

## **Approaching a New Multimodal GIS-Interface**

**Ingmar Rauscher and Rajeev Sharma**

Department of Computer Science and Engineering  
The Pennsylvania State University  
University Park, PA 16802, USA  
{rauscher, rsharma}@cse.psu.edu

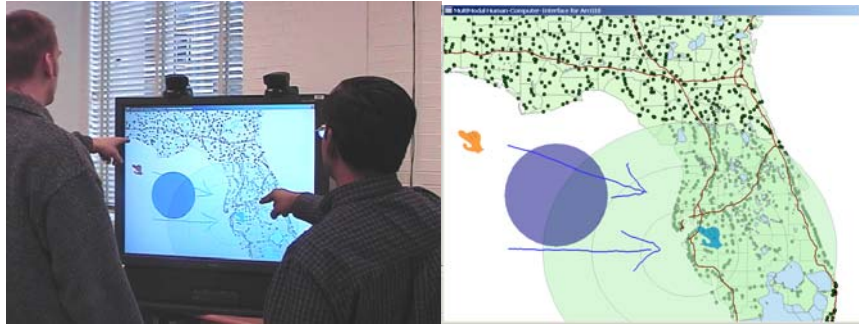
**Sven Fuhrmann, Isaac Brewer and Alan MacEachren**

GeoVISTA Center  
The Pennsylvania State University  
University Park, PA 16802, USA  
{fuhrmann, isaacbrewer, maceachren}@psu.edu

Geospatial information is used in emergency management centers, architectural bureaus, environmental protection offices and daily life. Not only are trained experts using Geographic Information Systems (GIS) but novices, e.g. government employees, rescue workers and the general public, also use GIS. Despite the success of various GIS applications, their current use and usability is still constrained by overly complex user interfaces. While keyboard and command line shortcuts can make use of GIS applications more efficient for expert users, often some functions can only be accessed through repetitive menu and wizard tool usage. These functions are “hidden” for novice users, which remain tethered by the WIMP (windows, icons, menus, pointers) model of human computer interaction. Another, more often overlooked limitation of current GIS interfaces is that GIS user interfaces are designed for single-user interaction and thus do not support aspects of collaborative group work [1].

In order to overcome these limitations, a paradigm shift (and associated new approaches) towards multimodal interface design for GIS are needed. Current research in multimodal user interfaces [2-7] suggest that such interfaces are more efficient and more usable than the traditional interaction mode. Our goal is to develop a Dialogue-Assisted Visual Environment for Geoinformation (DAVE\_G) that uses different interaction modalities, domain knowledge and task context for a meaningful dialog management that supports easy, efficient and collaborative group work with GIS. In combination, speech and gesture interaction modalities indicate many advantages over systems that are using only one input modality (e.g. speech) or standard input devices (keyboard and mouse). While speech provides an effective and direct way of expressing actions, pronouns and abstract relations, it often fails when spatial relations or locations have to be specified. Despite belief to the contrary, speech is not self-sufficient [8]. Therefore gestures can provide an effective second input modality that is more suitable for expressing spatial relations and is less error prone than if expressed in words [9].

The research reported upon here presents a first prototypical GIS interface that builds upon the multimodal interface framework [10] by Sharma et al. The set-up of the initial DAVE\_G uses a large screen display, non-intrusive microphone domes (attached to the ceiling) and active cameras that allow multiple users to move freely in front of the system and issue queries to the GIS. Supported interaction modalities are speech and natural



**Figure 1. Collaborative interaction with the multimodal GIS interface, and actual snap shot of screen**

gestures, whereas spoken commands could be chosen freely within the definition of an annotated grammar. Natural gestures such as pointing and outlining areas of interest on the large screen display help to specify the spatial concept of many geographical queries. Figure 1 shows the prototype system when used during a collaborative work session with two users. Currently, the functionality of the prototype is limited to simple drawing and display commands, such as draw circle, scroll or zoom plus selected more complex commands, such as spatial selection or selection by attribute (see Table 1). The prototype however, is easily extendable and therefore new utterances and commands can be incorporated into the existing.

However, before extending the system to its full functionality, the usability of the prototypical interface design needs to be investigated. Only limited empirical research has been conducted on the design of integrated speech-gesture GIS interfaces and no guidelines exist to determine interface designs that are efficient and natural to users. Thus, a human centered design approach has been chosen for the DAVE\_G project that focuses on emergency management as the exemplary task domain. A cognitive systems engineering approach is being developed that incorporates domain expertise throughout the initial design, implementation, assessment and refinement phase. Focusing on domain users and tasks, we have started to analyze emergency management tasks through onsite visitations of operations centers and administration of questionnaires and interviews. Some of the emergency management tasks include transportation support, search and rescue efforts, environmental protection, and firefighting. The questionnaires and interviews have allowed us to identify GIS-based response activities and operations during disaster events and helped to determine the range of possible GIS functionalities within DAVE\_G. The resulting prototypical multimodal interface of DAVE\_G serves as a test-bed for user studies and helps to assess users gestures and speech patterns.

Presently, information retrieval in GIS remains complicated, time consuming and often distracts users from their actual tasks. In addition, current GIS are not designed for collaborative work. If users were able to express their requests using speech and gestures, GIS queries such as *“show me all hotels close to this shoreline ”* would become much easier to express and the result of this query would appear instantly on the display. An instant response would allow an efficient interaction process with the GIS, even with multiple individuals, supporting time-critical tasks as they occur in emergency

Command Type	Data Query	Viewing	Drawing
Functionality	show/hide layers, buffers spatial, w/ gestures attribute, w/o gestures	scroll left/right/up/down zoom in/out/full extend center at zoom area	circle line free-hand

**Table 1. Supported User Commands**

management centers. In addition, novice users could apply such a multimodal GIS interface more naturally and overcome initial inhibitions much easier. A first step towards such “easy” GIS interactions has been realized with the development of the multimodal and multi-user interface prototype of “DAVE\_G”. In subsequent research, we will focus on improving the interaction modalities by conducting realistic user studies, developing methods that allow more natural gesture and speech expressions and incorporating a knowledge-based interaction dialog between users and a GIS. Our goal is to approach a more “natural” communication between users and a GIS, and as a result to utilize the full power of GIS that today often remains unused.

### Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. BCS-0113030, PI: Alan M. MacEachren, CoPIs: Rajeev Sharma and Guoray Cai. ESRI’s software contribution to the project is gratefully acknowledged.

### References

- [1] I. Brewer, A. M. MacEachren, H. Abdo, J. Gundrum, and G. Otto, "Collaborative Geographic Visualization: Enabling shared understanding of environmental processes," Proc. of the IEEE Information Visualization Symposium, Salt Lake City, Utah, 2000.
- [2] P. Cohen, D. McGee, S. Oviatt, L. Wu, J. Clow, R. King, S. Julier, and L. Rosenblum, "Multimodal Interaction for 2D and 3D Environments," *IEEE Computer Graphics and Applications*, vol. 19, pp. 10-13, 1999.
- [3] M. Egenhofer, "Query Processing in Spatial-Query-by-Sketch," *Journal of Visual Languages and Computing*, vol. 8, pp. 403--424, 1997.
- [4] S. Kettebekov, N. Krahnstoeber, M. Leas, E. Polat, H. Raju, E. Schapira, and R. Sharma, "i2Map: Crisis Management using a Multimodal Interface," in *ARL Federate Laboratory 4th Annual Symposium, College Park, MD*, 2000.
- [5] D. R. McGee, P. R. Cohen, and L. Wu, "Something from Nothing: Augmenting a Paper Based Work Practice Via Multimodal Interaction," Proc. of the Proceedings of the ACM Designing Augmented Reality Environments DARE, Helsinor, Denmark, 2000.

- [6] S. Oviatt and P. Cohen, "Multimodal interfaces that process what comes naturally," *Communications of the ACM*, vol. 43, pp. 45-53, 2000.
- [7] R. Sharma, I. Poddar, E. Ozyildiz, S. Kettebekov, H. Kim, and T. Huang, "Toward Interpretation of Natural Speech/Gesture: Spatial Planning on a Virtual Map," in *Proceedings of the Advanced Display Federated Laboratory Symposium*, 1999, pp. 35-39.
- [8] S. Oviatt, "Ten myths of multimodal interaction," *Communications of the ACM*, vol. 42, pp. 74--81, 1999.
- [9] S. Oviatt, "Multimodal interfaces for dynamic interactive maps," Proc. of the Conference on Human Factors in Computing Systems (CHI'96), 1996.
- [10] N. Krahnstoeber, S. Kettebekov, M. Yeasin, and R. Sharma, "A Real-Time Framework for Natural Multimodal Interaction with Large Screen Displays," (in press), Proc. of the Fourth IEEE International Conference on Multimodal Interfaces (ICMI 2002), Pittsburgh, USA, 2002.