Key Aspects in 3D File Format Conversions

KENTON MCHENRY AND PETER BAJCSY

Abstract: This presentation will address the problem of understanding 3D file format conversions. The motivation of our work is to provide (a) basic understanding of 3D file format conversions and (b) web-based conversion services to assist archivists when confronted with 3D information stored in more than 140 different 3D file formats.

The problem has been approached by (a) automatically finding multiple conversion paths, (b) performing conversions on batches of 3D files, and (c) estimating information loss over file format and conversion path. A conversion path consists of a sequence of software packages used for importing and exporting files from the source file format to the target file format. Moving from one format to another sometimes involves dropping information and/or converting the data representation itself. In addition, software vendors have unique implementations of file importer/exporters which must be considered when converting between formats.

In order to measure the change in the content of 3D files, we introduce a directed graph of conversions called an I/O Graph. This data structure captures the inputs and outputs from various software packages and drives an extensible web-based conversion system called NCSA Polyglot. Polyglot calls the relevant third party packages in order to perform a conversion from a source format to a target format according to the shortest path found in the I/O Graph. Given a set of 3D files we can then assign numerical values to conversion paths based on comparisons of the original and resulting content.

About the authors:

Dr. Kenton McHenry is with the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, working as a research programmer on problems related to 3D content creation, conversion and preservation. Dr. McHenry's research interests include computer vision, pattern recognition and automation.

Peter Bajcsy's research focuses (a) on building bridges from raw data to information and to knowledge where the raw data come from ubiquitous multi-instrument measurement systems, and (b) on understanding computational and algorithmic challenges for automated data-centric operations. Peter Bajcsy has authored more than 16 papers in peer reviewed journals that have been cited more than 200 times on scholar.google.com, and co-authored six books and more than 75 conference papers. His research could also be described as X-informatics, where the X stands for document, hydro, geo, bio, medical image, or sensor.