VISUALLY-ENABLED GEOCOLLABORATION TO SUPPORT DATA EXPLORATION & DECISION-MAKING

Alan M. MacEachren, Isaac Brewer, Guoray Cai, and Jin Chen

GeoVISTA Center (www.geovista.psu.edu)
302 Walker, Department of Geography
Penn State University
University Park, PA 16802
maceachren@psu.edu

Abstract:

Current mapping and related geospatial technologies are not designed to support group work and we have a limited theoretical or practical basis from which to extend (or reinvent) technologies for group use of geospatial information. To address the challenge of supporting group work with geospatial information, we have developed a comprehensive conceptual approach to *geocollaboration* and are applying that approach to a range of prototype systems that support both same- and different-place group activities.

Our focus in this paper is on same-time, same-place group work environments that mediate distributed thinking and decision-making through use of large-screen displays supporting multi-user, natural interaction. Two environments will be described and compared. Both make use of hand gestures as a mechanism for specifying display locations. One adopts a white board metaphor while the other adopts a drafting table metaphor. We also consider two use cases: group data exploration (by scientists and analysts) and group decision-making (by crisis managers and planners).

1 INTRODUCTION

Visual displays of geospatial information in the form of maps and images have long served as enabling devices for group work. Urban and regional planners, for example, often gather around large paper maps to discuss master plans or specific development choices and these same large format maps are used as the object of discussion at subsequent public meetings. Similarly, teams involved in crisis management use large maps in both situation-assessment and response activities and earth scientist (e.g., geologists, ecologist) often work collaboratively on development of map categories and on planning for field research activities. These are rudimentary examples of what we label *geocollaboration.* As an activity, we consider geocollaboration to be group work about geographic scale problems facilitated by geospatial information technologies. As a field of research, we consider geocollaboration to be the study of these group activities, together with the development of methods and tools to facilitate them.

Recent technological advances in display hardware and multimodal interfaces are making it possible to merge the advantages of large format representations that facilitate group work with those of dynamic, interactive displays (applied over the past decade to desktop mapping and GIS applications designed for individual use). This merger is likely to have a substantial impact on group productivity. In addition, dynamic, large-format displays having natural interfaces designed specifically to support group work have the potential to dramatically (and qualitatively) change the way groups work with geospatial data, thus to create fundamentally new kinds of geocollaboration.

This paper provides an update on two ongoing projects to develop methods and tools that support visually-enabled geocollaboration – among humans and between human and computer agents. The research builds on a human-centered conceptual approach to both design of geocollaboration environments and evaluation of environment usability. For details, see: [3, 4]. The overall approach integrates perspectives from cognitive science (particularly distributed cognition), semiotics (particularly the mechanisms through with representations are devices for sharing meaning), and usability studies (particularly cognitive systems engineering). Here, we focus on different metaphors for support of group work with large screen display and on some of the key design decisions that underlie the natural, multi-user interfaces we have implemented.

We begin below (in section 2) with a brief overview of recent research on large-screen, map-based displays and their use in facilitating group work. In section 3, we describe and compare two environments that we are developing. Both make use of large displays and natural interaction to enable same-time, same-place group work with geospatial information. One environment supports joint use of exploratory geovisualization tools. The second is directed toward crisis response facilitated by GIS. Section 4 provides discussion and plans for future research.

2 BACKGROUND

The advantages of large format maps as group situation-assessment and decision-making tools have prompted multiple authors to consider the potential of dynamic, large-format, map-based displays for group work with geospatial information. Florence, et. al [5], for example, proposed (but did not implement) the *GIS wallboard*, an electronic white board envisioned to support sketch-based gestures (of the sort implemented by Oviatt [6] and Egenhofer [7] for tablet displays). In the precursor to our multiuser Dialogue-Assisted Visual Environment for Geoinformation (DAVE_G) system (discussed in section 3) our colleague Rajeev Sharma and his research team successfully implemented a natural multimodal (speech-gesture) interface to a large screen dynamic map [8, 9] and extended the system to support a crisis response scenario used to test robustness of the interface methods [10].

The environments above all adopt a white board (or wall map) metaphor. This kind of interface is likely to be useful in a context such as a public planning meeting or emergency operations center briefing in which one or two individuals take a lead role in presenting information and steering a group discussion. This kind of interface (like the traditional white board or black board) affords the action of walking up and drawing or writing, then giving way to another actor.

An alternative metaphor is the drafting/work table. This metaphor affords group activity around (rather than in front of) the map display (creating an environment similar to one with a large map on a drafting table). This format is typical of work by military and emergency management personnel in the field or urban planners in the office (where they may conduct extended work prior to its presentation with a wall display at a public meeting). Hopkins and colleagues [11] as well as Arias, Fischer and colleagues [12, 13] have implemented large, table-like group work displays to map-based planning activities. The later group has merged virtual and physical space in a system that allows users to create a shared model of a planning problem by manipulating 3D physical objects that provide a “language” for interacting with a computer simulation.

In some contexts, such as military planning and crisis response, large paper maps retain a distinct advantage in their combination of high resolution and portability. McGee [14, 15] has studied military planners working with such maps. Based on this research, he proposed an approach to augmenting paper maps through digital Post-it notes (physical notes for which the position and content could be sensed by the system). The goal was to create a robust system that did not require users to learn new work routines and that would continue to work even when technological or power failures occurred.

A third metaphor used in group work environments is activity (or geographic) space itself. Activity spaces afford entering and behaving within them; and that is what immersive environments for group work attempt to support. Neves and colleagues [16] developed an immersive virtual workspace based on a GIS room metaphor (a room in which maps can be mounted on the wall or placed on a digitizing tablet for encoding in the system). They implemented the environment only individual users but, conceptually, the metaphor could support multiple users. One of the first collaborative, immersive environments using a geographic space as the underlying metaphor is the *Round Earth Project*, developed to enable children’s learning about the shape and size of the earth [17]. While that effort focuses on same-place collaboration, there have been several Cave and ImmersaDesk-based demonstration projects that support collaboration within 3D, geographic-scale environments representing real and modeled spatio-temporal processes, see: [3, 18, 19]. Recently, Armstrong [20] identified teleimmersive environments (different-place, collaborative, immersive environments that rely on high performance computing and distributed geo-processing) as a grand challenge to the research communities in geographic and information sciences.

3 NATURAL, MAP-BASED INTERACTION WITH GEOSPATIAL INFORMATION

Here, we discuss two geocollaborative system development efforts, emphasizing the role of maps as a primary interface component in each. The first system uses a horizontal display that functions much like traditional drafting tables that multiple participants in a group activity can gather around. The second system uses a vertical display that functions more like an electronic white board. Both differ from most other large screen environments in their use of hand gestures in place of mouse, pen, or wand as a primary interface method for specifying display location.

3.1 HI-SPACE

The HI-SPACE (Human Information Workspace) environment offers a platform for enabling groups of analysts to interact with each other and with geospatial data in new ways, remedying some of the inefficiencies involved in group use of visualization tools on traditional displays. This prototype, collaborative virtual environment (CVE) is an experimental, hands-free, untethered, enhanced reality system developed by Richard May [1]. The goal of developing this HI-SPACE environment was to promote more natural interaction between groups of users and modern computing displays.

HI-SPACE has specific attributes that have the potential to significantly alter collaborative interaction for decision making, exploration and command and control situations. First, the size of the display enables groups of individuals to work in a comfortable round-table fashion, rather than dispersed on separate personal computers or clustered around smaller, vertical displays. Second, untethered gesture recognition (not requiring a data glove or other device) allows group members to use natural forms of communication to share ideas (such as pointing to indicate emphasis). Third, the table supports phicon (physical object) recognition so that users can utilize real world objects on the display as they would on a traditional table or desk top to augment and enhance collaborative discussions. Each of these features is discussed below, and the context in which these functionalities have an impact for users of geospatial information is considered.

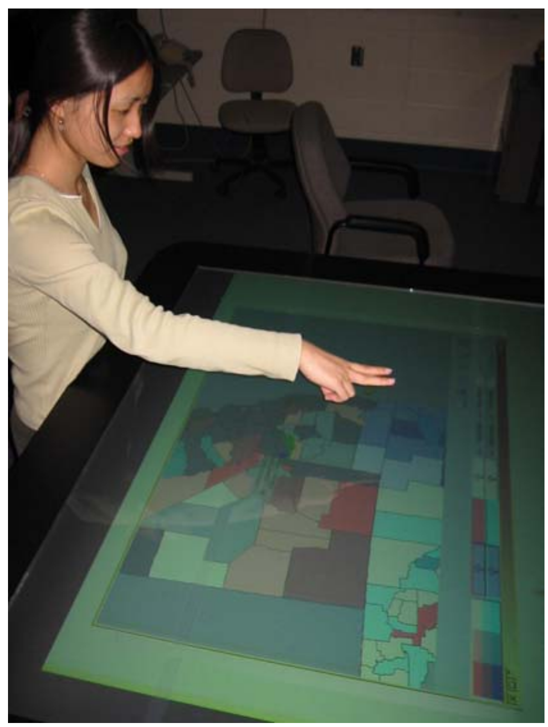

Figure 1. Gesture interface to the HI-SPACE Table. Demonstration of collaboration with interactive map component in GeoVISTA *Studio*. HI-SPACE *Table developed by Richard May[1], on loan to the GeoVISTA Center from the Pacific Northwest National Laboratory.*

Data gloves, head mounted displays, data wands, and other tools for interacting with virtual data have not been widely adopted by practitioners. There is, thus, a need for untethered interaction that reflects the natural interaction among collaborators, the surrounding environment, and the CVE. The HI-SPACE environment has the potential to comfortably support 3-6 simultaneous collaborators using relatively natural (untethered) gestures and the software provides an individual cursor icon for each of the participants. This form of interaction is likely to improve group communication through eye contact, gaze, and the ability of each person to experiment with their individual cursor.

Our work addresses this need by building on the neural network gesture recognition developed by May [1]. Currently, the HI-SPACE table can track the hand position and identify individual gesture poses (e.g. two fingers extended). Modern Operating Systems (OS) are designed to support interaction with single users. That means there is only one mouse available for interaction between a user and a computer. In order to support *geocollaboration*, in which multiple users work concurrently on a single platform (computer), multiple mice or channels are needed in single computer. Our extensions to the HI-SPACE environment address this issue.

Here, we introduce, briefly, how these extensions to HI-SPACE support interactions between multiple users and a Java application platform. Understanding multi-user interaction requires a brief discussion of how a single user interacts with a Java application. As shown in figure 2, a mouse click is translated by the operating system into an OS-level event. The event is sent to the Java Virtual Machine (JVM) where it is translated into a JVM mouse event. Java applications actually respond to JVM (rather than OS) events. In order to enable multiple-users interaction, we can generate virtual mouse events, either OS-level or JVM-level, for each user.

HI-SPACE collects interactive information from multiple users by capturing user gestures. Different gestures indicate different mouse behaviors. For example, we have implemented two simple actions: stretching out one finger indicates a mouse move action and using two fingers indicates a mouse press action. The gestures of each user are translated into virtual mouse events which are fed into the OS, sequentially, thus, the establishing a direct link between the users and the computer through HI-SPACE. In practice, as JVM mouse events are generated they are recognized, processed, and fed to the Java Virtual Machine. Figure 2 shows how this procedure works.

Integrating HI-SPACE with a Java application is relatively easy. From the perspective of the JVM, the mouse events generated by HI-SPACE are not different from those generated by the real mouse. Thus a Java application responds to HI-SPACE events in the same way as it does to real mouse events. This means, theoretically, we do not have to change the Java application except by attaching an adapter to accept HI-SPACE events.

Concurrent users of HI-SPACE are not limited to same-place work; they can be in distributed places. For distributed users, virtual mouse information can be transmitted via the network. Priorities for virtual mouse events can be established so that interference among users' operation can be avoided.

Specifically, our work is focusing on the development of a coordinator or arbitrator that helps determine which user has control of the system at any given time, while storing other related events into a queue for later processing. Long term efforts are aimed at merging voice recognition software to identify the person who is in control of the collaborative discussion, and subsequently provide that individual with highest priority for interaction.

May [1] envisioned a HI-SPACE environment that would minimize attention shifting from data work to collaborative work by providing seamless interaction between collaborators and the information through the use of physical objects, or phicons. His plan was to merge the workstation

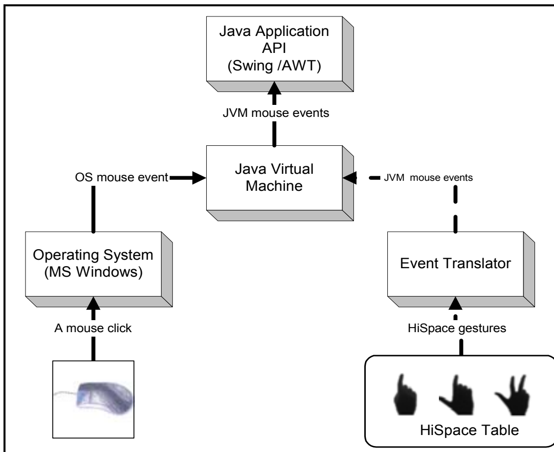

Figure 2. Implementation strategy for supporting multiple participants, using gesture to initiate mouse events.

and the typical desk environment together into a seamless coupling of shared information. For example, if a collaborator were to place pen at a location on the table to indicate something, but then get distracted by a discussion with a fellow collaborator, the table should recognize the pen as a place holder that assists in guiding the discussion away from the tangent and back to its original focus. To this end, we are exploring the use of pens, erasers, markers, magnifying glasses and other physical objects that can be used to not only facilitate human-human collaboration, but also be recognized by the HI-SPACE display for interaction with the underlying geospatial data.

We are also building on this base to provide complex gesture support that does not require individual hand poses, but instead, uses the gestures that come naturally when indicating information on the table top display. Our approach to natural interaction with geographic applications is also expanding from a gesture-only approach to the inclusion of voice commands. We are in the process of adding simple voice commands that complement and augment gesture commands to create a flexible, easy to learn, and easy to use interface. Both natural, free-hand gesture interpretation and its integration with speech input are central components of DAVE_G (described below).

3.2 DAVE_G – Dialogue-Assisted Visual Environment for Geoinformation

Development of our initial DAVE_G prototype has been made tractable by narrowing the potential application domain from collaborative work generally to support for collaborative work on geospatial data in crisis management. To accommodate the need for a large format map as a shared context for collaborations among different domain experts, DAVE_G (figure 3) uses a large screen display where maps are served from geographical information servers, and users can stand freely in front of the display (implementing a white board metaphor). In order to make the collaborative decision-making more effective, DAVE_G addresses two challenging problems commonly found in the traditional use of geographical information in emergency management centers. First, there is a need to relieve emergency managers from the burden of having to use keyboard and mouse to formulate well-structured commands. Here, we offer the ability to interact with the system using natural modalities (spoken language and natural gestures). Second, emergency managers should be able to interact directly with geographical information instead of interacting with a GIS operator who can be a bottleneck to (rather than facilitator of) communication in a collaborative work environment. To deal with the first challenge, DAVE_G uses microphones and active cameras to capture spoken language and natural gestures as direct input that drives the system's response on the map display. To deal with the second challenge, an intelligent dialogue agent is employed to process ill-structured, incomplete, and sometimes incorrect requests, and to facilitate task-oriented, extended interactions and collaborations. For a detailed explanation of the architecture for our initial DAVE_G prototype see [2].

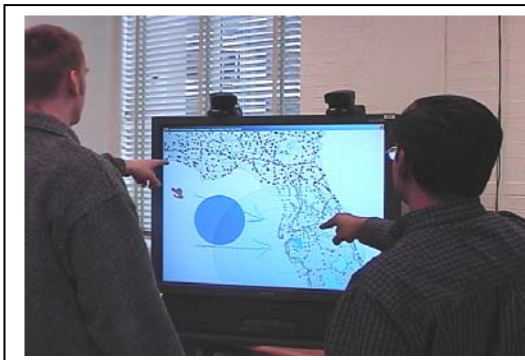

Figure 3. Two-person, gesture-speech interface to DAVE_G. Demonstration of a collaboration scenario focused on analyzing potential hurricane impacts. *figure reproduced from* [2].

DAVE_G is based on the interaction framework initially developed in *iMap* [21] XISM [8, 9, 22]. We have added substantial extensions to support multiple user interaction (by duplicating modules for speech and gesture recognition for each additional participant) as well as human-system collaboration (through addition of a human collaboration manager). To capture and process speech, DAVE_G utilizes a speaker dependent voice recognition engine (ViaVoice from IBM) that allows reliable speech acquisition after a short speaker training procedure. The set of all possible utterances is defined in a context free grammar with embedded annotations. This constrains the available vocabulary but retains flexibility in the formulation of speech commands.

Hand gestures are captured using computer vision-based techniques, and are used to keep track of the user's spatial interest and spatial attention. For reliable recognition of hand gestures, a number of vision-related components (face detection, palm detection, head and hand tracking) are engineered to cooperate together under tight resource constraints. The results of speech recognition and gesture recognition each provide partial information for intended actions. To achieve a complete and coherent understanding of a user's request, verbal utterances from the speech recognition have to be associated with co-occurring gestures observed by the gesture recognition process. Currently, DAVE_G can understand speech/gesture requests for most commonly used map display functions such as "show a map of population within Pennsylvania", "zoom here{gesture}", "highlight these{gesture} features", "make a one-mile buffer around these features", and more.

In DAVE_G, dialogue is neither user-led nor system-led, but rather is a mixed-initiative process controlled by both the system and the users in collaboration. It allows complex information needs to be incrementally specified by the user while the system can initiate dialogues anytime to request missing information for the specification of GIS query commands. This is important since the specification of required spatial information can be quite complex, and the input from multiple people in several steps might be needed to successfully complete a single GIS query. Therefore, the HCI can not require the user to issue predefined commands, but needs to be flexible and intelligent enough to allow the user to specify requested information incompletely and in collaboration with other users and the system.

Information requests are provided to the system in fragments of spoken utterances and gestures that can not be understood without taking into account the shared context established by previous discussions (interactions). Furthermore, information requests that come from different users may be incoherent, or even conflicting with each other, and such problems must be handled carefully to avoid 'breakdowns' in the collaboration process. The dialogue manager in DAVE_G is able to understand and guide the user through the process of querying the system for information and acts to verify and clarify the dialog with the user when there is missing information or recognition errors. To provide such behavior, the dialog manager requires a deep understanding regarding the current discourse context and task progress, and also must maintain a model of users in terms of their intention, attention and information pool. To handle complex human-GIS-human dialogues in geocollaborative use of map information, DAVE_G uses the SharedPlan theory [23] to guide the development of a model of rational behavior in group spatial decision making. It models the map-mediated geocollaborative environment as a system of multiple agents that plan and act cooperatively.

3.3 Discussion

Our approach to designing, developing, and creating multimodal systems is yielding promising results. For example, lessons learned about work domains, work tasks, collaboration, and technological challenges from work in the HI-SPACE environment often carry over to work on the DAVE_G system (or the reverse). This robust, simultaneous development cycle has yielded new insights not only into the nature of collaboration with geospatial information, but also into the design of complex systems themselves.

4 CONCLUSIONS

The two system development efforts discussed above are part of a larger effort to develop a theoretical framework that supports the design, implementation, assessment, and application of technologies to support geocollaboration as well as the study of geocollaboration as a process. Technology-enabled geocollaboration is a relatively new domain of research and practice. As such, there are many unanswered questions and the platforms detailed above provide an opportunity to investigate a subset of them. Specifically, we are focusing on: the impact of different metaphors to enable collaboration in different problem domains and with different kinds of geoinformation technologies, alternative methods for making interfaces more natural (and whether this does, in fact, make them easier to use), and how visual displays enable (or might enable) human-system and human-human dialogue and joint work.

Supporting group work with geospatial information is a challenging task. Maps have played a substantial role in collaborative activities for centuries, but cartographers seem to have given little thought to the design of maps (or map-based interactive displays) to specifically support group work. Similarly, while there has been considerable attention given to group spatial decision support [24-26], only limited attention has been given to visually-enabled group work. We view this gap in our knowledge and understanding as a substantial opportunity for cartography to make an impact on GIScience and information science more generally and on the application of that science in a range of contexts for which group work with geospatial information is critical. We encourage cartographers to this engage this opportunity.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grants No. BCS-0113030 and EIA-0306845 and on work supported by the U.S. Geological Survey. Many colleagues have contributed to development of these ideas; in particular we would like to acknowledge Rajeev Sharma, Ingmar Rauschert, Levent Boelli, Sven Fuhrmann, Benyah Shaparenko, and Hongmei Wang.

References

[1] R. A. May, "HI-SPACE: A Next Generation Workspace Environment." Pullman, WA: Washington State University, 1999.

[2] I. Rauschert, P. Agrawal, S. Fuhrmann, I. Brewer, H. Wang, R. Sharma, G. Cai, and A. MacEachren, "Designing a human-centered, multimodal gis interface to support emergency management," presented at ACM GIS'02, 10th ACM Symposium on Advances in Geographic Information Systems, Washington, DC, USA, 2002.

[3] A. M. MacEachren and I. Brewer, "Developing a conceptual framework for visually-enabled geocollaboration," *International Journal of Geographical Information Science*, in press.

[4] A. M. MacEachren, R. Sharma, G. Cai, I. Rauschert, P. Agrawal, and I. Brewer, "Enabling collaborative geoinformation access and decision-making through a natural, multimodal interface," submitted.

[5] J. Florence, K. Hornsby, and M. J. Egenhofer, "The GIS wallboard: interactions with spatial information on large-scale displays," in *International Symposium on Spatial Data Handling*, vol. 7, M. J. Kraak and M. Molenaar, Eds. Delft, The Netherlands: Taylor and Francis, 1996, pp. 449-463.

[6] S. L. Oviatt, "Multimodal Interactive Maps: Designing for Human Performance," *Human-Computer Interaction*, vol. 12, pp. 93-129, 1997.

[7] M. J. Egenhofer, "Query processing in spatial-query-by-sketch," *Journal of Visual Languages and Computing*, vol. 8, pp. 403-424, 1997.

[8] S. Kettebekov and R. Sharma, "Understanding gestures in a multimodal human computer interaction," *International Journal of Artificial Intelligence Tools*, vol. 9, pp. 205-224, 2000.

[9] S. Kettebekov and R. Sharma, "Toward multimodal interpretation in a natural speech/gesture interface," presented at Proceedings, International Conference on Information Intelligence and Systems, 1999.

[10] S. Kettebekov, N. Krahnstöver, M. Leas, E. Polat, H. Raju, E. Schapira, and R. Sharma, "i2Map: Crisis Management using a Multimodal Interface," presented at ARL Federate Laboratory 4th Annual Symposium, College Park, MD, 2000.

[11] L. D. Hopkins, R. Ramanathan, and R. V. George, "Interface for a Planning Workbench," vol. 2001: Department of Urban and Regional Planning, University of Illinois at Urbana-Champaign, 2001.

[12] E. Arias, H. Eden, G. Fischer, A. Gorman, and E. Scharff, "Transcending the individual human mind - creating shared understanding through collaborative design," *ACM Transactions on Computer-Human Interaction*, vol. 7, pp. 84 - 113, 2000.

[13] G. Fischer, "Articulating the task at hand by making information relevant to it," *Human-Computer Interaction (Special issue on context-aware computing)*, vol. 16, pp. 243-256, 2001.

[14] D. R. McGee, P. R. Cohen, and L. Wu, "Something from nothing: Augmenting a paper based work practice via multimodal interaction," presented at Proceedings of the ACM Designing Augmented Reality Environments DARE 2000, Helsinor, Denmark, 2000.

[15] D. R. McGee, P. R. Cohen, R. M. Wesson, and S. Horman, "Comparing paper and tangible, multimodal tools," presented at Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves, Minneapolis, Minnesota, USA, 2002.

[16] N. Neves, J. Silva, P. Goncalves, J. Muchaxo, J. M. Silva, and A. Camara, "Cognitive spaces and metaphors: A solution for interacting with spatial data," *Computers and Geosciences*, vol. 23, pp. 483-488, 1997.

[17] A. Johnson, T. Moher, S. Ohlsson, and M. Gillingham, "The round earth project - Collaborative VR for conceptual learning," *IEEE Computer Graphics and Applications*, vol. Nov/Dec, pp. 60-69, 1999.

[18] G. H. Wheless, C. M. Lascara, A. Valle-Levinson, D. P. Brutzman, W. L. Hibbard, B. Paul, and W. Sherman, "The Chesapeake Bay Virtual Ecosystem: Initial results from the prototypical system," *International Journal of Supercomputer Applications and High Performance Computing*, vol. 10, pp. 199-210, 1996.

[19] A. M. MacEachren, R. Edsall, D. Haug, R. Baxter, G. Otto, R. Masters, S. Fuhrmann, and L. Qian, "Virtual environments for geographic visualization: Potential and challenges," presented at Proceedings of the ACM Workshop on New Paradigms in Information Visualization and Manipulation, Kansas City, KS, Nov. 6, 1999 (www.geovista.psu.edu/publications/NPIVM99/ammNPIVM.pdf), pp. .

[20] M. P. Armstrong, "The Four Way Intersection of Geospatial Information and Information Technology," presented at White paper prepared for NRC/CSTB Workshop on Intersections between Geospatial Information and Information Technology, 2001.

[21] R. Sharma, I. Poddar, E. Ozyildiz, S. Kettebekov, H. Kim, and T. S. Huang, "Toward Interpretation of Natural Speech/Gesture: Spatial Planning on a Virtual Map," presented at Proceedings of ARL Advanced Displays Annual Symposium, Adelphi, MD, February, 1999, 1999.

[22] R. Sharma, M. Yeasin, N. Krahnstoever, I. Rauschert, A. MacEachren, G. Cai, K. Sengupta, and I. Brewer, "(invited) Role of Speech/Gesture Driven Multimodal Interfaces for Crisis Management," *Proceedings of the IEEE*, in press.

[23] B. J. Grosz and S. Kraus, "Collaborative plans for complex group action," *Artificial Intelligence 86: 269-357*, vol. 86, pp. 269-357, 1996.

[24] M. P. Armstrong, "Requirements for the development of GIS-based group decison-support systems," *Journal of the American Society of Information Science*, vol. 45, pp. 669-677, 1994.

[25] T. L. Nyerges and P. Jankowski, "Enhanced adaptive structuration theory: A theory of GIS-supported collaborative decision making," *Geographical Systems*, vol. 4, pp. 225-259, 1997.

[26] P. Jankowski and T. Nyerges, *Geographic Information Systems for Group Decision Making: Towards a participatory, geographic information science*. New York: Taylor & Francis, 2001.