Visualization system for analyzing collaborative learning interaction

Yuki Hayashi\textsuperscript{a,}\textsuperscript{*}, Yuji Ogawa\textsuperscript{a}, Yukiko I. Nakano\textsuperscript{a}

\textsuperscript{a}Seikei University, 3-3-1 Kichijoji-kitamachi, Musashino-shi, Tokyo 180-8633, Japan

Abstract

In collaborative learning, participants interact with other members by exchanging not only verbal information but also nonverbal information, such as looking at other participants and note taking, which plays an important role in facilitating effective learning. In order to utilize such non-verbal information to analyze collaborative learning interaction, we conducted experiments to collect non-verbal information (gaze direction, speech intervals, and writing actions of participants) using multimodal measurement devices. In this paper, we propose a visualization method of collaborative learning based on a multimodal corpus of collaborative learning. The method intuitively visualizes the learning interaction with respect to collaborative and learning attitudes of each participant. We introduce the prototype system and show an example of the interface.

Keywords: Collaborative learning; Multimodal interaction; Visualization; Collaborative and learning attitudes.

1. Introduction

While understanding the process of interaction is a key to understanding collaborative learning, the interaction process is too complex for instructors or system designers to analyze \cite{1, 2}. In order to understand learning interactions effectively, many researchers have focused on analyzing and modeling the argument process using dialogue data because participants learn through conversation. Many of them use the annotated dialogue data such as utterance types \cite{2} and sentence openers \cite{3}. Although contents of arguments is important for analyzing collaborative learning processes, there is little research that focuses on analyzing the interaction process of collaborative learning in terms of non-verbal interaction. In collaborative learning, non-verbal behaviors such as looking at other participants and note taking also plays an important role in facilitating effective collaborative learning \cite{4–6}. In addition to analyzing the contents of arguments, to provide data for

\* Corresponding author. Tel.: +81-422-37-3756; fax: +81-422-37-3871.
\ E-mail address: hayashi@st.seikei.ac.jp
the non-verbal learning interactions that observers are liable to pass over will promote the efficient analysis of the collaborative learning process.

In this research, we focus on utilizing non-verbal information for analyzing the interactions of collaborative learning. So far, we constructed a learning environment to collect non-verbal data such as participants’ gaze target and writing actions [7]. By integrating this primitive information we can infer sophisticated collaborative learning interactions, such as who is gazing at the speaker and who is speaking to another participant who is also taking notes. As a first step to analyze collaborative learning, this research introduces a visualization system that represents the attitudes of each participant; whether or not a participant tackles their learning in a cooperative manner (collaborative attitude) and his/her attempts to understand the learning subject (learning attitude). To our knowledge, although studies have been made on capturing and visualizing the interaction among participants based on non-verbal information in the research field of computer-human interaction (e.g., visualization of participants’ turn-talking patterns [8] and empathy/antipathy emotions [9]), there is no literature that deals with the collaborative and learning attitudes of participants in a research field of collaborative learning. We quantify the participants’ attitudes based on multimodal data, and represent each participant in the interface as a circle object that changes its size, location, and color density to intuitively visualize the learning situations. We believe that the system can be used by the researchers/designers who analyze collaborative learning interactions to facilitate group collaboration.

2. Non-Verbal Information in Collaborative Learning

2.1. Target collaborative learning and non-verbal information

In this research, we deal with collaborative learning among participants in small groups (three participants) who study/discuss in face-to-face environments. Each participant has notepaper for writing answers/ideas freely. Through collaborative learning, they discuss and try to share their knowledge of the subject.

In collaborative learning, participants progress their learning, not only individually, by writing down the answers or notions in their notes, but also by looking around at others, listening to what someone is saying, and sometimes expressing his/her own ideas. In order to analyze these various types of interactions in learning, we conducted experiments to collect non-verbal information in collaborative learning using multimodal measurement devices [7]. The non-verbal information we extracted consists of the gaze targets, speech intervals, and writing actions of participants. This primitive and exhaustive information becomes a clue to interactions such as mutual gaze [10]. Fig. 1 shows a snapshot of a data collection experiment in our learning environment. In this situation, three participants (A, B, and C) discuss and learn together while wearing multimodal measurement devices.

The following describes the process of collecting non-verbal information.

Fig. 1. Data collection experiments in collaborative learning
• **Gaze target**

The gaze direction of each participant is obtained by eye-tracking glasses [11]. These wearable glasses have a camera that can capture the scene (640*480 pixels, 30Hz). It records what the participant is looking at as two-dimensional scene coordinates on an assistant recording device. In addition, when IR markers exist in the scene, the glasses recognize the identity numbers of the markers and their coordinates based on an IR sensor.

To acquire a participant’s gaze information, we detect the gaze target (another participant or their notes) by putting IR markers on each participant’s neck and on his/her notes. According to the aggregated eye-tracking data, the gaze targets were annotated by calculating the distance between tracked eye coordinates and IR coordinates. In this research, when the distance is lower than certain threshold, the gaze target is detected as the target marked by the nearest IR marker.

• **Speech interval**

In our learning environment, each participant wears a microphone to detect utterance duration and the identity of speakers at any particular time. The audio data stream is transmitted via an audio interface device and recorded independently. In order to detect speech interval information, we use software that segments the audio data based on the amplitude level [12]. The start and end times of an utterance gives us the number of utterances for each participant and the length of these intervals.

• **Writing action**

We introduce a digital pen device [13] to extract the writing actions of participants. The pen has ultrasonic and infrared-ray transmitters to detect movement and press/release actions according to writing motions. Participants can write down his/her notes in the usual manner because the pen uses a normal ballpoint-type refill. In order to judge whether participants were writing or not, we implemented a digital pen capturing tool which sends the pen pressed/released signal based on the pressure-sensitive information of the pen. The writing data of each participant is synchronized in the server based on timestamp information.

2.2. **Multimodal corpus of collaborative learning**

We conducted experiments to collect multimodal information during collaborative learning. For the experiments, 30 participants (20 males and 10 females) participated in collaborative learning groups. Each learning group consisted of three participants, and 10 groups were created for the experiments. Each participant was asked to wear eye-tracking glasses and a microphone (headset or tiepin type) and write using the digital pen. To record the overall learning scene, a high-definition video camera was placed above the learning environment. In the experiment, each group was asked to study with the others for two sessions. We set two types of exercises: Exercise 1, in which participants took notes frequently to derive a unique answer, and Exercise 2, in which participants mainly discussed and shared knowledge with others.

The average times taken for Exercises 1 and 2 were 839 and 817 s, respectively. The results of the extracting methods for each non-verbal information mode showed that while speech intervals were detected with high accuracy, eye-targets and writing actions were not adequately annotated because of failed eye-tracking detection and IR signal interference. After the experiment, in order to create the plenary annotated data, we annotated the correct eye targets and the writing action intervals based on the extracted eye-coordination points using a generic video annotation tool [14]. The annotated data files are formatted in xml.

The following section describes the method for visualizing a collaborative learning situation based on this multimodal corpus.
3. Visualization Method of Collaborative Learning Interaction

In collaborative learning, participants get educational benefits mainly through the interaction process [15]. Since the learning is conducted by multiple participants, it is desirable for them to discuss with others by collaborative means. In addition, they need to progress the learning itself to understand the subjects. In order to analyze such a learning situation effectively, this research proposes a method to visualize participants’ collaborative and learning attitudes based on nonverbal information. In the method, to intuitively represent the learning situation, each participant is expressed as a colored circle. The following sections discuss in detail the method of the visualization.

3.1. Visualization of collaborative attitudes

3.1.1. Circle size

Participants share knowledge with others through conversation. The more a participant takes a cooperative attitude to the others, the more he/she may catch their attention. In order to represent the participants who develop a collaborative attitude toward others, each participant’s circle size changes, increasing at a rate proportional to the amount of speech and gaze time of the participant. For an interval consisting of $s$ seconds before the present time $t$, $U_n(t, s)$ is the cumulative length of participant $n$’s speech intervals, and $Ged_r(t, s)$ is the amount of time that the other participants gazed at $n$. The proportional size increase of $n$’s circle, $p_n(t, s)$, is represented by the following expression.

$$p_n(t, s) = \frac{U_n(t, s) + Ged_r(t, s)}{s} \quad (1)$$

Here, we set $p_n = 1$ when $U_n(t, s) + Ged_r(t, s) \geq s$. Participant $n$’s circle radius $r_n(t, s)$ increases proportionally to $p_n$. We calculate $r_n(t, s)$ by

$$r_n(t, s) = MinRad + (MaxRad - MinRad) \times p_n(t, s) \quad (2)$$

where the allowed maximum radius ($MaxRad$) and minimum radius ($MinRad$) of a displayed circle are predefined.

Fig. 2 shows details of how the circle radius may change. The size of participant $n$’s circle is determined according to $r_n$. In this manner, based on the amount of time each participant talks and is looked at, we intuitively represent whether each participant’s study attitude is collaborative or not.
3.1.2. Distance of circles

In addition to representing the collaborative attitude of participants by changing circle radius, we also represent the relationship between participants. Mutual gaze often indicates the nature of the relationship between two participants. It has been reported that affinity between participants is related to gaze direction during interactions [16]. In order to represent this relationship visually, we focus on mutual gaze information and move participants’ circles closer together if mutual gazing time increases. Representing the relationship as relative distance allows us to find a participant who is not effectively communicating with others.

Assume that \( g(n_1, n_2) \) is the amount of time that participant \( n_1 \) looks at \( n_2 \) over an interval of \( s \) seconds before the present time \( t \). Then the total time \( G(n_1, n_2) \) that the two participants looked at each other is expressed by

\[
G(n_1, n_2) = g(n_1, n_2) + g(n_2, n_1)
\]

The relationship between \( n_1 \) and \( n_2 \) is close when \( G(n_1, n_2) \) becomes large. In order to visually display this relationship, the distances between the three circles is changed by rotating the angles from the weighted center of the circles; circle \( n_1 \) and \( n_2 \) approach when \( G(n_1, n_2) \) is large. We set \( f (=1/G) \) as the relative ratio between two participants, and calculate the angle between \( n_1 \) and \( n_2 \) \( \theta(n_1, n_2) \), according to the following expression.

\[
\theta(n_1, n_2) = 2\pi \times \frac{f(n_1, n_2)}{f(n_1, n_2) + f(n_2, n_3) + f(n_3, n_1)}
\]

The angles \( \theta(n_2, n_3) \) and \( \theta(n_3, n_1) \) are calculated in the same fashion.

Fig. 3(a) shows the initial placement of the participants’ circles. At this stage, the circles are located at regular intervals, that is, every 120 degrees around the circle. In order to change the circles according to the calculated angles, one circle \( (n_3) \) remains fixed on the y-axis as a reference point and the other circles \( (n_1 \text{ and } n_3) \) are rotated based on the calculated \( \theta \) for each frame. Fig. 3(b) shows the location of the circles at a certain point in time. This situation indicates that \( n_2 \) and \( n_3 \) have progressed their learning by looking at each other while \( n_1 \) is not interacting with the others as much.

![Fig. 3. Changing circle locations](image)
3.2. Visualization of learning attitude

3.2.1. Color density of circles

During collaborative learning, participants also use their notes for writing down such as answers, ideas and also dubious points about the subject through discussion with others. If such a learning attitude that the participants attempt to progress their subject can be observed, the situation where a participant is not actively engaged in the learning process can be analyzed. In order to represent such an attitude in a participant, the color of their circle becomes darker than normal in proportion to their writing time and gazing time at the speaker and speaker’s notes.

Let $W_n(t, s)$ be participant $n$’s writing time during the interval of $s$ seconds before the present time $t$, and let $n$’s gazing time at the speaker and the speaker’s notes be $GtoU_n(t, s)$ and $GtoUN_n(t, s)$, respectively. The proportional density of $n$’s circle color $b_n(t, s)$ is calculated by

$$b_n(t, s) = \frac{W_n(t, s) + GtoU_n(t, s) + GtoUN_n(t, s)}{s}$$

where $b_n(t, s) = 1$ when the numerator of this expression greater than $s$. In order to make it easier to observe the differences in circle color, the density is adjusted by the following expression.

$$d_n(t, s) = MinDens + (MaxDens - MinDens) \times b_n(t, s)$$

where $MinDens (=0.2)$ and $MaxDens (=1.0)$ are predefined constants. Fig. 4 illustrates the changing color density according to $d_n(t, s)$. We can intuitively grasp who is studying hard and who is not from observing the circle colors.

![Fig. 4. Changing color density of circles](image)

4. Prototype System

We implemented a prototype system that visualizes learning situations in collaborative learning. Fig. 5 shows a snapshot of our visualization system. The interface consists of four areas.

In the setting area, users can load a target learning video file (flv format) and its annotated data file (xml format). After clicking the start button, the video plays in the video area. Users can play, stop, and change the playing position by moving a sequence bar. The visualization area is divided into two areas. One is the area that represents the real-time learning situation (upper right). The other is the area that displays the total learning data information (lower right). In the visualization area (real-time), circles that represent participants $A$, $B$, and
C are displayed so as to correspond with the video image. Each circle in this area has a unique color and changes its size, position, and color density in response to the situation in the video. Users can adjust the threshold time $t$, which represents the interval of visualization time by setting the position of a slide bar in the area. In addition, when a participant is looking at another participant in the video, an arrow appears between the circles. When a participant makes an utterance or writes on his/her notes, the corresponding label “Utterance” or “Writing” appears on the participant’s circle. In addition, the visualization area (total) represents the ratio of participant action (as a pie chart) for each action type: gaze, speech, and writing time. The pie chart colors correspond to circle colors, and the numerical ratio appears on the graph when the user moves the mouse over it. In this way, the visualization system graphically displays the participants’ collaborative and learning attitudes, their real-time actions, and their total performance during collaborative learning.

![Fig. 5. Interface of our visualization system](image)

Fig. 6 shows an example of a changing visualization area. Fig. 6(a) shows the learning situation at a certain time $t$, and Fig. 6(b) shows the situation after 30 seconds has elapsed. During these 30 seconds, participant $A$ tried to explain the solution of an exercise to $B$ while $C$ worked on the exercise alone. In Fig. 6(b), it can be seen that $A$’s and $B$’s circles become larger than in Fig. 6(a) because $A$ and $B$ exchanged some teaching/learning utterances with each other. We understand that some interaction has occurred between $A$ and $B$ because of the relative closeness of their circles when compared to $C$. In addition, we visually grasp that $A$ and $C$ might be progressing their learning by listening to the explanation or writing in his/her notes because their circle colors are darker than in Fig. 6(a). In this area, we can also observe the current actions from the labels on the circles; $B$ is making an utterance while both $B$ and $C$ are writing on their notes in Fig. 6(b).
5. Conclusion

This research proposed a visualization system for briefly analyzing interaction sequences during collaborative learning. In the proposed interface, each participant is represented by a colored circle object. The circle’s size, position, and color density change according to the participant’s non-verbal interaction in order to represent his/her collaborative and learning attitudes. The system also displays a participant’s real-time actions such as making utterances and writing to notes, and total performance is tracked throughout the collaborative learning session.

Currently, we validate our visualization method in terms of whether the method can estimate and display the learning situation correctly according to subjective human judgment. For future work, we intend to analyze the interactions in more detail, e.g., the interactions when a participant is not actively engaged in the learning process, the differences in interactions between exercise types, also the learning styles where our system can appropriately visualize the participants’ collaborative and learning attitudes. By modeling and integrating such interaction to our visualization method, the system will be able to intelligently tell more specific learning situation for the people such as researchers who analyze collaborative learning interaction and system designers.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 25280076.

References


