Estimating Collaboration through Variation in Time Series from Members in Group Works

[1] **Yuto Hattori**, [2] **Fumiko Harada**, [3] **Hiromitsu Shimakawa**

[1] *Graduate School of Information Science and Engineering Ritsumeikan University, Japan;*
[2] *Connect Dot Ltd., Japan;*
[3] *College of Information Science and Engineering Ritsumeikan University, Japan;*
yutoooo10@gmail.com; harada@cs.ritsumei.ac.jp; simakawa@cs.ritsumei.ac.jp

ABSTRACT

In successful group works, members engage in monitoring where they guess what other members are thinking. Since monitoring promotes self-regulated learning, many teachers try to introduce group works.

This paper proposes a method to inform teachers on the fly of the monitoring status of students in group works using inexpensive and less invasive sensors. In the field of education, the method enables a teacher who supervises many student groups to direct their group works to successful ones.

The method collects accelerations of multiple body parts and pulse waves from each member using sensors. The method uses the singular spectral transformation (SST) to detect the relative changes of each signal in the course of time. When a significant change takes place, it is considered the members get arousal. The method considers group members engage in monitoring when significant changes appear simultaneously among them.

In an experiment to detect successful discussion to integrate ideas from the members into one feasible solution, we have obtained the accuracy of over 0.7. It indicates that the simultaneity of significant changes detected by SST is effective to estimate the monitoring state. In the experiment result, acceleration often worked better than pulse wave signals. It was also found that each member has a different role in the group. The behavior of each member varies with the role of the member. These results show that the method has potential to estimate the monitoring state. They also imply examination of the acceleration of each body parts would enable us to estimate the role of the members. The method allows one teacher to lead many groups to successful group works.

Keywords: Group work; Collaboration; Time series data; Singular spectral transformation

1 Introduction

Many teachers are paying special attention to the realization of *self-regulated learning* because it emphasizes the autonomous and independent efforts of students. Self-regulated learning directs students to *active learning* [29]. Group works are effective to bring self-regulated learning to students [23]. Especially, Malmberg et al. [17] focused on *monitoring*, which means students in a group observe their co-workers during group works to find out what their co-workers consider. The consideration of others fosters their own learning and factors affecting it.

Traditionally, many teachers adopted teaching styles where they support individual students learning independently. Reviewing the individual teaching styles, they have started to shift their style into *collaborative learning* using group works [11]. To enhance the quality of the collaborative learning, it is essential for teachers to always grasp collaboration of students. In this study, the collaborative state is defined as one where students actively participate in discussions. To achieve it, however, we need a method specialized for collaborative learning. Unlike individual one, we should provide a method to constantly recognize how students monitor with each other in group works. Since one teacher supervises many groups at the same time, we should automate the recognition.

In this study, we investigate a method that leads group works to success in educational fields where one teacher supervises a large number of student groups. The method should use off-the-shelf, inexpensive, low-invasive sensing devices to automatically grasp the states of monitoring of each student group on the fly, which should be informed of the teacher.

The study of the self-regulated learning [39] tries to clarify the actual situation and teaching plan, regarding the autonomous and independent efforts of students themselves as the key to learning. Students engaged in a group work try to share information, search for meanings and solutions, and gain a common understanding among members of the group [23]. A study by Su et al. [30] examined groups learning a foreign language for one semester. Successful groups repeated content monitoring, organizing, and process monitoring. On the other hand, monitoring did not appear in unsuccessful groups. Each member of the successful groups compared their understanding against their assigned tasks with what they have gained during the group work. Through this process, they consider what they are learning and what influences it, without being forced [17]. Individual students monitor their own learning and the factors influencing it, while they find consensus and conflicts during the group works. In collaborative learning using group works, it is the monitoring that enables each student to develop the ability of self-regulated learning.

According to Roschelle's framework [23], collaboration is the process of building a common meaning for conversation, concept, and experience. Baker et al. [1] state that building common meanings in group works is associated with an increased cognitive burden brought by the transition from learning for group members to understand each other to learning to understand the meaning of the symbols mediating them. Increased cognitive load results in physiological arousal.

In successful collaborative learning, members within a group tend to refer to something in common. They tend to monitor each other's attention to external entities such as objects, people, and events. This is called Joint Visual Attention (JVA) [32]. Their arousal gets simultaneous because the members of the group engage in activities to focus on something at the same time. That is, if monitoring becomes active, it is expected that symptoms of arousal will be simultaneous. Existing studies [17,21,26] have shown that physiological arousal and physiological simultaneity are significative to know how each student is trying to monitor group work. Haataja et al. [7] investigated how physiological arousal temporally occurred together with events in monitoring. They are reported to have statistically significant positive association in some groups.

The collaborative learning scenario for the proposed method to be applied is building a solution, through discussion in a group, for a given subject related to software development or data analysis. This scenario assumes that group members discuss through face-to-face communication along with gestures, which is common with group works in many educational institutes. In the group work, members in the group seek for candidates of the solution using the Internet as many as possible. From them, the members integrate one feasible solution common to all members through discussion.

In the group work, members engage in monitoring to estimate the consideration of other members in order to make their ideas accepted. Interactions between the members such as idea sharing and persuasion will occur, when they are engaged in monitoring. In desired group works, interaction becomes active, which develops into discussion. Group works with active discussion are full of activities to integrate various opinions, which brings self-regulated learning to the members. This study focuses on discussions in which more than one member participate while engaging in monitoring under this scenario. Such discussions cause arousal for each member.

Unlike existing studies, we believe that arousal can be measured not only with physiological signal sensors but also with accelerometers. Suppose a discussion heats up on a specific item, which excites all the members participating it actively. Some people would use gestures to explain their ideas. which might excite others sitting in chairs with their arms crossed. Features appear as the movement changes of arms in the former, while heart rate would indicate characteristic change in the latter as symptom of physiological arousal.

The purpose of the study is to establish a method to find periods during which arousal of students working on group works simultaneously occurs, with off-the-shelf, inexpensive, low-invasive sensors, so that the method can be used in general classes.

We want to enable a small number of faculty members to bring self-regulated learning to many students in actual class as soon as possible. As the means of measuring the monitoring of each student, the study examines 2 indicators, acceleration of body part movement and pulse waves. They are superior to others at present in terms of accuracy, cost, and low-invasiveness. This study develops a means for detecting the simultaneous arousal among members using these indicators, to distinguish groups failing to monitoring from others.

Schneider et al. [26] stated in a discussion of their experiments that the physiological arousal of members presenting good performance in a group work shows more simultaneity to over time. It motivates us to apply Singular Spectral Transformation (SST) for investigation of the relative changes of arousal over time. As Palumbo et al. [21], Malmberg et al. [17], and Schneider et al. [26] reports, states of monitoring in group works can be known with the simultaneity of arousal. However, all members do not necessarily present large arousal. The increase in simultaneity over time would more clearly express the quality of group work. To catch the increase in simultaneity over time, the level of arousal should be checked relatively, not absolutely, in the course of the time.

The proposed method regards humans are aroused when body movements and pulse waves take different behavior from the one so far. The relative difference is detected with SST in the study. Only significant one is selected according to Hotelling's rule. It would enable us to grasp the occurrence of monitoring more correctly.

In this study, we focus on the axis of arousal in Russell's model [24], but do not measure the degree of comfort-discomfort of the Valence axises. However, the proposed method is expected to measure the quality of collaborative learning because the simultaneity of arousal among members suggests the level of their monitoring.

The research questions of the study are as follows.

1. How effective are the relative changes SST detects from acceleration and pulse waves varying over time in distinguishing successful group work?
2. From sensors for the acceleration and the pulse wave, we can obtain various quantities to predict the success of group works. Which of them play a vital role in the prediction?
3. What kind of actual behavior of members causes value changes of the quantities obtained during collaborative work?

Schneider et al. [26] state that groups of poor performance may exhibit a "free-rider effect" under low simultaneity. It suggests we can detect poor group works, examining the level of their simultaneity of arousal. The detection would enable teachers to give appropriate guidance which directs students of poor groups to join discussion actively. The proposed method would allow a small number of teachers to concentrate their limited teaching resources on the detected groups. It would let a few teachers to efficiently improve the quality of group works for the entire class, which bring a significant benefit both to teachers and students.

The remains of the paper are organized as follows. Section 2 introduces the related works. Section 3 explains the proposed method to estimate the collaborative states. Section 4 presents the experiment and estimation accuracy. It turns out the accuracy varies with the density of collaboration inside the group. Section 5 explains important features playing vital roles to detect collaboration. Section 6 summarizes this study.

2 Interpersonal Interaction in Group Works

2.1 Estimation of Student Monitoring States

Intelligent Tutoring Systems (ITS) in the next-generation should be equipped with mechanisms that respond to student's conditions, such as providing encouragement, modifying materials into suitable ones for each student, and giving another task when a student is distracted [3]. Winne [37] argues that students recognize that learning does not proceed as expected at the metacognitive level. However, they do not always modify the way they learn, but often keep learning in the way. In order to properly modify the method of learning, it is important to detect, on the fly, important behaviors that characterize the monitoring at the metacognitive level of individuals and groups. The detection of such behavior would enable us to provide timely support for students to take reactively regulate their own learning to adjustment themselves for the current conditions around them.

In order to facilitate group works of students, it is effective to continuously observe the monitoring states of each student in the group. However, one teacher supervises many student groups in actual educational institutes. It is difficult for the teacher to observe how students monitor relationships between themselves and other members during group works.

It is impossible to collect detailed information on arousal and interaction by interviews and questionnaires conducted after group works. It is a reasonable way to let sensors record the information on the fly. Note that students cannot concentrate on group works unless the information is collected with as little awareness of students as possible. It needs to be obtained from the students with less invasive sensors.

2.2 Existing Researches

Collaborative learning is more effective than lectures and individual learning [11]. In collaborative learning, students can obtain or reinforce their knowledge, while they teach with each other. Students who teach others can acquire a better understanding [6]. The interaction between students is regarded as an important activity in collaborative learning. Through interaction, students can acquire not only the learning effects but also communication abilities that leads to social skills [10,14]. Even during group works, it may get close to the styles of individual learning if students have a low sense of belonging to the group. The group would run into the non-collaborative state due to students who lack enthusiastic participation in group discussions. To deal with these issues, the methods have been proposed to promote interaction through group formation based on student prior knowledge [4] and student preferences [13]. Liu et al. [16] propose a method of communication support using a shared display. However, teachers evaluate only the results of group learning, neglecting the progress of collaboration in the group.

To analyze learning process, Voyiatzaki et al. [34] made students use a dialogue tool on computers. The dialogue tool makes it possible to grasp the interactions between students engaging in groupwork at a distance. The work analyzed the interaction among students represented with logs consisting of the group activity time and the number of messages obtained with the dialogue tool. On the other hand, this work cannot evaluate direct communication in the classroom.

Lien et al. [15] recognized emotions from facial expressions based on Facial Action Coding System proposed by Ekman [5]. Shi et al. [28] detected the face orientation of each student from video images in face-to-face group learning. They extracted student interaction based on the face orientation. Students strongly avoid being photographed in video. This is due to the mental repression of being monitored.

In recent years, some research has been conducted to analyze group works from multimodal information. However, there are few studies that have focused on collaborative learning. Okada et al. [20] estimated the communication skills of group members from utterance turns, prosodic information, head activities, and part-of-speech of words. Hung et al. [8] estimated the group cohesion from patterns of turn-taking of utterance and body motion. These works suggested that non-verbal information is effective to estimate the state of the group work. Most of the information is extracted from audio information. However, audio information can be acquired only in an environment with little noise. Actual classrooms involving many groups are too noisy to grasp learning states of students using microphone. Bosch et al. [3] integrates interaction patterns which are derived from learning environments on computers with features of face expression proposed by Ekman. Interactions in group work are not performed on computers, but among humans. It is difficult to log them through tools on computers. Verbal interactions are not enough to detect arousal. This is because it is more common for one person to agree silently with utterance from another than for them to agree with utterance from both parties. In case of disagreement, the situation is same.

Malmberg et al. [18] analyzed various data including videos during collaborative learning and extracted the learner's features actions in the interactions among students. This study suggested the importance of grasping body motion, including gestures, and postures to recognize interactions. It has also been shown that not only body motions but also emotional change is effective to grasp the states of collaboration in a group. Emotions also change through interactions [9,36] in discussion. The change of emotions causes physiological arousal. The change occurs because students in a group engage in monitoring which comes from the association or the interdependency among more than one member [21]. The physiological arousal happens as a result of emotion change. The simultaneity [19] is a phenomenon in which group members take nonverbal actions linked through interaction at the same time. The simultaneity of body motions has been observed in arm, head, facial expressions, and postures [12,38]. The simultaneity is also found in physiological arousal of members. The simultaneity has been observed in various communication channels. Emotion change appears as motion of body parts in one person, while electrodermal activity or heart rate variation in another. We should investigate various combination of signals from students to know the simultaneity.

We can depend physiological arousal and its simultaneity to know psychological activities during group work. It is difficult to visually determine whether or how an individual in a group is aroused for the same event, unless verbally described as aroused. The sensor seems to be effective [16].

Malmberg et al. [16] used an electrodermal activity (EDA) sensor to investigate arousal and physiological simultaneity in group works. Events in monitoring activities showed a significant correlation with arousal. Students in the group with physiological simultaneity encountered problems to be solved. On the other hand, in the group without physiological simultaneity, there were no problems that students had to solve. This indicates that physiological entrainment may be an indicator of important events in group work. In group work, detection of arousal of students in the group leads to knowing the occurrence of monitoring. Furthermore, detection of arousal occurring in multiple students at the same time enables the development of ITS that direct group works toward success. Schneider et al. [27] also used EDA sensors to examine the differences of high performance groups from low performance ones in collaborative works, which showed the cycle of the simultaneity density was highly correlated with the quality of collaboration.

Eye trackers were used as a tool to quantify the structure of JVA [32], which shows the commonality and simultaneity of group member references in collaborative learning [25,26]. In these studies, augmented cross recurrence graphs are used to visualize JVA.

2.3 Sensors to Be Used

In recent years, the development of inexpensive wearable sensors that collect human physiological signals has progressed. Various sensors have been put on the market. Students may wear them like wristwatches.

Eye-trackers and EDA sensors are indeed now available as commercial tools, but they are still expensive to use in general classes; the former and the latter costs more than 5000 USD and 2000 USD, respectively. On the other hand, accelerometers and pulse wave sensors are available for less than 50 USD.

Pulse wave analysis techniques work as a dependable detector to find physiological arousals like EDA [33]. Many medical experiments have shown the association of the heart rate with functions of the autonomous nervous system. Especially, they pay attention to interval of the R waves (RRI), where the R waves correspond to the largest peak in the heart rate wave. Excitement is successfully detected through the strength rate of low frequency parts with the high frequency parts of RRI [22,31].

On the other hand, it is effective to combine gestures and head movement with eye movement and facial expression in recognizing the emotional states of students [2]. As a means for measuring monitoring, the simultaneity of motion of student body parts is also considered to be an effective means, because excitement and arousal often appear in behavior.

These sensors make it possible to recognize the monitoring status of individual students even if they are not accompanied by teachers.

3 Estimation of Collaborative State in Group Works

3.1 Simultaneous Changes of Motion and Emotion

Malmberg et al. [16] states that monitoring is a process that makes students aware of the need to change the current status of collaborative learning. It implies it is effective for self-regulated learning. They also state that physiological simultaneity correlates with the tension and negative utterance in groups that encounter difficulties. We can say physiological simultaneity is a dependable means to measure monitoring. On the other hand, changes in behavior are also considered to be an effective means to measure monitoring. This is because, in face-to-face communication, mental excitement and arousal often appear as non-verbal body movements. Students experience arousal when they recognize the need to change their current status in collaborative learning. The study aims at detecting it from the simultaneity of changes in physiological signals and changes in body movements, to find a group that has succeeded in collaborative learning. In this study, physiological arousal of students is measured with wearable heart rate sensors, while acceleration sensors are used to sample the motion of specific body parts of students.

In this study, engagement to monitoring is detected using Singular Spectral Transformation (SST) unlike the existing studies. SST examines the time series of a specific signal to detect the timing when the variation tendency of the signal is different from the one so far. Because of the examination of relative changes in the tendency in the course of time, the proposed method is expected to be more sensitive to the change of physiological arousal and behavior of students caused by their monitoring than the existing methods that ignore temporal changes of the signal.

In group works successful in collaboration, when students discuss lively with each other, their body motions and emotions would change. On the other hand, body motions and emotions do not often change when the group has little interaction where each member works in a form close to individual learning. It is considered that there are many periods in which the timing of change of body motions and emotions among members are similar in the collaborative state, while the changes seldom happen simultaneously in the non-collaborative state. The collaborative states can be estimated from the occurrence frequency of simultaneous changes of motions and emotions.

Figure 1 shows an overview of our method. SST picks up time points where the variation is different from the one so far. From them, the Hotelling's Rule extracts ones whose variation is significantly larger than others. These points are referred to as change points. If the change points obtained from the members in a group are dense enough within a specific period, the collaborative state is regarded as high in the period.

When the group is collaborative, there are many simultaneous changes. On the other hand, when it is not collaborative, there are few. It indicates the number of simultaneous changes in periods within group works can be a feature to classify group works in terms of the intensity of collaboration. Teachers who know the intensity of collaboration use it for not only the evaluation of group works but also the encouragement of groups with poor collaboration to discuss more.

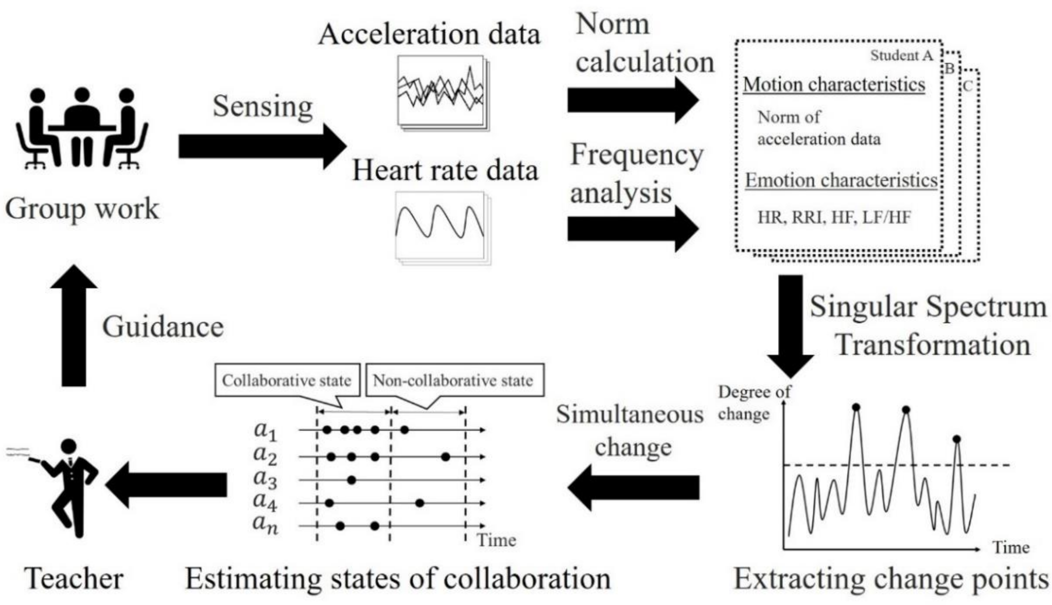

Figure 1. The proposed methods

3.2 Characteristics of Motion and Emotion

The method uses acceleration sensors and heart rate sensors to acquire time series data representing the transition of motion and emotion from students.

It measures the motion at an arm, the neck, and the waist. Students usually move their arms to explain something with gestures. In other students listening it, the motion of the neck occurs to look at the direction of the explaining students. They nod convincing ideas. In general group works for education, students are usually seated. In the seated state, left and right movements would occur to look at the direction of talking students on both sides. In order to grasp those movement, an acceleration sensor is attached to students.

3-axis acceleration sensors are used in the study. For each sensor, the acceleration of axis r at time t is defined as $\alpha_r(t)$. In the proposed method, the Euclidean norm for 3 axes defined as $\| \boldsymbol{\alpha}(t) \|_2$ are calculated by the following equation.

$$||\boldsymbol{\alpha}(t)||_2 = \sqrt{\sum_{r=1}^{3} \alpha_r(t)^2} \tag{1}$$

The time-series of the acceleration in each of the 3 body parts mentioned above is obtained to represent the features of the motion during group works. The motions of each body part are complicated. During group works, the motion may occur in any direction. The method examines the four types of time series data of the acceleration for each sensor: the acceleration of 3 axes and their Euclidean norm. It means 4 acceleration time series are obtained from each of the 3 body parts; 12 kinds of acceleration time series are used in total. Let us refer to them as the motion features.

The changes of emotion are obtained by HRV (Heart Rate Variability) which is one of the vital signs. HRV reflects variations in the balance of the sympathetic and the parasympathetic nerve that form the autonomic nervous system. Since the emotion of people affects the autonomic nervous system, HRV is considered to reflect an emotional state [31]. Exciting at a lively discussion in the group arouses emotions. It is also considered that emotions are aroused by empathy with other students.

HR (Heart Rate) and RRI (R-R Interval), which are indexes of HRV, are calculated from acquired pulse wave measurements. Furthermore, the frequency analysis of RRI figures out the HF (High Frequency) part, and the LF (Low Frequency) part of it. The former ranges from 0.15 to 0.40 Hz, while the latter from 0.04 to 0.15 Hz. The HF part is used to quantify the parasympathetic nerve fluctuation, while the LF part indicates the fluctuation of both of the sympathetic nerve and the parasympathetic nerve [22]. It means the sympathetic nerve fluctuation is examined with LF/HF. In the proposed method, 4 kinds of time series data, HR, RRI, HF, and LF/HF are referred to as emotion features.

3.3 Extracting Change Points by SST

SST is applied to extract the change points from features of motion and emotion. Let us consider 2 periods overlapped with each other. Suppose all kinds of values are sampled in a fixed sampling rate. Time series for a specific measured value in the periods represent 2 groups of features: former features and latter features. SST is a method to calculate the difference of the latter features from former features as a degree of change at time t for one-dimensional time series data. Figure 2 shows an overview of SST.

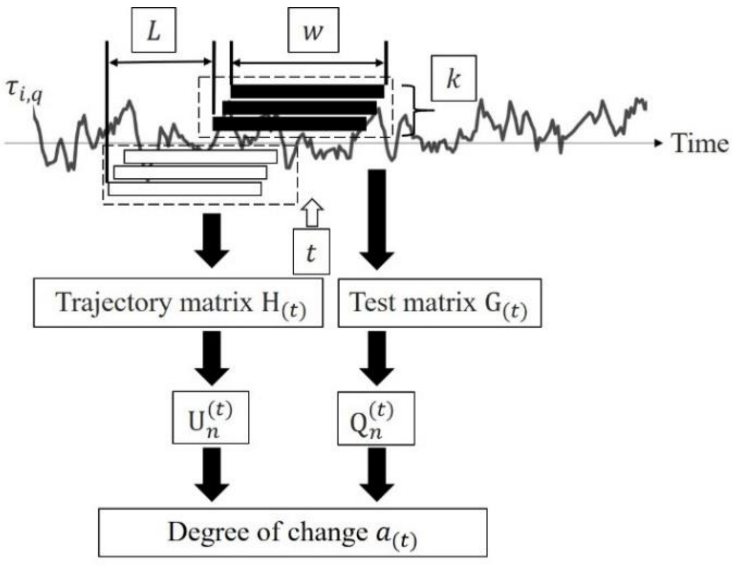

Figure 2. Calculation of degree of change

The time series data of one variable q among the 16 kinds of features of a student i is referred to as $\tau_{i,q}$. First, the trajectory matrix $\mathrm{H}_{(t)}$ and the test matrix $\mathrm{G}_{(t)}$ at time t of $\tau_{i,q}$ are generated as follows. The data from $t-w$ to $t-1$ is extracted from $\tau_{i,q}$ to make column vector $\boldsymbol{x}^{(t-w)}$ of length w, where the unit of time is the sampling rate. The column vector is called as a partial time series. The trajectory matrix $\mathrm{H}_{(t)}$, representing to the former features, is generated by the following equation using k partial time series.

$$\mathrm{H}_{(t)} \equiv [\boldsymbol{x}^{(t-k-w+1)}, \dots, \boldsymbol{x}^{(t-w-1)}, \boldsymbol{x}^{(t-w)}] \tag{2}$$

Similarly, the test matrix $\mathrm{G}_{(t)}$ representing to the latter features, is generated with k partial time series of $\mathrm{H}_{(t)}$ shifted by L in terms of the time. $\mathrm{G}_{(t)}$ is defined by the following equation.

$$\mathrm{G}_{(t)} \equiv [\boldsymbol{x}^{(t-k-w+1+L)}, \dots, \boldsymbol{x}^{(t-w-1+L)}, \boldsymbol{x}^{(t-w+L)}] \tag{3}$$

The singular value decomposition of the trajectory matrix $\mathrm{H}_{(t)}$ and the test matrix $\mathrm{G}_{(t)}$ is performed to get their left singular vectors. Singular value decomposition is a generalization of eigenvalue decomposition. Let $\boldsymbol{u}^{(t,n)}$, $\boldsymbol{q}^{(t,n)}$ be the left singular vectors corresponding to the top n singular values of $\mathrm{H}_{(t)}$ and $\mathrm{G}_{(t)}$, respectively. The matrices $\mathrm{U}_\mathrm{n}^{(\mathrm{t})}$ and $\mathrm{Q}_\mathrm{n}^{(\mathrm{t})}$ are constructed from $\boldsymbol{u}^{(t,n)}$, $\boldsymbol{q}^{(t,n)}$.

$$\mathrm{U}_n^{(t)} \equiv [\boldsymbol{u}^{(t,1)}, \boldsymbol{u}^{(t,2)}, \dots, \boldsymbol{u}^{(t,n)}] \tag{4}$$

$$\mathrm{Q}_n^{(t)} \equiv [\boldsymbol{q}^{(t,1)}, \boldsymbol{q}^{(t,2)}, \dots, \boldsymbol{q}^{(t,n)}] \tag{5}$$

The left singular vectors in $\mathrm{U}_{\boldsymbol{n}}^{(t)}$ and $\mathrm{Q}_{\boldsymbol{n}}^{(t)}$ can be regarded as feature patterns of the trajectory matrix, $\mathrm{H}_{(t)}$, and the test matrix, G(t), respectively. The larger the singular value, the more significant the feature pattern.

If the feature patterns change, the n-dimensional spaces represented by $\mathrm{U}_{\boldsymbol{n}}^{(t)}$ and $\mathrm{Q}_{\boldsymbol{n}}^{(t)}$ get separated. Based on the distance between these spaces, the intensity of the change of the time series at time t is defined as the difference of the feature pattern of test matrix, $\mathrm{G}_{(t)}$, from that of the trajectory matrix, $\mathrm{H}_{(t)}$. The change intensity $a_{(t)}$ is defined by the following equation.

$$a_{(t)} = 1 - ||\mathrm{U}_n^{(t)\top} \mathrm{Q}_n^{(t)}||_2{}^2, \tag{6}$$

where $||\mathrm{A}||_2$ is 2 norm of matrix A, which is equal to the maximum singular value. The calculated $a_{(t)}$ satisfies

$0 \le a_{(t)} \le 1$. If the features of the time-series change significantly at time t, $a_{(t)}$ gets close to 1. In contrast, if the change is small, $a_{(t)}$ is close to 0.

The degree of change $a_{(t)}$ is repeatedly calculated while moving the time t at equal intervals within the period in which the trajectory matrix and the test matrix can be generated. A threshold is set to detect the time points at which the change is large. Let the average value and the standard deviation of $a_{(t)}$ be m and s, respectively. The proposed method calculates the threshold T by the following equation.

$$T = m + 0.5s \tag{7}$$

If $a_{(t)}$ is larger than the threshold T, time t is regarded as a time point at which the motion or the emotion of the student change more than usual. It is extracted as a change point of the motions or the emotions.

3.4 Estimating Collaborative State

The change points among students are compared. A simultaneous change is regarded to occur in the period in which change points of all students are located within a short time. From section 3.3, 16 kinds of change points of motion features and emotion features are extracted per student. The change points of the 3 students are compared for $16 \times 16 \times 16 = 4096$ cases. Figure 3 shows an example of extracting simultaneous changes.

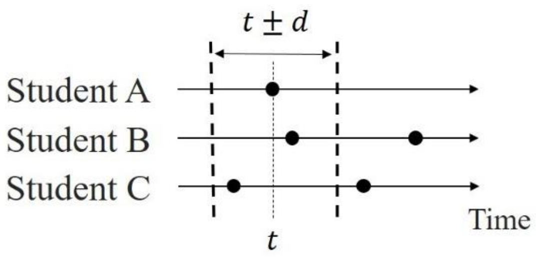

Figure 3. Example of extracting simultaneous changes

Let us consider an arbitrary change point of the student whose total number of change points is the smallest as a base point. Suppose t is one of the base points, if the change points of the other students are within $t \pm d$, the time t is extracted as a simultaneous change. The discussion seems to be active during the period in which simultaneous changes are dense. In contrast, when there are few or no simultaneous changes, it is considered that they are not collaborative. The total number of simultaneous changes in a certain period plays a role of feature quantities in classifying the collaborative state of the group.

From the above, 4096 kinds of feature quantities are obtained. However, it is not always possible to obtain simultaneous changes for all combinations of feature quantities. For this reason, the proposed method excludes feature quantities whose total number of simultaneous changes is 0 during the entire group works. The number of feature quantities differs depending on the group.

In the method, the collaborative state of a group for a certain period is discriminated with a classifier based on supervised machine learning. The classifier takes the number of simultaneous changes for each of the 4096 kinds of combinations in a specified period as the feature quantities to detect collaborative states.

The method adopts Random Forest (RF) and Gradient Boost Tree (GBT) as the supervised machine learning algorithm.

4 Interpersonal Interaction in Group Works

4.1 Experimental Overview

The experiment examines the ability of the proposed method to estimate collaborative states of group works from features in motion and emotion obtained from each member student in the group. In the experiment, the participants worked on practices in which the members of a group carry out from idea proposal to integration. In this group work, the interaction between group members occurs when they engage in monitoring. It variates the motion and emotion of the members. The proposed method detects the successful monitoring state from the simultaneity of relative changes in motion and emotion in the course of time. Collaborative states very with the length of discussion. The experiment also evaluates the effects of the length of discussion on the ability to estimate collaborative states.

The subjects are 9 male undergraduate and graduate students. In this experiment, subjects are divided into 3 groups, each of which consists of 3 subjects. In order to acquire the motion from each subject, each subject wear 3-axis acceleration sensors of TWELITE 2525A manufactured by Mono-wireless on the neck, the right arm, and the waist. The sample rate is 50Hz. In order to measure the heart rate variation (HRV), a heart rate sensor, H10 manufactured by Polar, is attached to the subject’s chest. The activities during the group works are recorded on video.

Each subject can talk freely with the other subjects in the group during group works. They describe generated ideas on paper. They can use a PC to collect the necessary information. Each group worked on group works for 1 hour, consisting of the former 30 minutes and the later 30 minutes, on the following content.

Group A: Discussion on programming methods for beginners

Group B: Discussion on data analysis in robot development

Group C: Discussion on development of smartphone applications

Note that the contents correspond to the assumption mentioned in section 1. The proposed method assumes to be applied to the group work scenario where members seek for candidate solutions, expecting to unite them into one feasible solution through discussion.

The group works was divided into the first half and the second half. The difference between them is the goal to be achieved. In the first half, each member should enumerate ideas as many as possible. In the first half, it is expected that individual working gets active, to increase the frequency of non-collaborative activities. In the second half, the group is requested to produce a united idea. In this half, discussions within the group are expected to get active so that the frequency of collaborative states should increase.

4.2 Calculating Characteristics and Change Points

The sample rate of time series acquired in the experiment varies with data items. The proposed method needs time series with an identical sample rate, in order to compare the change point of the motion features and the emotion features. The sampling rate of each time series gets unified to 2Hz. The acceleration data acquired from the acceleration sensor of each body part is averaged every 0.5 seconds. The norm in each axis and the norm of 3 axes are calculated from the averaged data. Since the heart rate sensor measures HR and RRI in unequal periods, the cubic spline interpolation is applied to the acquired time series, to resample time series of 2Hz whose sample rate is identical. HF and LH/HF are calculated through frequency analysis of resampled RRI time series in the following way. First, the power spectral density of the RRI time series is calculated by Welch's method [35]. In the calculated power spectrum density, LF is the integral from 0.04 to 0.15 Hz is LF, while HF is that from 0.15 to 0.4 Hz.

The degree of change is calculated using SST on the time series on the above features of motion and emotion. The parameters of SST in Figure 2 are $w = 48$, $k = 10$, $L = 16$, and $n = 2$. The threshold, T, to extract change points is set in the way explained in section 3.3.

4.3 Calculating Simultaneity

Change points of one member occurring around a change point of another are extracted as simultaneous changes. In this experiment, d described in section 3.4 is set to 5. Namely, suppose arbitrary change point t of a certain student is a base point. If change points of other students occur within 5 seconds before and after of t, time point t is extracted as a simultaneous change.

Figure 4 shows the average and the standard deviation of 4096 kinds of simultaneous changes in the first half and the second half for each of the 3 groups. The left box plot of each group indicates simultaneous changes in the first half, while the right box plot indicates those in the second half. The average in Group A has little difference between the first half and second half. In Group B, the average is higher in the first than in the second. Conversely, in Group C, the average is higher in the second. Note the large standard deviation in all groups.

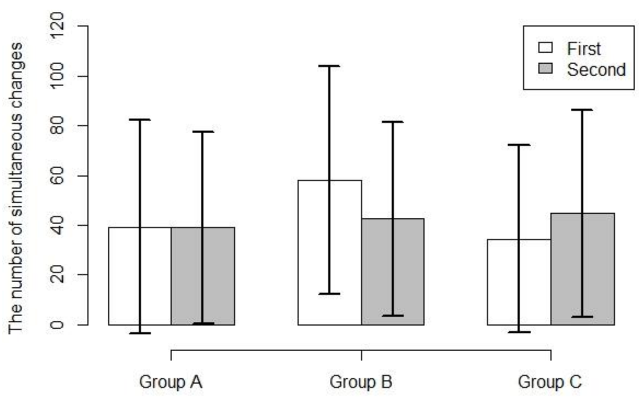

Figure 4. The number of simultaneous changes in group works

4.4 Estimation Accuracy of Collaborative State

The collaborative state of each group is classified with the occurrence frequency of simultaneous changes in a specific period. Two classifiers are generated based on supervised machine learning methods, Random Forest (RF) and Gradient Boost Tree (GBT). The collaborative state is examined for a period of 30 seconds shifted every 15 seconds in the learning time of each group. Since a time series of a certain length is necessary to apply Welch's method and SST, note that the classification is not possible for the whole duration of the learning time. The number of samples of the time series is 109 in each half of the group works, that is, 218 in total.

Let us regard 1 second as the unit of the time. Suppose any period P consisting of 30 seconds is regarded as a sequence of 30 basic periods. The size of the sequence P is $|P| = 30$. A faculty watched the whole video of the experience in advance to determine the criteria for distinguishing active discussion from others. The faculty determined the criteria from discussions in several periods where groups were productive. According to the criteria, the video is manually examined to label each basic period with regard to whether an active discussion is taken place in it. Suppose the number of basic periods labeled with discussion denoted by $s(P)$. In the experiment, the collaborative status for P is determined as follows.

Collaborative state: if $s(P) \geq \frac{|P|}{2}$

Non-collaborative state: otherwise

Table 1 shows the number of samples of non-collaborative state (NS) and collaborative state (CS) in each group.

Table 1. The number of samples

	Group A	Group B	Group C
NS	121	84	82
CS	97	134	136

The models based on RF and GBT for each group are generated from the samples shown in Table 1 For each student, we collect 16 kinds of change points of motion features and emotion features. There could be $16 \times 16 \times 16 = 4096$ combinations for simultaneous changes. Some of them occurs frequently, while others seldom happen. Table 2 gives the number of combinations that showed at least one simultaneous change in each group. The number of the combination varies with each group, as explained in section 3.4. From the density of the simultaneous changes, the generated model estimates the collaborative states.

Table 2. The number of combinations for which simultaneous changes occur

Group A	Group B	Group C
3754	4011	3709

Let us examined the performance of the models. Table 3 and Table 4 show the results of classifications with RF and GBT, respectively. For fair evaluation, 3-fold cross-validation is carried out to calculate the recall, the precision, the F1-score, the accuracy, and the AUC.

Table 3. Result of Random Forest

	Group A		Group B		Group C	
	NS	CS	NS	CS	NS	CS
Recall	0.61	0.84	0.57	0.76	0.59	0.93
Precision	0.83	0.63	0.60	0.73	0.84	0.79
F1-score	0.70	0.72	0.58	0.74	0.69	0.85
Accuracy	0.71		0.68		0.80	
AUC	0.80		0.72		0.80	

Table 4. Result of Gradient Boost Trees

	Group A		Group B		Group C	
	NS	CS	NS	CS	NS	CS
Recall	0.78	0.71	0.61	0.73	0.88	0.79
Precision	0.76	0.72	0.57	0.77	0.58	0.95
F1-score	0.77	0.71	0.59	0.75	0.70	0.86
Accuracy	0.75		0.69		0.81	
AUC	0.80		0.69		0.84	

In case of RF, the F1-score is over 0.70 in all cells except the non-collaborative state of Group B. The lowest accuracy is 0.68 of group B. For the most parts, the F1-score and the accuracy are around 0.7 or more in all groups. As for the AUC, the figures are better. Almost similar results are presented in case of GBT. In this case, the accuracy and the AUC are improved for almost all groups, because GBT generally generates a stronger learner than RF.

4.5 Effects of Length of Collaborative State

The proposed method aims to detect groups whose discussions are inactive, to encourage them more collaborative learning. It is preferable for the performance of the method to be better in the first half, where individual works are dominant. Section 4.4 regards samples as collaborative ones if group members have discussions in more than half of the periods. Let us examine the performance, changing the threshold of the length of discussions.

In the first half of the group works, where individual results are required, there have been many short discussions. By contrast, in the second half, which demands group results, long discussions have taken place. Accordingly, the criteria for the collaborative state is changed into the following one in the first half of the group works.

$$\text{Collaborative state: if } s(P) \geq \frac{|P|}{3}$$

Non-collaborative state: otherwise

Whereas, in the second half, the conditions for the collaborative state is set as follows.

$$\text{Collaborative state: if } s(P) \geq \frac{2|P|}{3}$$

Non-collaborative state: otherwise

Group C is excluded in the examination because an excessive bias toward one state is found in each half. Table 5 presents the numbers of samples in each state in the first and second half of groups A and B. Table 6 shows the number of features in each group.

Table 5. The number of samples in the first and second half

	First		Second	
	Group A	Group B	Group A	Group B
NS	73	39	55	53
CS	36	70	54	56

Table 6. The number of combinations for which simultaneous changes occur in the first and second half

First		Second	
Group A	Group B	Group A	Group B
2920	3805	3252	3489

The precision, the recall, the F1-score, the accuracy, and the AUC are calculated according to 3-fold cross-validation. Table 7 and Table 8 show the results of classifications by RF and GBT, respectively. Note that the accuracy is basically improved in the first half and decreased in the second half, compared to Table 3 and Table

4. In case of GBT, though there is only one exception in the second half of Group B, they are nearly equal to each other. As for the AUC, the tendency is almost same with RF, except the performance is improved in the second half of Group B. The results indicate the performance for the model to detect low collaboration is higher when the length of the collaboration is shorter. It is suggested the shorter threshold makes the proposed method well estimate the collaboration where the discussion does not last long.

Table 7. Result of Random Forest

	First				Second			
	Group A		Group B		Group A		Group B	
	NS	CS	NS	CS	NS	CS	NS	CS
Recall	0.84	0.63	0.82	0.75	0.70	0.65	0.57	0.78
Precision	0.82	0.67	0.65	0.88	0.67	0.68	0.74	0.63
F1-score	0.83	0.65	0.72	0.81	0.69	0.66	0.64	0.70
Accuracy	0.77		0.77		0.68		0.67	
AUC	0.78		0.85		0.76		0.77	

Table 8. Result of Gradient Boost Trees

	First				Second			
	Group A		Group B		Group A		Group B	
	NS	CS	NS	CS	NS	CS	NS	CS
Recall	0.84	0.66	0.65	0.83	0.66	0.70	0.72	0.68
Precision	0.83	0.68	0.71	0.80	0.74	0.62	0.69	0.71
F1-score	0.84	0.67	0.68	0.81	0.70	0.66	0.70	0.69
Accuracy	0.78		0.76		0.68		0.70	
AUC	0.80		0.79		0.71		0.75	

5 Discussion

5.1 Factors Affecting Accuracy

Section 4.4 indicates not only values of the F1-score but also ones of the accuracy and the AUC are generally over 0.7 in group A and group C, while group B presents a low value in the F1-score, 0.58, in the non-collaborative state in case of RF. Both the recall and the precision of group B are much worse in the noncollaborative state than in the collaborative state. As one of the reasons for the erroneous classifications in group B, it might be that many simultaneous changes are extracted in the non-collaborative state.

The method examines simultaneous changes for all characteristic combinations of motion and emotion to classify all periods in terms of collaboration. Figure 5 compares the number of simultaneous changes for each characteristic combination among the three groups. In Figure 5, the horizontal axis and the vertical axis show the total number of simultaneous changes detected in the whole experiment period and the number of combinations, respectively. Group B has significantly fewer simultaneous changes for 0 to 10 than the other two groups, while it has a peak around 100. In group B, the characteristic combinations of motion and emotion are different from the other groups.

Examining the video footages, Group B has many short discussions, which might cause the low performance in Group B. Table 9 shows the average and the standard deviation of the number of labels of discussion in periods where each group is not collaborative. It indicates Group B had many short discussions than the other groups.

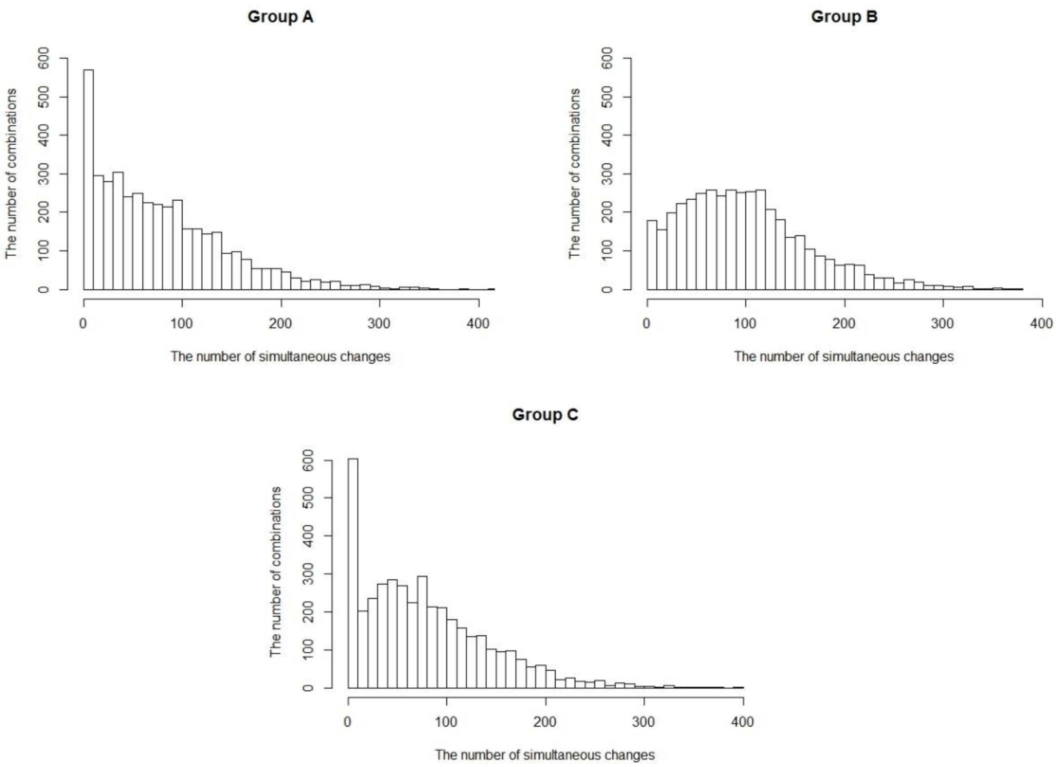

Figure 5. The number of simultaneous changes in each group

Table 9. The average and the standard deviation of the number of labels of discussion

	Group A	Group B	Group C
Mean	3.5	5.2	1.9
SD	5	4.8	3.9

In the experiment shown in Section 4.4, discussions for shorter than 15 seconds, which is half of 30 seconds, are not regarded as collaboration. Short discussions that are not regarded as collaboration could change the features of motion and emotion of the members of Group B. We can assume it leads to the misclassification in group B. From the analysis in Section 4.5, the classification performance for group B is improved for the first half of group works, where they have short discussion. It supports the assumption above.

As this example shows, the length of the discussion depends on group members and topics. It is necessary to define multiple collaborative states according to the length of discussion. Rather than binary classifications, it would be effective to grasp collaborative states in multiple stages using a multi-class classifier.

In the analysis in section 4.5, the classification performance gets worse if we try to find longer discussions. As the reason, the number of simultaneous changes in collaborative states might be reduced, which leads to the erroneous classification. The SST has possibilities to detect fewer changes in features varying in the same way for a long time. The SST calculates the degree of change through comparison of the features of two partial time series slightly shifted around a specific time point. The SST detects the time point at which the major directions of value variation are different in the 2 partial time series. Suppose the motion and the emotion of group members change due to a short discussion, and then calm down immediately. Since one partial time series may be affected while the other may not, a time point is extracted because of the large difference between the two partial time series. Conversely, a long discussion is highly likely to make group members active both in motion and emotion for a long time, which causes the two partial time series to change in the same way. The time point is less likely to be extracted as a change point. In order to extract the change points in the discussion over a long time, it is necessary to appropriately set the parameters of the SST, such as the period length and the lag from one partial time series to the other.

In section 4.2, the parameters, w and L, are set to 24 seconds and 8 seconds, respectively. They are considered suitable to detect short discussions than long discussions. If the duration of partial time series is long, both of two partial time series have many possibilities to be affected by a specific collaboration event. If the lag from one partial time series to the other is large, it is likely that a specific collaboration event will affect either of them. Change points can be extracted with high accuracy by the adjustment of the parameters of the SST. The adjustment to collaborative states appearing frequently contributes to the high performance in the classification. From the experimental results in section 4.5, note that it is desirable to set window size w to a smaller value if the goal is to find a non-collaborative state.

5.2 Consideration of Feature Quantities by Variable Importance

The number of combinations used to classify each group in section 4.4 is over 4000. Let us examine the variable importance figured out in the training process in RF to see what combination is effective to classify collaborative states. Table 10 shows 10 combinations of high importance in the classifications of each group. In the table, features with regard to body parts are represented with 'Arm', 'Waist', and 'Neck', while those from heart rate sensors with 'RRI', 'HR', and 'LF/HF'. The former corresponding to motion occupies 72% of the whole cells. The latter corresponding to emotion does not appear in less than half the cells of the former. We can say motional changes are more dominant than emotional changes. Motion seems to be more effective than emotion to classify collaborative states.

Table 10. Top 10 significant characteristic of each subjects

	Group A			Group B			Group C		
	Subject1	Subject2	Subject3	Subject4	Subject5	Subject6	Subject7	Subject8	Subject9
1	Arm	LF/HF	Waist	Arm	Neck	Waist	Neck	Waist	Neck
2	Arm	Waist	Neck	Neck	LF/HF	HR	Neck	Waist	Neck
3	Arm	Neck	Arm	Neck	Waist	Neck	Waist	Waist	Neck
4	Arm	Neck	Waist	Neck	LF/HF	RRI	LF/HF	Neck	HF
5	Arm	Waist	Neck	Neck	Neck	Arm	HF	LF/HF	RRI
6	Arm	Neck	Arm	Arm	Waist	Waist	Neck	Waist	Neck
7	Arm	HF	Arm	Neck	Waist	HR	Neck	Waist	Neck
8	Arm	Neck	Neck	Neck	LF/HF	LF/HF	RRI	Arm	LF/HF
9	RRI	Neck	LF/HF	Waist	Arm	HF	Neck	Neck	Neck
10	HR	Neck	LF/HF	HF	Waist	LF/HF	HR	Arm	LF/HF

Using SST, the proposed method finds the tendency of relative changes of a specific signal in the course of time. The consciousness of students changes due to monitoring in direct communication during group works. It arouses students. The existing studies [16,27] detect arousal with physiological signal sensors such as EDA. However, in group works where direct communication is intensive, student arousal seems to appear not only in physiological signals but also in motion signals from various body parts. At arousal time, their motion signals take variation different from one so far. If the actions taken by the students are relatively different from the previous ones, it is possible that the difference could be detected by SST. Since inexpensive acceleration sensors is equipped with many body parts of students in the experiment, the variation is more likely to be captured in motion than in emotion.

We should pay special attention to the bias in Table 10. The Subject1, Subject2, Subject4, Subject7, Subject8, and Subject9 have more than 5 motion features of the same part in the table. The bias implies a body part specific to each subject takes characteristic motion in the experiment. For each subject, let us investigate the behavior of the subjects, using the video in the experiment.

In Group A, the top 8 features in Subject1 are arms, and 6 features in Subject2 were necks. In the video, Subject1 had a larger amount of speech than other subjects in the group. He often explained by gestures, pointing PCs. Due to the behavior, it is considered that many changes in arm movement occurred. Subject2 often switched his line of sight among his PC, documents, and other members, which should bring changes in the neck movement. In addition, it seems that the movement of not only the neck but also the whole body changed because his tenth characteristic is the waist.

In Group B, 6 slots of Subject4 are occupied by necks. The video examination turns out he was often a listener with little utterance. It is considered that the neck movement appeared due to the high frequency of nods.

In Group C, 5 features in Subject7 and 7 features in Subject9 are necks, while 5 features in Subject8 are waist. According to the Video, Subject7 had the largest amount of utterance in the group. He often summarized the opinions. He did not speak along with arm movement such as gestures, but moved his head many times. Subject 9 had a little amount of utterance but responded instantly to the utterances of the other subjects. The behavior is convincing for the neck movement to be changed. Subject 8 had the second largest amount of utterance. He leaned forward many times with his upper body swaying during utterance. It is plausible that changes in movements mainly appeared in his waist.

In emotion features, LF/HF is most common over the groups. LF/HF generally reflects changes in the sympathetic nervous system. Its increase is known to state that the person is uplifted due to excitement and pleasure, or stress and tension. It is conceivable that changes in emotions appeared due to interest and concern with the opinions of others during discussions. As the collaboration went well, more conversation appeared together with laughter. The mood seemed to be elevated, which activated the sympathetic nervous system.

It is significant that members have roles such as speaker and listener in a successful group work. The analysis above brings findings that collaboration occurs when body parts characteristic to the roles change its movements, along with mood elevation. It coincides with the importance of LF/HF. The findings suggest that we should pay attention to the simultaneous occurrence of changes in various features over group members, rather than changes within individuals. In collaboration, each member is considered to have their own roles. Collaboration is regarded to occur when each member changes their features in motion and emotion according to their role. This method, which examines the simultaneous changes for all combinations of features in motion and emotion among group members, is suitable for the classification of collaborative states.

During the active discussion, it is interesting that the members play different roles, which makes their behaviors different. Given active discussion, we can identify the role of each member there. The identification of roles in successful groups has possibilities to enable a teacher to direct unsuccessful groups towards active discussion. For the teacher, it is worthy trying to make the members of groups failing collaboration take on roles similar to that of successful groups to stimulate group works.

5.3 Availability for actual classes

There are concerns that wearing the sensor may cause discomfort to students. The experiment result shows we can detect a group in which monitoring is stagnant if we have the participants of group works wear acceleration sensors and pulse wave sensors. Moreover, acceleration is more important than pulse wave. If sensors for acceleration and pulse wave are attached to the wrist like wristwatches, students will hardly feel any discomfort. The wireless acceleration sensor, TWELITE 2525A, used in the experiment weighs 6.5g. It is almost the same size as it smaller than the US quarter coin. It is currently commercially available for less than 50 USD including a transmitter and a receiver. The sensors attached to the collar of clothes would provide little inconvenience for students.

We also need physiological sensors, because some members in successful groups might get excited staying in chairs without moving their bodies. With the rise of health consciousness in the world, inexpensive pulse wave sensors have become available. They are designed so that we do not mind wearing them all the time in our daily life. We use such pulse wave sensors to examine changes in emotions.

With these sensors, teachers can detect groups that are failing to improve their activities in group work, even if they do not accompany the group all the time. Teachers can supervise the groups so that monitoring occurs in the members. Sensors deserve to be worn by group members because the effective use of less teaching resources is obtained with the sensors that are less costly and less invasive.

Some may question that collaboration can be measured by the number of discussions. In the group works which are the target of this research, group members are requested to present ideas to others. The ideas should be exchanged to be put together into a feasible solution. When the quality of monitoring is high, each member listens to the ideas of others to integrates them with their own ideas. The process increases the number of discussions.

The generality of methods for diversity such as age, gender, and ethnicity is a matter of debate. However, it is necessary to collect various samples to discuss it. It also takes a huge amount of time. The first application target of the method is actual group works in Japanese universities, where a group consists of two or three students. The age of the members is also biased to a narrow range of around 20 years old. Since Japan is a single ethnic group, it may be postponed to investigate the robustness of ethnic diversity. In response to the above, we designed the experiment, paying particular attention to the members.

6 Conclusion

In group works to integrate ideas into a feasible solution, members of successful groups would have active discussion. During the discussion, they engage in monitoring which is effective to foster self-regulated learning.

The paper proposes a method to find groups failing to have successful discussion among the members, using the singular spectrum transformation (SST). Since the method automates notifying teachers of unsuccessful groups, they can concentrate their power on directing the groups to active discussion.

In the experiment presented in the paper, the followings have turned out.

- The proposed method gives the accuracy of more than 0.7 in most of cases in an experiment. The SST seems to work well to distinguish unsuccessful discussion, because it can detect differences of a change tendency of a specific signal in the course of time.
- In the detection, the acceleration of body parts of group members works well in more examinees than the pulse wave signals. The behavior of group members is significant to measure the state of monitoring in group works where direct communication is intensive.
- In groups, the behavior of each member varies with his/her role. During active discussion, one member would present ideas with dynamic behavior, while others might show approvals with static behavior. It indicates investigation of the acceleration of body parts for each member enables us to identify the role of the member in the discussion. It is important to take each member role in group works into account for the detection. Introduction of roles are also expected to be effective to improve active discussion in unsuccessful groups.

As a future task, we will experiment with many samples with different genders to confirm the effect of the method in actual lessons. After that, we plan to investigate the effectiveness of the method for other diversity.

REFERENCES

[1]. Baker, M., Hansen, T., Joiner, R., & Traum, D. (1999). The role of grounding in collaborative learning tasks. Collaborative Learning: Cognitive and Computational Approaches, 31, 63., p.31

[2]. Behera, A., Matthew, P., Keidel, A., Vangorp, P., Fang, H., & Canning, S., (2020). Associating Facial Expressions and Upper-Body Gestures with Learning Tasks for Enhancing Intelligent Tutoring Systems, International Journal of Artificial Intelligence in Education, Volume 30, pp.236–270

[3]. Bosch, N., D ’Mello, S., Baker, R.S., Ocumpaugh, J, Shute, V., Ventura, M., Wang, L., Zhao,W. (2016). Detecting Student Emotions in Computer-Enabled Classrooms. IJCAI 2016.

[4]. Deibel, K. (2005). Team formation methods for increasing interaction during in-class group work. In ACM SIGCSE Bulletin (Vol. 37, No. 3, pp. 291-295). ACM.

[5]. Ekman, P., and Friesen, W.V. (1975). Unmasking the face. Englewood Cliffs, N.J., Prentice-Hall, 1975

[6]. Garfield, J. (1993). Teaching statistics using small-group cooperative learning. Journal of Statistics Education, 1(1).

[7]. Haataja, E., Malmberg, J., J¨arvela¨, S., (2018) Monitoring in collaborative learning: Co-occurrence of observed behavior and physiological synchrony explored, Computers in Human Behavior, Volume 87, Pages 337-347

[8]. Hung, H., & Gatica-Perez, D. (2010). Estimating cohesion in small groups using audiovisual nonverbal behavior. IEEE Transactions on Multimedia, 12(6), 563-575.

[9]. Ito,T.(2018). The Psychological Changes of Learners in the Active Learning Style Class. Japan journal of educational technology 41(Suppl.), 061-064.

[10]. Johnson, D. W., & Johnson, R. T. (1987). Learning together and alone: Cooperative, competitive, and individualistic learning. Prentice-Hall, Inc.

[11]. Johnson, R. T., & Johnson, D. W. (2008). Active learning: Cooperation in the classroom. The annual report of educational psychology in Japan, 47, 29-30.

[12]. Jokinen, K. (2009, July). Gaze and gesture activity in communication. In International Conference on Universal Access in Human-Computer Interaction (pp. 537-546). Springer, Berlin, Heidelberg.

[13]. Joseph, N., Pradeesh, N., Chatterjee, S., & Bijlani, K. (2017, September). A novel approach for group formation in collaborative learning using learner preferences. In 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI) (pp. 1564-1568). IEEE.

[14]. [Lie 2002] Lie, A. (2002). Cooperative learning: Mempraktikan cooperative learning di ruang-ruang kelas. Jakarta: Grasindo.

[15]. Lien, J.J., Kanade, T., Cohn, J. F., & Li, C. C. (1998, April). Automated facial expression recognition based on FACS action units. In Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition (pp. 390-395). IEEE

[16]. Liu, C. C., & Kao, L. C. (2005, November). Handheld devices with large shared display groupware: Tools to facilitate group communication in one-to-one collaborative learning activities. In IEEE International Workshop on Wireless and Mobile Technologies in Education (WMTE'05) (pp. 128-135). IEEE.

[17]. Malmberg, J., Haataja, E., Sepp¨anen, T., J¨arvela¨, S., (2019) Are we together or not? The temporal interplay of monitoring, physiological arousal and physiological synchrony during a collaborative exam. ijcscl 14 (4), pp. 467-490

[18]. Malmberg, J., J¨arvel¨a, S., Holappa, J., Haataja, E., Huang, X., & Siipo, A. (2019). Going beyond what is visible: What multichannel data can reveal about interaction in the context of collaborative learning? Computers in Human Behavior, 96, 235–245.

[19]. Nagaoka, C., Komori, M., & Yoshikawa, S. (2005, May). Synchrony tendency: interactional synchrony and congruence of nonverbal behavior in social interaction. In Proceedings of the 2005 International Conference on Active Media Technology, 2005.(AMT 2005). (pp. 529-534). IEEE.

[20]. Okada, S., Ohtake, Y., Nakano, Y. I., Hayashi, Y., Huang, H. H., Takase, Y., & Nitta, K. (2016, October). Estimating communication skills using dialogue acts and nonverbal features in multiple discussion datasets. In Proceedings of the 18th ACM International Conference on Multimodal Interaction (pp. 169-176). ACM.

[21]. Palumbo, R. V., Marraccini, M. E.,Weyandt, L. L., Wilder-Smith, O., McGee, H. A., Liu, S., & Goodwin, M. S. (2016). Interpersonal autonomic physiology: A systematic review of the literature. Personality and Social Psychology Review, 21(2), 99–141

[22]. Pomeranz, B., Macaulay, R. J., Caudill, M. A., Kutz, I., Adam, D., Gordon, D., Kilborn, K.M, Barger, A.C., Shannon, D.C., Cohen, R. J. & Benson, H.(1985). Assessment of autonomic function in humans by heart rate spectral analysis. American Journal of Physiology-Heart and Circulatory Physiology, 248(1), H151-H153.

[23]. Roschelle, J. (1992). Learning by collaborating: Convergent conceptual change. The Journal of the Learning, Sciences, 2(3), 235–276

[24]. Russell, J.A., (1980). A circumplex model of affect. Journal of Personality and Social Psychology, 39(6), 1161–1178

[25]. Schneider, B. & Pea, R. (2014) Toward collaboration sensing. ijcscl 9 (4), pp. 371-395

[26]. Schneider, B., Sharma, K., Cuendet, S. et al. (2018) Leveraging mobile eye-trackers to capture joint visual attention in co-located collaborative learning groups. ijcscl 13 (3), pp. 241-261

[27]. Schneider, B., Yong, D., Radu, I. (2020) Unpacking the relationship between existing and new measures of physiological synchrony and collaborative learning: a mixed methods study. ijcscl 15(1)

[28]. Shi, W., Pattichis, M. S., Celed´on-Pattichis, S., & L´opezLeiva, C. (2018, October). Dynamic Group Interactions in Collaborative Learning Videos. In 2018 52nd Asilomar Conference on Signals, Systems, and Computers (pp. 1528-1531). IEEE.

[29]. Smith, B. L., & McGregor, J. T. (1992). What is collaborative learning., Procedia - Social and Behavioral Sciences, 31, 491 – 495

[30]. Su, Y., Li, Y., Hu, H., & Ros´e, C. P. (2018). Exploring college English language learners ’ self and social regulation of learning during wiki-supported collaborative reading activities. International Journal of Computer-Supported Collaborative Learning, 1–26.

[31]. Task Force of the European Society of Cardiology. (1996). Heart rate variability, standards of measurement, physiological interpretation, and clinical use. circulation, 93, 1043-1065.

[32]. Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. J. Dunham (Eds.), Joint attention: Its origins and role in development (pp. 103–130). Hillsdale: Lawrence Erlbaum

[33]. Tsunoda, K., Chiba, A., Chigira, H., Yoshida, K., Watanabe, T., Mizuno, O., (2016), Online estimation of a cognitive performance using heart rate variability, 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp.761-765

[34]. Voyiatzaki, E., Margaritis, M., & Avouris, N. (2006). Collaborative interaction analysis: The teachers' perspective. In Sixth IEEE International Conference on Advanced Learning Technologies (ICALT'06) (pp. 345-349). IEEE.

[35]. Welch, P. (1967). The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. IEEE Transactions on audio and electroacoustics, 15(2), 70-73.

[36]. Wheatley, T., Kang, O., Parkinson, C., & Looser, C. E. (2012). From mind perception to mental connection: Synchrony as a mechanism for social understanding. Social and Personality Psychology Compass, 6(8), 589-606.

[37]. Winne, P. H. (2017). Cognition and metacognition within self-regulated learning. In Handbook of self-regulation of learning and performance (pp. 52-64). In D. Schunk, & J. Greene (Eds.). Handbook of self-regulation of learning and performance(2nd ed.). Routledge.

[38]. Xiao, B., Georgiou, P., Baucom, B., & Narayanan, S. S. (2015). Head motion modeling for human behavior analysis in dyadic interaction. IEEE transactions on multimedia, 17(7), 1107-1119.

[39]. Zimmerman, B.J., & Schunk, D.H., (2013) Self-Regulated Learning and Academic Achievement: Theoretical Perspectives, Lawrence Erlbaum Associates, 2nd edition