

Common Ground Can be Efficiently Achieved by Capturing a Screenshot in Handheld-Based Learning Activity

Kibum Kim, Motorola Labs, Schaumburg, IL, USA, kibum.kim@motorola.com
Deborah Tatar, Steve Harrison, Virginia Tech, Blacksburg, VA, USA,
Email: tatar@cs.vt.edu, shr@cs.vt.edu

Abstract: Constructing common ground and the associated convergent conceptual change is critical to collaborative learning. Convergent conceptual change is achieved as participants in a conversation update common ground through presentations, repairs, and acceptances of utterances. Many previous studies in human-computer interaction show that face-to-face communication is more effective than other forms of technology-mediated collaboration, such as video conferencing or telephoning, primarily because such forms of communication cannot fully replicate the context so vital to common understanding. To meet these concerns while enabling the use of handhelds, we devise and test empirically the value of shared visual context in creating common ground by examining communication efficiency.

Introduction

Wireless connecting of computers to the Internet from a variety of everyday locations—from coffee shops to libraries, from airports to hotel rooms—has become so commonplace that it no longer attracts special attention. New, increasingly common opportunities for ubiquitous wireless connectivity exist and so, it seems only logical that many schools have begun adding wireless network capabilities into their traditional classrooms. By creating wireless “hotspots” on the fly, a teacher can set up an instant computer lab, auditorium, and virtual classroom, even outside of buildings. The import of wireless technology to K-12 education is meaningful because it saves money, provides flexibility, and supports expandability on the already-made wired infrastructure. However, in a handheld-enhanced classroom activity, traditional face-to-face communication does not provide an effective enough context, because handhelds—with their small and truly individual screens—do not naturally support the sharing of a workspace. When people can see where each person is looking, it is easier to establish common ground (Kraut et al. 2003). To augment face-to-face communication in handheld-mediated joint activities, we present a new network service that enables learners to share a screenshot during collaborative activities. We tested empirically the value of shared screenshots of handhelds in order to create common ground by examining communication efficiency.

Achieving Common Ground in Communication

As indicated by Clark (Clark & Marshall 1992), common ground is created with evidence, assumption, and induction schema:

$$\text{Equation (1)} \quad \text{Evidence} + \text{Assumption} + \text{Induction schema} = \text{Common Ground}$$

Physical co-presence (such as looking at the same objects), linguistic co-presence (such as referring to the same objects), and community co-membership (such as well-known terms among society members) provide evidence, the first component of common ground. Each occurrence of evidence requires several auxiliary assumptions about the situation, which provide the second component of common ground. For example, when two students discuss a math graph as it appears on their screens, evidence of physical co-presence is used to create their common ground. Common ground involves the assumption that both students are paying attention to their screens at the same time (i.e., simultaneity, attention, locatability, associability, and rationality assumptions). Other kinds of assumptions involve recallability, understandability, and universality of knowledge. The third component of common ground, induction schema, can be defined as follows (Clark & Marshall 1992):

A and B mutually know that p if and only if some state of affairs G holds such that:

1. A and B have reason to believe that G holds.
2. G indicates to A and B that each has reason to believe that G holds.
3. G indicates to A and B that p .

Because the induction schema is fixed, “evidence” exists in inverse proportion to “assumption” in Equation 1 to create the same amount of common ground. In other words, weaker assumptions require stronger evidence. For example, in a classroom environment that exhibits weak levels of attention, understandability and universality of knowledge, students who desire to create common ground will receive more benefits from physical co-presence

(i.e., strong evidence) such as visual-aids than they would from linguistic co-presence (i.e., weak evidence), such as audio-aids. Our new network service, “Look,” permits a student to capture the view from another’s handheld onto his handheld in real time. This synchronized physical co-presence can establish that the items or concepts indexed are within the students’ joint range of attention.

Efficient Creation of Common Ground in Handheld Mediated Communication

In an experiment contrasting “Look” and “No-Look” conditions, 32 groups of three engaged in a task involving substantial shared reference for a peripheral participant. During the experiment, all sessions were recorded on digital video and audiotapes. In the video analysis, we focused on two factors: the mean number of turn-takings and the mean number of overlaps in the conversation. A turn was defined as a stretch of talk contributed by a single speaker; an overlap was defined as occurring when two or more people spoke at a time. Previous literature shows that more effort in achieving common ground is indicated by more turns of talk (Clark & Krych 2004). The act of grounding between participants in a conversation requires that *A* presents an action and/or a signal *s* for *B* to understand, and that *B* in turn eventually validates that action and/or signal as having been recognized or understood. When these two phases are accomplished properly, they constitute the shared basis for the mutual belief that *B* understands what *A* means by signal *s* (Clark 1996).

Displaying understanding gives partners the opportunity for such validation or correction. Using the “Look” network service, participants make displays and exemplifications of understanding practicable for the purpose of validation. Although speakers tend to avoid verbal overlap in primary talk, utterances with visual presentation of understanding are usual enough in everyday settings (Sacks *et al.*, 1974). Therefore, “Look” can be used to allow that an addressee’s presentations overlap a speaker’s verbal descriptions, and, thus, that they continue the conversation without separate turns. When the workspace is made visible by “Look,” the addressee will be able to continuously reformulate his/her tryouts without forming turn-taking. However, without “Look,” the workspace is not visible, so the addressee will seek validation from the speaker, which requires both parties to take more turns. This difference was reflected in the mean number of turns by the discourse participants. In the first trial, student groups without “Look” took over five times as many turns as the groups with “Look” (67 turns vs. 13 turns, $F(1, 30) = 13.66$, $p < .001$). A similar pattern of result was shown in the second trial. Without “Look,” there was an average of 37 turns but with “Look,” an average of 10 turns occurred ($F(1,30)=12.35$, $p < .001$).

The second test used for communication efficiency involved the overlapping of utterances, which is defined as simultaneous speech by participants. Previous findings for audio-only and video-mediated conversations show more interruptions when visual cues are reduced (Argyle *et al.* 1968). Simultaneous speech may be taken to indicate a problem in floor control of a conversation (Sellen 1992). Participants may miss the timing for floor control, or may bid for the floor and fail. Studies which label simultaneous speech as “interruptions” make this tacit assumption. Overlapping speech should be tolerated only if both participants can be attended to well enough for current purposes (Clark 1996). As predicted, the occurrence of overlap by students in the “no-Look” condition was larger than the occurrence in the group that had access to “Look.” In the first trial, an analysis of variance yields a significant difference: 16.4 occurrences for the “no-Look” group vs. 2.4 occurrences for the “Look” group, $F(1, 30) = 22.47$, $p < .001$. In the second trial, the result also shows a significant difference, 12.8 for the “no-Look” group vs. 2.2 for the “Look” group, $F(1, 30) = 13.45$, $p < .001$. Taken together with the measure of turn-taking, these results indicate the effect of sharing visual context for collaborative activity and show the impact of “Look” on creating common ground easily and efficiently.

References

- Argyle, M., Lalljee, M., & Cook, M.(1968). The effects of visibility on interaction in a dyad. *Human Relations*, 21, 3-17.
- Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.
- Clark, H.H., & Krych, M.A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, 50, 62-81.
- Clark, H. H. & Marshall, C. R. (1992). Definite Reference and Mutual Knowledge. In H.H.Clark, (Ed.), *Arenas of Language Use* (pp. 9-59). Chicago: The University of Chicago Press.
- Kraut, R.E., Fussell, S.R., & Siegel, J.(2003). Visual Information as a Conversational Resource in Collaborative Physical Tasks. *HUMAN-COMPUTER INTERACTION* 18, 13–49.
- Sacks, H., Schegloff, E. A. & Jefferson G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language* 50(4), 696-735.
- Sellen, A.(1992). Speech Patterns in Video-Mediated Conversations. *Proc. of human factors in computing systems* (CHI '92) (pp. 49-59). New York: ACM Press.