The Value of Shared Visual Space for Collaborative Physical Tasks

Darren Gergle
Human Computer Interaction Institute
Carnegie Mellon University
Pittsburgh, PA 15213
dgergle+@cs.cmu.edu

ABSTRACT

The goal of this research is to elucidate the ways shared visual space supports group communication and performance. This work involves three stages: a series of empirical studies that decompose the features of shared visual space and task, a methodology for assessing the sequential structure of how visible actions serve to augment discourse, and the development of a computational model of discourse to further our theoretical understanding of the ways in which shared visual information serves communication in collaborative physical tasks.

Categories and Subject Descriptors: H.5.3 [**Information Interfaces and Presentation**]: Group and Organization Interfaces – *computer-supported cooperative work.*

General Terms: Theory, Experimentation, Performance, Design.

Keywords: Shared visual space, language, and communication.

INTRODUCTION

Consider the following scenarios. An automotive design team jointly develops a 3D model for a new chassis, only the materials processing engineer is located in Detroit while the structural engineer is in Stuttgart. A team of surgeons performs an operation while a world expert monitors the progress from his office overseas. An architecture student gets timely help on her mechanical simulation from an engineering tutor across campus. Each of these scenarios is an example of a distributed collaborative task in which at least one person is physically remote from the primary task site. Although there have been recent advances in computing technologies to support such collaborative tasks, a large body of empirical work suggests they are often more difficult and less successful when carried out remotely than when collocated (for a review, see [5]).

In each of these scenarios, shared visual information plays an important role in coordinating communication and joint activities. Yet, we lack a detailed theoretical understanding of the ways in which features of shared visual space interact with task features to impact collaboration [6]. How, for example, does seeing a partner's actions alter a person's speech? How does a small field of view affect the ability of pairs to plan subsequent actions? How does the linguistic complexity of the shared objects affect performance?

This paper summarizes my work on understanding the ways in which shared visual space affects communication and collaboration. I begin by discussing a series of laboratory studies aimed at understanding how features of shared visual space interact with task features to affect performance. Then I overview a sequential analysis method demonstrating how visual information transforms discourse over time. Finally, I propose the development of a computational model detailing how shared visual information affects communication.

METHODOLOGY

Studies of Shared Visual Space

In my work, I have run a series of studies using a dyadic puzzle-building paradigm to explore the value of shared visual space (see Figure 1; for details see [2,4]). The findings demonstrate that the availability of shared visual space supports communication in two major ways. First, it provides an awareness of the task state. Second, it serves as an efficient resource for conversational grounding [1].

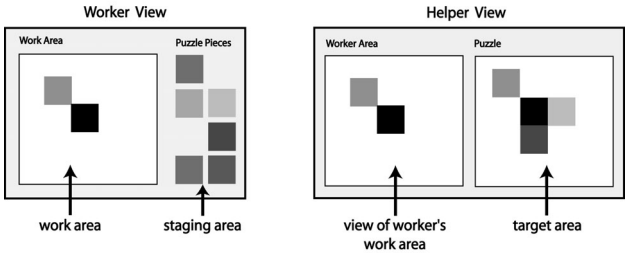

Figure 1. Puzzle task paradigm demonstrating the Worker's view (left) and the Helper's view (right).

The results of these studies show that pairs are approximately 30% faster [2,4] when they have an immediately available shared view of the complete workspace. However, these benefits fluctuate depending on the features of the shared visual space. For example, performance linearly decreases with a coinciding reduction in the proportion of the shared field of view. Offsetting the spatial perspective, delaying the visual update, or allowing

CHI 2005, April 2–7, 2005, Portland, Oregon, USA.
ACM 1-59593-002-7/05/0004.

manual control over the viewable portion of the shared space all lead to degraded performance. In addition, the shared visual space is more beneficial when the task environment is visually complex, dynamically changing or when the objects in the display are difficult to describe linguistically (for details of these studies see [2,3,4]).

In addition to task performance differences, changes in the features of the shared visual space lead to adaptations on the part of the participants in their verbal communication. The pairs make use of the shared visual space in order to ground their conversations more efficiently (e.g., by using linguistic shortcuts such as deictic pronouns [2]), and the pattern of linguistic contributions change in accordance with the theory of least collaborative effort [1].

Sequential Analysis of Language and Action

In order to assess whether the discourse structure changes with the addition of visual information, we applied sequential analysis methods to assess the temporal structure of the discourse, both with and without shared visual information [3].

These analyses reveal that pairs use the visual information in two major ways. First, they use the visual information to serve as a more efficient and less ambiguous source of confirmation. With a shared visual workspace the pairs are less likely to explicitly verify their actions with speech. Rather, they rely on more accurate visual information to provide the necessary communicative and coordinative cues. Secondly, the overall structure of their discourse changes. For example, pairs are able to detect errors earlier on in the course of their work and remedy the situation in a timely fashion before their actions become nested and they need to revert through several previous task states in order to fix any problems.

Computational Modeling of Shared Visual Information

The empirical studies and sequential analyses provide a foundation for the final stage of my dissertation work: the development of a computational model of communication that accounts for shared visual information. While current frameworks mention the utility of visual information in discourse they do not provide a principled way of accounting for it in their models. This work will serve several purposes:

- It will constrain and refine further theory development by forcing a deeper understanding of the actual mechanisms at play.
- It will provide a theoretical baseline from which to make assessments regarding the development of novel technologies aimed at supporting distributed collaborative tasks that use visual information.
- It will provide a theoretical foundation from which software agents, automated tutors, etc. can reason and make use of visual information.

IMPACT FOR HCI

This work impacts the field of HCI at both theoretical and applied levels. At a theoretical level, it leads to an improved understanding of how features of tasks and media, both alone and in combination, affect communication and coordination. It also adds to our knowledge of how task features influence people's use of visual space, and how language and actions are coordinated in team performance. This work will therefore add to general knowledge in a number of disciplines, including HCI, CSCW, communications theory, and social and cognitive psychology.

As the opening examples illustrated, distributed tasks play important practical roles in medical, educational, and industrial domains. My research is building a theoretical framework that will help maximize the fit between technologies and tasks in these and other critical domains. The findings benefit the public by allowing us to identify technologies that enable specialists to work remotely to the best of their capabilities, and by providing guidelines for new technologies that will allow greater numbers of individuals to participate in these domains from a distance. Ultimately, it is the goal of this work to provide a foundation and rationale for the future development, design and deployment of systems to support distributed collaborative physical tasks.

ACKNOWLEDGMENTS

This research is funded through the generous support of an IBM PhD Fellowship. I would like to thank Robert Kraut, Susan Fussell, Carolyn P. Rosé, David Millen, Susan Brennan, and Judy Olson for their valuable thoughts and guidance on this work.

REFERENCES

1. Clark, H.H. & Brennan, S.E. (1991). Grounding in communication. In L.B. Resnick, R.M. Levine, & S.D. Teasley (Eds.). *Perspectives on socially shared cognition*.
2. Gergle, D., Kraut, R.E., & Fussell, S.R. (2004). Language efficiency and visual technology: Minimizing collaborative effort with visual information. *Journal of Language and Social Psychology, 23*, 491-517.
3. Gergle, D., Kraut, R.E., & Fussell, S.R. (2004). Action as language in a shared visual space. *CSCW 2004*, pp. 487-496.
4. Kraut, R.E., Gergle, D., & Fussell, S.R. (2002). The use of visual information in shared visual spaces: Informing the development of virtual co-presence. *CSCW 2002*, pp. 31-40.
5. Olson, G.M., & Olson, J.S. (2003). Distance matters. *Human Computer Interaction, 15*, 139-179.
6. Whittaker, S. (2003). Things to talk about when talking about things. In *Human Computer Interaction, 18*, 149-170.