

Imagery as Data: Structures for Visual Model Building

Brian K Smith & Erik Blankinship

MIT Media Laboratory

Abstract: Scientific visualization tools render quantitative data into visual representations that can be analyzed and manipulated by students. But certain types of problems require more qualitative, observational data that cannot be easily represented with current visualization software. This paper describes an approach to using photographs and video as a primary data source for observational inquiry. We describe a framework for students to collaborate around photographs and video, collaboration that leads to inquiry and the development of explanatory models. We also describe two of our learning environments to illustrate how students can begin to develop predictive theories from image data.

Keywords: digital imaging, explanation, modeling, problem-based learning

Introduction

Techniques and tools from scientific visualization hold great promise for changing the ways that students conduct classroom inquiry (Gordin & Pea, 1995). By representing large quantities of quantitative data visually, students can begin to grasp complex ideas that were once foreign to classroom activities. For many problem domains, the visual representations provided by scientific visualizations are simpler to master and manipulate than symbolic and numerical ones, allowing students to more easily detect patterns and make sense of complex data. Instead of simply memorizing and applying formulas and algorithms, students can use static and dynamic visualizations to gain a deeper understanding of domain phenomena through authentic activities.

Most applications of scientific visualization map quantitative data into visual representations. For instance, temperature may be displayed as a color spectrum, with red being the hottest points, blue being the coolest points. Such mappings work when numerical data can be collected, but many problems require more qualitative, observational data in order to be explained. For instance, census values can tell us a great deal about the population of migrating animals, but we may need other sorts of observational data to understand why their numbers are increasing or decreasing. In this paper, we discuss ways to use observational data in visualization and modeling activities. Rather than map numerical quantities into graphical forms, we use ordinary photographs and video clips as primary sources of data. Imagery can provide evidence for claims about observable processes and behaviors, and our research provides students with tools to analyze and generalize explanations from still and moving images.

In this paper, we describe a class of activities where students use photographs and video to investigate and explain complex phenomena. By removing the narratives and captions that generally accompany images and film, we create situations where students must discover visual patterns for themselves. Once patterns are discovered, they develop qualitative models to generalize their findings and make predictions about additional, related visual data. In this way, we shift from the use of imagery as an information source to *imagery as data* to facilitate model building and explanation.

We describe a framework for developing models from image data and two learning environments that embody the framework. The first, *Animal Landlord*, presents digitized nature films to students exploring issues in behavioral ecology. These students annotate and compare film clips to produce qualitative models of predator-prey behaviors. The second application, *Image Maps*, presents students with historical images of their communities that are used to develop understandings of urban planning and cultural change. In both cases, students collaborate around image data to construct qualitative models of observable processes (*i.e.*, predation behaviors, community change). The remainder of the paper describes how students use images in inquiry tasks.

Why imagery?

Visual events are rich with opportunities for students to pose questions and reflect on behaviors and processes (Bransford, Franks, Vye, & Sherwood, 1989; Collins & Brown, 1988; Wetzel, Radtke, & Stern, 1994). Teachers can encourage students to think about events and anomalies present in the imagery, including issues not explicitly mentioned in a film's narration or a photograph's caption. Being able to view moments in time with photographs and video can also facilitate learning — watching a lion chase its prey or seeing the styles of dress in the 1940's is dramatically different than simply reading about animal behavior or history. And because imagery maintains a historical record, students can return to it, reexamine it, and continue to learn from it. As a result, imagery establishes a context for problem solving, for generalizing explanations from pictorial evidence.

Imagery as data

However, photographs and video are not typically thought of as artifacts for problem solving and inquiry; they are more likely to be treated as "visual aids" to accompany text or audio information. For video, we can draw distinctions between video as information and video as data (Nardi, Kuchinsky, Whittaker, Leichner, & Schwarz, 1996; Whittaker & O'Conaill, 1997). The former treats video as an additional sensory channel accompanied by text and/or audio information. In educational films, narratives provide most of the content, creating self-contained "lectures" complemented with visuals (Erdman, 1990).

While there can be useful content in such films, getting students to generate their own questions and hypotheses around video can also be instructive. Video as data assumes that viewers will analyze and interpret the visual information for themselves. More so, video as data provides opportunities for collaboration between teachers and students, as problem solving becomes the central task. For instance, there are learning environments that allow students to analyze properties of motion (Cappo & Darling, 1996; Rubin, Bresnahan, & Ducas, 1996; Rubin & Win, 1994) and kinesiology (Gross, 1998) with digital video. With these, students measure physical phenomena directly from video clips to develop quantitative stories about distance, rate, and time relationships or how muscle movements relate to human actions. Educational television formats can also present contexts for using imagery as data. For instance, *The Adventures of Jasper Woodbury* provide problems to be solved and clues to their solution (Barron et al., 1998), rather than giving away solutions as many educational films do (Ellsworth, 1990).

Similarly, Goldman-Segall's *Learning Constellations* (1997) allows students and teachers to become "multimedia ethnographers" as they study video of their classroom practices. Nardi et al.

(1996) use video to facilitate real-time collaboration during neurosurgery. Medical students use the same video clips after surgery to understand the demands of operating room practices. And collaboration around video also plays a role in Media Fusion, a system integrating digitized video with quantitative data tools (e.g., spreadsheets) for students to back imagery with numerical evidence for problems (Bellamy, 1996).

Investigating and modeling with image data

In all of these applications, students use video as data to conduct authentic inquiry around a problem. We build upon this previous work by stressing the role of photographs and video in our learning environments. Expert practitioners often use imagery in their research as a basis for developing models and theories about phenomena. For instance, behavioral ecologists might use videos of animals to study and analyze behaviors and patterns. Urban planners often use historical images to make decisions about future zoning and construction issues. Students can also engage in this use of imagery as data to develop models and predictions of visual events. Just as students can learn by analyzing and modeling quantitative data (e.g., Jackson, Stratford, Krajcik, & Soloway, 1995; Mandinach & Cline, 1994), we suspect that much can be learned through similar activities with image data.

Students need support to become active observers and investigators of visual data. In particular, if we want them to develop causal explanations and models from visual data, they will need task structures to facilitate the inquiry process. Students can learn a great deal from generating questions and hypotheses for themselves, but they also need to understand what makes a *good* question, a *reasonable* hypothesis. More so, they need to understand how to assemble photographs and video into explanatory models. Our imaging applications are designed to explicitly model expert strategies for investigation (Tabak, Smith, Sandoval, & Reiser, 1996), articulating strategic knowledge needed to explain complex events and processes with image data.

Strategic activity

Assembling a causal story about complex behavior means organizing observational data into coherent structures or models for explanation. It means thinking about the types of actions or events involved in the process and understanding how they influence final outcomes. Through consulting with experts in several domains, we developed a structure for constructing explanatory models from image data. That structure defines four steps that assist in building models from casual observations of still and moving images:

- 1) *Decompose*. Complex processes consist of many constituent, related actions. Interactions between predators and prey, for instance, must progress through stages of detection, stalking, chasing, and finally capturing. The changes in a city's major modes of transport may progress from horses to trains to automobiles. Identifying these components provides the building blocks for the remaining strategic steps.
- 2) *Compare*. It is not enough to analyze a single film or photograph of a complex process. Our students investigate libraries of video and images and compare them to look for similar events. By looking for variations in a routine or across time, students can identify patterns that may prove critical to explaining the process.

- 3) *Identify factors.* Once variations are detected through comparison, students need to perform additional analyses to determine the factors influencing the variance. For example, one might observe that trees are disappearing over time in a collection of urban photographs. To explain why this is the case, it is necessary to look deeper at the images, to identify additional factors that may account for the disappearance (*e.g.*, the number of electrical poles are increasing).
- 4) *Model.* With variations and influencing factors identified, students can generalize causal models that explain the phenomenon under investigation. These may take the form of decision trees explaining the flow of an event or causal chains describing changes over time. Regardless of the form, the modeling step creates an explanatory framework that can be used to predict and design future configurations of the problem space.

This investigation model provides structure for analyzing complex, observable processes, whether that means field observations or observations of imagery. While students are accustomed to looking at photographs and films, they are not necessarily accustomed to making fine-grained observations and explanations with imagery. The investigation model helps them move from raw image data to predictive theories about observable phenomena. We also provide domain-specific heuristics to help students understand the types of questions to ask during their investigations. For instance, in behavioral ecology, asking about costs and benefits of particular behaviors is a good strategy when trying to explain how and why an action has evolved.

Explaining animal behavior with video

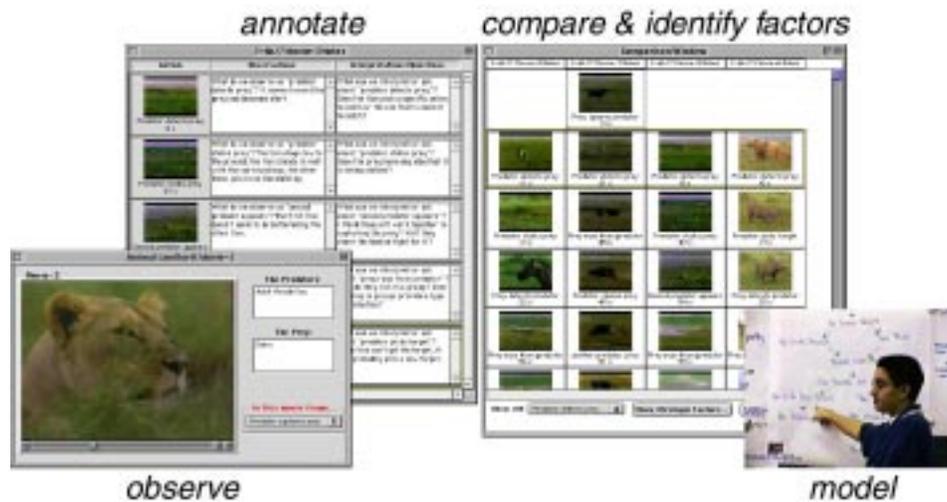
To understand how the investigation model is instantiated in curricular materials, we provide two examples of learning environments that we have developed. The first concerns itself with animal behavior. Nature documentary films are commonly used in biology classrooms to introduce concepts in animal behavior, but they tend to provide descriptive overviews of behavior, neglecting many interesting causal processes in favor of straightforward outcomes. For instance, a film might mention that a creature performs a particular behavior without explaining the complexities of how and why it does so. Narration is the primary source of knowledge in such educational films (Ellsworth, 1990; Erdman, 1990), yet there is a great deal of implicit information in the video that students can observe and explain for themselves.

We developed a video environment, called Animal Landlord, for high school students to investigate the hunting behaviors of the Serengeti lion (Smith & Reiser, 1997; Smith & Reiser, 1998). Only 15-30% of all hunts attempted by lions result in successful capture (Bertram, 1979) and understanding the reasons for this requires investigating the causal interactions between the lion, its prey, and the environment. Students become field researchers with digitized nature films, working to understand how and why lions and their prey interact during the hunt. In conducting their investigations, they explore topics in behavioral ecology such as social organization, resource competition, variation between individuals and species, and environmental pressures.

Animal Landlord provides support for the investigation structure described earlier (Figure 1). Students decompose hunting sequences into important actions using an annotation tool. For each action identified, they also record observations and interpretations that are later used to construct behavioral models. There is also a tool for comparing these collections of annotated videos, allowing students to detect similarities and differences across filmed events. Students use the

comparison tool to identify reasons for variation in the films. For instance, by lining up all actions marked "Predator stalks prey", students can visually inspect all films to see how stalking actions differ (*e.g.*, the type of prey, the amount of ground cover). These factors can then be used to describe evolutionary reasons for the variation.

Figure 1: Activities in Animal Landlord. From left to right, students observe nature films, annotate them with significant events, compare across films to identify variations and factors leading to them, and, finally, produce decision tree models to explain predator-prey interactions.



Students create models of the possible predator-prey interactions that can occur during a hunting episode. They currently do this by creating decision trees on posters using their annotations of the video clips. These posters are displayed around the classroom and whole-class discussions are led by teachers to get students thinking about the evolutionary reasons for the paths through the tree. More importantly, their decision trees are used when viewing nature films after the computer intervention. That is, whenever additional hunting films are shown in class, students use their decision trees to make predictions about the behaviors of the animals and to refine their models if needed. In this way, all nature film viewing after using Animal Landlord becomes a model testing and refinement exercise.

Explaining community change with photographs

Our second example explores urban planning and community change with historical photographs. In most K-12 classrooms, students are exposed to historical issues through the writings and narrative accounts of others. They generally lack primary data sources to complement these writings and allow them to form their own interpretations of the past. By making historical images accessible to students, we hope to develop new ways for them to collaborate and investigate how and why local communities have evolved over time.

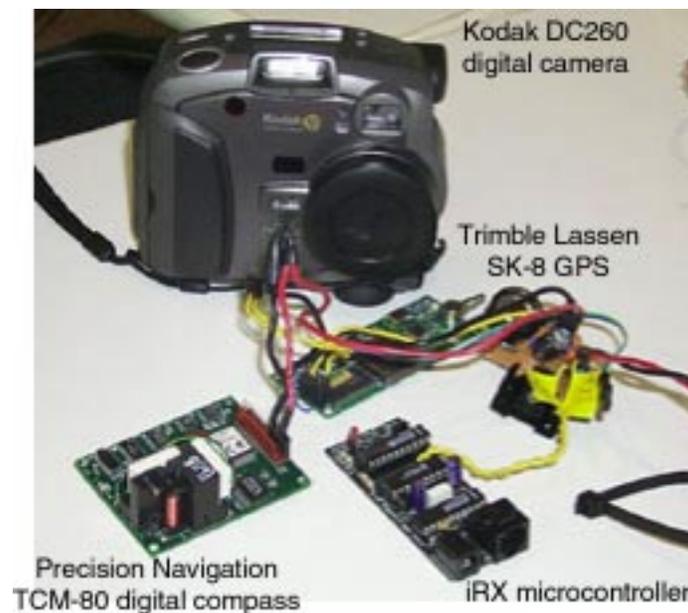
The seemingly ordinary landscapes that we live in contain a great deal of history that goes unnoticed by most residents of a community. When students are taught to explore their outdoor surroundings, they become more aware of the intricacies of man-made environments (Stilgoe, 1998). Not only can they begin to appreciate architectural patterns, they may begin questioning and posing hypotheses about historical and social aspects of their communities. For instance, the

high rent district of Cambridge, Massachusetts still holds evidence of its industrial past, and observant students may begin to wonder when the area shifted to high-technology. A key to answering such a question lies in the historical images of Cambridge.

Camera historica

Careful observation of the present can yield interesting questions, but we also need to provide students with a glimpse of a community's past. The photographs that one finds in a local, historical archive are well suited for developing a story of a community's past, but these are typically difficult for the general public to access. To gain access to these image collections, we send students into their communities to become urban researchers armed with a digital camera augmented with a global positioning system (GPS) and a digital compass (Figure 2). When students take photographs, the position and orientation of the camera are recorded directly into the digital image file. By integrating geographical information systems (GIS) with multimedia (Kraak, 1996; Spohrer, 1998), we can record a geo-referenced trail of students' photographs that can be used for inquiry. And, as opposed to Animal Landlord, students are now responsible for collecting field data as well as producing models from imagery.

Figure 2: An "out of the box" look at the Image Maps hardware. A Kodak DC260 digital camera is attached to a Trimble Lassen SK-8 GPS and a Precision Navigation TCM-80 digital compass.

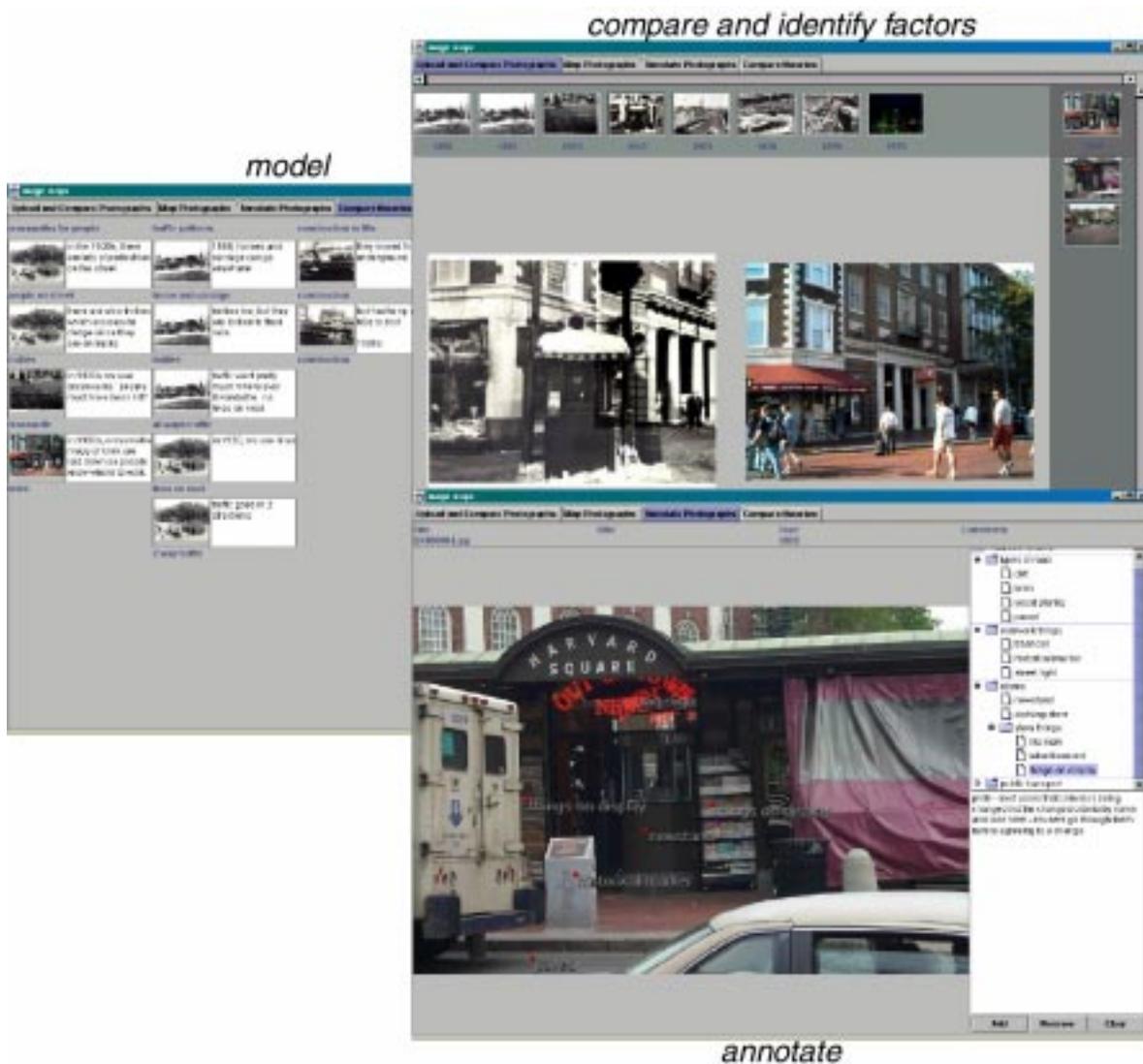


When students return to their classrooms, they download their images into an application we have developed. The program parses each image, extracts the position metadata, and performs a polygonal search (Tsui, 1998) of a Cambridge GIS map to return the name of the photographed building. By identifying the current building, we can retrieve and display historical images of the photographed location (Figure 3). In this way, our camera provides a window into the past: students photograph the present and receive historical images for their investigations of community change.

Once historical images are retrieved and displayed, students can begin annotating and comparing them. More importantly, they can begin to build models of how and why their local communities

have changed over time. The models that they construct are based on the architectural patterns described in (Alexander, Ishikawa, & Silverstein, 1977). A problem/theme is chosen (*e.g.*, "Crosswalks for people"), the context for the problem is described (*e.g.*, pedestrian traffic is conflicting with transportation), and solutions are provided in the form of historical images. In the crosswalk case, students would construct a causal chain illustrating the progression from unmarked pavement to marked crosswalks. After constructing a number of chains, they can return to the field to see how well their generalizations hold up in unexplored parts of the city.

Figure 3: Modeling community change with Image Maps. Students photograph their communities and annotate the images with features that appear to be changing. They compare their images against historical photographs retrieved by our software. They then construct causal, models of patterns in the community.



Conclusion

Both Animal Landlord and Image Maps have similar structures. In both, students use imagery as

data to construct explanatory models of complex processes — the interactions of predators and their prey and the changes in a community's architecture. They also share the same investigation model, the process of annotating, comparing, identifying factors, and creating predictive models to explain the image data. 300+ students in 12 Chicago-area classrooms have used Animal Landlord, and it is being redesigned for a new unit on conservation biology. Our studies of Animal Landlord (Smith & Reiser, 1997; Smith & Reiser, 1998) have provided an understanding of how students can improve their causal reasoning by using imagery as data. The Image Maps project hopes to build on that knowledge, and it will deploy to students in July 1999.

We are working towards a new class of visualization and modeling applications that use imagery as a primary data source for inquiry. Rather than simply looking at photographs or watching videos, we would like to see students arguing and debating over differences in image data. While most scientific visualization tools map quantitative data into visual representations, our students work directly with observational, image data, constructing qualitative models to predict future outcomes and events. Because students often lack an understanding for the importance of modeling (Grosslight, Unger, Jay, & Smith, 1991; Schauble, Glaser, Duschl, Schulze, & John, 1995; Snir, Smith, & Grosslight, 1995), we imagine that the immediacy and concrete qualities of imagery may be an appropriate way to scaffold students into more complex, quantitative modeling tasks. The applications described in this paper represent a first step towards reusing existing photographs and video for inquiry learning and model construction.

Acknowledgments

We would like to thank the Cambridge Historical Commission for their gracious donation of many decades of historical photographs. Michael Baker, Alfred Ashford III, and Timothy Hirzel assisted in the development of the Image Maps hardware and software. Brian J Reiser advised Animal Landlord, while Aggelici Agganis and Pamela Lentine contributed to its development. The James S. McDonnell Foundation and a Patricia Roberts Harris fellowship from the United States Department of Education funded the original implementation of Animal Landlord. The Image Maps are supported by the MIT Media Laboratory's News in the Future Consortium and kind donations from Eastman Kodak.

Bibliography

- Alexander, C., Ishikawa, S., & Silverstein, M. (1977). *A Pattern Language: Towns, Buildings, Construction*. Oxford: Oxford University Press.
- Barron, B.J.S., Schwartz, D.L., Vye, N.J., Moore, A., Petrosino, A., Zech, L., Bransford, J.D., & the Cognition and Technology Group at Vanderbilt. (1998). Doing with understanding: Lessons from research on problem- and project-based learning. *The Journal of the Learning Sciences*, 7(3&4), 271-311.
- Bellamy, R.K.E. (1996). Designing educational technology: Computer-mediated change. In B.A. Nardi (Ed.), *Context and Consciousness: Activity Theory and Human-Computer Interaction* (pp. 123-146). Cambridge, MA: The MIT Press.
- Bertram, B.C.R. (1979). Serengeti predators and their social systems. In A.R.E. Sinclair & M. Norton-Griffiths (Eds.), *Serengeti: Dynamics of an Ecosystem* (pp. 223-248). Chicago:

University of Chicago Press.

Bransford, J.D., Franks, J.J., Vye, N.J., & Sherwood, R.D. (1989). New approaches to instruction: Because wisdom can't be told. In S. Vosniadou & A. Ortony (Eds.), *Similarity and Analogical Reasoning* (pp. 470-497). New York: Cambridge University Press.

Cappo, M. & Darling, K. (1996). Measurement in Motion. *Communications of the ACM*, 39(8), 91-93.

Collins, A. & Brown, J.S. (1988). The computer as a tool for learning through reflection. In H. Mandl & A. Lesgold (Eds.), *Learning Issues for Intelligent Tutoring Systems* (pp. 1-18). New York: Springer-Verlag.

Ellsworth, E. (1990). Educational films against critical pedagogy. In E. Ellsworth & M.H. Whatley (Eds.), *The Ideology of Images in Educational Media: Hidden Curriculums in the Classroom* (pp. 10-26). New York: Teachers College Press.

Erdman, B. (1990). The closely guided viewer: Form, style, and teaching in the educational film. In E. Ellsworth & M.H. Whatley (Eds.), *The Ideology of Images in Educational Media: Hidden Curriculums in the Classroom* (pp. 27-42). New York: Teachers College Press.

Goldman-Segall, R. (1997). *Points of Viewing Children's Thinking: A Digital Ethnographer's Journey*. Mahwah, NJ: Lawrence Erlbaum Associates.

Gordin, D.N. & Pea, R.D. (1995). Prospects for scientific visualization as an educational technology. *Journal of the Learning Sciences*, 4(3), 249-279.

Gross, M.M. (1998). Analysis of human movement using digital video. *Journal of Educational Multimedia and Hypermedia*, 7(4), 375-395.

Grosslight, L., Unger, C., Jay, E., & Smith, C.L. (1991). Understanding models and their use in science: Conceptions of middle and high school students and experts. *Journal of Research in Science Teaching*, 28(9), 799-822.

Jackson, S.L., Stratford, S.J., Krajcik, J., & Soloway, E. (1995). Making systems dynamics modeling accessible to pre-college science students. Paper presented at the *Annual Meeting of the National Association on Research on Science Teaching*. San Francisco, CA.

Kraak, M.-J. (1996). Integrating multimedia in geographical information systems. *IEEE Multimedia*, 3(2), 59-65.

Mandinach, E.B. & Cline, H.F. (1994). *Classroom Dynamics: Implementing a Technology-Based Learning Environment*. Mahwah, NJ: Lawrence Erlbaum Associates.

Nardi, B.A., Kuchinsky, A., Whittaker, S., Leichner, R., & Schwarz, H. (1996). Video-as-data: Technical and social aspects of a collaborative multimedia application. *Computer Supported Collaborative Work*, 4, 73-100.

Rubin, A., Bresnahan, S., & Ducas, T. (1996). Cartwheeling through CamMotion. *Communications of the ACM*, 39(8), 84-85.

- Rubin, A. & Win, D. (1994). Studying motion with KidVid: A data collection and analysis tool for digitized video. In *Conference Companion to CHI '94* (pp. 13-14), New York: ACM Press.
- Schauble, L., Glaser, R., Duschl, R.A., Schulze, S., & John, J. (1995). Students' understanding of the objectives and procedures of experimentation in the science classroom. *The Journal of the Learning Sciences*, 4(2), 131-166.
- Smith, B.K. & Reiser, B.J. (1997). What should a wildebeest say? Interactive nature films for high school classrooms. In *ACM Multimedia 97 Proceedings* (pp. 193-201), New York: ACM Press.
- Smith, B.K. & Reiser, B.J. (1998). National Geographic unplugged: Classroom-centered design of interactive nature films. In *Proceedings of CHI 98* (pp. 424-431), New York: ACM Press.
- Snir, J., Smith, C., & Grosslight, L. (1995). Conceptually enhanced simulations: A computer tool for science teaching. In D.N. Perkins, J.L. Schwartz, M.M. West, & M.S. Wiske (Eds.), *Software Goes to School: Teaching for Understanding with New Technologies* (pp. 106-129). New York: Oxford University Press.
- Spohrer, J. (1998). *Worldboard: What comes after the WWW?* Available: <http://www.worldboard.org/pub/spohrer/wbconcept/default.html>.
- Stilgoe, J.R. (1998). *Outside Lies Magic: Regaining History and Awareness in Everyday Places*. New York: Walker and Company.
- Tabak, I., Smith, B.K., Sandoval, W.A., & Reiser, B.J. (1996). Combining general and domain-specific strategic support for biological inquiry. In *Proceedings of the Third International Conference on Intelligent Tutoring Systems* (pp. 288-296), London: Springer Verlag.
- Tsui, C. (1998). *Multimedia Data Integration and Retrieval in Planning Support Systems*. Unpublished M.S. thesis, Massachusetts Institute of Technology, Cambridge, MA.
- Wetzel, C.D., Radtke, P.H., & Stern, H.W. (1994). *Instructional Effectiveness of Video Media*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Whittaker, S. & O'Connell, B. (1997). The role of vision in face-to-face and mediated communication. In K.E. Finn, A.J. Sellen, & S.B. Wilbur (Eds.), *Video-Mediated Communication* (pp. 23-49). Hillsdale, NJ: Lawrence Erlbaum Associates.

Authors' addresses

Brian K Smith (bsmith@media.mit.edu)

MIT Media Laboratory; 20 Ames Street; Cambridge, MA 02139; Tel. (617) 253 6537. Fax (617) 258-7444.

Erik Blankinship (erikb@media.mit.edu)

MIT Media Laboratory; 20 Ames Street; Cambridge, MA 02139; Tel. (617) 253 1401. Fax (617) 258-7444.