

## Extraction of Students' Events for a Distant Lecture Using a Computer Room

Suganuma, Akira  
Department of Intelligent Systems, Kyushu University

Inaki, Hiroaki  
Department of Intelligent Systems, Kyushu University

<https://hdl.handle.net/2324/5359>

---

出版情報 : Proceedings of the Third IASTED International Conference on Communications,  
Internet, and Information Technology. 2004, pp.228-232, 2004-11. International Association of  
Science and Technology for Development

バージョン :

権利関係 :

## EXTRACTION OF STUDENTS' EVENTS FOR A DISTANT LECTURE USING A COMPUTER ROOM

Akira Suganuma and Hiroaki Inaki  
Department of Intelligent Systems  
Kyushu University  
6-1, Kasuga-koen, Kasuga, 816-8580  
Japan  
email: {suga, inaki}@limu.is.kyushu-u.ac.jp

### ABSTRACT

Distance Education Course (DEC) has been becoming more and more popular in many universities with the fast development of the Internet technology. The communication between the teacher and the students in DEC environment is, however, a more difficult problem than that of conventional courses. This paper proposes a real-time detecting method of students' attitude to enhance the communications between the teacher and the students who take part in the lecture in a computer room. This method detects the events on the student's personal computer (PC), and detects the existence of the student in front of PC and the direction of the student's face by the image processing. The teacher can judge whether each student takes part in the lecture positively. Finally, our experimental results have illustrated the feasibility of this method.

### KEY WORDS

Distance education, Supporting tool, Detecting method of student's attitude, Application of the image processing, Communication between a teacher and students

## 1 Introduction

The growth of a communication network technology enables people to take part in a distant lecture. There are mainly two kinds of method to hold such a lecture. One is a web-based method; the other is a method of sending video and audio of lecture scene. We are studying some supporting systems for a distant lecture. In the web-based method, we have designed and developed a supporting system named CACCE: a Computer Aided Cooperative Classroom Environment[1]. It consists of a teacher's browser and a student's browser, which synchronously display a web page. On the other hand, we are developing a camera control system for the distant lecture held with the method sending a lecture scene. We call it "ACE": Automatic Camera control system for Education[2].

In a distant lecture, a teacher cannot watch his/her students in the distant classroom or he/she can watch them only by a monitor of the medium size. In the later case, his/her students in the monitor are too small to find their action. Furthermore, a camera stand on the platform cannot

take students' face because it is occluded by the student's PC. The teacher cannot consequently judge the attitude of his/her students in the distant classroom as well as he/she judges that of students in the local classroom. Some student portfolios assessments and behavior analysis methods are studied to solve this problem[3][4]. However, they are for supporting the lecture when students learn the HTML course-ware by themselves, and cannot timely support the teacher who is holding his/her lecture.

We designed and developed a function to inform a teacher of the attitude of his/her students in the distant classroom with cooperation between ACE and CACCE. Figure 1 illustrates a supporting system for the distant lecture we envisage. ACE is a system videoing a lecture. It takes a suitable shot even if the lecture is traditional style. In the lecture of the traditional style, a teacher writes something on a blackboard and explains them. The device used by the teacher is only the blackboard. In figure 1, ACE runs on the top of the blue-colored PC in the local classroom. ACE controls an active video camera according to the teacher's behavior and focuses on the area explained by teacher. Once ACE captures the explained areas, the connective subsystem stores their still frames on the web server.

When we designed CACCE, we assumed that each student in the distant classroom was assigned one personal computer (PC) to watch the teaching material. CACCE consists of teacher's browser and student's browser. The student's browser runs on each student's PC in figure 1. CACCE cannot only help the students to brows the teaching material to display the same page as the teacher's browser display but also arrow them to brows the still frames on demand. The student's browser always sends the URL of the displayed page to the teacher's browser.

We proposed a real-time analysis method of students' behavior with our cooperative system[5]. However, this method is useful for the lecture either using a browser or not using some other application. The method can detect difficult points in the lecture but cannot detect an earnest attitude and/or a do-nothing one of the students. We have designed, therefore, an additional system for spying students' attitude. In this paper, we describe the real-time detecting method of students' attitude with the additional system.

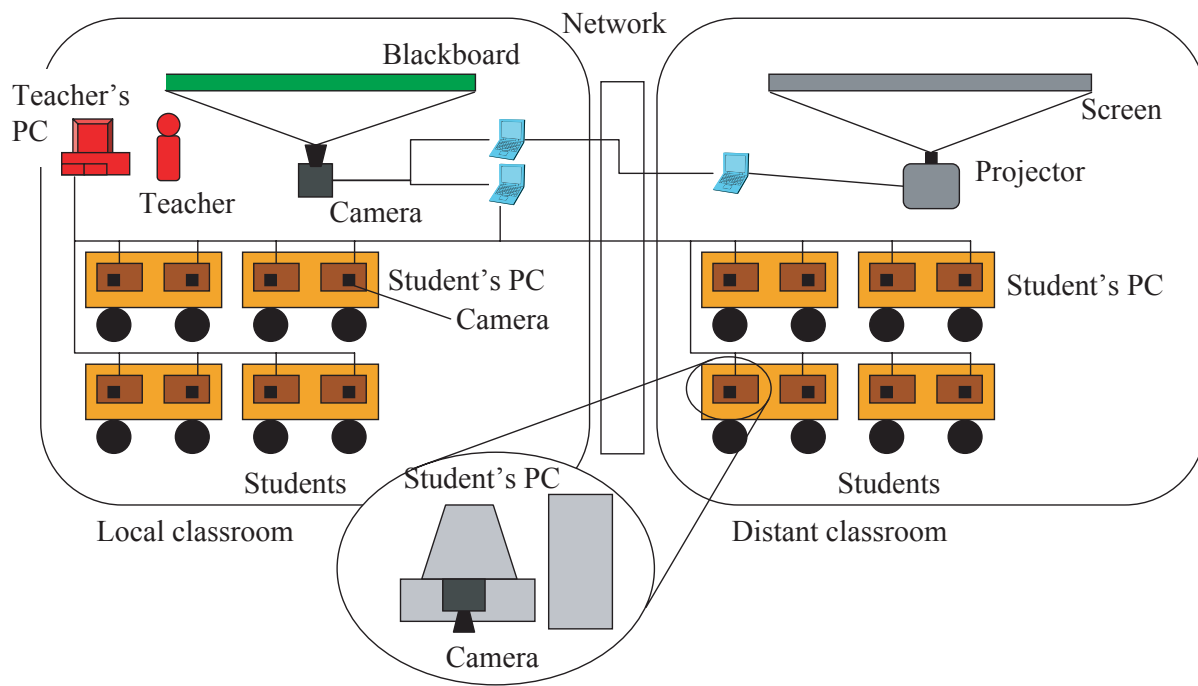


Figure 1. Supporting system for the distant lecture we envisage

## 2 Detection of a student's attitude about a lecture

### 2.1 In case of using browser

We designed and implemented the system combined ACE and CACCE. The connective subsystem supplies students the feature that they can refer a scene captured before with a browser on their own pace because it automatically makes a Web page of the captured image. We guessed that student's operation to his/her browser probably denotes the student's attitude. If the student uses the student's browser of CACCE, it can inform the teacher's subsystem of the URL of the displayed page. The transmitted information probably helps the teacher to adjust the progress of his lecture.

The student's browser of CACCE usually kept displaying the latest page captured by ACE by sending an URL of the page from the teacher's browser. The student's browser reports a URL of the displayed page and the time of beginning of displaying to teacher's browser whenever the student watches a page which is different from the page indicated by the teacher's browser.

We classified the student's reference into the following five types: (1) the reference under the teacher's direction, (2) the reference because of the narrow image projected on the screen, (3) the reference because of the difficult contents, (4) the reference because of the fast progress, and (5) the reference because of confirmation of contents of the lecture. It is important for a teacher to be informed the information whether his/her students understand his/her

explanation. What his students don't understand is also important. Furthermore, information about the progress of the lecture helps the teacher. We consequently designed a real-time analysis method to detect the reference of type (3) and (4) by using the URLs of the referred page and the indicated page and the time of beginning of referring.

Our analysis method detects type (3) and (4) references by eliminating other type of reference from all references. In the first phase, our method eliminates the reference to the page out of the still image of the lecture and the reference to the latest page of the lecture. In the second phase, it eliminates the reference of type (5) by using the feature of a short time reference. It can eliminate, furthermore, the reference of type (1) and (2) of reference because these references have the synchronism among almost all students. Finally, our method classifies the remains into type (3) or (4).

### 2.2 Detection of the event in the student's PC

There are two pattern of student's attitude in the lecture: one is an earnest attitude such as following the teacher's instructions, the other is a do-nothing attitude such as net-surfing and/or operating an application program not using in the lecture. The students operate the application in the foreground window in any case. Therefore, we designed our system to investigate the name of the foreground window on the student's PC. The student's subsystem, which runs on the student's PC as a background process, always compares the name with the list of the window name which was sent from the teacher's subsystem. If the student's sub-

system finds an unlisted name, it guesses that the student is operating another application. And it informs the teacher's subsystem of the fact and the name of the window.

The student's subsystem also investigates a student's operation of both a mouse and a keyboard. The student isn't always earnest even if he/she looks at the window designated by his/her teacher. The teacher may not instruct his/her students to watch the window, but instruct them to operate something. For example, when the teacher instructs his/her students to make a program on the PC, the student doesn't positively take part in the lecture if he/she only looks at the window of the programming environment. So the student's subsystem gets events on the student's PC such as the mouse operations and/or the keyboard operations, sends the information to the teacher's subsystem. Consequently, the teacher can or cannot amplify an explanation by adding details according to confirming whether his/her students following his/her instruction.

### 2.3 Detection the student's motion by the image processing

Our system has a function to detect the student's motion with the image captured by video camera connected to the student's PC. The function mentioned above enable a teacher to find whether his/her students operate the designated application. Some students may, however, only display the designated window on their PC, be actually sleeping or chatting with friends sitting the next seat. It is not impossible that they aren't in front of their PC. The teacher cannot check that with the function mentioned in the previous subsection. Our system requires a video camera on the display of the student's PC (see figure 1) in order to detect both the existence of a student in front of his/her PC and the direction of the student's face. The teacher can judge that his/her students positively take part in the lecture, if they are facing forward and their PCs are displaying the designated window.

Our system guesses the existence of a student by using the subtraction method between the successive frames and the skin color detection method. A student involves no little motion in the frames captured by a video camera if he/she sits in front of his/her PC. So our student's subsystem detects the differential pixels between two successive frames. Since the image has some noises, some differential pixels are detected even if no objects are moving. If the number of the differential pixels is greater than the threshold we preconcerted, the subsystem infers that a moving object is extracted. On the other hand, since the video camera takes a scene of the front of the PC, the skin color pixels, which the image of a student's face consists of, are detected in the captured image if a student sits in front of his/her PC. For the same reason as above, if the number of the pixels detected by the skin color detection is greater than the threshold, the subsystem infers that a skin color object is extracted. Our system infers that a student

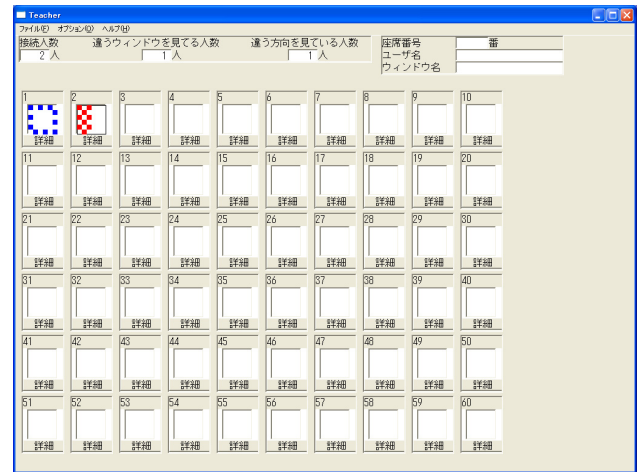


Figure 2. Supporting system for the distant lecture we envisage

isn't in front of the PC, if it extracts either a moving object or a skin color object.

Our system guesses the direction of a student's face by using the template matching method. There are some estimating methods of facial direction[6][7]. We selected the template matching method because the computing time can be decreased by using the eigenspace method[8]. We prepared the images of the full face, the left profile, the right profile and the slouch face. We established the eigen bases of the template image. In matching, the system extracts the facial area from the image captured by the camera with the skin color detection method. Since a student's face isn't always snapped by the equidistance, the facial area is normalized at 150x150 size. After the preprocess, our system converts the facial area to eigen coefficients with the same eigen bases and calculates the distance between the facial area and each template. Our system estimates the direction of a student's face from the template whose distance is the smallest.

## 3 Interface to a teacher

When a teacher gives a lecture, he/she has to get behavior of both students in the local classroom and a distant classroom. And the teacher would like to focus his/her attention on his/her explains. Then the interface must be easy to get information. Figure 2 shows the window of our teacher's subsystem.

The window of our system consists of three types of information. First information is the numbers of students who is connected to this system, students who displays another window on their PC and students who doesn't look at the display. This information is displayed at top of the window, informs the teacher how many students are earnest.

Second information is signs of each student's attitude. This information is displayed almost all window. Our sys-

Table 1. Experimental results of estimating the facial direction

	Student A	Student B	Student C	Student D	Student E	Average
Full face	84.5%	76.6%	79.8%	71.4%	76.4%	78.0%
Left profile	91.5%	80.4%	85.9%	83.9%	84.3%	85.1%
Right profile	72.2%	77.1%	74.2%	75.8%	74.4%	74.7%
Slouch face	88.9%	70.1%	77.0%	87.0%	75.0%	80.4%
Average	84.6%	75.8%	79.2%	78.7%	76.9%	79.1%

tem displays a blue sign when it infers that the student watches the designated window. A red sign denotes that the student watches another window. And it displays a black sign when the student isn't in front of his/her PC. There are many shapes of the sign. The checked sign denotes that the student types any keys. The sign of the bold colored frame denotes that the student clicks his/her mouse. Our system displays a left (right) half of the sign when it infers that the student look at the left (right) hand side. It displays a bottom half of the sign when the student look down.

The last information is the student's name and the name of the foreground window which the student displays on his/her PC. This information is displayed at the top of the window if the teacher clicks a particular button. For example, the top-left sign is a checked-bold-full-frame one colored blue. This means that the student looks at the designated window, types any keys and clicks his/her mouse. The teacher may estimate he/she positively take part in the lecture.

## 4 Experiment

We have tested our system to connect among one teacher's subsystem and 53 student's subsystems. The PCs used in our experiment are not a brand-new one; their CPU is Intel Celeron 500MHz, the capacity of their main memory is 128MB and the OS is Windows 2000. The student's subsystem runs as a background process. Each student's subsystem sends some information to the teacher's subsystem every five seconds. The network in the classroom is ordinary one (100Base-T). In this experiment, the function of estimating the direction of the student's face wasn't validated because we couldn't set a camera for each PC. We confirmed that the student's subsystem worked well during full 90 minutes lecture and it didn't affect another application which the student operated in the lecture. The interface to a teacher also got a good evaluation from the teacher.

We have done another experiment on the function of estimating the student's facial direction. We investigated the precision of estimating method. In this experiment, we used the PC whose CPU is Athlon<sup>TM</sup> XP 1800+ 1.53GHz. The capacity of its main memory is 512MB, WindowsXP Professional runs on the PC. A video camera was connected to the PC via usb protocol. We had prepared the templates using the facial image of four students. Our measurement was made with five students, which consisted of a

student whose facial image had been registered as the templates and four students whose facial images hadn't been registered. We made them make an ordinary operation of their PC. We observed and notated their facial direction, compared observed direction with the direction estimated by our system. The experimental results are shown in table 1. Student A is a student whose facial image was registered. The precision of student A is greater than any other students. We confirmed that our system could estimate all direction of all students' face with a probability of more than 70 percent even if the student's facial image isn't registered to the templates.

## 5 Conclusion

In this paper, we described our system which spies students' attitude. It enables the teacher to get some information about his/her students in the distant classroom. We plan to improve the graphical user interface. The interface to the teacher digitizes information how many students are watching their display. The teacher doesn't, however, know in his/her bones how high the ratio of the students who is earnest. We will make the interface easier to know in the teacher's bones with a pie chart and so on. Our system has the function estimating the direction of the student's face. The precision of this function, however, is not so good. The precision is decreased when the image from the camera includes a skin color object except student's face such as a hand. We try to make our estimating method more robust and increase the precision.

## Acknowledgments

This research was partly supported by a Grant-in-Aid for Scientific Research (Category C) from Japan Society for the Promotion of Science (JSPS) No. 14580224, 2002–2004.

## References

- [1] A. Suganuma, R. Fujimoto and Y. Tsutsumi, An WWW-based Supporting System Realizing Cooperative Environment for Classroom Teaching, *Proceedings of World Conference on the WWW and Internet*, 2000, pp.830–831.

- [2] A. Suganuma and S. Nishigori, Automatic Camera Control System for a Distant Lecture with Videoing a Normal Classroom, *Proceedings of World Conference on Educational Multimedia, Hypermedia & Telecommunications*, 2002, pp.1892–1897.
- [3] C. Liu, G. Chen, K. Ou, C. Lee and C. Lu, An instrument for on-line learning performance analysis by using decision tree technology on web based portfolios, *Proceedings of Advanced Research in Computers and Communications in Education*, 1999, pp.452–459.
- [4] H. Satoh, T. Jiang and S. Watanabe, HyperClassroom: The Relation between Learning History and Learning Environment on the Synchronously Learning Environment, *JSET Technical Reports*, 1998, pp.343–346 (in Japanese).
- [5] A. Suganuma, A Real-time Analysis Method of Students' Behavior for a Supporting System of a distance Lecture, *Proceedings of IADIS International Conference WWW/Internet*, 2003, pp.1183–1186.
- [6] Y. Araki, N. Shimada and Y. Shirai, Face Detection and Face Direction Estimation Using color and Shape Features, *International Conference on Machine Automation*, 2000, pp.351–355.
- [7] S. Ishii, D. Arita and R. Taniguchi, Real-time Head Pose Estimation with Stereo Vision, *Proceedings of the 9th Korea-Japan Joint Workshop Frontiers of Computer Vision (FCV 2003)*, 2003, pp.79–83.
- [8] S. Horimoto, D. Arita and R. Taniguchi, Real-time Hand Shape Recognition for a Human Interface, *Proceedings of the 9th Korea-Japan Joint Workshop Frontiers of Computer Vision (FCV 2003)*, 2003, pp.106–111.