

# Interacting with multimedia: 3D navigation and media collections

Ethan V. Munson

Published online: 6 April 2014  
© Springer Science+Business Media New York 2014

## 1 Editorial process

Papers were solicited via an open call for papers and a number of submissions were received from research groups active in the area of interactive multimedia. Papers were rigorously reviewed on a double-blind basis by researchers recognized in the field. Each accepted paper was reviewed multiple times by multiple reviewers and six were selected for inclusion in this special issue.

## 2 Papers on interacting with multimedia

The six accepted papers fall into two broad categories, each with three papers: 3D graphics navigation and multimedia collection organization.

### 2.1 3D graphics navigation

Trindade and Raposo's paper is titled "Improving 3D Navigation Techniques in Multiscale Environments: A Cubemap-based Approach." It presents three novel techniques for navigation in multiscale 3D virtual worlds, such as may be used in visualization of large scientific or engineering data sets. All are based on the use of the Cubemap [4] approach to provide a multiscale grid for orientation to location and scale and allows users to interact at different granularities of action. Trindade and Raposo provide automatic speed adjustment when *flying* in the virtual world, plus automated collision detection so that users avoid objects while flying. Finally, they automatically position the pivot point for camera rotations in order to avoid unintuitive rotation behaviors associated with pivot points outside the field of view or behind the object of interest.

A second paper addressing the 3D navigation domain is "Efficient Visualisation of 3D Models on Hardware-Limited Portable Devices" by Ramos, Ripolles, and Chover. Like the previous paper, they are interested in multi-resolution graphics representations, but for this paper, the focus is on mobile devices with limited computational and memory resources. Their

---

E. V. Munson (✉)

Department of EECS, University of Wisconsin – Milwaukee, P.O. Box 784, Milwaukee, WI 53201, USA  
e-mail: munson@uwm.edu

approach can generate a variety of mesh resolutions in order to provide a level of detail suitable for the computational resources of the user's device. Their approach merges triangles using the half-edge collapse technique [3] in order to avoid increasing memory requirements. They describe an implementation for mobile devices based on OpenGL ES [2].

A quite different problem is considered by Ripolles, Simó, Vivó, and Benet in their article titled "Smart Video Sensors for 3D Scene Reconstruction of Large Infrastructures." They are interested in providing remote observers with coherent information from a set of multiple video cameras placed at various fixed points in a large environment. Their project focuses on issues that might arise in a large airport, such as following people and their luggage as they move through the airport, but the system was tested in a university setting. Their system uses cameras with embedded processing that can identify objects in the video images and provide appropriate metadata about the objects to a central server. The server is then responsible for reconstructing the information from the cameras into a coherent whole and can track people as they move from one location to another and automatically change point of view for a human monitoring the security system.

## 2.2 Organizing multimedia collections

In the area of multimedia collection organization, Kusama and Itoh describe the MusCat system in the paper titled "Abstract Picture Generation and Zooming User Interface for Intuitive Music Browsing." Their goal is to provide an interface to a music collection that automatically organizes music files into clusters and provides a useful visualization via which users can choose files matching their current listening goals. To meet this goal, they first generated profiles of music files by analyzing five signal features (e.g. tempo, RMS energy) of a randomly chosen 15-s segment of each file. Then, they generated an abstract image for each file based on these five features. For example, the level of RMS energy determined the gradation of the color background of the abstract image, while tempo mapped to the number of circles in the image. Their user interface organized files into clusters of similar files. They describe a user test that provided a number of insights that can motivate future versions of MusCat.

Ryu and Cho have similar goals for the research described in "A Summarized Photo Visualization System with Maximal Clique Finding Algorithm." Instead of music files, they want to automatically organize digital photo collections, which are widely considered to be difficult to organize manually due to both the volume of images and the difficulty of organizing them meaningfully. A key insight of Ryu and Cho is that personal photo collections often contain many Nearly Identical Photos (NIPs) since people take multiple photos in rapid succession in order to get the best view or pose of their subjects. Their interface presents NIPs by overlapping them, since they convey similar information. Photos are then clustered based on estimates of quality from blur and depth of field analysis.

The paper titled "An Experimental Evaluation of Ontology-Based User Profiles" by Hopfgartner and Jose looks at improving video recommendation through the use of ontologies. Their system observes user preferences for news videos over time. Based on the users' viewing choices are matched to ontologic information drawn from the DBpedia representation of Wikipedia articles [1] in order to construct a user profile that can be used for later content-based recommendations. They validate their results via a relatively long-term user test involving 10 days of system use by each subject. Subjects consistently preferred the ontology-based recommender system. Hopfgartner and Jose's research shows that an ontology-based approach can be used to construct long-term user profiles for suitable classes of media data and that the ontology-based approach applies to video information in addition to textual information.

### 3 Final thoughts

Given the polished applications that many of us use everyday to interact with multimedia, it would be easy to think that there is little room for advanced research in the area. But the papers in this special issue show that this is not true. First, they remind us that multimedia artifacts and environments are so varied, so large, and so numerous that they are hard for users to organize and integrate. Second, they have features that are either hard to understand (e.g. metrics of music, video or photos) or that impose constraints on our interaction that are unintuitive (e.g. pivot points for camera rotation). Thus, research is still needed to find better ways to interact and this special issue presents some steps in that direction.

### References

1. Auer S, Bizer C, Kobilarov G, Lehmann J, Cyganiak R, Ives ZG (2007) DBpedia: a nucleus for a web of open data. In: Proc. 6th Int. Semantic Web Conf. Springer Berlin/Heidelberg, 11, pp 722–735
2. Baek N, Lee H (2012) OpenGL ES 1.1 implementation based on OpenGL. *Multimedia Tools Appl* 57:669–685
3. Luebke D, Reddy M, Cohen J, Varshney A, Watson B, Huebner R (2003) *Level of detail for 3D graphics*. Morgan-Kaufmann, Inc, San Francisco
4. McCrae J, Mordatch I, Glueck M, Khan A (2009) Multiscale 3D navigation. In: I3D'09: Proceedings of the 2009 Symposium on Interactive 3D graphics and games, ACM, New York, NY, USA, pp 7–14, doi:[10.1145/1507149.1507151](https://doi.org/10.1145/1507149.1507151)



**Ethan V. Munson** is Professor and Co-Chair of Computer Science in the Department of Electrical Engineering and Computer Science at the University of Wisconsin-Milwaukee, where he is also the Director of the Multimedia Software Laboratory. He received the M.S. (1989) and Ph.D. (1994) in Computer Science from the University of California, Berkeley.

His research focuses on configuration management for software and documents and on human-computer interaction. Dr. Munson is a recipient of a National Science Foundation CAREER award, as well as four NSF educational grants, and a variety of industrial funding. He is a Senior Member of ACM and a member of the Brazilian Computer Society (SBC). He was Chair of ACM SIGWEB from 2006 to 2011.