (but not all) courses participated in the assessment of videos.

The evaluation described below adds to a small, but expanding, literature on the value of capturing classroom lectures as videos. Most studies show that students willingly use video lectures to their advantage when they are available [1,5,12,15,16].

We have assessed ICS Video use to develop an understanding of the overall perceived value of the video lectures and the value added, if any, by providing indexing and search. Surveys were administered over three years in six semesters; all but the original pilot survey were administered online. The original pilot survey was developed in part based on the open-ended questionnaire in our previous work [16] that motivated this project. The surveys asked about the nature and frequency of usage, strength of need, value, preparation for the class, expected grade, and several individual difference variables, including hours worked to earn income, credit hours taken, hours of studying per week, commute time, marital status, number of dependents, and demographic information. In this paper we report the key findings from the two most recent terms, Spring 2010 and Spring 2011 in regard to general video usage, and specifically on indexing and search tools. The procedures and detailed results from the survey are made available online through [icsvideos.cs.uh.edu](http://icsvideos.cs.uh.edu).

### 4.1 Survey Procedures and Sample

Surveys were administered in the last week of the semester. Participating faculty representing 18 courses sent a link to their students, resulting in 1167 usable surveys (Spring 2010 N=612, Spring 2011 N=555). About half of students in the sample were in freshman biology courses (53%), though not primarily freshmen themselves. The remaining courses were in computer science, chemistry, geology, and mathematics. Undergraduates were widely distributed across levels, with 17% freshmen, 23% sophomores, 31% juniors, and 29% seniors and post-baccalaureate. A majority (60%) of the undergraduates sampled were female, because of high percentage of women in the biology courses (67%). The sample is also ethnically diverse: 27% were Eastern or Southeastern Asian; 24% were White; 16% were Asian Indian; 14% were Latin American; and 7% were African American. In addition, although the majority of the students are

American citizens or permanent residents (96%), about 30% speak English as a second language. Most of them rate their English competence as good.

The survey was conducted at the University of Houston, where there are many demands on students' time. Students reported a mean one-way commute time to campus between 30-45 minutes. Most of the students (57%) work to support themselves, with about 7% working 36 or more hours per week. A larger fraction of students in this sample work as compared to students nationally: 46% of full-time students, compared to 41% nationally.

The surveyed students are committed to their education. The majority of respondents felt that going to class was important and were very concerned about their course grade and their grade point average in general.

## 4.2 Summary of Overall Findings

Instructors posted anywhere between 8 and 30 videos per course. About 84% of students responded that they had used the videos.

### 4.2.1 Videos Strongly Valued for Learning, Grades

A series of survey items asked students to agree or disagree, or to rate the importance of certain aspects of the videos. Responses indicated strikingly high value for the videos. Table 1 shows evaluation by the students who viewed at least one video. Students found the videos helpful for clarifying material, for reviewing, and for studying for quizzes and tests. Additionally, most of the students agreed that having access to videos was important to them for the individual class (72% agreed strongly) and rated the videos very important for getting the grade they hoped for (64% felt it was very important; mean of 3.5 on a 4-point scale for this item).

Table 1. Perceived value of videos

Survey Item	N	Mean (6-pt scale)	Std. Dev.
Lecture videos help me to clarify material that was not clear in class.	825	5.45	.826
Lecture videos are useful for reviewing.	841	5.63	.693
Having access to lecture videos for this class is important to me.	837	5.61	.774
The lecture videos helped me to study for quizzes or tests.	829	5.51	.842

Students were asked to rate several content sources for their importance in their learning for the class, and these ratings are shown Figure 7. The most highly rated were "professor's lecture notes" (84% of students) and lecture videos (63% of students), followed by notes they themselves take. The lowest in relative importance were the textbook and other reading sources.

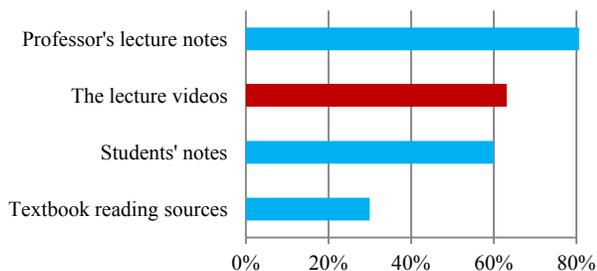


Figure 7. Student ratings of learning resources: Percentage of students that consider the resource "very important"

### 4.2.2 Index Use Correlated with Increased Value of Videos

A test of significance was conducted to compare index users' perceptions of the value of the lecture videos to non-index users' perceptions of value. In order to conduct the test, a scale variable, Overall Attitude Toward Videos, was computed by combining the six variables discussed above (Lecture videos help me to clarify material that was not clear in class; Lecture videos are useful for reviewing; Having access to lecture videos for this class is important to me; The lecture videos helped me to study for quizzes or tests; How important was use of the lecture videos for this class for getting the grade you hoped for? Rating of the importance of lecture videos for learning in this class). Students who had used the indexing feature scored higher on average on the Overall Attitudes Toward Video scale than students who had not used the indexing feature ( $t(616) = -3.284, p < .001$ ). To estimate the magnitude of the value added by the index, the effect size was calculated using Cohen's  $d$ . A small to moderate effect ( $d = 0.265$ ) was found.

### 4.2.3 Index was Helpful, Easy to Use

A series of forced-answer questions were intended to provide an understanding of the usefulness of the indexing tool. Among the 299 students who had used the indexing feature, response was very positive:

- 97% felt that the index was helpful
- 89% said they knew immediately what to do with the index (a proxy for the construct "intuitive")
- 96% felt the placement of information and images made the index easy to use
- 90% thought the time intervals were appropriate for the lecture

In open-ended comments, students stated that using the indexing feature saved them time (e.g., "I didn't have to wade through the rest of lecture just to answer one question") and helped them to get back to studying after interruptions (e.g., "sometimes I would have to pause the lecture to take care of other responsibilities that I had to attend too, and when I was ready to come back to the lecture I'd pick up exactly where I was at. It was great!").

### 4.2.4 Search Tool Less Used, But Valued

Relatively few students (46) responding to the 2011 survey had used the search tool. In hindsight, perhaps this is not surprising: the students were never informed that search existed or what it meant, the search icon did not stand out in the interface, and the search feature was not usable for part of the semester for technical reasons. A series of forced-answer statements asked students to evaluate the search tool. Most respondents found the tool easy to use (96%), found the search results easy to use (97%), and felt that the search tool helped them find the part of the video they were looking for most of the time (70%). Students usually knew which words to enter into the search box to find the segment of the video they were looking for (72%). Although these results are strong, they are still preliminary, given the small sample size.

### 4.2.5 More Class Attendance, More Video Use

A fear often reiterated in literature on making lecture videos available is that the students will be less likely to go to class, using the videos as a substitute [1,2,5,7,12,13]. That fear was not substantiated in our findings for STEM coursework, in any of the terms. Most (86%) students strongly agreed that it was important for them to go to class, whether or not the video lectures were available, and 67% reported that they had attended 75-100% of classes. In addition, class attendance ( $r(1019) = .234, p < .000$ ).

## 5. DISCUSSION AND FUTURE WORK

The underlying significance of this and related research is that video is a dynamic and versatile resource, not a poor substitute for a classroom. The videos were valuable for students who attended classes, were primarily used for reviewing and clarifying concepts, and had little impact on class attendance. In this regard, we hope some of the following major directions of work will be pursued by the community or educational vendors:

- Enabling instructors to customize videos with easy interfaces for selection of indexes and other basic editing and manipulation.
- One touch operation from starting of recording a lecture to it being posted online for students.
- With mobile devices becoming increasingly common, understanding what modifications need to be made to make use on mobile devices attractive.
- Enhancement and automation of captioning to make video material more useful to the deaf students.
- Making infrastructure for ICS or similar video technology widely available to STEM instructors.

## 6. CONCLUSIONS

This paper reports on technologies developed for Indexed, Captioned, and Searchable videos and their usage for STEM coursework. We demonstrated that automated indexing and search frameworks are effective and efficient. The ICS Videos framework underwent large scale deployment and assessment and was a success in all respects. The videos were judged by students to be very valuable and employed for diverse purposes such as clarifying and reviewing material and preparing for exams. Indexing and search features were considered very helpful and easy to use. This work points to a new and innovative direction for effective use of videos in STEM coursework. The framework developed is freely available to educational institutions.

## 7. ACKNOWLEDGMENTS

Partial support for this work was provided by the National Science Foundation's Division of Undergraduate Education under the Course, Curriculum, and Laboratory Improvement (CCLI) program with Award No. DUE-0817558. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

We would also like to acknowledge the contributions of several members of the ICS Videos group: Gautam Bhatt, Zhigang Deng, Salvador Baez-Franceschi, Pradeep Krishnan, and Joann Li.

## 8. REFERENCES

- [1] Abowd, G. D. 2010. Classroom 2000: An experiment with the instrumentation of a living educational environment. *IBM Systems Journal*, 38(4), 508–530.
- [2] Bell, T., Cockburn, A., McKenzie, B., & Vargo, J. 2001. Flexible delivery damaging to learning? Lessons from the Canterbury Digital Lectures Project. In *ED-MEDIA 2001: Proceedings of the World Conference on Educational Multimedia, Hypermedia and Telecommunications* (pp. 117–122).
- [3] Bianchi, M. 2004. Automatic video production of lectures using an intelligent and aware environment. In *Proceedings of the 3rd international conference on Mobile and ubiquitous multimedia (MUM '04)*. ACM, New York, NY, USA, 117–123. DOI=10.1145/1052380.1052397  
<http://doi.acm.org/10.1145/1052380.1052397>
- [4] Brecht, H. D., & Ogilby, S. M. 2008. Enabling a comprehensive teaching strategy: Video lectures. *Journal of Information Technology Education*, 7.
- [5] Chandra, S. 2007. Lecture video capture for the masses. *ACM SIGCSE Bulletin*, 39(3), 276–280.
- [6] The Condition of Education 2011. (2011, May 26). National Center for Education Statistics.  
<http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2011033>
- [7] Keengwe, J. 2007. Faculty integration of technology into instruction and students' perceptions of computer technology to improve student learning. *Journal of Information Technology Education*, 6(1), 169–179.
- [8] Lienhart R and Effelsberg, W. 2000. Automatic text segmentation and text recognition for video indexing. *Multimedia Syst.* 8, 1 (January 2000), 69-81. DOI=10.1007/s005300050006  
<http://dx.doi.org/10.1007/s005300050006>
- [9] Ma, W., Lee, Y., Du, D. H. C., and McCahill, M. P. 1998. Video-Based Hypermedia for Education-on-Demand. *IEEE MultiMedia* 5, 1 (January 1998), 72-83. DOI=10.1109/93.664744  
<http://dx.doi.org/10.1109/93.664744>
- [10] Merler, M. and Kender, J.R. 2009. Semantic keyword extraction via adaptive text binarization of unstructured unsourced video, *IEEE International Conference on Image Processing, Cairo, Egypt*, Nov. 2009 .
- [11] Microsoft Project Tuva,  
<http://research.microsoft.com/apps/tools/tuva/>
- [12] Odhabi, H., & Nicks-McCaleb, L. 2011. Video recording lectures: Student and professor perspectives. *British Journal of Educational Technology*, 42(2), 327-336.
- [13] Phillips, R., Gosper, M., McNeill, M., Woo, K., Preston, G., & Green, D. 2007. Staff and student perspectives on web based lecture technologies: Insights into the great divide. In *ICT: Providing choices for learners and learning. Proceedings of the Australasian Society for Computers in Learning in Tertiary Education (ASCILITE) Conference. Singapore, Dec* (pp. 2–5).
- [14] Simple Image Statistics Thresholding,  
<http://www.aforgenet.com/framework/docs/html/39e861e0-e4bb-7e09-c067-6cbda5d646f3.htm>.
- [15] Soong, S. K., Chan, L. K., Cheers, C., & Hu, C. 2006. Impact of video recorded lectures among students. *Who's Learning*, 789–793.
- [16] Subhlok, J., Johnson, O., Subramaniam, V., Vilalta, R., & Yun, C. 2007. Tablet PC video based hybrid coursework in computer science. *Proceedings of the 38th SIGCSE technical symposium on Computer science education - SIGCSE '07* (p. 74). Presented at the Proceedings of the 38th SIGCSE technical symposium, Covington, Kentucky, USA.
- [17] Tuna, T., Subhlok, J., Shah, S. 2011. Indexing and Keyword Search to Ease Navigation In Lecture Videos, *Applied Imagery Pattern Recognition Workshop, Washington, DC, Oct 2011*.
- [18] Yuan, J., Duan L., Tian Q., and Xu C. 2004. Fast and robust short video clip search using an index structure. In *Proceedings of the 6th ACM SIGMM international workshop on Multimedia information retrieval (MIR '04)*. ACM, New York, NY, USA, 61-68. DOI=10.1145/1026711.1026722  
<http://doi.acm.org/10.1145/1026711.1026722>