Regular Articles

SceneKnowledge: Knowledge Sharing System Using Video-scene-linked Bulletin Board

Satoshi Shimada†, Ken Tsutsuguchi, Akira Kojima, Hiroshi Konishi, and Suguru Higashino

Abstract
We present an overview and describe trial demonstrations of SceneKnowledge, a powerful new system based on fully exploiting the potential of video for discovering and sharing knowledge.

1. Introduction

Many companies and other entities have become increasingly interested in creating knowledge within their own organizations that can be shared and put to practical use by using network systems. The objective of skill transfer and knowledge management is to foster person-to-person exchanges of opinions, thoughts, ideas, information, and knowledge using an assortment of tools such as bulletin boards, social network services, blogs, and twitter.

However, relying on text alone to convey the practical knowledge and skills required by technical vocations and social activities is extremely limiting. Moreover, it is not easy to compile knowledge using conventional communication tools because knowledge is typically very unevenly distributed among the members of an organization.

One way around this problem is to harness the audiovisual impact and expressive power of video. Knowledge can be represented in a very easy-to-understand, intuitive way by video, so it should be relatively easy to accumulate expertise and knowhow by sharing background information. Of course, video has its own limitations—one may have to watch the video for some time to grasp its message, and the knowledge contained in the video may be elusive—so

using simple video by itself is not really appropriate.

Considering these inherent limitations, we propose a method of dividing video into work-related and action-related scenes, then providing bulletin board functions linked to the video scenes. This scene-linked bulletin board approach expresses the knowledge lurking within video and generates a knowledge-creation spiral that can be readily shared and further developed.

2. Knowledge creation by video-scene-linked bulletin boards

2.1 Knowledge discovery and sharing processes

We achieve continuous knowledge creation based on the SECI (socialization, externalization, combination, and internalization) model [1] by using the following method. The spiral of knowledge discovery and sharing within an organization is illustrated in **Fig. 1** (for simplicity, only the first loop is shown). Video sharing takes place when the behavior or pattern of performing a skill is filmed, and the implicit knowledge is then shared with others when they view the video.

(1) Knowhow and expertise are then accumulated by setting up a bulletin board to discuss each video scene, and the knowledge provided by all the users—in the form of comments, sketches and drawings, questions and answers, and discussion on the bulletin board—is converted to

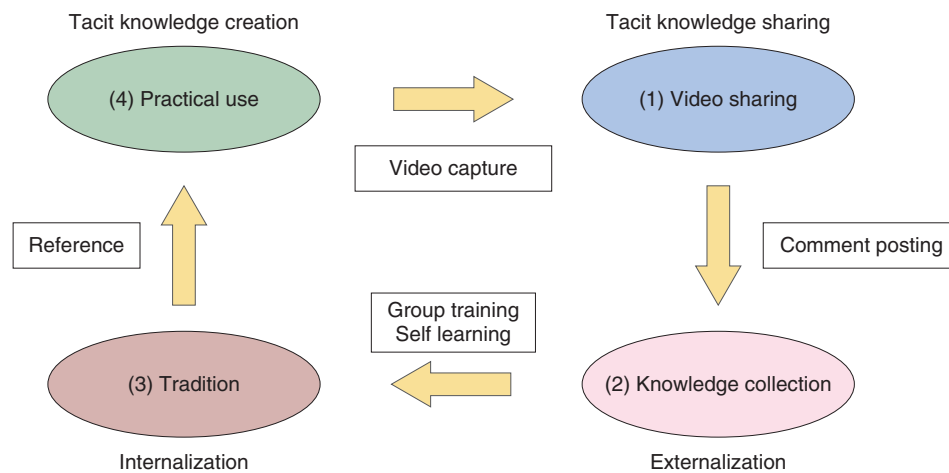† NTT Cyber Solutions Laboratories
  Yokosuka-shi, 239-0847 Japan

Fig. 1.   Spiral of knowledge discovery and sharing within an organization.

explicit knowledge.

(2) Skills are transferred through the acquisition of knowledge obtained by perusing video scenes tagged with comments, and new technologies and ways of doing things are learned through practice. Posted comments and video scenes are organized by topic to enable the bits of knowledge contributed by each user to be combined to form an integrated comprehensive body of knowledge.

(3) This system enables new knowledge to be created by practicing the knowledge that has been mastered. This tacit knowledge that is created can then be shared through the same process.

As this cycle continues, an ongoing knowledge-creation spiral is set in motion that creates new knowledge at the organizational level.

2.2 Developed system

To support this knowledge-creation spiral, we have developed a web application system called SceneKnowledge. As can be seen in **Fig. 2**, the system consists of a SceneKnowledge server that provides video/user registration management and scene-linked bulletin boards, a video indexing server that detects registered video events and supports scene definition, and a video distribution server. SceneKnowledge can be run on a web browser on a personal computer, but can also be viewed on mobile phones (i-mode or Flash Lite) and smartphones (Android 2.2 and higher). Below, we briefly describe the scene definition and scene-linked bulletin board features of the system.

2.2.1 Registration of video with defined scenes

When you access the SceneKnowledge server via your browser and upload videos, the video indexing server automatically detects major scene changes in the content, camera work sequences, speech and sound sequences, and other events in the video. The timeline on the scene definition screen shows the times at which such events are detected. Scene definition is very simple: an interface lets you manually sort and group video scenes by theme while referring to the timeline event data. Bulletin boards are established by adding titles and summaries to each scene and selecting representative images.

2.2.2 Video-scene-linked bulletin boards

A screenshot of a video-scene-linked bulletin board is shown in **Fig. 3**. If you have a question or comment while viewing the video, you simply click on the text box to post comments. This pauses the playback and allows you to type in comments. The comments are then linked to the video on the basis of the playback time. When you move the cursor over the video scene display bar, the title and key images associated with the scene are displayed. Then, a particular scene of interest is chosen and the video is played back beginning with that scene, and posted comments associated with the scenes are displayed on the right. Clicking on the title of a previously posted comment takes you directly to the scene where the comment was originally posted and begins playback from that point. Users can also discuss or respond to comments using the comment reply function. Other functions permit you to draw on the video; bookmark video playback times, posted comments, and scenes; vote for posted
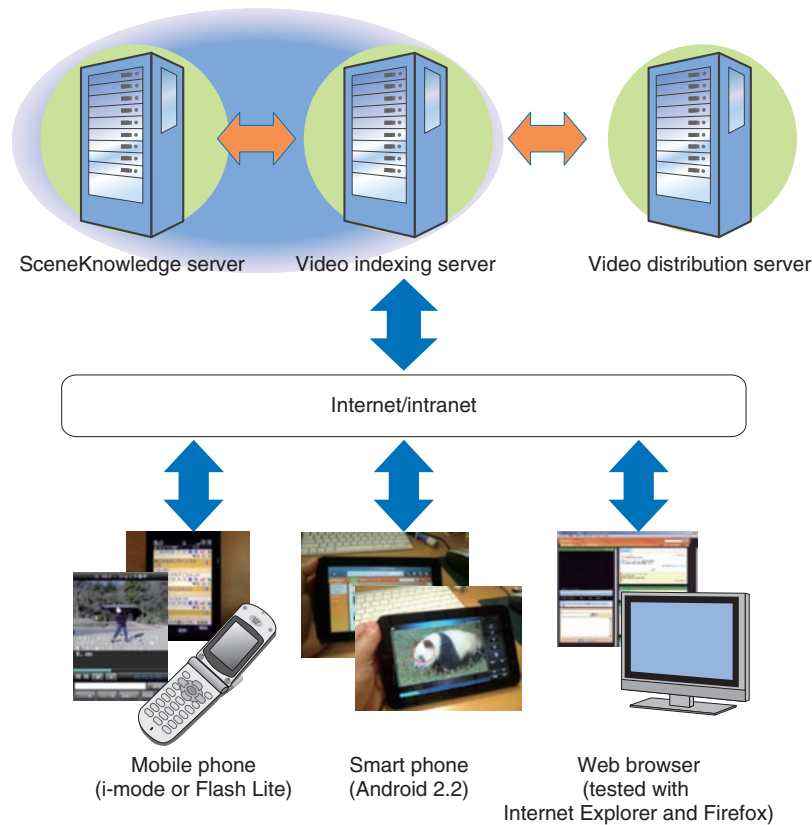
Fig. 2.   System configuration of SceneKnowledge.

comments, record private notes, perform text searches, and so on. Using this novel video bulletin board approach, we streamlined three basic capabilities: viewing video, reading comments, and posting comments. These capabilities not only support the aggregation of knowhow and expertise, but also foster introspection and reflection.

3.   Field trials

3.1   Pilot study in a practical skill-based field

We conducted a trial to evaluate the system's ability to collect and consolidate expertise in the field of skilled nursing practice, a field that clearly requires technical expertise [2]. For the experimental environment, we registered an instructional video of nursing practice and recruited 1718 nurses employed in four general hospitals in Osaka for the trial, which lasted several weeks. We set up the system so that nurses could access it during their regular shifts as much as possible, and we asked them to post comments offering tips, techniques, and pointers regarding the nurs-

ing practices illustrated in the video. A total of 189 comments were posted during the trial as the nurses went about caring for their patients. Analysis of the comments revealed that a good deal of detailed content regarding practical nursing practice was collected over this relatively short trial period, and it confirmed that a lot of nursing knowledge was generated on the basis of the video contents.

Meanwhile, we conducted a one-year trial at another general hospital involving new nursing trainees. The trainees prepared for their actual onsite training at the hospital in advance by using the SceneKnowledge system to air their concerns or address problems by posting comments, and they continued to use the system periodically even after the training period was over. The trial demonstrated that the system worked just as expected: a good deal of practical knowledge and knowhow was accumulated as the more experienced veteran nurses responded to questions raised by the new trainees, and this proved remarkably effective for the transfer of practical techniques and skills to the trainees.

Text search for video metadata and comments

Comment input form

Video scene | Indication of number of comments | Bulletin board linked to each scene | Comments about a specific video time
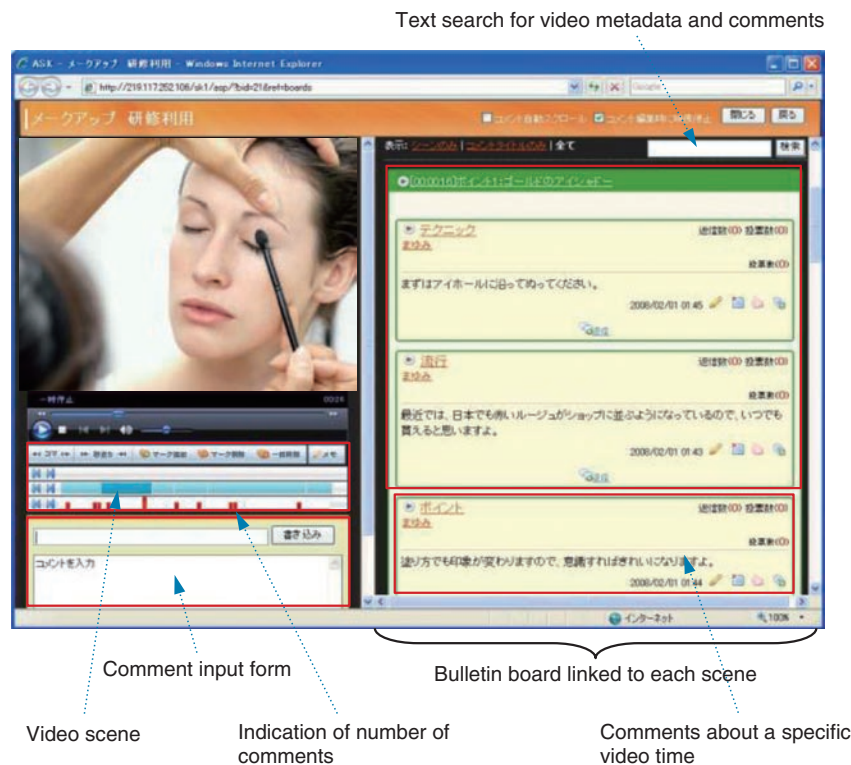
Fig. 3. Screenshot of a video-scene-linked bulletin board.

Turning to a totally different field, we conducted a collaborative trial with Sakuho Township, an agricultural community in Nagano Prefecture. Here, we found that the system helped rejuvenate the local farming community through the use of the intuitive easy-to-understand nature of video to share knowledge and address the basic issue of succession planning in farming communities: how to pass down technical skills to the younger generation.

3.2 University lecture field trials

3.2.1 New approach for creating lecture video content

Shooting and editing video to produce lecture-related content is extremely laborious and time-consuming. This led us to develop a new scheme using a fixed high-definition television (HDTV) camera to film lectures and automatically create clippings from the video stream of key areas in the lecture hall and deliver small quantities of HDTV clipping content for learning purposes [3]. The key areas for creating clippings are (1) the area around the instructor or lecturer that show his or her expressions and behavior, (2) the resource area showing screens, blackboards, or other presentation devices, and (3) the lecture hall area, which represents the overall venue. As illustrated in **Fig. 4**, the three key areas were extracted from the HDTV feed and then edited and synthesized using a predefined editing template to produce a polished video of the lecture. Any perspective distortion in the resource area or brightness adjustments around the lecturer were automatically corrected when the feed was synthesized. Note that scene changes in the resource area are automatically defined, so when a lecture is accompanied by slides or similar electronic media, the output is automatically divided into different scenes as the slides are presented.

3.2.2 Instructional design for better learning efficiency

We conducted a series of trials to see if learning efficiency might be improved by having students prepare weekly assigned reports and final reports on our system [4]. The trial procedure is summarized as follows. First, a class video was recorded; then, students posted their own views and opinions in the form of comments about the content until the next week. During the first half of the class, the instructor reviewed and addressed the students' comments from the
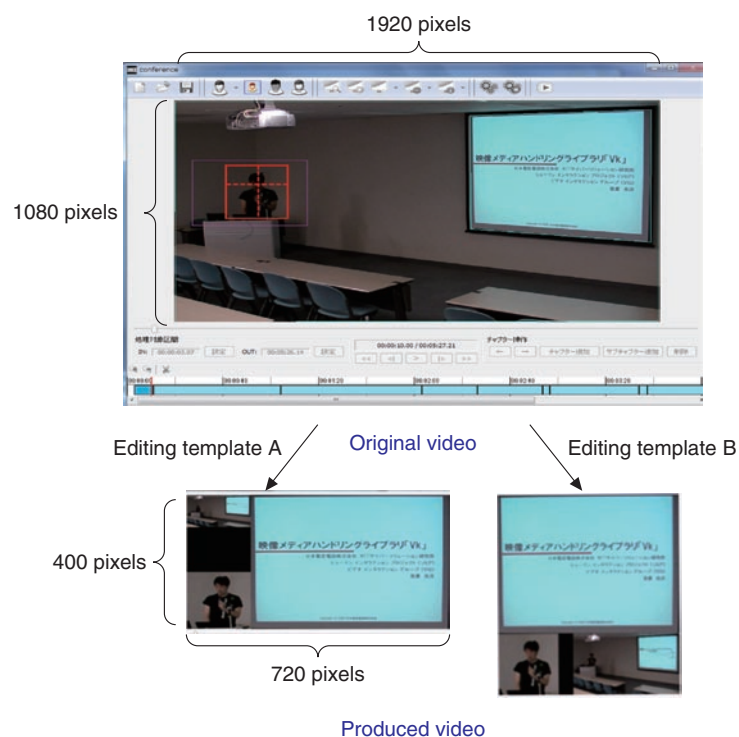
Fig. 4.   Principle of video editing.

previous session. Finally, the students were asked to select 3–10 scenes from the previous 12 classes that they found particularly interesting and to write a brief report describing what they learned from the class.

Continuing this approach for six months, we verified a number of interesting findings.

(1) Breaking the class video up into scenes proved to be effective for reading and writing comments, and the summary report of the class video was effective.

(2) Writing comments and reading other students' comments caused students to look back over the course, reflect on their own views, and compare their own ideas with those of their classmates; this clearly raised their interest in the class.

(3) Having the students prepare summaries of the class led them to discover new insights and organize their own thinking and thus deepened their understanding of the class content.

The system clearly has great potential for open education using publicly available lecture videos [5].

3.3   Application to e-learning

In order to assess the system's practicality and effectiveness in a commercial e-learning environment, we teamed up with NTT Knowledge Square Inc. (NTTKS) to conduct a joint trial beginning in October 2010 (**Fig. 5**). During the trial, several free lectures available on the NTTKS e-learning N-Academy website [6] were provided using the video-scene-linked bulletin boards provided by Scene-Knowledge. We were confident that the system's features—the ability to add comments to key scenes, quickly locate required scenes, make notations on video images, and so on—would not only deepen student understanding of the subject matter, but also facilitate a two-way dialog between the instructor and the students.

The system's effectiveness was confirmed by a follow-up questionnaire in March 2011; about 75% of the participants said that the system helped them understand the course. The trial was still underway at the time of writing. Anyone can sign up to participate, and we would encourage any interested readers to visit the site and give SceneKnowledge a try.

4.   Concluding remarks

SceneKnowledge is a system that partitions video into scenes and then uses scene-linked bulletin boards

Application to golf lesson course

Fig. 5.   Application to e-learning.

to enable discussions and comments about the scenes. Field trials demonstrated that the system works well for compiling knowhow and expertise and as an efficient learning aid. Building on the results that we have achieved so far, in future work we will focus on upgrading the video indexing technology, video display and processing technologies that can aid in discovering the knowledge that lies within video, and schemes for reorganizing posted video and comments.

References

[1]   I. Nonaka, "Knowledge-creating Company," Harvard Business Review, pp. 162–171, 1991.

[2]   Y. Majima, S. Shimada, and Y. Maekawa, "Field Experiments in Social Networking Service Using a Knowledge Sharing System with Nursing Skill Videos," Proc. of the 15th International Conference on Knowledge-based and Intelligent Information and Engineers Systems, Lecture Notes in Computer Science, Vol. 6884/2011, pp. 280–287, 2011.

[3]   S. Shimada, T. Nakanishi, A. Kojima, and Y. Fukuhara, "Semi-automatic and Easy Creation of Learning Friendly OCW Video Content," OCWC Global 2011.

[4]   M. Murakami, T. Marutani, K. Kakusho, S. Azuma, S. Shimada, and M. Mino, "Development of Making Course Summarization Using Lecture Video and Practice at University," Transactions of JSET, Vol. 34, No. 3, pp. 299–307, 2010 (in Japanese).

[5]   T. Nakanishi, S. Shimada, A. Kojima, and Y. Fukuhara, "Lecture Video and Scene-related Knowledge Sharing Common Platform: Design and Prototyping," Proc. of the 18th International Conference on Computers in Education (ICCE 2010), Putrajaya, Malaysia.

[6]   NTTKS e-learning N-Academy (in Japanese). http://n-academy.jp/

Satoshi Shimada

Senior Research Engineer, Human Interaction Project, NTT Cyber Solutions Laboratories.

He received the B.S., M.S., and Ph.D. degrees from the University of Kanazawa in 1983, 1985, and 1998, respectively. Since joining NTT in 1987, he has been engaged in research on image recognition, video handling, and knowledge management. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) and the Institute of Image Electronics Engineers of Japan (IIEEJ).

Ken Tsutsuguchi

Senior Research Engineer, Visual Media Communications Project, NTT Cyber Space Laboratories.

He received the B.E. and M.E. degrees and the Dr. of Informatics degree from Kyoto University in 1989, 1991, and 2001, respectively. In 1991, he joined NTT Human Interface Laboratories and engaged in research on computer graphics, computer animation, computer vision, and video/image processing. He is a member of IEEE Computer Society, the Association for Computing Machinery (ACM), the Information Processing Society of Japan (IPSJ), and IEICE.

Akira Kojima

Senior Research Engineer, Supervisor, Visual Media Communication Project, NTT Cyber Space Laboratories.

He received the B.E. and M.E. degrees in mathematical engineering and information physics from the University of Tokyo in 1988 and 1990, respectively. Since joining NTT in 1990, he has been engaged in research on video databases, multimedia information retrieval, video surveillance, and high-reality visual communications. He is a member of IEICE, IIEEJ, and ACM.

Hiroshi Konishi

Senior Research Engineer, Human Interaction Project, NTT Cyber Solutions Laboratories.

He received the B.E. and M.E. degrees in computer science and information mathematics from the University of Electro-Communications, Tokyo, in 1991 and 1993, respectively. He joined NTT Network Information Systems Laboratories in 1993. From 1995 to 2003, he worked on R&D of e-learning systems. Since 2003, he has been engaged in R&D of audio and video processing, and multimedia handling. He is a member of IEICE and the Acoustical Society of Japan.

Suguru Higashino

Senior Research Engineer, Supervisor, Human Interaction Project, NTT Cyber Solutions Laboratories.

He received the B.S. degree in applied physics from Tokyo University of Science in 1986 and the M.E. degree in information processing from Tokyo Institute of Technology in 1988. Since he joined NTT Human Interface Laboratories in 1988, he has mainly been engaged in R&D of image coding and human interfaces. From 2005 to 2008, he worked in the R&D Strategy Department. Since 2008, he has been engaged in the development of new IPTV services and new digital signage services. He is a member of IEICE, the Japanese Society for Information and Media Studies, and IIEEJ.