Journal of Information Processing Vol.0 1–12 (??? 1992)

[DOI: 10.2197/ipsjjip.0.1]

Regular Paper

An IoT System with Business Card-Type Sensors for Collaborative Learning Analysis

Shunpei Yamaguchi[1,a)] Shusuke Ohtawa[2,b)] Ritsuko Oshima[2,c)] Jun Oshima[2,d)] Takuya Fujihashi[1,e)] Shunsuke Saruwatari[1,f)] Takashi Watanabe[1,g)]

Received: xx xx, xxxx, Accepted: xx xx, xxxx

Abstract: Collaborative learning practices foster the ability to solve creative problems in collaboration with other learners. The collaboration enables learners to learn new ideas from other learners and enhances the social ability of the learners through interaction with other learners. Although the learning science field now uses qualitative analysis to analyze the effects of the collaborative discourse, qualitative analysis requires much human and time costs to analyze the collaborative discourse with dozens of students. This study proposes Sensor-based Regulation Profiler to reduce the analysis costs. The proposed scheme consists of the business card-type sensors that acquire sensor data from each learner with a precise time synchronization as well as learning analysis methods that analyze the collaborative discourse from the acquired sensor data. Experimental evaluations using the proposed scheme showed that the proposed business card-type sensors realized the time synchronization error of 7.7 μs on average across the sensors. In addition, the proposed learning analysis could extract and visualize the collaborative activity of each learner in the collaborative discourse through the social graph extraction, learning phase extraction, speaker identification, and activity estimation by using the sensor data from the proposed business card-type sensors.

Keywords: Collaborative learning, Human activity recognition, Sensor-based learning analysis, Time synchronization

1. Introduction

Collaborative learning fosters the ability to solve challenging problems in collaboration with other learners. Each learner can learn new skills and improve social skills through the collaboration with other learners. The existing studies on collaborative learning find that distinctive interaction patterns increase the effectiveness of the collaborative learning [17–21, 25]. However, the existing studies require a much time to find the patterns during the collaborative learning activity. In the existing studies, the researchers in learning science manually analyze each learner's activity from the recorded video and transcribed voice information. Such qualitative analysis is difficult to be carried out for the collaborative learning activity with a large number of the learners due to the amount of data and manual analysis, and thus the results of the qualitative analysis are difficult to be immediately fed back to the learners.

In this study, we propose Sensor-based Regulation Profiler to quantitatively analyze the activity of learners in collaborative learning and automatically extract and visualize the key points of the activity to support the qualitative analysis by the researchers in learning science. The proposed Sensor-based Regulation Profiler consists of business card-type sensors, namely, Sensor-based Regulation Profiler Badges and sensor-based learning analysis. The Sensor-based Regulation Profiler Badges acquire sensor data from the learners in collaborative learning. In addition, each Sensor-based Regulation Profiler Badge has a RF-based time synchronization module to achieve high-precision synchronization of the sensor data between all the sensors. The proposed sensor-based learning analysis automatically analyzes and visualizes the key points of the collaborative learning activity from the acquired and synchronized sensor data. The proposed Sensor-based Regulation Profiler allows researchers in learning science to analyze the collaborative learning activity in more detail through the automatic extraction and visualization of the activity from the acquired sensor data. We envisage the use of automatic extraction results to find groups whose collaborative learning activity is not working well and to help the researchers who navigate the learning activity of those groups in real-time.

From experimental evaluations using the proposed Sensor-based Regulation Profiler, we found 1) the average time synchronization error in the proposed Sensor-based Regulation Profiler Badges is approximately 7.7 µs, 2) the proposed scheme automatically extracts social graphs that represent face to face interactions between learners, and 3) the proposed method also extracts learning phases, speakers, and learner's activity during collaborative learning. For example, the proposed scheme accurately identifies

[1] Graduate School of Information Science and Technology, Osaka University, Suita, Osaka 565–0871, Japan
[2] Graduate School of Integrated Science and Technology, Shizuoka University, Hamamatsu, Shizuoka 432–8011, Japan
[a)] yamaguchi.shunpei@ist.osaka-u.ac.jp
[b)] ohtawa.shusuke.15@shizuoka.ac.jp
[c)] roshima@inf.shizuoka.ac.jp
[d)] joshima@inf.shizuoka.ac.jp
[e)] fujihashi.takuya@ist.osaka-u.ac.jp
[f)] saru@ist.osaka-u.ac.jp
[g)] watanabe@ist.osaka-u.ac.jp

speakers under different numbers of users, environmental noises, and reverberation conditions as well as short utterances.

The remainder of this paper is as follows: Section 2 describes the requirements for introducing quantitative analysis to collaborative learning environments. In Sec. 3, we present the overview of our proposed scheme, namely, Sensor-based Regulation Profiler. Sections 4 and 5 also present the details of the proposed Sensor-based Regulation Profiler Badge and sensor-based learning analysis. Experimental evaluations are carried out in Secs. 6 and 7 to demonstrate the performance of the proposed Sensor-based Regulation Profiler Badge and sensor-based learning analysis. Section 8 discusses the related studies and Sec. 9 finally concludes our paper.

2. Requirements

There are two requirements to deploy a quantitative analysis tool into collaborative learning activity.

(1) Precise time synchronization across each learner's sensor data
(2) Automatic extraction of the key points during collaborative learning

The first issue is required for the extraction of interpersonal collaboration. A main objective of this study is to analyze collaborative learning activity from the business card-type sensors deployed on various targets such as learners and learning environments. There are several business card-type sensors such as Hitachi's business microscope [29, 30] and Massachusetts Institute of Technology (MIT)'s Sociometric Badge [31], Open Badges [11], and Rhythm [12]. For example, the business microscope enabled the quantitative analysis of the communication in the organization to solve communication issues by analyzing the social networks between the members of the organization. However, the conventional business card-type sensors did not realize precise time synchronization across the business card-type sensors. However, the studies have a drawback in precise analysis of human collaboration in terms of synchronization across sensor data. The studies attempt to synchronize sensor data by means of software correction. For example, the study [30] finds similar patterns in each sound pressure and synchronizes sound pressure data sampled at 8 kHz within 100 ms. Pattern recognition further decreases the synchronization accuracy across the sensors at lower sampling rate for low power consumption. Such error causes inaccurate and meaningless analysis of collaborative learning activity. To prevent the meaningless analysis owing to the synchronization error, the synchronization accuracy should be less than one-tenth of the sensor's maximum sampling rate. For example, if the maximum sampling rate of the sensor data is 100 Hz, the sensors should be synchronized with the time within 1 ms or less each other.

The second issue is required to connect qualitative and quantitative analysis in collaborative learning. Ideally, quantitative analysis should automatically extract all of the same results as those obtained in qualitative analysis by the researchers in learning science. However, it is difficult to replace all of the qualitative analysis with the quantitative analysis at present because there is a significant gap between the information that can be acquired by machines and humans. To fill the gap between the qualitative and quantitative analysis, our study focuses on automatic extraction of learning phases, social graph, speakers, and learner's activity. The extracted results can be used to detect speech segmentation during collaborative learning activity. The detected speech sections can help the part of the conversation analysis, i.e., the part of the qualitative analysis, by the researchers in learning science.

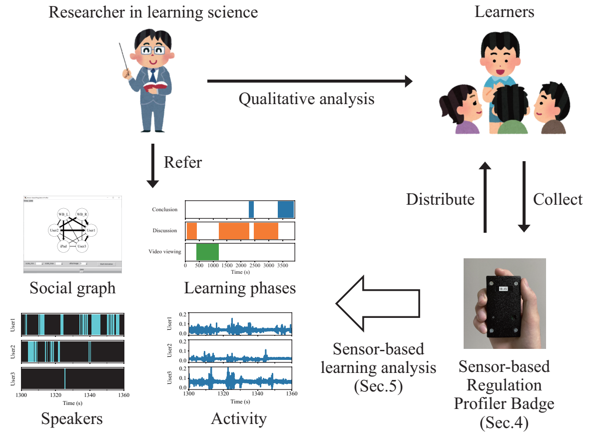


Fig. 1: Overview of collaborative learning analysis using the proposed Sensor-based Regulation Profiler.

3. Proposed Scheme: Sensor-based Regulation Profiler

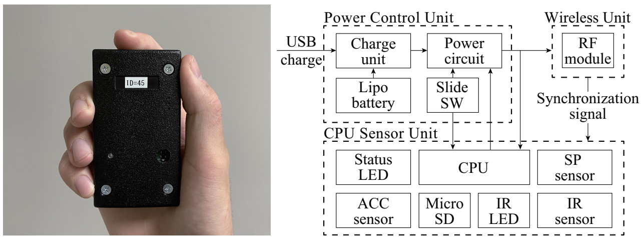
In order to reduce the cost of qualitative analysis in collaborative learning, we propose Sensor-based Regulation Profiler. The proposed scheme consists of Sensor-based Regulation Profiler Badge that acquires sensor data from learners and sensor-based learning analysis that extracts collaborative learning activity from the acquired sensor data. Figure 1 shows the overview of the collaborative learning analysis using the proposed Sensor-based Regulation Profiler. The Sensor-based Regulation Profiler supports qualitative analysis of collaborative learning by the researchers in learning science as following steps:

(1) Distribute the proposed Sensor-based Regulation Profiler Badges to learners
(2) Conduct collaborative learning activity between the learners
(3) Collect the distributed Sensor-based Regulation Profiler Badges from the learners
(4) Extract sensor data from the collected Sensor-based Regulation Profiler Badges
(5) Automatically extract and visualize the social graph, learning phases, and each learner's speech and activity from the sensor data
(6) Carry out qualitative analysis by learning science researchers using the visualized results of social graph extraction, learning phase extraction, speech detection, and activity estimation

The proposed Sensor-based Regulation Profiler Badge is described in Sec. 4 and the proposed learning analysis methods are described in Sec. 5, respectively.

4. Sensor-based Regulation Profiler Badge

Figures 2 (a) and (b) show the proposed Sensor-based Regula-



(a) Sensor node (b) Block diagram

Fig. 2: Sensor-based Regulation Profiler Badge.

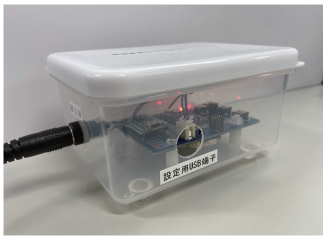


Fig. 3: Sync node.

tion Profiler Badge (Sensor node) and the block diagram of the Sensor-based Regulation Profiler Badge. Fig. 3 shows Sensor-based Regulation Profiler Badge Synchronizer (Sync node). The Sensor-based Regulation Profiler Badge is a business card-type sensor designed to be worn on the chest of each learner. The sensor node consists of a power control unit, a CPU sensor unit, and a wireless unit.

The power control unit has a lithium-ion battery to drive the sensor node. The lithium-ion battery supplies power to the power switch and the Micro Controller Unit (MCU). The sensor node can continuously run for 24 hours.

The CPU sensor unit is equipped with STM32L476RGT6 from STMicroelectronics as the MCU, ADXL362 accelerometer from ANALOG DEVICES, OSI5LAS1C1A infrared light emitting diode (LED) from OptoSupply, PIC79603 infrared receiver from KODENSHI CORP., and INMP510 analog microphone from TDK. The 3-axis accelerometer samples 12 bits at 100 Hz and the sound pressure sensor samples 12 bits at 100 Hz. The microSD card connector of DM3AT-SF-PEJM5 from Hirose Electric is used to record the sensor data. The acceleration data, infrared data, and sound pressure data can be recorded in a microSD card.

The wireless unit uses CC2650 from Texas Instruments which contains a wireless synchronization module. The wireless synchronization module transfers a synchronization signal transmitted every 10 ms from the sync node to other sensor nodes to realize time synchronization between the sensor nodes. The CC2650 uses UNISONet, which is also known as Choco [7, 28], to realize precise time synchronization between the sensor nodes. In Choco, an arbitrary sensor node forwards a time-synchronous packet to the neighboring sensor nodes and then propagates the received time-synchronous packet to the destination node. When a sensor node receives a new time-synchronous packet from the neighboring sensor node, it immediately forwards the packet to the neighboring sensor nodes. Each sensor node repeatedly receives and forwards time-synchronous packets by flooding, resulting in fast propagation of time-synchronous packets throughout the sensor nodes.

5. Sensor-based Learning Analysis

We also propose sensor-based learning analysis methods to automatically and precisely extract information of collaborative learning using sensor data obtained from Sensor-based Regulation Profiler Badges. The proposed learning analysis supposes to support researchers in learning science to post analysis of collaborative learning. The proposed method realizes social graph and learning phase extractions from the infrared LED sensors, speaker identification from the sound pressure sensors, and activity estimation from the three-axis accelerometers.

Table 1: Notation

Variable / Function	Description
U	Set of all the sensor IDs
L	Set of the infrared data obtained from all the sensors
l_d	Infrared data of sensor d
t_0	Target time for social graph extraction
G	Social graph matrix with the size of $\|U\| \times \|U\|$
W	Window size (s)

Algorithm 1 Social graph extraction

Require: L, U, t_0
Ensure: G

```
Insert zeros into all elements of G
for all d ∈ U do
    S ← all received IDs in l_d ∈ L between t_0 to t_0 + W
    for all s ∈ S do
        Increment G[s][d]
    end for
end for
return G
```

5.1 Social Graph Extraction

The social graph extraction visualizes social graphs that represent the network of the learners in collaborative learning from the face to face relationship between the Sensor-based Regulation Profiler Badges. The face to face relationship can be measured from the infrared data of each Sensor-based Regulation Profiler Badge. We note that the social graph extraction detects not proximity but face to face across the users. The infrared data contain the IDs of the other Sensor-based Regulation Profiler Badges detected every second. Algorithm 1 shows the procedure of our social graph extraction and Table 1 shows the notation of the algorithm. Algorithm 1 outputs the matrix G which represents the social graph at t_0 for L from the set of all sensor IDs U, the set of infrared data from all the sensors $L = \{l_1, l_2, \ldots, l_{|U|}\}$, and time instant t_0. The matrix G counts the number of infrared data across all the sensor nodes between t_0 and $t_0 + W$, where the row is the source sensor ID and the column is the destination sensor ID.

For example, we consider that the infrared data of learner 1's Sensor-based Regulation Profiler Badge are as follows.

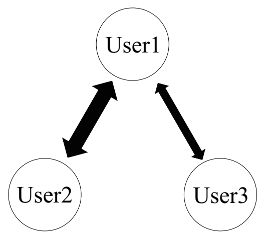


Fig. 4: Example of social graph corresponding to the matrix G.

```
900000000, 1, 2
900000001, 2, 3, 2
900000002, 1, 2
```

Here, the infrared data consist of the time stamp, the number of other sensors detected by the infrared Light Emitting Diode (LED) sensor since the last time stamp, and ID of each detected sensor. We also consider that the infrared data of learner 2's Sensor-based Regulation Profiler Badge are as follows.

```
900000000, 1,1
900000001, 1,1
900000002, 1,1
```

In addition, the infrared data of learner 3's Sensor-based Regulation Profiler Badge are as follows.

```
900000001, 1,1
```

In this case, the matrix G is given as follows:

$$G = \begin{pmatrix} 0 & 3 & 1 \\ 3 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

The resultant matrix represents the directed graph shown in Fig. 4. Here, a larger value of the matrix indicates a higher frequency of face to face interaction between the learners and we use darker arrows in Fig. 4.

5.2 Learning Phase Extraction

The learning phase extraction automatically divides a collaborative learning period into learning phases based on the time variation of the learners' network during collaborative learning. The appropriate number of learning phases depends on each case of collaborative learning [2, 4, 9, 13, 22, 24, 26, 27, 38, 39]. For example, the study [13] classifies learning activity into two phases: introduction and discourse and analyzes each phase. Our learning phase extraction supposes collaborative learning activities composed of three phases named as video viewing, discussion, and conclusion based on the study [2] which uses the same educational material called the Adventures of Jasper Woodbury [5]. The video viewing phase obtains information from the video to solve the problem. The discussion phase is that the learners discuss the problem and bring their ideas to solve the problem. The conclusion phase decides one answer from the ideas brought in the discussion phase.

To divide the collaborative learning period into the learning phases, we regard the time variation of the learners' network as the transition of the learning phase. At the transition of the learning phase, distinctive points appear in the face to face relationship between the learners. For example, some learners may increase their attention to the different learners/objects or concentrate on a large number of the learners rather than the ones they were facing before. We capture the distinctive points from the time variation of the learners' network by using the infrared data of each learner's Sensor-based Regulation Profiler Badge.

The time variation of the learners' network is quantified from the time variation of the matrix G. The matrix G of each time t_0 is extracted according to Algorithm 1. We consider that the matrices are extracted every 3 seconds and the window size W of each matrix is set to 60 seconds. After created the matrix G for each window, it regards the residual sum of squares between the matrices of one window and the next window as the time variation of the network.

We use AutoPlait [14] to quickly and automatically extract each learning phase from the network time variation. AutoPlait detects the features from large time series data containing various patterns and regards the features as the groups of time series data. The proposed method discovers each learning phase as each group extracted by AutoPlait.

5.3 Speaker Identification

Figure 5 shows an overview of the proposed speaker identification algorithm. There are three steps for speaker identification: 1) pre-processing of sound pressure data, 2) speech section estimation, and 3) speaker identification.

1) Pre-Processing: The first step extracts the sound pressure detection for each user. The algorithm calculates the minimum sound pressure value for each user and subtracts the minimum value from all the sound pressure data to make a zero-point correction. The algorithm labels whether each user speaks with sliding windows for the sound pressure data of each user obtained by zero-point correction for each window. Algorithm 2 exhibits the labeling procedure in Figure 5, and Table 2 lists the algorithm notation. Algorithm 2 outputs the array $\mathbb{A}$, which represents "the 1–0 data for each user" from the set of all sensor IDs U and the set of the sound pressure data from all the sensors $\mathbb{S} = \{S_1, S_2, \ldots, S_{|U|}\}$. We find the maximum of the sound pressure m for each user in each window W in line 6. If the maximum m in window W does not exceed the speech threshold η_s across all users, it is assumed that the speech of the user is not detected in window W, and the window slides in line 16. If the maximum m in window W exceeds the speech threshold η_s, the algorithm updates a threshold η_m as $m * 0.1$ in line 8. The algorithm compares the sound pressure of a user with the threshold η_m and assigns 1 if the sound pressure is higher than the threshold and 0 if the sound pressure is lower than the threshold in lines 9–13. The labels w in window W overwrite the corresponding elements of array A_d in line 14. We call the data obtained through pre-processing "the 1–0 data for each user."

2) Speech Section Estimation: The second step extracts the presence or absence of a user's speech from the 1–0 data for each user. The algorithm fills the data using the 1–0 data for each

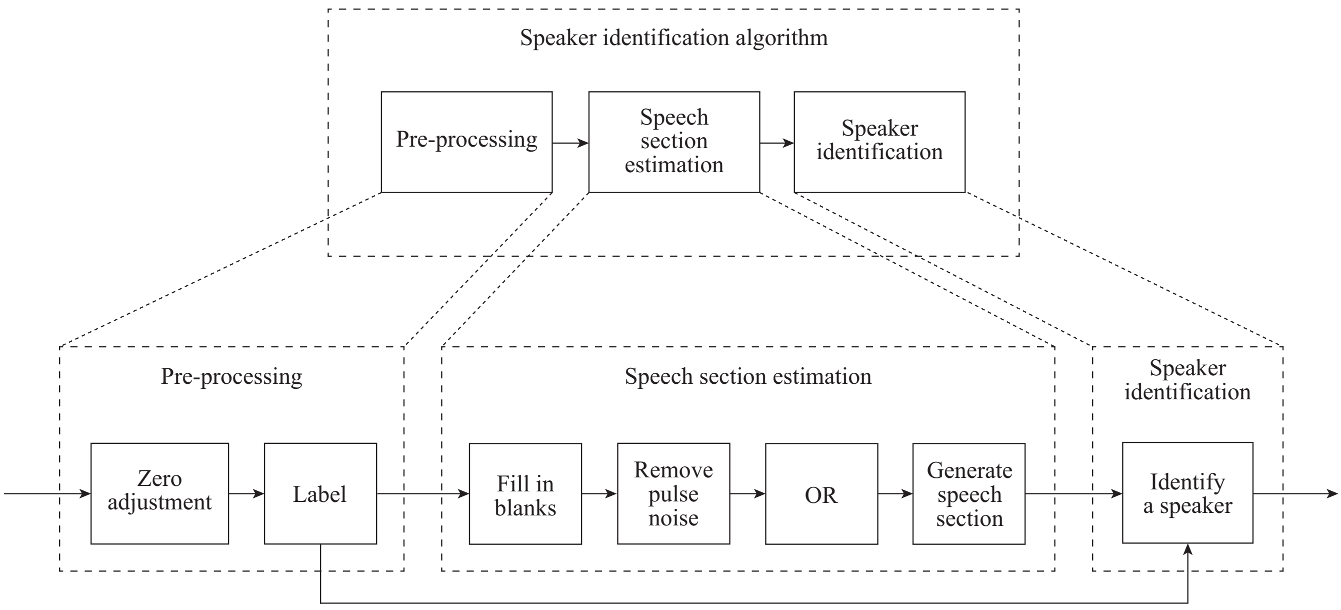


Fig. 5: Overview of the speaker identification algorithm.

Table 2: Notation

Variable / Function	Description
U	Set of all sensor IDs
d	Sensor ID
$\mathbb{S}$	Set of the sound pressure data obtained from all the sensors
S_d	Sound pressure data for sensor d
$\mathbb{A}$	Set of 1 bit arrays with speech labels
A_d	1 bit arrays with speech labels of sensor d
ξ	Top index of window
D	Window size
η_s	Speech threshold for all users
η_m	Speech threshold based on maximum sound pressure in the window
$\max(X)$	Calculate the maximum of all the elements of X

Algorithm 2 Labeling in pre-processing

Require: $U, \mathbb{S}$
Ensure: $\mathbb{A}$

for all $d \in U$ **do**
  Insert zeros into all elements of A_d
  $\xi \Leftarrow 0$
  while $\xi <$ length of A_d **do**
    $W \Leftarrow S_d \in \mathbb{S}$ between ξ to $\xi + D$
    $m \Leftarrow \max(W)$
    if $m > \eta_s$ **then**
      $\eta_m \Leftarrow m * 0.1$
      if $w \in W > \eta_m$ **then**
        $w \Leftarrow 1$
      else
        $w \Leftarrow 0$
      end if
      Insert $w \in W$ into elements of A_d with OR
    end if
    $\xi \Leftarrow \xi +$ slide width
  end while
  Insert A_d into $\mathbb{A}$
end for
return $\mathbb{A}$

user. The algorithm complements labels 1 in a section with consecutive labels 0 within 90 ms between labels 1 considered in the middle of speech in the 1–0 data for each user. The algorithm removes pulse noise using 1–0 data for each user with complements. The algorithm replaces a short interval with continuous labels 1 within 150 ms by labels 0, assuming that the section is where speech is falsely detected by ambient noise. The algorithm takes the logical summation of the 1–0 data for each user with pulse noise removal. We call the binary data obtained through the speech section estimation “the speech section data.”

3) Speaker Identification: The third step determines who speaks in each speech section by combining the 1–0 data for each user and speech section data. The algorithm focuses on each section where a user is considered to speak based on the speech section data. The algorithm extracts a user with the most labels 1 in each speech section and regards the user as a speaker in the speech section on the basis of the 1–0 data for each user.

5.4 Learner’s Activity Estimation

Each learner’s activity can be estimated from the accelerometer’s data during collaborative learning. We first take L2-norm across three-axis accelerometer’s data every sample in each Sensor-based Regulation Profiler Badge for the activity estimation motivated by the study on the human activity estimation using the accelerometer [23]. Our proposed sensor mounts an acceleration sensor ADXL362. ADXL362 quantizes and records acceleration within twice the gravitational acceleration. We subtract the offset in the data sheet of ADXL362 from all acceleration data to make a zero-point correction. We convert quantized acceleration to relative values from 0 to 1 and visualize the values as learners’ activity.

6. Experimental Evaluation: Time Synchronization Preciseness

We experimentally evaluated the time synchronization accu-

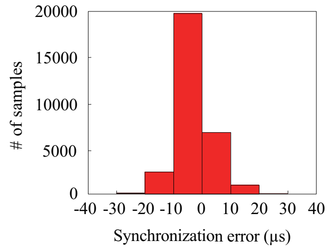


Fig. 6: Time synchronization accuracy between the sync and sensor node.

Fig. 7: Experimental environment of collaborative learning.

racy between the sync node and the sensor node in the proposed Sensor-based Regulation Profiler Badge. We set up a sync node and a sensor node at a short distance on a desk and measured the time deviation between the nodes based on the synchronization signals sent from the sync node. We used an oscilloscope to measure the clock rise time at each node to accurately get the time deviation between the nodes. We assumed that the number of samples was 30,003 and the wireless synchronization module of each Sensor-based Regulation Profiler Badge transmits a synchronization signal every 10 ms.

Figure 6 shows the time synchronization accuracy between the nodes. The horizontal axis shows the deviation of the time synchronization and the vertical axis indicates the number of the samples corresponding to the deviation. Fig. 6 shows that the time synchronization error is kept within ± 30 µs. Here, the mean and maximum synchronization errors are −7.7 µs and 30 µs, respectively. Since the sampling rate of both the sound pressure sensor and the acceleration sensor on the Sensor-based Regulation Profiler Badge is 100 Hz, the synchronization error is sufficient to meet the required synchronization accuracy of less than 1 ms. The proposed synchronization structure contributes to accurately analyze sensor data for collaborative learning.

7. Experimental Evaluation: Sensor-based Learning Analysis

We carried out collaborative learning experiments and evaluated the feasibility of the automatic extraction of social graph, learning phases, speakers, and activity with each learner's Sensor-based Regulation Profiler Badge during the activity. Figure 7 shows a snapshot of our experiments in the collaborative learning activities. We carried out five cases of collaborative learning activities with three learners and monitored each activity who possessed Sensor-based Regulation Profiler Badges. We set a sync node at the center of the learners' desk for synchronization between the sensor nodes. Each learner mounted a Sensor-based Regulation Profiler Badge on his/her chest in case 1 and his/her head in cases 2 through 5 during the activity. A whiteboard was set up to assist the learners in their discussions. Two Sensor-based Regulation Profiler Badges were put on both edges of the whiteboard. In addition, an iPad was placed on the desk to present learning tasks to the learners and one Sensor-based Regulation Profiler Badge was attached to the top of the iPad.

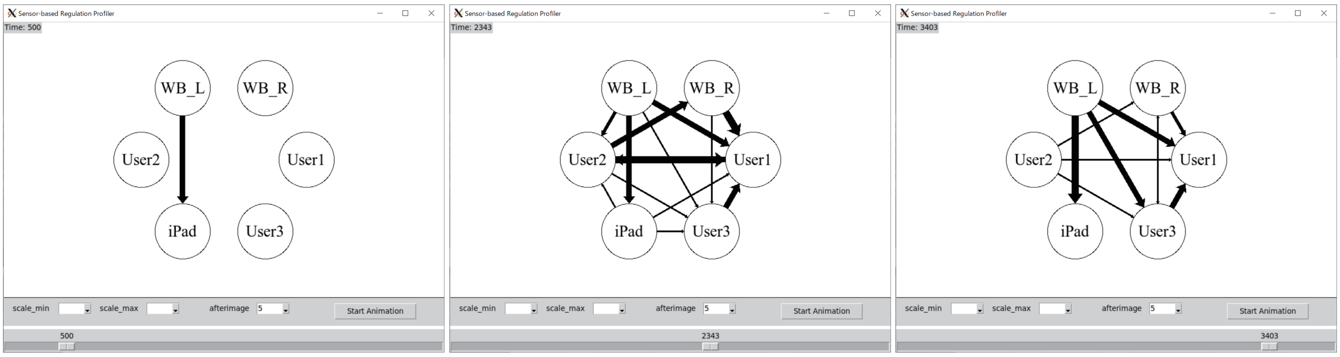
7.1 Results on Social Graph Extraction

We evaluated the accuracy and validity of face to face detection with social graph extraction. We did an experiment with infrared sensors in our proposed business card-type sensors attached to three users to calculate the accuracy of face to face detection. We prepared the room with the dimensions of 10.6 m × 7.05 m × 2.65 m. The room furnished multiple LED recessed ceiling lights. Each user stood 1.50 m away from the other users and two out of three users spoke face-to-face for 60 s. A non-speaker faced the middle between two speakers during the conversation. We tried all combination of speakers in the conversation and calculated the accuracy of face to face detection.

We found that 1) the infrared sensors detect face to face with an accuracy of 75.3 %, 78.0 %, and 78.0 % in each combination of speakers and 2) our proposed face to face detection sufficiently supports researchers in learning science to reduce qualitative analysis cost of face to face in each experiment case. Figures 8 (a) through (c) show social graphs in the learning phases in case 1 as an example. The learning elapsed time (s) is displayed at the bottom. In addition, Users 1 through 3 are the sensor nodes mounted on each learner, WB_R and WB_L represent the sensor nodes placed on the right and left edges of the whiteboard, iPad is the sensor node attached on the iPad, and the arrows show the face to face relationship across the sensor nodes. In Fig. 8 (a), we can see the learners did not face each other during the video viewing phase because the face to face relationship is scarce. Figure 8 (b) shows that User 1, User 2, and the right edge of the whiteboard faced each other and User 2 also faced the left edge of the whiteboard. Since the learner closest to the right edge of the whiteboard was User 1, User 1 used the whiteboard for discussion and User 2 saw User 1's writing. In Fig. 8 (c), all the users faced the right edge of the whiteboard and Users 1 and 2 faced each other. The result said that User 1 used the whiteboard to conclude the activity and Users 2 and 3 saw User 1's writing. Qualitative analysis requires researchers in learning science to repeat the recorded video and carefully annotate who and when learners met face-to-face. The proposed face to face detection reduces the process of watching the video and automatically extracts learners' face to face.

7.2 Results on Learning Phase Extraction

We evaluated the accuracy and validity of learning phase extraction in the collaborative learning activities. We simulated all combinations of window size and slide width for sliding windows in learning phase extraction by seconds and chose the best parameters to calculate face to face difference across the users for learning phase extraction. We calculated the accuracy of learning



(a) Video viewing (b) Discussion (c) Conclusion

Fig. 8: Extracted social graph in each learning phase.

phase extraction based on the qualitative analysis result as ground truth. Based on the design of learning phases in Sec. 5.2, we extracted the best combinations of parameters for sliding windows in learning phase extraction from all combinations which output three phases with AutoPlait.

Table 3 shows the best combinations of parameters and qualitative/quantitative phase transitions in the learning phase extraction. Cases 1 through 5 accurately extract learning phases with the accuracy of 86.9 %, 100 %, 99.8 %, 91.1 %, and 90.9 % and predict the transition between learning phases within the error of about 1 min on average. We found that the results sufficiently support researchers in learning science to reduce qualitative analysis cost of learning phase. Figure 9 shows the results of quantitative and qualitative extraction of learning phases. The top figure in Fig. 9 shows the time variation of face to face difference across the learners during the learning activity. The horizontal axis represents the elapsed time and the vertical axis represents the normalized time variation. The three figures in the middle of Fig. 9 show the results of our learning phase extraction. The three figures show the duration of video viewing, discussion, and conclusion phase. The bottom figure in Fig. 9 shows the result of the qualitative extraction of the learning phases by the researchers in learning science. The result of the quantitative extraction indicates that 1) the learners do not often turn around owing to watching the video in the video viewing phase, 2) the learners start to turn around to discuss the problem in the discussion phase, and 3) the learners often turn around to conclude the solution for the problem in the conclusion phase. The result shows the transitions between three phases appear in 1,202 s and 3,326 s. On the other hand, the result of the qualitative extraction shows that the transitions between three phases are from 1,173 s to 1,213 s and 3,335 s to 3,360 s. Although there are some deviations between the qualitative and quantitative extraction results from 51 s to 403 s and from 2,289 s to 2,459 s, the automatic extraction accurately extracts the transitions between three phases.

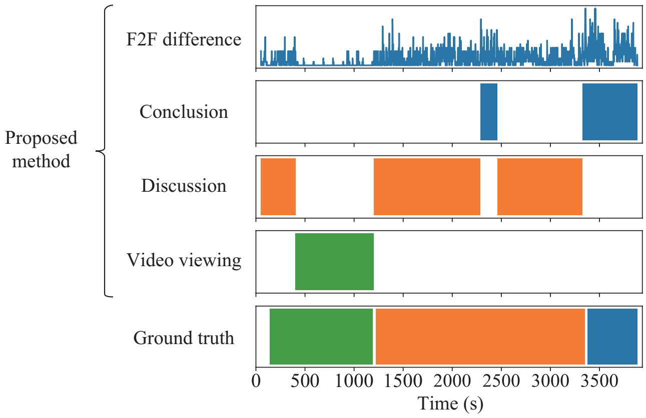


Fig. 9: Automatic extraction results of the learning phases.

7.3 Results on Speaker Identification

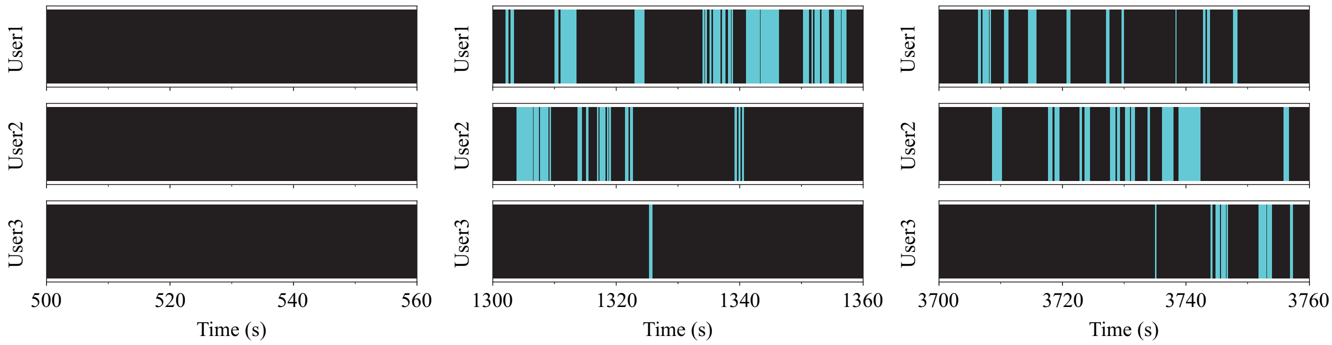
We evaluated the accuracy and validity of our proposed speaker identification. Our paper [37] shows the proposed system accurately identifies speakers under different numbers of users, environmental noises, and reverberation conditions as well as for long or short utterances. The experimental evaluations of collaborative learning show that the proposed speaker identification enables to streamline transcription of learners' utterance for collaborative learning analysis in each experimental case. Figures 10 (a), (b), and (c) show the results of speaker identification in parts of the video viewing phase, the discussion phase, and the conclusion phase in case 1 as an example. The horizontal axis represents the elapsed time and the blue bars regard as the each learners' speech. Tables 4 (a) and (b) show the results of qualitative speech transcription in the sections of Figs. 10 (b) and (c). We extracted 60 seconds of the speaker identification result in each learning phase for simplicity. Figure 10 (a) accurately detects no speech from 500 s to 560 s in the video viewing phase. The result shows that the learners did not speak to watch the video. Figures 10 (b) and (c) accurately extract each speech section from 1,300 s to 1,360 s in the discussion phase and from 3,700 s to 3,760 s in the conclusion phase. Qualitative analysis requires researchers in learning science to repeat the recorded video and observe who and when spoke as Tables 4 (a) and (b). The proposed speaker identification reduces the process of watching the video and automatically extracts speakers as Figs. 10 (a), (b), and (c).

7.4 Results on Activity Estimation

We evaluated the activity estimation of three learners in the collaborative learning activity. Figures 11 (a), (b), and (c) show the estimated results of each learner's activity. The horizontal axis

Table 3: Best combinations of window size and slide width, accuracy, and phase transitions in learning phase extraction

Case	Window size (s)	Slide width (s)	Accuracy	Transit from video viewing to discussion (s)		Transit from discussion to conclusion (s)	
				Qualitative	Quantitative	Qualitative	Quantitative
1	86	1	100 %	1,356 to 1,445	1,361	3,166 to 3,167	3,167
2	571	1	99.8 %	1,386 to 1,502	1,386	3,011 to 3,012	3,003
3	554	2	91.1 %	1,283 to 1,334	1,403	2,609 to 2,610	2,483
4	127	1	90.9 %	1,275 to 1,343	1,262	2,541 to 2,542	2,259



(a) Video viewing (b) Discussion (c) Conclusion

Fig. 10: Speaker identification results in each learning phase.

shows the elapsed time and the vertical axis indicates the relative acceleration. Tables 5 (a) and (b) show the result of qualitative records of learners' activity in the sections of Figs. 11 (b), and (c). We extracted 60 seconds of the same section as the speaker identification in each learning phase. Figure 11 (a) accurately detects small movements from 500 s to 560 s in the video viewing phase. The result shows that the learners did not move to watch the video. Figures 11 (b) and (c) accurately extract each particular movement from 1,300 s to 1,360 s in the discussion phase and from 3,700 s to 3,760 s in the conclusion phase. Qualitative analysis requires researchers in learning science to repeat the recorded video and observe who and when moved as Tables 5 (a) and (b). The proposed activity estimation reduces the process of watching the video and automatically extracts learners' particular behaviors as Figs. 11 (a), (b), and (c).

8. Related Works

Our study relates to the studies on collaborative extraction using business card-type sensors, sensor-based activity recognition, and collaborative learning analysis.

8.1 Collaborative Extraction Using Business Card-Type Sensors

Several studies have tackled to extract the collaboration between users using business card-type sensors on the users. Hitachi proposes a business card-type sensor called Business Microscope [29, 30] equipped with an infrared sensor which each user wears. The use of face to face information from the infrared sensors has shown that the appropriate frequency of meetings has an impact on work efficiency. MIT also proposes a business card-type sensor called Sociometric Badge [31] equipped with an accelerometer, a sound pressure sensor, a position sensor, Bluetooth, and an infrared sensor which each user wears. Sociometric Badge collects face to face information between the users, conversation tone changes, and proximity. The study [31] found that the face to face information between the users affects the work productivity and efficiency of the users. MIT extends Sociometric Badge to a small and low-energy sensor called Open Badges [11] with a sound pressure sensor and Bluetooth around the neck. Open Badges visualize face to face information between the users based on the sound pressure data and the Received Signal Strength Indicator (RSSI) from Bluetooth. MIT also develops hybrid software-hardware platform called Rhythm [12] with Open Badges. The platform measures face to face interaction in co-located contexts with Open Badges and in distributed contexts with their designed online applications.

However, the studies have a drawback in precise analysis of human collaboration in terms of synchronization across sensor data. The studies attempt to synchronize sensor data by means of software correction. For example, the study [30] finds similar patterns in each sound pressure and synchronizes sound pressure data sampled at 8 kHz within 100 ms. Pattern recognition further decreases the synchronization accuracy across our proposed sensors sampling sound pressure at 100 Hz for low power consumption. Such error causes inaccurate and meaningless analysis of collaborative learning activity.

Considering the abovementioned drawbacks, we develop a novel business card-type sensor based on our initial study [1] to realize accurate synchronization across the sensors and learning analysis algorithm with the acquired sensor data. The proposed sensor mounts the hardware structure to realize precise synchronization across the sensors. The proposed sensor receives and forwards synchronization packets from its synchronizer across the other sensors. The proposed sensor accurately acquires sound pressure, acceleration, and infrared data with synchronization across the sensors. We have shown that 1) the proposed sensors achieve synchronization accuracy less than the error of 1 ms for acquired sensor data sampled at 100 Hz [32], 2) the learning analysis algorithm extracts social graph, learning phases, and speakers [33–35], and 3) the algorithm improves the accuracy of

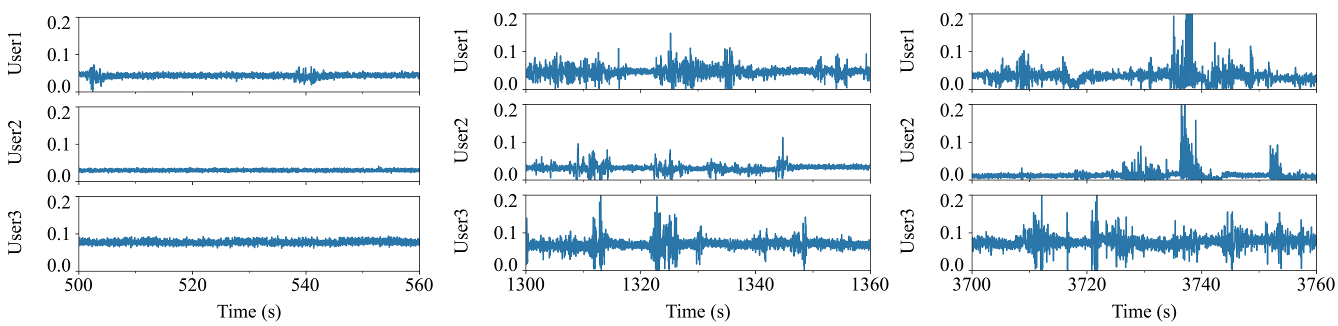
Table 4: Qualitative transcription in each phase for case 1

(a) Discussion

Number	Start (s)	End (s)	Speaker	Speech content (in Japanese)
1	1302	1303	User 1	Then two thousand feet are... Ah, I see.
2	1303	1309	User 2	One foot is one-third yard so three feet are two thousand-third yards.
3	1310	1314	User 1	Really... I learn something new.
4	1310	1311	User 2	Ha ha.
5	1310	1311	User 3	Ha ha.
6	1314	1316	User 2	I'm not confident...
7	1314	1315	User 3	Ha ha.
8	1317	1319	User 2	Six pounds...
9	1322	1323	User 2	Fifteen pounds.
10	1323	1325	User 1	Fifteen pounds.
11	1325	1326	User 3	Pound...
12	1332	1333	User 2	Ten...
13	1334	1339	User 1	I know that the normal plane is two thousand feet long, but...
14	1339	1441	User 2	They used this plane?
15	1441	1347	User 1	Didn't the video say that the fuel is half?
16	1342	1343	User 2	Yes, the video said.
17	1350	1357	User 1	At the end of the video... Well, as I said before, the part of the normal plane is two thousand feet long...
18	1352	1353	User 3	At the end?

(b) Conclusion

Number	Start (s)	End (s)	Speaker	Speech content (in Japanese)
1	3706	3711	User 1	Yes, yes, yes, fifteen plus sixty, the fuel is loaded here and fully used...
2	3708	3710	User 2	Ah, I see.
3	3710	3711	User 3	(Whispered)
4	3714	3716	User 1	About six gallons.
5	3717	3719	User 2	One gallon is six pounds, right?
6	3720	3721	User 1	Yes, yes, yes, yes.
7	3722	3724	User 2	Then eight gallons are...
8	3727	3728	User 1	Forty eight?
9	3728	3729	User 2	Forty eight pounds.
10	3729	3730	User 1	I see.
11	3730	3731	User 2	Can they load the fuel of forty eight pounds?
12	3731	3732	User 3	Forty eight pounds are bad.
13	3732	3733	User 2	Bad?
14	3733	3734	User 3	Less than forty five.
15	3736	3737	User 2	Oh my!!!
16	3737	3738	User 1	Ah...
17	3738	3742	User 2	Ha ha ha, and they also have to load the eagle.
18	3742	3743	User 1	The eagle, guy.
19	3743	3754	User 3	But they use fifteen so reduce one gallon when the eagle, the eagle arrives.
20	3755	3756	User 2	Hmm... ha ha ha.
21	3756	3757	User 3	So...



(a) Video viewing (b) Discussion (c) Conclusion

Fig. 11: Activity estimation results in each learning phase.

speaker identification under various environments [36, 37]. This paper finally develops an IoT system with business card-type sensors for collaborative learning analysis. The proposed scheme extracts social graph extraction, learning phase extraction, speaker identification, and activity estimation from the acquired sensor data. Experimental evaluations show the validity of each learning analysis algorithm in the collaborative learning activities. Our proposed sensor and algorithms contribute to support qualitative analysis of collaborative learning by researchers in learning science.

8.2 Sensor-based Activity Recognition

Some studies have been carried out to recognize a user's behavior by using multiple sensors attached to the user [3, 6, 8, 10,

Table 5: Qualitative recodes of movement in each phase for case 1

(a) Discussion

Number	Start (s)	End (s)	Learner	Movement
1	1300	1316	User 1	He wrote on the whiteboard.
2	1303	1314	User 3	She watched the iPad and whiteboard in turn.
3	1308	1314	User 2	She spoke moving the chair back and forth.
4	1323	1336	User 1	He wrote on the whiteboard.
5	1323	1326	User 2	She watched the iPad and whiteboard in turn.
6	1323	1326	User 3	She manually replayed the video on the iPad.
7	1329	1330	User 3	She turned her head toward the whiteboard from the iPad.
8	1339	1342	User 3	She manually replayed the video on the iPad.
9	1343	1344	User 2	She pulled away from the desk.
10	1346	1348	User 3	She manually replayed the video on the iPad.
11	1350	1356	User 1	He turned his head toward the whiteboard from the iPad.

(b) Conclusion

Number	Start (s)	End (s)	Learner	Movement
1	3705	3717	User 1	He wrote on the whiteboard.
2	3708	3713	User 3	She pointed out to the whiteboard.
3	3720	3722	User 3	She scratched the side of her nose.
4	3723	3730	User 3	She nodded repeatedly.
5	3726	3734	User 2	She gestured in thinking.
6	3734	3740	User 1	He swang the body with laughing.
7	3735	3739	User 3	She laughed.
8	3736	3740	User 2	She swang the body with laughing.
9	3742	3749	User 1	He wrote on the whiteboard.
10	3744	3749	User 3	She pointed out to the whiteboard.
11	3750	3758	User 3	She swang the body with putting hand on her hip.
12	3752	3754	User 2	She wondered scratching her head.

15, 16]. In the literature [6], the sensor data obtained from attaching accelerometers to the user's wrist, ankle, and chest are transmitted to the cloud. The cloud uses decision tree analysis to classify user's six activities: lying down, sitting, standing, walking, running, and riding a bike. In the literature [15], the user wears a wristwatch-type wearable device with built-in accelerometer, light sensor, thermometer, and sound sensor to classify the user's six activities: sitting, standing, walking, going up stairs, going down stairs, and running. They demonstrated that the classification could be realized in real-time with 92.5 % accuracy by using decision tree analysis. Literature [10] uses Zephyr BioHarness Bluetooth to collect acceleration and biometric information on each user and then classify three activities of running, walking, and sitting by using the decision tree analysis. In addition, they showed that the classification can cope with new users without re-learning by using the data of various users. Literature [8] used fuzzy basis functions for 3-axis accelerometer's values worn on the wrist of the user's dominant arm to classify seven activities: brushing teeth, tapping a person, tapping a desk, working on a computer, running, waving, and walking.

Our proposed Sensor-based Regulation Profiler uses sensor data obtained from the proposed Sensor-based Regulation Profiler Badges to extract and visualize the key points during the collaborative learning activity. For example, the proposed Sensor-based Regulation Profiler automatically extracts the variations of the learning phases by measuring the network variation among the learners from the infrared sensor data mounted on each Sensor-based Regulation Profiler Badge. The automatic extraction of the learning phases can reduce the qualitative analysis cost of collaborative learning as well as proper navigation in collaborative learning by the researchers in learning science.

8.3 Collaborative Learning Analysis

A general way to evaluate the effects of collaborative learning is to analyze the captured video and audio data corresponding to the activity of the collaborative learning [17, 19, 20]. On the other hand, the above-mentioned way has the following two issues.

(1) The cost of detailed learning analysis

(2) The inclusion of the researcher's subjectivity in the analysis

The first issue stems from the transcription of the learners' conversations from the captured video and audio data and the analysis of the conversations considering collaboration using nonverbal behaviors. The second issue stems from the misalignment of notation between the researchers when they analyze the learners' conversations.

To solve the aforementioned issues, our study quantitatively analyzes the collaborative learning activity by using the proposed sensor-based learning analysis. The quantitative analysis of collaborative learning realizes the detailed and quick analysis of the interaction between the learners. In addition, our Sensor-based Regulation Profiler Badge can regard the learner's nonverbal behaviors as the value of each sensor. The proposed sensor-based learning analysis enables the analysis of collaborative learning activity without the researcher's subjectivity.

9. Conclusion

We proposed Sensor-based Regulation Profiler to automatically extract and visualize the key points that researchers in learning science pay attention to during collaborative learning activity. Specifically, the proposed sensor-based learning analysis obtained social graph, learning phases, speakers, and activity from the acquired data of the proposed Sensor-based Regulation Profiler Badges. Experimental evaluations show that our proposed

Sensor-based Regulation Profiler Badge achieves the synchronization error across the sensors within $\pm 30\,\mu s$. In addition, our proposed sensor-based learning analysis extracts social graph, learning phases, speakers, and activity in collaborative learning. Each quantitative analysis reduces qualitative analysis cost of collaborative learning by researchers in learning science.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number 19H01714 and JST PRESTO Grant Number JPMJPR2032, Japan.

References

[1] Akimoto, Y., Ohtawa, S., Oshima, R., Saruwatari, S. and Oshima, J.: A Preliminary Evaluation of Collaborative Learning Analysis Tool with IoT (in Japanese), *National Convention of Information Processing Society of Japan*, pp. 1–2 (2019).

[2] Barron, B.: When Smart Groups Fail, *Journal of the Learning Sciences*, Vol. 12, No. 3, pp. 307–359 (2003).

[3] Barshan, B. and Yurtman, A.: Classifying Daily and Sports Activities Invariantly to the Positioning of Wearable Motion Sensor Units, *IEEE Internet of Things Journal*, Vol. 7, No. 6, pp. 4801–4815 (2020).

[4] Chen, G., Chan, C. K. K., Chan, K. K. H., Clarke, S. N. and Resnick, L. B.: Efficacy of Video-based Teacher Professional Development for Increasing Classroom Discourse and Student Learning, *Journal of Learning Analytics*, Vol. 29, No. 4–5, pp. 642–680 (2020).

[5] Cognition and at Vanderbilt, T. G.: The Jasper Series as an Example of Anchored Instruction: Theory, Program Description, and Assessment Data, *Educational Psychologist*, Vol. 27, No. 3, pp. 291–315 (1992).

[6] Ermes, M., Pärkkä, J. and Cluitmans, L.: Advancing from Offline to Online Activity Recognition with Wearable Sensors, *International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS)*, pp. 4451–4454 (2008).

[7] Ferrari, F., Zimmerling, M., Thiele, L. and Saukh, O.: Efficient Network Flooding and Time Synchronization with Glossy, *Proceedings of the ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, pp. 73–84 (2011).

[8] Kao, T.-P., Lin, C.-W. and Wang, J.-S.: Development of a Portable Activity Detector for Daily Activity Recognition, *IEEE International Symposium on Industrial Electronics (ISIE)*, pp. 115–120 (2009).

[9] Kirschner, F., Paas, F., Kirschner, P. A. and Janssen, J.: Differential Effects of Problem-Solving Demands on Individual and Collaborative Learning Outcomes, *Learning and Instruction*, Vol. 21, No. 4, pp. 587–599 (2011).

[10] Lara, Ó. D. and Labrador, M. A.: A Mobile Platform for Real-Time Human Activity Recognition, *IEEE Consumer Communications and Networking Conference (CCNC)*, pp. 667–671 (2012).

[11] Lederman, O., Calacci, D., MacMullen, A., Fehder, D., Murray, F. and Pentland, A.: Open Badges: A Low-Cost Toolkit for Measuring Team Communication and Dynamics, *International Conference on Social Computing, Behavioral-Cultural Modeling & Prediction and Behavior Representation in Modeling and Simulation (SBP-BRiMS)*, pp. 1–7 (2016).

[12] Lederman, O., Mohan, A., Calacci, D. and Pentland, A. S.: Rhythm: A Unified Measurement Platform for Human Organizations, *IEEE MultiMedia*, Vol. 25, No. 1, pp. 26–38 (2018).

[13] Lee, A. V. Y. and Tan, S. C.: Promising Ideas for Collective Advancement of Communal Knowledge Using Temporal Analytics and Cluster Analysis, *Journal of Learning Analytics*, Vol. 4, No. 3, pp. 76–101 (2017).

[14] Matsubara, Y., Sakurai, Y. and Faloutsos, C.: AutoPlait: Automatic Mining of Co-Evolving Time Sequences, *Proceedings of the ACM SIGMOD International Conference on Management of Data*, pp. 193–204 (2014).

[15] Maurer, U., Smailagic, A., Siewiorek, D. P. and Deisher, M.: Activity Recognition and Monitoring Using Multiple Sensors on Different Body Positions, *International Workshop on Wearable and Implantable Body Sensor Networks (BSN)*, pp. 113–116 (2006).

[16] Nandy, A., Saha, J., Chowdhury, C. and Singh, K. P.: Detailed Human Activity Recognition using Wearable Sensor and Smartphones, *International Conference on Opto-Electronics and Applied Optics (Optronix)*, pp. 1–6 (2019).

[17] Oshima, J., Ohsaki, A., Yamada, Y. and Oshima, R.: Collective Knowledge Advancement and Conceptual Understanding of Complex Scientific Concepts in the Jigsaw Instruction, *The International Conference on Computer Supported Collaborative Learning (CSCL)*, pp. 57–64 (2017).

[18] Oshima, J., Oshima, R. and Fujii, K.: Student Regulation of Collaborative Learning in Multiple Document Integration, *The Proceedings of the International Conference of the Learning Science (ICLS)*, Vol. 2, pp. 967–971 (2014).

[19] Oshima, J., Oshima, R. and Fujita, W.: A Multivocality Approach to Epistemic Agency in Collaborative Learning, *The Computer Supported Collaborative Learning (CSCL) Conference*, pp. 62–69 (2015).

[20] Oshima, J., Oshima, R. and Fujita, W.: A Mixed-Methods Approach to Analyze Shared Epistemic Agency in Jigsaw Instruction at Multiple Scales of Temporality, *Journal of Learning Analytics*, Vol. 5, No. 1, pp. 10–24 (2018).

[21] Oshima, J., Oshima, R. and Matsuzawa, Y.: Knowledge Building Discourse Explorer: A Social Network Analysis Application for Knowledge Building Discourse, *Educational Technology Research and Development*, Vol. 60, No. 5, pp. 903–921 (2012).

[22] R., J. Z., Kirschner, F., Sweller, J. and Kirschner, P. A.: Effects of Prior Knowledge on Collaborative and Individual Learning, *Learning and Instruction*, Vol. 63, p. 101214 (2019).

[23] Ravi, N., Dandekar, N., Mysore, P. and Littman, M. L.: Activity Recognition from Accelerometer Data, *Proceedings of the Conference on Innovative Applications of Artificial Intelligence (IAAI)*, pp. 1541–1546 (2005).

[24] Rummel, N., Mullins, D. and Spada, H.: Scripted Collaborative Learning with the Cognitive Tutor Algebra, *International Journal of Computer-Supported Collaborative Learning*, Vol. 7, pp. 307–339 (2012).

[25] Sawyer, R. K.: *Cambridge Handbook of the Learning Sciences, Second Edition*, Cambridge University Press (2014).

[26] Schoor, C. and Bannert, M.: Motivation in a Computer-Supported Collaborative Learning Scenario and Its Impact on Learning Activities and Knowledge Acquisition, *Learning and Instruction*, Vol. 21, No. 4, pp. 560–573 (2011).

[27] Shin, Y., Kim, D. and Jung, J.: The Effects of Representation Tool (Visible-Annotation) Types to Support Knowledge Building in Computer-Supported Collaborative Learning, *Educational Technology & Society*, Vol. 21, No. 2, pp. 98–110 (2018).

[28] Suzuki, M., Liao, C.-H., Ohara, S., Jinno, K. and Morikawa, H.: Wireless-Transparent Sensing, *Proceedings of the International Conference on Embedded Wireless Systems and Networks (EWSN)*, pp. 66–77 (2017).

[29] Tsuji, S., Sato, N., Yano, K., Broad, J. and Luthans, F.: Employees' Wearable Measure of Face-to-Face Communication Relates to Their Positive Psychological Capital, Well-Being, *IEEE/WIC/ACM International Conference on Web Intelligence - Companion Volume (WI)*, pp. 14–20 (2019).

[30] Wakisaka, Y., Ohkubo, N., Ara, K., Sato, N., Hayakawa, M., Tsuji, S., Horry, Y., Yano, K. and Moriwaki, N.: Beam-Scan Sensor Node: Reliable Sensing of Human Interactions in Organization, *International Conference on Networked Sensing Systems (INSS)*, pp. 1–4 (2009).

[31] Wu, L., Waber, B., Aral, S., Brynjolfsson, E. and Pentland, A.: Mining Face-to-Face Interaction Networks Using Sociometric Badges: Predicting Productivity in an IT Configuration Task, *International Conference on Information Systems (ICIS)*, p. 127 (2008).

[32] Yamaguchi, S., Ohtawa, S., Oshima, R., Oshima, J., Fujihashi, T., Saruwatari, S. and Watanabe, T.: Examination of Time Synchronization Accuracy for Realizing a Sensor Network to Extract Collaboration (in Japanese), *The Institute of Electronics, Information and Communication Engineers Society Conference*, pp. 1–1 (2019).

[33] Yamaguchi, S., Ohtawa, S., Oshima, R., Oshima, J., Fujihashi, T., Saruwatari, S. and Watanabe, T.: A Study on Collaborative Learning Analysis Methods using Business Card-Type Sensors (in Japanese), *Information Processing Society of Japan Multimedia, Distributed, Cooperative, and Mobile Symposium*, pp. 1194–1203 (2020).

[34] Yamaguchi, S., Ohtawa, S., Oshima, R., Oshima, J., Fujihashi, T., Saruwatari, S. and Watanabe, T.: Fundamental Evaluation of Collaborative Learning Analysis Tools for Realizing IoT to Extract Collaboration (in Japanese), *National Convention of Information Processing Society of Japan*, pp. 1–2 (2020).

[35] Yamaguchi, S., Ohtawa, S., Oshima, R., Oshima, J., Fujihashi, T., Saruwatari, S. and Watanabe, T.: Collaborative Learning Analysis Using Business Card-type Sensors, *International Conference on Quantitative Ethnography*, pp. 1–15 (2021).

[36] Yamaguchi, S., Oshima, R., Oshima, J., Fujihashi, T., Saruwatari, S. and Watanabe, T.: A Preliminary Study on Speaker Identification Using Business Card-Type Sensors, *IEEE International Conference on Consumer Electronics*, pp. 1–3 (2021).

[37] Yamaguchi, S., Oshima, R., Oshima, J., Shiina, R., Fujihashi, T., Saruwatari, S. and Watanabe, T.: Speaker Identification for Business-Card-Type Sensors, *IEEE Open Journal of the Computer Society*,

Vol. 2, pp. 216–226 (2021).
[38] Zambrano, J., Kirschner, F., Sweller, J. and Kirschner, P. A.: Effects of Group Experience and Information Distribution on Collaborative Learning, *Instructional Science*, Vol. 47, pp. 531–550 (2019).
[39] Zhou, X., Chen, L.-H. and Chen, C.-L.: Collaborative Learning by Teaching: A Pedagogy between Learner-Centered and Learner-Driven, *Sustainability*, Vol. 11, No. 4, pp. 1–14 (2019).

Shunpei Yamaguchi received his Bachelor of Engineering from Osaka University in 2020. He joined the Information Processing Society of Japan in 2018. He is currently a student at Graduate School of Information Science and Technology, Osaka University. His research interest is wearable sensors.

Shusuke Ohtawa received his Degree of Bachelor of Informatics from Shizuoka University in 2019. He is currently a student at the Graduate School of Integrated Science and Technology, Shizuoka University. His research focuses on behavioral analytics of collaborative problem solving.

Ritsuko Oshima is a professor at the Faculty of Informatics, Shizuoka University, Japan. She has been involved in a research project to develop a project-based learning curriculum at an engineering department for several years. Her current interest lies in the development of a scenario-based questionnaire to evaluate students' collaboration skills.

Jun Oshima is a professor at the Faculty of Informatics, Shizuoka University, Japan. His research interest lies in the development of new methodologies to evaluate students' collective knowledge advancement. In his recent work, he developed a social network analysis of discourse from the perspective of knowledge-building.

Takuya Fujihashi received Ph.D. degree from the Graduate School of Information Science and Technology, Osaka University, Japan, in 2016. He is currently an assistant professor at the Graduate School of Science and Engineering, Ehime University since Jan. 2017. He was research fellow (PD) of Japan Society for the Promotion of Science in 2016. From 2014 to 2016, he was research fellow (DC1) of Japan Society for the Promotion of Science. From 2014 to 2015, he was an intern at Mitsubishi Electric Research Labs. (MERL) working with the Electronics and Communications group. He selected one of the Best Paper candidates in IEEE ICME (International Conference on Multimedia and Expo) 2012. His research interests are in the area of video compression and communications, with a focus on multi-view video coding and transmission over high and low quality networks.

Shunsuke Saruwatari received the B.E. degree from The University of Electro-Communications, Japan, in 2002, and the M.S. and Ph.D. degrees from The University of Tokyo, Japan, in 2004 and 2007, respectively. In 2007, he was a Visiting Researcher with the Illinois Genetic Algorithms Laboratory, University of Illinois at Urbana–Champaign. From 2008 to 2011, he was a Research Associate with the Research Center for Advanced Science and Technology, The University of Tokyo. From 2012 to 2015, he was a Tenure-Track Assistant Professor with the Graduate School of Informatics, Shizuoka University, Japan. He is currently an Associate Professor with the Graduate School of Information Science and Technology, Osaka University, Japan. His research interests include the areas of wireless networks, sensor networks, and system software. He is a member of ACM, IPSJ, and IEICE.

Takashi Watanabe is a Professor of Graduate School of Information Science and Technology, Osaka University, Japan. He received his B.E. M.E. and Ph.D. degrees from Osaka University, Japan, in 1982, 1984 and 1987, respectively. He joined Faculty of Engineering, Tokushima University as an assistant professor in 1987 and moved to Faculty of Engineering, Shizuoka University in 1990. He was a visiting researcher at University of California, Irvine from 1995 through 1996. He has served on many program committees for networking conferences, IEEE, ACM, IPSJ, IEICE (The Institute of Electronics, Information and Communication Engineers, Japan). His research interests include mobile networking, ad hoc networks, sensor networks, ubiquitous networks, intelligent transport systems, specially MAC and routing. He is a member of IEEE, IEEE Communications Society, IEEE Computer Society as well as IPSJ and IEICE.