The Potential for Scientific Collaboration in Virtual Ecosystems

Brian Magerko
School of Literature, Communication and Culture
Georgia Institute of Technology
magerko@gatech.edu

This paper explores the potential benefits of creating *virtual ecosystems* from real world data. These ecosystems are intended to be realistic virtual representations of environments that may be costly or difficult to access in person. They can be constructed as 3d worlds rendered from stereo video data, augmented with scientific data, then deployed online for use. The application of virtual ecosystems stretches from interdisciplinary scientific research that may not occur otherwise to providing science students with an environment to conduct studies and virtual field trips in that they would otherwise not have access to.

Introduction

There are significant meta-level issues in scientific research that have beset scientists since the founding of their fields. It is common to find in specializations of scientific fields that information sharing and collaboration decreases as fields specialize. For example, it is unlikely that artificial intelligence researchers are aware of the work that their colleagues in VLSI design are doing, despite being in the same broad field of Electrical Engineering and Computer Science. More significantly, scientists often depend on expensive equipment (e.g. using a particle collider) or costly site visits (e.g. visiting a remote Pacific island) to conduct their research. These problems exist at the cost of potential collaborations and discoveries that could be made by individuals within and across scientific disciplines.

Virtual environments (VEs) have the potential for addressing the issues described above. VEs have been explored as tools for conducting ethnographic research, economics studies, and scientific experiments with virtual equipment (Bainbridge, 2007). VEs have several features that make them enticing for such academic work. First, VEs are relatively *inexpensive*. Once the time and work has been put into building an environment, only maintenance work need be put in. Contrast this cost to the ongoing funds put into individual researchers or labs funding their own excursions or equipment, and the savings in cost (for our society) over time becomes evident. Second, virtual environments that provide tools for users to *contribute* their own content (e.g. *Second Life* users can build their own persistent objects in the world) can add content without significant cost to developers. Third, VEs provide a medium that allows for the significant presentation of scientific data *in situ*. As opposed to presenting content in decontextualized forms (e.g. in spreadsheets), scientists can use VEs to directly associate the presentation of data to the phenomenon they are studying. The field of scientific visualization exists to explore the potential of representing data in a visual, and often contextualized, form (McCormick et al., 1988).

Our observation that VEs can be used to visualize scientific data in situ leads us to consider that there are specific fields, such as environmental or ocean sciences, where they visually observe and collect data from real 3D environments, such as coral reef systems, wetlands, or forests. We hypothesize that VEs are an appropriate fit for the aforementioned research meta-level issues within the field of environmental sciences. The development of *virtual ecosystems* (virtual environments that represent real ecosystems) for scientific discovery and collaboration has the potential to create powerful research tools for environmental scientists that both mitigate the high cost of field work and encourage interdisciplinary work in related fields that are otherwise disparate.

Virtual Ecosystems

Virtual ecosystems that are created from real world data to represent an actual reality-based ecosystem present a powerful application for scientific research within virtual worlds. While other research studies human behaviors in existing virtual environments, the creation of virtual ecosystems allows researchers (or students) to study the environment itself. Scientists can therefore visually and spatially examine a space of interest without the cost of transporting themselves and their equipment to that space.

The question of how to build a digital copy of a real environment is a major issue with the creation of virtual ecosystems. 3D artists could effectively do the work, but that is an incredibly large, expensive, and timely approach. An alternative is to visit the environment, collect data on that environment, and then use that data to automatically (or semi-automatically) reproduce the environment on a computer. Our research group's process of creating a virtual ecosystem based on real world 3D data involves processing the captured data using a class of algorithms to build a 3D geometric model called *Structure from Motion* or *SFM* (Ni, Steedly and Dellaert, 2007; Steedly, Essa and Dellaert, 2003; Dellaert et. al, 2000). These algorithms take either a video sequence or a collection of images and produce a 3D model of the scene. Other approaches, such as image-based rendering and video-rendering methods (Schödl et al., 2000; Schödl and Essa, 2002), can be used to then render the data into a usable 3D environment.

Research has shown that the fidelity of a digital experience is not necessarily positively correlated with the desired effects. For example, the *uncanny valley* (e.g. MacDorman, 2006) is well known as a point at which the increasing fidelity of a digital character becomes negatively correlated with the believability of that character. "Very lifelike " might be much more disturbing than a more abstract representation. In terms of virtual environments, studies have shown that there may be a negative correlation with high fidelity experiences and learning environments. Mayer asserts "…people learn more deeply from a multimedia message when extraneous material is excluded rather than included," (p. 184, 2005. However, this is unlikely to be an issue with virtual ecosystems since scientists commonly conduct research in situ where the fidelity of the experience is higher than we could hope to capture with current technologies.

We are currently creating a virtual coral reef ecosystem based on video data collected at the Andros Barrier Reef. A scuba diver collected data by making crisscross traversals across an area of interest with two high definition video cameras mounted on a frame. The resulting VE will allow researchers to place their own work, and that of others, into a large-scale ecosystem framework to aid in the interpretation of their results (thus freeing them from an over-dependence on results gathered over a handful of research dives, for example). Furthermore, there are many aspects of reef health that can be studied using only the virtual environment itself, including species diversity, coral colony size, algal cover, degree of bleaching, etc., at scales that are representative of an entire reef system. The VE will provide a way to collate, index, and contextualize a wealth of interdisciplinary scientific data that would be difficult to glean from the scientific literature. For example, a biologist would certainly benefit from temperature logging data collected by a climate scientist some years earlier at her research site. Likewise, the climate scientist might be able to use the biologist's salinity data to look at current patterns on the reef. Many such datasets go unpublished, yet would be perfect candidates for a virtual coral reef laboratory.

Previous work has been done on Collaborative Virtual Environments (CVEs) (e.g. Barab et al., in press). However, little work has been done on how realistic environments can be used for work that incorporates scientific data as a key element for collaboration and research. We are creating an environment that not only provides features that are common to CVEs, such as *embodiment* (i.e. representations of users as animated avatars), *virtual spaces*, and *group telecommunication*, but also allows group interactions to take place in a realistic virtual ecosystem that contains relevant scientific data that can be visualized and manipulated by researchers.

The visualization of scientific data (e.g. ocean currents, temperature change, acidification, fauna populations, etc.) within a virtual ecosystem has three main purposes. The first is to provide access to relevant local and global data within the space being studied. With the ability to visualize, compare, manipulate, and annotate data within the target ecosystem, scientists will have a powerful tool for online scientific collaboration and experimentation. This would help facilitate online scientific discovery and collaboration in a space that is either remote or no longer in existence. A strong consequence of this approach is that virtual spaces could be used as a portal for scientific data that visualizes spaces that do not exist anymore. The second purpose would be to present data that would be of interest to science students (which is discussed further below), making them aware of the key scientific issues as it relates to natural and human ecosystems. The third is to foster new ways of thinking about complex systems by providing access to both local and global environmental data in a contextualized manner (e.g. of the kind of relationships and data relevant to the health of coral reefs, from everyday human activity to specific health issues that reefs have). We contend that by bringing all of this knowledge to bear within the context of the target environment, scientists and students will be able to directly reason about the relationship between far away effects (e.g. human electricity use) and the health of the ecosystem.

Future Work and Issues

Once the virtual ecosystem is created, we will evaluate whether we should use currently available 3D game technology or create our own custom application to allow scientists to remotely log in and explore the space with a virtual avatar. Users will be able to contribute their own data in an XML-defined format, browse and select different data sets, and chat with other users online. A key feature will be to visualize data in both a private (only that user can see it) and public (everyone can see it) manner so scientists can do work on their own while others are online as well as work as a group. However, we will not understand how users should be able to interact with each other and the data within a group setting until we can do an evaluative study of a completed prototype.

An open question regarding the practical use of VEs for scientific research of this nature is related to the buy-in needed from researchers to contribute data. Consistent use of data repositories has had a very mixed history across academic fields; therefore, ease of use and visibility will be crucial to the success of using virtual environments for the sharing of scientific data. Organizing conferences to be held in the space and requiring a submission of data is one feasible approach to increasing the visibility and relevance of an environment. An alternative approach is for the maintainers of a virtual ecosystem to coordinate with research groups to encourage them to share their data, even if only in raw form, for uploading.

The process of data collection is also an additional concern. Whether it is a land-based or marine ecosystem, there is a difficult in collecting a large amount of data over a space using handheld cameras. Future work will explore the possibility of using multiple robots to swarm over an area, greatly reducing the time cost in data collection.

Aside from applications in scientific collaboration, an accessible, data-rich virtual ecosystem could potentially also be used as an educational tool for higher education science students. For example, Barab and others (in press) have already begun using virtual cultural heritage sites as platforms for educational computer games. Students could log in to a virtual ecosystem and virtually visit a locale that they would likely never visit on a field trip. Students, with guidance from their teacher, could have access to the same data sets used by scientists and can conduct experiments by manipulating that data. Additional development would need to be done to provide students with the appropriate interface to the environment, tools to operate on the data, etc.; however, the potential application has definite potential. Our future work will focus on developing a more structured, game-based curriculum incorporated into an educational version of the application that students can make use of with or without a teacher's guidance (Magerko et al., 2006). As mentioned earlier, there is the question as to whether or not the high fidelity nature of what is discussed here would have a negative impact on learning. However, the focus of the presentation needs to be on the scientific-related data (e.g. the visualization of the environment and related scientific data) with a subsequent de-emphasis on unrelated features of the learning experience to address the earlier mentioned concerns about the possible negative effect of high fidelity virtual experiences on learning (e.g. do not worry about making a hyper realistic model of how to navigate underwater since we are not concerned with teaching about how to scuba dive).

Acknowledgements
Thanks to my collaborators , Dr. Kim Cobb, Dr. Frank Dellaert, Tanyoung Kim, and Carlos Nieto for their efforts on collecting data at Andros and to the GVU Center at Georgia Tech for funding this work.

References

Bainbridge, W. 2007. The Scientific Research Potential of Virtual Worlds. *Science* 317: pp. 472-476.

Barab, S.A., M. Thomas, T. Dodge, R. Carteaux, and H. Tuzun. 2005. Making learning fun: Quest Atlantis, a game without guns. *Educational Technology Research and Development* 53(1): 86-108.

Dellaert, F., S.M. Seitz, C.E. Thorpe, and S. Thrun. 2000. Structure from Motion without Correspondence. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*.

MacDorman, K. F. 2006. Subjective ratings of robot video clips for human likeness, familiarity, and eeriness: An exploration of the uncanny valley. In Proceedings of the ICCS/CogSci-2006 Long Symposium: Toward Social Mechanisms of Android Science, Vancouver.

Magerko, B., B. Stensrud, and L. Holt. 2006. Bringing the Schoolhouse Inside the Box – A Tool for Engaging, Individualized Training. In Proceedings of the 25th Army Science Conference. Orlando, FL.

Mayer, R.E. (2005). *The Cambridge Handbook of Multimedia Learning*. New York: Cambridge University Press.

McCormick, B.H., T.A. DeFanti, and M.D. Brown. 1988. Visualization in Scientific Computing. *ACM SIGBIO Newsletter* 10(1).

Ni, K., D. Steedly, and F. Dellaert. 2007. Out-of-Core Bundle Adjustment for Large-Scale 3D Reconstruction. In Proceedings of the IEEE International Conference on Computer Vision (ICCV).

Schödl, A., R. Szeliski, D.H. Salesin, and I. Essa. 2000. Video textures. In Proceedings of the 27th Annual Conference on Computer Graphics and interactive Techniques International Conference on Computer Graphics and Interactive Techniques.

Schödl, A. and I. Essa. 2002, Controlled animation of video sprites. *In Proceedings of the 2002 ACM SIGGRAPH/Eurographics symposium on Computer animation*, July.

Steedly, D., I. Essa, and F. Dellaert, 2003. Spectral Partitioning for Structure From Motion. In Proceedings of the IEEE International Conference on Computer Vision (ICCV).