A Framework For Automated Analysis of Videos of Informal Classroom Educational Settings

Naveen Kumar*, Vikram Ramanarayanan[*] **& Shrikanth S. Narayanan**
Ming Hsieh Department of Electrical Engineering
University of Southern California
Los Angeles, CA 90089
`<komathnk,vramanar>@usc.edu, shri@sipi.usc.edu`

Abstract

We present a framework to record and analyze audiovisual data of participant engagement behavior observed during instructional classroom sessions. We used a single-camcorder setup to capture audio-visual data in a setting offering informal, after-school science learning experiences for young children and their families. The recording setup has the advantage that it is simple, portable and easily available (and hence easily replicable) allowing ecologically-valid observations. Using this setup, we have recorded approximately 10 sessions of audiovisual classroom data. The classroom sessions consisted of both an instructive (where teachers give a lecture) as well as an interactive component (which involved discussion among the students and teachers). In this study we primarily consider the instruction component since it is far less unstructured as compared to the interaction sessions.

We split the task of video analysis into two phases – an initial data annotation phase followed by an automated analysis phase. For the first phase, we asked experts to annotate the videos for different examples of engagement (or any kind of constructive behavior that educators would like to encourage in classroom interactions) and disengagement behavior. We further set up a YouTube-based system for different experts to annotate video, and software code to automatically collate this information for further analysis. Since these videos contain a "bird's eye" view of the classroom, our tools were designed to allow the annotators to pin-point the event of interest both in space and time.

The second phase involves using the annotated data as a testbed for two specific tasks: a detection task and a classification task. More specifically, the detection task addressed the question: given a video of classroom interaction, can we design a system that can automatically detect unusual behavior? Recall that annotations provided by experts extend both spatially and temporally. This makes the detection / localisation task two fold – temporal and spatial. The temporal aspect involves finding where in time the event of interest occurs, while the spatial aspect requires finding the appropriate region of the classroom space where the event of interest occurs. The classification task then requires us to classify these examples of "unusual" behavior as engagement, disengagement or neither.

We envision such an effort will aid the process of classroom instruction, by providing instructors with a tool to assess classroom response (for e.g., engagement or disengagement) to the instruction, thus facilitating the teaching process and increasing the efficacy of the learning process.

---

[*]Joint first authors, listed in alphabetical order