Multimodal Assessment of Teachers in Immersive Learning Environment

Roghayeh Barmaki
Department of Computer Science
University of Central Florida
barmaki@eecs.ucf.edu

ABSTRACT

People send and receive messages with their body, especially with their hands in interpersonal communications. Gestures play an important role in teacher-student interactions in classrooms. In the domain of education, many research projects have focused on the study of gestures in real classrooms or in tutorial settings with practicing teachers. Novice teachers need to hone their required skills, including communication, management and content delivery skills, and prior to entering the classroom. To assist in this process, we have developed a virtual classroom environment - TeachLivE™, and used it for teacher practice, reflection and assessment. Using semi-automated multimodal annotation procedure, the data collection will be completed. The study tries to perceive the correlations between objective measures of teaching preparedness with nonverbal communication signals. Providing the assessment from the collected data will help teachers in their professional development.

1. INTRODUCTION

A good communication between students and the teacher introduces successful steps for both the learning and teaching. Communication is more than words, and it is important for teachers to understand the nonverbal messages they are sending and receiving in the classroom [6, 16]. It is important for teachers to learn to apply nonverbal communication signals in the classroom for different purposes. For instance, teachers’ gestures influence student comprehension and student learning, especially in instructional discourses [14, 20]. Gestures are a form of nonverbal communication in which visible bodily actions are used to communicate important messages, either in place of speech or together and in parallel with spoken words [10].

In this research the multimodal behaviors of biology teachers in an immersive learning environment - TeachLivE are being investigated. I am trying to analyze the behavior of teachers from their communication aspect including body language/ gesture/ facial expression coding and their correlations with teaching preparedness.

TeachLivE was developed at University of Central Florida to provide an interactive learning experience for teachers hone their management/communication and content delivery skills. It is digital puppetry software that consists of five student avatars in a virtual classroom setting [17]. Figure 1 shows more details of this avatar-mediated system.

In the TeachLivE environment, there is typically one student who is in focus and the others who are out of focus [8]. The student in focus is the one currently being addressed by the teacher. That student is inhabited by a human-in-the-loop, called an inter-actor, who controls behaviors and interactions. Students who are out of focus are controlled by agent-based software that can be influenced by the inter-actor who can choose a behavior genre. In general, that selection is influenced by the classroom management skills of the teacher. It is also possible to request the level of misbehavior and lesson plan of the classroom prior to teaching sessions in order to leverage teacher’s professional development [5].

Teacher assessment is provided in TeachLivE to reflect teacher’s performance. In order to facilitate the process of teaching assessment, TeachAARS, or TeachLivE after action review system, was designed and integrated into the TeachLivE system. TeachAARS does direct video/audio capturing that contains both the virtual classroom and the participant. Figure 2 represents TeachAARS environment.

In the preceding study [5], TeachAARS was used for manual annotation of nonverbal behaviors (body language and proximity) and also for teaching performance rating. Moving forward, in this research I intend to automate the nonverbal communication coding procedure using Microsoft Kinect V2 sensor and FACET.

I use Visual Gesture Builder (VGB) tool from Microsoft Kinect 2 SDK - a data driven machine learning solution for gesture detection - to track teachers’ body and hand movements in the virtual classroom. Visual Gesture Builder uses skeleton tracking data and AdaBoost ensemble classifier (for discrete gestures) for gesture detection. For every gesture, it is required to build the VGB by adding some recorded training clips with manual tagging to the frames indicating that specific gesture.

Gestures in the training data set for teaching evaluation are annotated with a subset of the attributes defined in the MUMIN annotation scheme [3]. The MUMIN scheme is a general framework for multimodal annotation that is useful for the study of gestures in interpersonal communication. Table 1 indicates used scheme for facial expressions and pose/gesture automatic annotation in the study. For facial expression annotation in sitted-sessions, FACET facial expression analysis engine will be used for automated face annotations. FACET reports intensity for different action units of the face muscles that combined with facial action coding system (FACS) will detect emotions of participant teachers during the interactive experiment [4].
Figure 1. TeachLivE as an avatar-mediated interactive training experience system [17]. Main stations, data flow, and roles are shown. SME indicates subject matter expert.

Figure 2. TeachAARS as a reflective tool. In the primary view, the left window shows the virtual classroom scene; the right window shows teacher participant while interacting with the classroom. An expert annotates tags associated with observed behaviors from teacher, e.g., the closed tag if the teacher exhibits a closed hand gesture [5].
Table 1. Facial displays, hand gestures and body posture: shape and dynamics [3]

<table>
<thead>
<tr>
<th>behavior attribute</th>
<th>behavior value</th>
</tr>
</thead>
<tbody>
<tr>
<td>General face</td>
<td>Smile, Laugh, Scowl, Other</td>
</tr>
<tr>
<td>Eyebrow movement</td>
<td>Frown, Raise, Other</td>
</tr>
<tr>
<td>Eye movement</td>
<td>Extra-Open, Close-Both, Close-One, Close-Repeated, Other</td>
</tr>
<tr>
<td>Gaze direction</td>
<td>Towards-Interlocutor, Up, Down, Sideways, Other</td>
</tr>
<tr>
<td>Mouth openness</td>
<td>Open mouth, Closed mouth</td>
</tr>
<tr>
<td>Head movement</td>
<td>Down, Down-Repeated, BackUp, BackUp-Forward, Back, Side-Tilt, Side-Tilt-Repeated, Side-Turn</td>
</tr>
<tr>
<td>Handedness</td>
<td>Both hands, Single hands</td>
</tr>
<tr>
<td>Hand movement trajectory</td>
<td>Up, Down, Sideways, Complex, Other</td>
</tr>
<tr>
<td>Body posture</td>
<td>Towards-Interlocutor, Up, Down, Sideways, Other</td>
</tr>
</tbody>
</table>

Some of most interested poses for the teaching study [6] is shown in Figure 3. These poses and their variations are trained in the VGB for automatic annotation, and are also compatible with MUMIN scheme.

![Figure 3. Three left-poses indicate closed posture and the three right-poses represent open posture in the interpersonal communication [6].](image)

Afterwards, the video recordings of participant teachers will be reviewed from teaching preparedness aspect by domain experts. There are some teaching frameworks [7, 13] as evaluation instruments of teachers in real classroom, but due to the specifications of TeachLivE as a virtual classroom, some domains may not be measurable/applicable. Therefore we need to extract the domains that are tractable in virtual classroom setting. There are many detailed indicators/ tags for teaching evaluation, but in TeachLivE reflection tool, they are categorized into two main buckets: academic and content-based skills and classroom management skills. Observers are requested to review the video sessions and annotate the corresponding objective measurements. Some examples of these measures are listed in table 2.

Table 2. Some measurements for teaching observation [7, 13] in TeachLivE reflection tool -TeachAARS

<table>
<thead>
<tr>
<th>content measures</th>
<th>management and communication measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demonstrating Knowledge of Content and Pedagogy</td>
<td>Establishing classroom routines</td>
</tr>
<tr>
<td>Chunking content into “digestible bites”</td>
<td>Managing response rates</td>
</tr>
<tr>
<td>Organizing students to practice and deepen knowledge</td>
<td>Demonstrating intensity and enthusiasm</td>
</tr>
<tr>
<td>Helping students practice skills, strategies, and processes”</td>
<td>Using verbal and nonverbal behaviors that indicate affection for students</td>
</tr>
<tr>
<td>Helping students revise knowledge</td>
<td>Organizing students to interact with new content</td>
</tr>
<tr>
<td>Using Questioning and Discussion Techniques</td>
<td>Noticing when students are not engaged</td>
</tr>
</tbody>
</table>

In my research, I try to understand if there are any correlations between objective measures of teaching and gestural signs that teachers use in their virtual classroom setting.

2. EXISTING SOLUTIONS

Some studies have shown that speakers’ gestures facilitate listeners’ comprehension of speech [2, 20]. In a survey by Roth in 2001, the role of gestures have been presented in teaching and learning, especially the role of gestures in knowing and learning scientific and mathematical concepts in school-aged children have been addressed [19].

Gesture movements are classified into four categories according to McNeill’s [15] study: iconic (the form is related to the semantic content of speech); metaphorics (linking to an abstract concept); deictics (pointing movements); and beats (do not present a describable meaning). In [12], Macedonia et al. explored the impact of iconic gestures in foreign language word learning. Their research indicated that iconic gestures on comparison to meaningless gestures helped the memorization of foreign language nouns in a significant fashion. In a similar work with Alibali et al. [1], gestures have been introduced as a means for teachers to scaffold students’ understanding. In this research, there was a scaffolding hypothesis based on the work of Lakoff [11] which states teachers use gestures to “ground” their instructional language, especially in abstract concepts. The analysis on selected video session of a mathematics lesson indicated that teacher’s gesture was used most frequently for new materials, for referents that were highly abstract, and in response to students’ questions and comments [1]. Finally, Pozzer-Ardenghi et al. in [18] explored the videos of science lectures (subject: human body parts in biology high school classes). Their research indicates multimodal resources and nonverbal aspects of teaching may help students to be able to better articulate their conceptions and understandings with peers.
3. RESEARCH CONTRIBUTIONS

There has been a flurry of research in education on gesture within the last decade. Most scholars studied the effect of use of gestures by teachers in real classroom and its impact in student learning. In virtual learning environments, the setting is different. One major difference in our virtual classroom setting with real classroom is that in TeachLivE environment, there are not real students. This is an opportunity for novice teachers to improve their skills without putting real children at risk [9]. For example, teachers can easily practice their targeted lesson plan several times until they master it.

In contrast, based on specific features of this virtual classroom, it is not possible to track and study the student learning similar to previous studies. Most research involving teacher’s gestures and students’ learning have been done either in real classroom or tutorial setting [1]. In real setting, students and their pre/posttests in the classroom have been used as an indicator for their teacher’s skills and their teaching preparedness, which is not useful in our virtual setting. One solution to resolve this concern is to use domain experts to review the teaching behavior of the teachers using the video. Combining the objective measures with automated nonverbal annotations and finding the correlations amongst them will be the main contribution of this research. Altogether, the collected data will help teachers by providing a comprehensive feedback and assessment form their teaching behaviors and will be useful in their professional development. As mentioned before, the classroom environment is not real, and student learning factor is absent in the study, which make the research domain distinguished from preceding researches.

4. METHODOLOGY

Fifteen Biology practicing teachers from Florida public schools are participating in the study. They were asked to teach virtual kids a lesson about the definition of technology and its applications in Biology in 10-minute sessions per month throughout one academic year. Some questions such as: “How do you define Technology?” or “In general, what does technology do for us?” were listed in the lesson plan for teachers. The lesson is designed to enhance science literacy. It also has been validated and field-tested in high school Biology classrooms as a part of a larger module from the NIH Curriculum Supplement Series “Using Technology to Study Cellular and Molecular Biology”. TeachAARS recorded the teaching session while interacting to the virtual classroom. A sample scene from TeachAARS was shown in figure 2. In TeachLivE, the simulation has been designed with Kinect SDK to enable the teacher to move close to the student avatar represented on the display screen. While moving, the visual perspective moves with the teacher, even allowing eye-to-eye communication [5]. One additional Kinect V2 sensor and a separate PC with FACET software installed are needed for recording the automated gesture detection and facial expression data set. One technical issue in previous version of Kinect sensors was the infrared signal interference. It means when multiple depth V1 sensors have overlapping field-of-view, the obtained depth data is very noisy and useless. I tested my system to see if this problem exists with multiple sensors, with Kinect V1 and V2. Fortunately, both depth sensors work properly since their infrared signals do not have interference with each other. Hence the proposed system can be used for gesture data collection.

At the next step, the video records of teachers from TeachAARS should be analyzed by domain experts from teaching preparedness point of view according to table 2. At this point, experts annotate existing objective behaviors that indicate the teacher’s proficiency and preparedness.

Finally all the collected data needs to combine for providing feedback for participant teachers. First work in data analysis will be finding the correlation matrix for the subjects and defined attributes. Afterwards, the dependent metrics could be predicted either by linear/logistic regression or other learning techniques. It may not be possible to build a predictive model for feedback at this point, but the collected data and the analysis with help to have a better understanding of how teachers use multimodal behaviors in virtual classroom.

5. CURRENT ACHIEVEMENTS

At this point the virtual classroom setting is tested for data collection and we are waiting for office of human research approval to run the study. I overviewed preceding videos from 15 participant teachers and extracted a library of their gestures in the virtual classroom environment. For example, folding the arms in front and back, hands crossed power-pose, etc. In particular I am mostly interested in closed and open gestures/ poses of teachers in the virtual classroom. This research is moving forward as following: improving the trained database for extracted gestures, with adding some negative samples in order to get a robust trained engine for gesture detection using VGB; adding the virtual environment’s log data and student to teacher talk time into current data collection in order to find more correlations.

The other aspect of this study that focus on objective measures of teaching in TeachLivE environment moved forward in collecting desired measures that are applicable for TeachLivE virtual setting. Current teaching frameworks are not appropriate for virtual setting and those measures are determined for real classrooms. Hence we explored different frameworks [7, 13] and extracted useful measures.

6. REFERENCES


