

Face Recognition-based Lecture Attendance System

Yohei KAWAGUCHI [†]

Tetsuo SHOJI ^{††}

Weijane LIN [†]

Koh KAKUSHO ^{††}

Michihiko MINOH ^{††}

[†] Department of Intelligence Science and Technology, Graduate School of Informatics, Kyoto University

^{††} Academic Center for Computing and Media Studies, Kyoto University

Abstract

In this paper, we propose a system that takes the attendance of students for classroom lecture. Our system takes the attendance automatically using face recognition. However, it is difficult to estimate the attendance precisely using each result of face recognition independently because the face detection rate is not sufficiently high. In this paper, we propose a method for estimating the attendance precisely using all the results of face recognition obtained by continuous observation. Continuous observation improves the performance for the estimation of the attendance. We constructed the lecture attendance system based on face recognition, and applied the system to classroom lecture. This paper first review the related works in the field of attendance management and face recognition. Then, it introduces our system structure and plan. Finally, experiments are implemented to provide as evidence to support our plan. The result shows that continuous observation improved the performance for the estimation of the attendance.

1 Introduction

Though the video streaming service of lecture archive is readily available in many systems, students have few opportunities to view the lecture in this service because lecture content is not summarized. If the attendance of a student of classroom lecture is attached to the video streaming service, it is possible to present the video of the time when he was absent. It is important to take the attendance of the students in the classroom automatically.

ID tag or other identifications such the record of login/out in most e-Learning systems are not sufficient because it does not represent students' context in face-to-face classroom. It is also difficult to grasp the contexts by the data of a single moment.

student's context such as presence, seat position, status, and comprehension are discussed in this paper. At the same time face images reflect a lot about these context information. It is possible to estimate automatically whether each student is present or absent and where each student is sitting by using face recognition technology. It is also possible to know whether students are awake or sleeping and whether students are interested or bored in lecture if face images are annotated with the students' name, the time and the place. We are concerned with the method to use face image processing technology.

By continuously observing of face information, our ap-

proach can solve low effectiveness of existing face detection technology, and improve the accuracy of face recognition.

We propose a method that take the attendance using face recognition based on continuous observation. In this paper, our purpose is to obtain the attendance, positions and images of students' face, which are useful information in the classroom lecture.

2 Related work

Cheng, et al. [1] developed the system to manage the context of the students for the classroom lecture by using note PCs for all the students. Because this system uses the note PC of each student, the attendance and the position of the students are obtained. However, it is difficult to know the detailed situation of the lecture. our system takes images of faces.

In recent decade, a number of algorithms for face recognition have been proposed [2], but most of these works deal with only single image of a face at a time. By continuously observing of face information, our approach can solve the problem of the face detection, and improve the accuracy of face recognition.

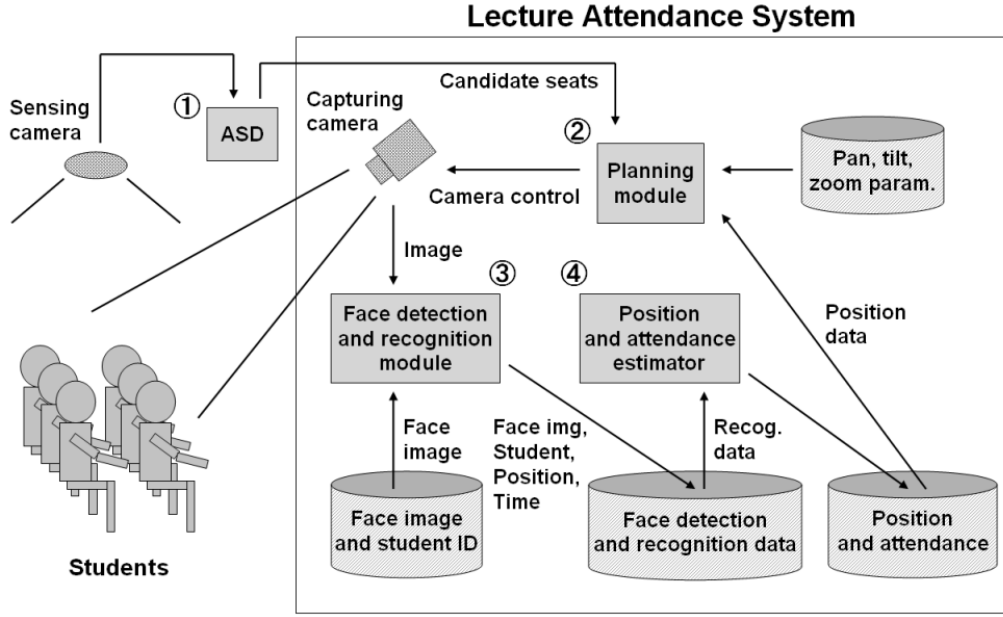


Figure 1: Architecture of the system

3 Lecture attendance system

3.1 Architecture

In this paper, our system consists of two kinds of cameras. One is the sensing camera on the ceiling to obtain the seats where the students are sitting. The other is the capturing camera in front of the seats to capture images of student's face. The procedure of our system consists the following steps (see Figure 1):

1. Seats information processing: this process determines the target seat to direct the camera. We adopt the approach called Active Student Detecting method (ASD) [3]. The idea of this approach is to estimate the existence of a student sitting on the seat by using the background subtraction and inter-frame subtraction of the image from the sensing camera on the ceiling.
2. Shooting plan: our system selects one seat from the estimated sitting area obtained by ASD, directs the camera to the seat and captures images.
3. The system processes the face images. the face images are detected from the captured image, archived and recognized. Face detection data and face recognition data are recorded into the database.
4. Attendance information processing: this process estimates the attendance by interpreting the face

recognition data obtained by continuous observation. The module obtains the most likely correspondence between the students and the seats under the constrained condition. The system regards a student corresponded to each seat as present. The position and attendance of the student are recorded into the database.

The procedure is repeated during lecture, and estimated the attendance of the students in real time.

3.2 Estimating students' existence

We use the method of ASD to estimate the existence of a student sitting on the seat. It is described in detail in [3]. In this approach, an observation camera with fish-eye lens is installed on the ceiling of the classroom and looks down at the student area vertically. ASD estimates students' existence by using the background subtraction and inter-frame subtraction of the images captured by the sensing camera (see Figure 2). In the background subtraction method, noise factors like bags and coats of the students are also detected, and the students are not detected if the color of clothes of them are similar the seats. ASD makes use of the inter-frame subtraction to detect the moving of the students.

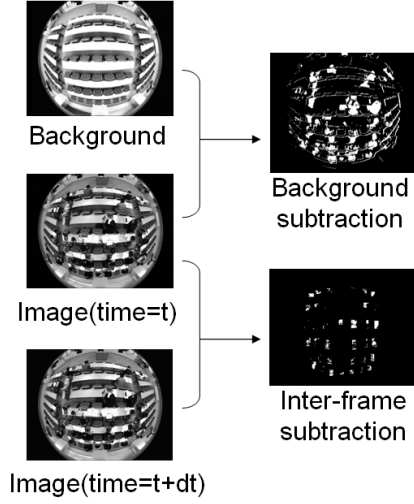


Figure 2: Active Student Detecting method

3.3 Shooting plan

Camera planning module selects one seat from the estimated sitting area in order to determine where to direct the front camera. Actually, in this paper, the module selects a seat by scanning the seats sequentially. This approach is insufficient because it wastes time directing the camera to where the student-and-seat the seats the students correspondence is already decided. In other words, if we direct the camera to each seat with the same probability, it is difficult to detect the faces according to the student or the seat, and the system judges the students who are actually present to be absent consequently. In order to solve this problem, it is important to the information of each student's position.

The camera is directed to the selected seat using the pan/tilt/zoom that have been registered in the database. The camera captures the image of the student.

3.4 Face detection and recognition

Face detection and recognition module detects faces from the image captured by the camera, and the image of the face is cropped and stored. The module recognizes the images of student's face, which have been registered manually with their names and ID codes in the database. Face detection data and face recognition data are recorded into the database.

3.5 Estimating the seat of each student

In order to solve the problem of ineffectiveness, we integrated students' seat information into the camera plan-



Figure 3: The face of the student on the back seat is detected.

ning. In this way, we can solve the problem such as mis-recognition of faces and seats by constraints of the correspondence relationship between them.

The face detected from the captured image may be another neighbor student's face (see Figure 3). Therefore, it is necessary to consider the possibility that the face image is the one of a neighbor student even if the camera is directed to the target seat.

Considering the points we mentioned above, we propose the following method. We assume that every seat has a vector of values that represent the relationship between the seat and each student. In the case that the module of face image processing recognizes Student A's face from the image of Seat B, our module votes for Student A's component of the vectors of the seats in the neighborhood of Seat B.

We assume the voting weights in Figure 4. Each cell means a seat, and the gray center cell means the focused seat. This assumption means that, when Student A is recognized at Seat B, 0.24 is voted to Seat B, and 0.11 is voted to the front seat of Seat B, and so on, for Student A's components. For example, Figure 5 shows Student A's components of each seat when Student A is recognized at the gray seat, and Figure 6 shows the case that Student A is recognized at the gray seat in the next step.

Considering the bipartite graph of the students and the seats, voting can be thought of as the addition to the scores of the edges between the students and the seats, and the cost of the edge is defined as the inverse of the score of the edge.

Before the seat information processing, we set two conditions as the premises:

- more than two students are not sitting on the same seat,
- the students do not move to different seats frequently.

The process of the seats information do not select independently the seat that has the highest score for each

Front		
0.06	0.11	0.06
0.11	0.24	0.11
0.06	0.18	0.06
Back		

Figure 4: An example of the voting weights

Front				
		0.06	0.11	0.06
		0.11	0.24	0.11
		0.06	0.18	0.06
Back				

Figure 5: 1) Student A's component of each seat when Student A is recognized at the gray seat

Front				
	0.06	0.11	0.06	
	0.11	0.30	0.22	0.06
	0.11	0.29	0.30	0.11
		0.06	0.18	0.06
Back				

Figure 6: 2) Student A's component of each seat when Student A is recognized at the gray seat after 1)

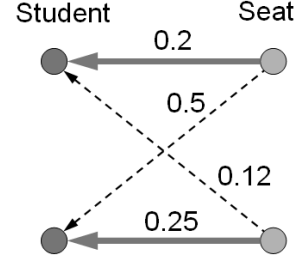


Figure 7: An example of 2 students and 2 seats

student but use the approach that find the matching in the bipartite graph such that the sum of the costs of the edges are minimized where the premises are satisfied. Figure 7 shows an example of the bipartite graph in the case that two students and two seats exist. In this case, our approach obtains the two thick arrows as the correspondence. Our process solves Linear sum assignment problem (LSAP) to estimate the correspondence. We assume the assignment of student i to seat j incurs a cost c_{ij} . The problem is formulated as follows:

$$\begin{aligned}
& \min \sum_{i=1}^n c_{ij} x_{ij} \\
& \sum_{i=1}^n x_{ij} = 1 \quad j = 1, \dots, n \\
& \sum_{j=1}^n x_{ij} = 1 \quad i = 1, \dots, n \\
& x_{ij} \in \{0, 1\} \quad i, j = 1, \dots, n
\end{aligned} \tag{1}$$

The least complexity of the best sequential algorithms for the LSAP is $O(n^3)$, where n is the larger one of the numbers of the students or the seats[4]. Thus, this problem is solved in real time.

In this procedure, the system regards the students corresponded to the seats as present.

4 Experiment

4.1 Result of Estimating the seat of each student

19 students existed in the center area, and we ran the process of camera control and detection for 20 minutes. We labeled the images of the detected faces with the name of the students manually. The system detected faces 186 times, and 15 students were detected. Table 1 shows the accuracy of seat estimation. We have compared the result of estimating the seat of each student

by using the method described in section 3.5. Method 1 is the method that corresponds each student to the seat where the most faces of the student are detected. Method 2 is the method that corresponds each student to the seat that has the lowest cost of the student. Method 3 is the method of section 3.5. Denominator of fractions in this table is the number of the face-detected students. This table shows that accuracy are improved by the method of section 3.5.

4.2 Result of Estimating the attendance based on continuous observation

We compared the results one cycle only and continuous observation. 12 students existed in the center area, and 2 of them did not have their faces registered. In this experiment of 79 minutes, 8 scanning cycles were completed during this period. Table 2 shows face detection rate, and Table 3 shows the result of estimating the attendance. In the case of 1 cycle only, we judge the recognized students to be present. In the case of continuous observation, the system estimates the attendance by the method of section 3.5 using the recognition data obtained during 79 minutes. This table shows that continuous observation improved the face detection rate and improved F-score of estimation of the attendance, which is the harmonic mean of precision and recall.

5 Conclusion and future directions

In this paper, in order to obtain the attendance, positions and face images in classroom lecture, we proposed the attendance management system based on face recognition in the classroom lecture. The system estimates the attendance and the position of each student by continuous observation and recording. The result of our preliminary experiment shows continuous observation improved the performance for estimation of the attendance.

Current work is focused on the method to obtain the different weights of each focused seat (in section 3.5) according to its location. We also need to discuss the approach of camera planning based on the result of the

Table 1: Result of estimating the seat of each student

Method	Accuracy
Method 1	60.0% (9/15)
Method 2	73.3% (11/15)
Method 3	80.0% (12/15)

Table 2: Face detection rate

Time	face detection rate
1 cycle only	37.5% (3.8/10)
79 min	80.0% (8/10)

Table 3: Result of estimating the attendance

Time	precision	recall	F-score
1 cycle only	89.2%	33.8%	48.3%
79 min	70.0%	70.0%	70.0%

position estimation in order to improve face detection effectiveness. In further work, we intend to improve face detection effectiveness by using the interaction among our system, the students and the teacher.

On the other hand, our system can be improved by integrating video-streaming service and lecture archiving system, to provide more profound applications in the field of distance education, course management system (CMS) and support for faculty development (FD).

Acknowledgements

The authors would like to thank Omron Corporation for their help to providing OKAO vision library used in face detection and recognition in our system.

References

- [1] K. Cheng, L. Xiang, T. Hirota and K. Ushijima, "Effective Teaching for Large Classes with Rental PCs by Web System WTS," in Proc. Data Engineering Workshop 2005 (DEWS2005), 2005, 1D-d3 (in Japanese).
- [2] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," ACM Computing Surveys, 2003, vol. 35, no. 4, pp. 399-458.
- [3] S. Nishiguchi, K. Higashi, Y. Kameda and M. Minoh, "A Sensor-fusion Method of Detecting A Speaking Student," IEEE International Conference on Multimedia and Expo (ICME2003), 2003, vol. 2, pp. 677-680.
- [4] R.E. Burkard and E. Çela, "Linear Assignment Problems and Extensions", In Handbook of Combinatorial Optimization, Du Z, Pardalos P (eds). Kluwer Academic Publishers: Dordreck, 1999, pp. 75-149.