

Singular Spectrum Transformation on Motion/Emotion Matrix to Measure Collaboration

Yuto Hattori, Fumiko Harada, Hiromitsu Shimakawa

Abstract—In this study, we propose the method estimating collaborative state by the similarity of motion characteristics and emotion characteristics changes. The group learning based on collaboration among students are introduced in many university classes. However, it may not function as a group and teachers evaluate only result of group learning. Our method focusses on similarity of motion and emotion characteristics change. The change points are extracted from these time series data by applying singular spectrum transformation (SST). In addition, the similarity points of change extracted by comparing the change points among students. The collaborative states are estimated in a certain period from the number of similarity points of change using machine learning. The experimental result indicates this method can identify between non-collaborative state and collaborative state. The method facilitates teachers to evaluate and guide the groups.

Keywords—Group learning, Time series data

I. Introduction

In traditional school education, teachers take the initiative in transferring knowledge unilaterally to students. However, the style of class has been reviewed in recent years, and collaborative learning has been focused on, which is a class method for urging learners to an active learning [1]. The interaction between students is regarded as important in collaborative learning and it is more effective than lecture and individual learning. Therefore, groups are formed and learning methods based on collaboration among students are introduced in many university classes. However, it may not function as a group in a field of education. In addition, the teachers evaluate only result of group learning, and the state of the group cannot be evaluated. Therefore, indexes for recognizing and evaluating the collaborative state of the group is needed.

Yuto Hattori
Graduate School of Science and Engineering Ritsumeikan University
Japan

Fumiko harada
Connect Dot Ltd.
Japan

Hiromitsu Shimakawa
College of Information Science and Engineering Ritsumeikan University
Japan

In this study, the collaborative state is defined a state in which all students participate in discussions and teaching each other. On the other hand, the non-collaborative state is close to the styles of individual learning, and the state does not function as a group. The body motion and emotion are changed in collaborative state. In addition, it is considered that the timing of each change is similar. Therefore, the collaborative state of the group is estimated from the similarity among the body motion and emotional change of students using machine learning. The teacher can recognize the state of groups based on the estimation results. They can give appropriate guidance according to the group state. In this paper, section II introduces the current state of group learning and related research. In section III explains our method to estimate the collaborative state. In section IV, we show the experiment and estimation accuracy. In section V summarizes this study based on the results.

II. Interpersonal Interaction in Group Learning

A. Current Status and Effects of Group Learning

Group learning can produce higher learning effects and outcomes than mass teaching or individual learning. Because students can reinforce and obtain knowledge by teaching each other. Students who teaches other students can themselves to acquire a better understanding [2]. Moreover, they can acquire not only the learning effect but also communication ability and social skills by interaction among students [3][4]. However, it may be close to the styles of individual learning because students have low sense of belonging to the group. Therefore, the groups become non-collaborative state because of student who is not enthusiastically participate in group discussions and teaching each other.

In dealing with these issues, the methods to promote interaction through group formation based on student's prior knowledge [5] and student preferences [6] were proposed. In addition, Liu et al. [7] proposed a method of communication support using a shared display. However, teachers evaluate only the results of group, so the state of the group is not recognized. In this case, poor groups from the viewpoint of collaboration always fail in group learning. Teachers should actively support to become collaborative. Therefore, an index that evaluate the collaborative state of the group is needed.

B. Existing Research on Estimating the State of Group

Voyiatzaki et al. [8] analyzed the interaction among students using logs that includes the group activity time and

the number of messages obtained by a dialogue tool. The interactions between students in group work at a distance can be grasped by a dialogue tool. On the other hand, this study cannot evaluate direct communication in the classroom. Shi et al. [9] determined the presence or absence of interaction based on the face orientation. In this study, it is possible to recognize a group in a face-to-face situation, but it is necessary to install multiple cameras for one group. In addition, devices such as cameras might prevent members from interacting as usual.

In recent years, some studies has been conducted to analyze a group from multimodal information. However, there are few studies that have focused on group learning. Hung et al. [10] estimated the group cohesion from patterns of turn taking of utterance and body motion. This study suggested that it is effective to use nonverbal information to estimate the state of the group. However, audio information can be acquired only in an environment with little noise. Therefore, automatic extraction cannot be performed accurately in a class with many groups.

C. Similarity of Body Motion and Emotion

Kudo et al. [11] analyzed the videos during group learning and extracted the learner's features actions in the interaction. This study suggests the importance of capturing head motion, gestures and postures to recognize interactions. It has also been shown that not only body motion but also emotional change [12]. Therefore, it is effective to capture changes in motions and emotions in order to recognize the collaborative state of the group.

In addition, we focus on the synchrony tendency as a feature that appears through interaction with others. The synchrony tendency [13] is a phenomenon in which nonverbal actions are linked through interaction and the occurrence patterns are similar. The synchrony tendency has been observed in various communication channels. The synchrony tendency of motion has been observed in motion of arms, head, facial expressions, and postures [14][15]. A lot of these studies aimed at detecting the synchrony tendency. However, few studies have focused on estimating the state of group. In addition, emotions are also synchronized through the interactions [16]. If it is possible to capture not only changes in observable body motion but also changes in unobservable emotions and to estimate the state of collaboration from their similarities, teachers will be able to recognize groups that the collaboration does not progress. Furthermore, teachers can make these groups produce results through discussions and teaching each other.

III. Estimation Method of Collaborative State in Group Learning

A. Estimation of Collaborative State

In this study, the motion and emotion characteristics of students in a group learning are calculated using wearable sensors, and collaborative state is estimated from the similarity of those changes.

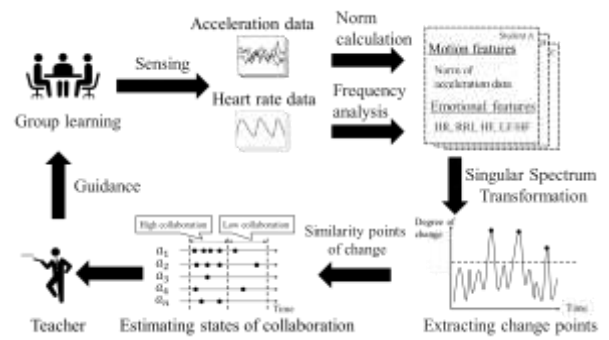


Figure 1. The proposed methods

In group learning, when students discuss or teach each other, their body motions and emotions may change. On the other hand, in a state where there is little involvement with other students and it is close to the form of the individual learning, changes of motion and emotions are unlikely to appear. Therefore, it is considered that there are many times when the timings of change of motions and emotions are similar in the collaborative state, and there are few time points of similarity in the non-collaborative state. Therefore, the state of collaboration can be estimated from the occurrence frequency of time points of similarity. Figure 1 shows an overview of our method.

When the group is collaborative, there are many similarities change points. On the other hand, when it is not collaborative, there are few similarity points of change. Therefore, the number of similarity points of change in period of group learning is used as a feature, and the collaborative state is estimated using classifier. Teachers can recognize degree of collaboration and use it for evaluation and guidance by estimating transitively over the entire time of group learning.

B. Acquiring Characteristics of Motion and Emotion

In this method, time series data as characteristics of motions and emotions is calculated from the data acquired from the acceleration sensor and the heart rate sensor.

The object parts of motion are an arm, a neck and posture changes. The motion of arm includes explanations with gesture and pointing at explanations written on paper.

The motion of the neck includes looking at the direction of the other person and explanations and nodding when other students are talking.

In addition, the posture changes same as motion of the arm and the neck. In general group work for educational, students are often seated. The left and right movements for looking at the direction of the talking partner, and the back and forth movements such as leaning forward when participating in the discussion are occur in the seated state. In order to capture this posture change, the waist is also object parts of motion.

The 3 axis acceleration sensor is used to measure the change in the intensity of the motions in the above 3 parts. The acceleration of axis r at time t is defined as $\alpha_r(t)$. The motions of each part are complicated, and the directions of motion are not a fixed. Therefore, the Euclidean norm for

each axis defined as $\|\alpha_r(t)\|_2$ and the Euclidean norm for 3 axes defined as $\|\alpha(t)\|_2$ are calculated by the following equation.

$$\|\alpha_r(t)\|_2 = \sqrt{\alpha_r(t)^2} \quad (1)$$

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

The time-series data representing the characteristics of the motions during group learning are calculated. Four types of time series data of acceleration in 3 axes and vectors generated from them are obtained from one part, so total of 12 types of time series data are obtained from 3 parts. These time series data are defined as motion characteristics.

The changes of emotion are captured by HRV (Heart Rate Variability) that is one of the vital signs. Since HRV reflects variations in the balance of the sympathetic and the parasympathetic nerve that form autonomic nervous system, it is considered to reflect emotional state [17]. The emotions are aroused such as exciting when a group has a lively discussion. In addition, it is considered that emotions are aroused by empathy with other students.

HR (Heart Rate) and RRI (R-R Interval), which is an index of HRV, are calculated from acquired pulse wave measurements. Furthermore, 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 sympathetic and parasympathetic nerve fluctuation. Thus, the sympathetic nerve fluctuation is examined with LF/HF. In this method, four time series data which is HR, RRI, HF, and LF/HF are defined as emotion characteristics.

C. Extracting Change Points by Singular Spectrum Transformation

SST (Singular Spectrum Transformation) is applied to extract the change points from characteristics of motion and emotion. SST [18] is a method of calculating the difference between past features and present features as a degree of change at time t for one dimensional time series data. Figure 2 shows an overview of SST.

The time series data of one variable q among the 16 kinds of characteristics of a student i is defined $\tau_{i,q}$. First, the trajectory matrix $H_{(t)}$ and the test matrix $G_{(t)}$ at time t of $\tau_{i,q}$ are generated by the following equation. The data from $t-w$ to $t-1$ is extracted from $\tau_{i,q}$ and the column vector $\mathbf{x}^{(t-w)}$ of length w is defined as the partial time series. The trajectory matrix $H_{(t)}$ is generated by the following equation using k partial time series.

$$H_{(t)} \equiv [\mathbf{x}^{(t-k-w+1)}, \dots, \mathbf{x}^{(t-w-1)}, \mathbf{x}^{(t-w)}] \quad (4)$$

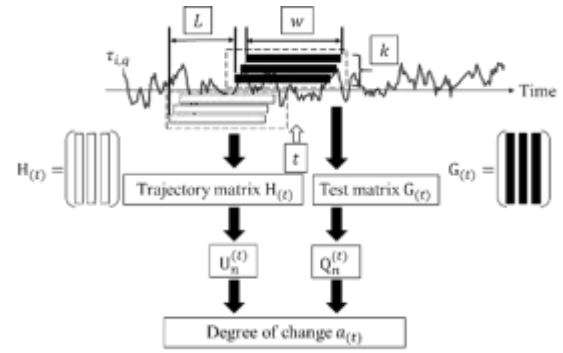


Figure 2. Calculation of degree of change

Similarly, the test matrix $G_{(t)}$ is generated by the following equation using k partial time series shifted in time L .

$$G_{(t)} \equiv [\mathbf{x}^{(t-k-w+1+L)}, \dots, \mathbf{x}^{(t-w-1+L)}, \mathbf{x}^{(t-w+L)}] \quad (5)$$

Next, the singular value decomposition of the trajectory matrix $H_{(t)}$ and the test matrix $G_{(t)}$ is performed to extract the past and present feature patterns. Singular value decomposition is a generalization of eigenvalue decomposition. The matrices $U_n^{(t)}$ and $Q_n^{(t)}$ as past and present feature patterns are constructed from the left singular vectors corresponding to the top n singular values of $H_{(t)}$ and $G_{(t)}$. The larger the singular value, the more regarded as the main feature pattern. In contrast, the smaller the singular value, the more regarded as noise. If the past and present feature patterns change, the n -dimensional spaces represented by $U_n^{(t)}$ and $Q_n^{(t)}$ are separated. Based on the distance between the spaces by $U_n^{(t)}$ and $Q_n^{(t)}$, the past and present change at time t is calculated. The degree of change $a_{(t)}$ is defined by the following equation.

$$a_{(t)} = 1 - \|U_n^{(t)T} Q_n^{(t)}\|_2^2 \quad (6)$$

$\|\cdot\|_2$ is the matrix 2 norm, which can be calculated by finding the maximum singular value. The calculated $a_{(t)}$ satisfies $0 \leq a_{(t)} \leq 1$. If the features of the time-series data change significantly before and after $a_{(t)}$ is 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 in the calculated degree of change of the time series in order to detect the time points at which the change is large. In this method, the average value of the time series changes $a_{(t)}$ is m , the standard deviation is s , and the threshold T is calculated by the following equation.

$$T = m + 0.5s \quad (7)$$

If $a_{(t)}$ is larger than the threshold T , time t is regarded as a time point when the motion or emotion more change than in the normal state and extracted as a change point.

D. *Estimating Collaborative State by Similarity Points of Change*

The change points among students are compared, and the period when all student change points exist within a short time is extracted as a similarity point of change. From 3.C, 16 kinds of change points of movement characteristics and emotion characteristics are extracted per student. Therefore, the comparison kinds of the change points of the 3 students is $16 \times 16 \times 16 = 4096$. Figure 3 shows an example of extracting similarity points of change.

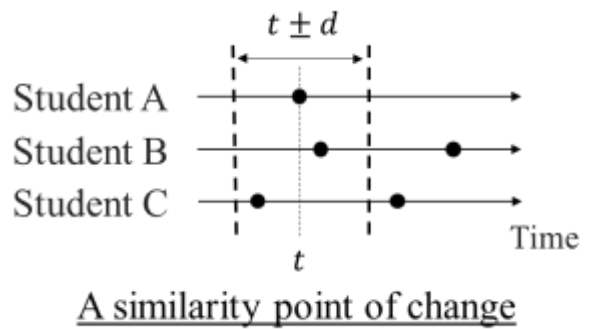


Figure 3. Example of extracting similarity points of change

This shows an arbitrary change point for each student. If the change point t of the student whose total number of change points is the smallest is the base point and the change points of the other students are within $t \pm d$, the time t is extracted as a similarity point of change. The discussion and teaching each other are actively conducted during the period when the similarity of change is dense during group learning. In contrast, when there are few or no similarity points of change, it is considered that they are not collaborative. Therefore, the total number of similarity points of change in a certain period is used as the feature quantities in identifying the collaborative state of the group.

From the above, 4096 kinds of feature quantities are obtained. However, it is not always possible to measure similarity points of change for all combinations of feature quantities. For this reason, feature quantities that total number of similarity points of change is 0 during the entire group learning are excepted. Therefore, the number of feature quantities differs depending on the group.

In this method, the collaborative state of a group for a certain period is estimated by inputting the feature quantities to a model that has been learned in advance. The model identifies two classes which is the collaborative state and the non-collaborative state.

The machine learning algorithm used in the construction of the model is Random Forest (RF). RF generates multiple decision trees from data extracted by bootstrap sampling. The collaborative state is identified by inputting features quantities to the constructed decision trees. The teacher can recognize the degree of group has been collaborating or not by continuously outputting the identification results for a certain period during the entire period of group learning.

IV. Experiment

A. *Experimental Overview*

In this experiment, we evaluated possibility to estimate the collaborative state of the group using the motion and emotion characteristics obtained from each student in the group learning.

The subjects were 9 males who were university students and graduate students. In this experiment, subjects were divided into 3 groups, each of which were constituted of 3 subjects. In order to acquire acceleration data from each subject, 3 axis acceleration sensors were attached to the subject's neck, arm, and waist. The 3-axis acceleration sensors were "TWELITE 2525A" manufactured by Mono-wireless and the sample rates were set to 50Hz.

In order to measure HRV data, the heart rate sensors were attached to the subject's chest. The heart rate sensors were "H10" manufactured by Polar was used. The states of the group during group learning were recorded on video.

Each group worked on group learning with the following content for 30 minutes in the first half and 30 minutes in the latter half for a total of 1 hour.

- Group A Discussion on the method for mounting program for beginning learners
- Group B Discussion on data analysis in robot development
- Group C Discussion on the method for Application Development

In order to equalize the state in which each group is collaborative and non-collaborative, group learning was divided into the first half and the second half. The difference between the first half and the second half was the result of group learning. In the first half of group learning, individual ideas were produced as results. It was considered that this condition increased the work time of individuals and the frequency of non-collaborative. In the second half, the ideas of the group were produced as results. In this case, discussions and teaching within the group become active, and the frequency of collaborative state was thought to increase.

B. *Calculating the Characteristics of Motion and Emotion and Extracting Change Points*

In this method, it is necessary to shape the time series data with the identical sample rate in order to compare the change point of the motion characteristics and the emotion characteristics. In this experiment, the sampling rate of each time series data was unified to 2Hz. The acceleration data acquired from the acceleration sensor of each part were averaged in 0.5 seconds. The norm of each axis and the norm of 3 axes were calculated from the averaged data. HR and RRI were measured from heart rate sensor. Since HR and RRI were unequally spaced data, cubic spline interpolation was applied and resampled to 2Hz equally spaced data. HF and LH/HF were calculated by frequency analysis of resampled RRI data. The power spectral density of the RRI data was calculated by the Welch's method. In the calculated power spectrum density, the integral from 0.04 to 0.15 Hz was LF, and the integral from 0.15 to 0.4 Hz was HF. The degree of change was calculated using SST for

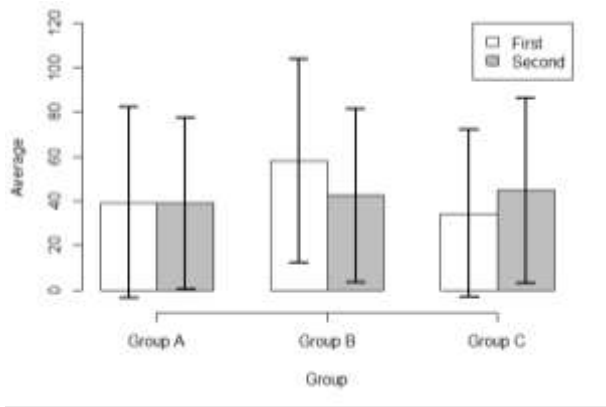


Figure 4. The number of similarity points of change

the characteristics of motion and emotion. The parameters of SST were $w = 48$, $k = 10$, $L = 16$, and $n = 2$. The threshold was set by the equation (7), and the change points were extracted.

C. Result of Calculating Feature Quantities

The similar time points between the extracted change points is calculated as similarity points of change. In this experiment, d described in 3.D was set to 5s. If an arbitrary change point t of a certain student is used as a base point and the change points of other students was within 5 seconds before and after, the time t is extracted as a similarity points of change.

The figure 4 shows the average and standard deviation of the 4096 kinds of the number of similarity points of change in the first half and the second half. Group A showed little difference between the first half and second half averages. In Group B, the average is higher in the first half than in the second half. Conversely, in Group C, the second half is higher. The large standard deviation in all groups indicates that the dispersions of total number of similarity points of change are large. Group A showed little difference between the first half and second half averages. In Group B, the average is higher in the first half than in the second half. Conversely, in Group C, the second half is higher. The large standard deviation in all groups indicates that the dispersions of total number of similarity points of change are large.

TABLE I. THE NUMBER OF SAMPLES

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

TABLE II. THE NUMBER OF FEATURE QUANTITIES

Group A	Group B	Group C
3754	4011	3709

Estimation Accuracy of Collaborative State

We evaluated the estimation accuracy of the collaborative state of each group for a certain period using RF which is supervised learning. In this experiment, we estimated from the feature quantities in 30 seconds shifted by 15 seconds for each group learning time. However, several data were needed in calculation of Welch's method and SST, so the period not calculated as a numerical value was excluded.

The number of samples of the time series was 109 in the first half and the second half of each group learning which was 218 in total.

The videos footage of the recorded experiment was observed and labeled the state of discussion and teaching or without these state for each basic period. If the number of labels of discussion or teaching in one period was 15 or more, the period was considered to be collaborative state. If the number of labels was less than 15, the period was considered to be non-collaborative state.

The table I shows the number of samples in each non-collaborative state (NS) and collaborative state (CS) in each group.

Next, table II shows the number of feature quantities in each group. For fair evaluation, 3 cross-validation was carried out.

Table III shows the results of classification by RF. The evaluation index of estimation accuracy was recall, precision, F-score, and accuracy.

The Recall were lower in non-collaborative states than collaborative state in all groups. The Precision were higher than or equal to 0.6 in the two states of all groups. The lowest F-score was 0.56 at non-collaborative state of Group B, and 0.7 or higher for the others. The lowest Accuracy was 0.69 of group B. The average of the F-score and the accuracy in all groups were around 0.7 or more.

v. Conclusion

In this paper, we proposed a method to estimate collaborative state by the similarity of motion characteristics and emotion characteristics changes. The change points are extracted from time series data of motion and emotion characteristics by applying SST. In addition, the similarity points of change extracted by comparing the change points among students. The collaborative states are estimated in a certain period from these number of similarity points of change. From result of Random Forest, the method can predict non-collaborative state and collaborative state with F-score and Accuracy of around 0.7 or more. We would like to find important feature quantities in the classification and improve accuracy.

TABLE III. THE RESULT OF RANDOM FOREST

	Group A		Group B		Group C	
	NS	CS	NS	CS	NS	CS
Recall	0.62	0.83	0.5	0.81	0.58	0.95
Precision	0.82	0.63	0.63	0.72	0.87	0.79
F-score	0.7	0.72	0.56	0.77	0.7	0.86
Accuracy	0.71		0.69		0.81	

References

- [1] Johnson, R. T., & Johnson, D. W. (2008). Active learning: Cooperation in the classroom. The annual report of educational psychology in Japan, 47, 29-30.
- [2] Garfield, J. (1993). Teaching statistics using small-group cooperative learning. *Journal of Statistics Education*, 1(1).
- [3] Lie, A. (2002). Cooperative learning: Mempraktikan cooperative learning di ruang-ruang kelas. Jakarta: Grasindo.
- [4] Johnson, D. W., & Johnson, R. T. (1987). Learning together and alone: Cooperative, competitive, and individualistic learning. Prentice-Hall, Inc.
- [5] Deibel, K. (2005, June). Team formation methods for increasing interaction during in-class group work. In *ACM SIGCSE Bulletin* (Vol. 37, No. 3, pp. 291-295). ACM.
- [6] 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.
- [7] 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.
- [8] Voyiatzaki, E., Margaritis, M., & Avouris, N. (2006, July). Collaborative interaction analysis: The teachers' perspective. In *Sixth IEEE International Conference on Advanced Learning Technologies (ICALT'06)* (pp. 345-349). IEEE.
- [9] Shi, W., Pattichis, M. S., Celedón-Pattichis, S., & LópezLeiva, C. (2018, October). Dynamic Group Interactions in Collaborative Learning Videos. In *2018 52nd Asilomar Conference on Signals, Systems, and Computers* (pp. 1528-1531). IEEE.
- [10] Hung, H., & Gatica-Perez, D. (2010). Estimating cohesion in small groups using audio-visual nonverbal behavior. *IEEE Transactions on Multimedia*, 12(6), 563-575.
- [11] Kudo, M., et al. (2016). Lerner's Characteristic Behavior Extraction in Collaborative Learning Based on Video Analysis. The 78TH National Convention of IPSJ, 2016(1), 653-654.
- [12] Ito, T. (2018). The Psychological Changes of Learners in the Active Learning Style Class. *Japan journal of educational technology* 41(Suppl.), 061-064.
- [13] 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.
- [14] 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.
- [15] 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.
- [16] 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.
- [17] Task Force of the European Society of Cardiology. (1996). Heart rate variability, standards of measurement, physiological interpretation, and clinical use. *circulation*, 93, 1043-1065.
- [18] Idé, T., & Inoue, K. (2005, April). Knowledge discovery from heterogeneous dynamic systems using change-point correlations. In *Proceedings of the 2005 SIAM International Conference on Data Mining* (pp. 571-575). Society for Industrial and Applied Mathematics.



Yuto Hattori

Received B.E from Ritsumeikan University in 2018. He advanced Graduate School of Ritsumeikan University. He advanced Graduate School of Ritsumeikan University. He engages in the research on data engineering.



Fumiko Harada

Received B.E. and M.E, and Ph.D degrees from Osaka University in 2003, 2004, and 2007, respectively. She joined Ritsumeikan University as an assistant professor in Ritsumeikan University in 2007, and is currently a lecturer. She engages in the research on real-time systems and data engineering. She is a member of IEEE.



Hiromitsu Shimakawa

Prof. Hiromitsu Shimakawa received Ph.D degree from Kyoto Univ. in 1999. He joined Ritsumeikan Univ. in 2002. Currently, he is a professor in Ritsumeikan Univ. His research interests include data engineering, usability, and integration of psychology with IT. He is a member of IEEE and ACM.