Parallel Realities? The Requirements of Web3D and Immersive VR

Nicholas F. Polys

Abstract— This workshop intends create shared undersatnding between the VR and Web3D Communities. While the two fields of Web 3D graphics and VR 3D graphics have tended to view their technology as different beasts, they have many of the same requirements in terms of virtual world content - modeling, lighting, animation, and interaction/direct manipulation. Indeed, commodity hardware and open-source software advances enable a spectrum of immersion- not an all-or-nothing proposition or 'just for national labs only'.

This workshop will highlight common interests in seamless access to network-based VR content and resources. We will look at the philosophies and feature sets of various scenegraph tools including ISO-based scenegraph standards such as X3D/VRML in immersive contexts. For example, what are common requirements for a data archive format that provides expressive runtime or application behavior? Through multiple tool and application examples, we will attempt to illuminate the commonalities and differences between the communities and develop a common roadmap of education and effort.

Index Terms— Computer Graphics, Data Storage Representations, Distributed Systems, Reusable Software

1 Introduction

THE delivery of interactive 3D content has undergone significant transformation in the last decade. Indeed, the pace of transformation seems to be accelerating recently. The rate of this change can be attributed to a number of factors:

- Increase in hardware speeds & rendering capabilities
- Decrease in hardware costs
- Increase in accessible network bandwidth
- Increase in capable Open Source tools
- Increase in capability of Standards representations
- Increased deployment of applications in industry
  - Visual Data Analysis for science and engineering applications
  - Education
  - Games

A significant result of this change is the interactive framerates and rich, networked graphics provided by commodity platforms. What was once an expensive immersive VR system can now be contructed with off-the-shelf parts. Immersion is no longer 'just for national labs only'.

In addition to the affordability and speed of hardware plaftorms, a number of capable toolkits have been released as Open-Source software. This has provided the developer community with the means to innovate features and expand deployment across many domains. With the recent demise of large VR companies, many organizations now see the value of keeping their data and applications vendor-independent.

The applications are demonstrated and the hardware is capable, so why is building 3D content so expensive and time-consuming?

One clear answer to this question is the fragmented nature of the data and software technology underpinnings of VR. Interactive virtual environments use components built on top of the standard graphics libraries such as OpenGL. These components include geometry and data loaders, scengraphs and an event model for runtime semantics. The numerous toolkits vary in their repesentation and evaluation of the scenegraph as well as the functionality that is exposed to programmers through its API.

This is a challenge not only for application developers to know all the nuances of a particular codebase, but it also inhibits reuse of virtual world content across the codebases. There is no doubt that these various tools may be unique in providing a particularly needed feature and developers should always apply the correct tool for the job. However, there is not much economy in re-inventing the wheel (or its rotational animation around an axis) for every toolkit.

It is my claim that from a scenegraph perspective, immersive virtual environments and Web3D virtual environments share at least 80% of the same requirements if not more. I will present the latest Web3D standards as a capable common denominator for storing, transmitting, and rendering virtual environments across platforms.

- *Nicholas F. Polys is with the Research Computing Group at Virginia Tech. He is an X3D Specification Co-author, Co-chair of the Web3D Consortium's User Interface Working Group, and Web3D representative to the DICOM standards committees. E-mail: npolys@vt.edu.*

2 Parallel Realities?

Market circumstances are most likely the cause of divergence. In the 1990s and early 00s, immersive systems were in an entirely different league than the machines available to the web consumer. Therefore the two technologies pursued separate priorities and paths. From today's vantage point, perhaps we can see that both systems require some common virtual world content. For example, 3D and 2D models, lighting, animation, and interaction with direct manipulation.

2.1 Immersive VR

Immersion can be defined as some technological aspect of sensory coupling with a virtual environment. As such, virtual environments may be rendered and experienced with varying 'Degrees of Immersion' such a viewing surround, stereoscopy, head-tracking, spatialized audio, or haptic feedback [e.g. 1,2,3].

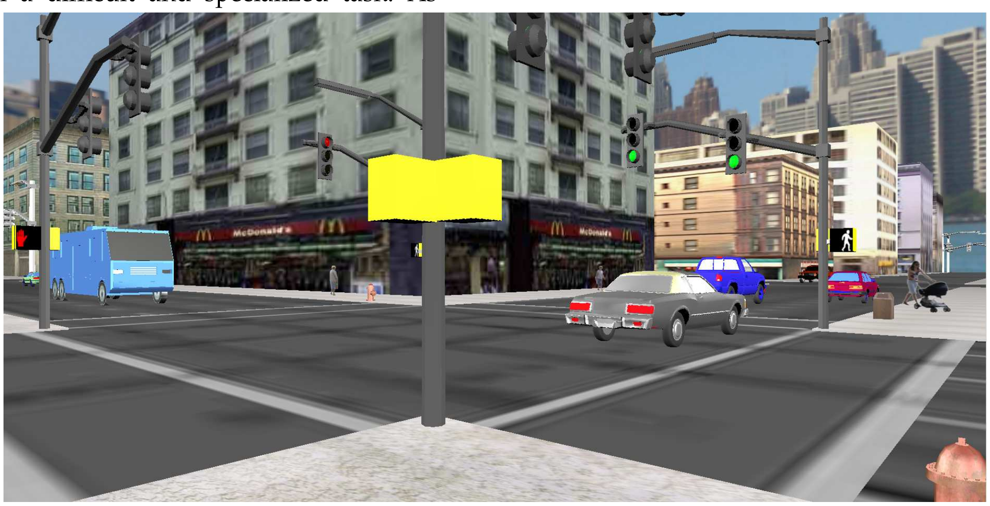
There has been significant research accomplished in the field as to the nature of perception, cognition, and interaction in these synthetic environments. Of course, much more remains unknown. Due to the research orientation of many of these systems, building content and applications is still a difficult and specialized task. As more commodity input and output devices become available to the mass audience, the principles from immersive VR and 3D User Interface must be put in practice.

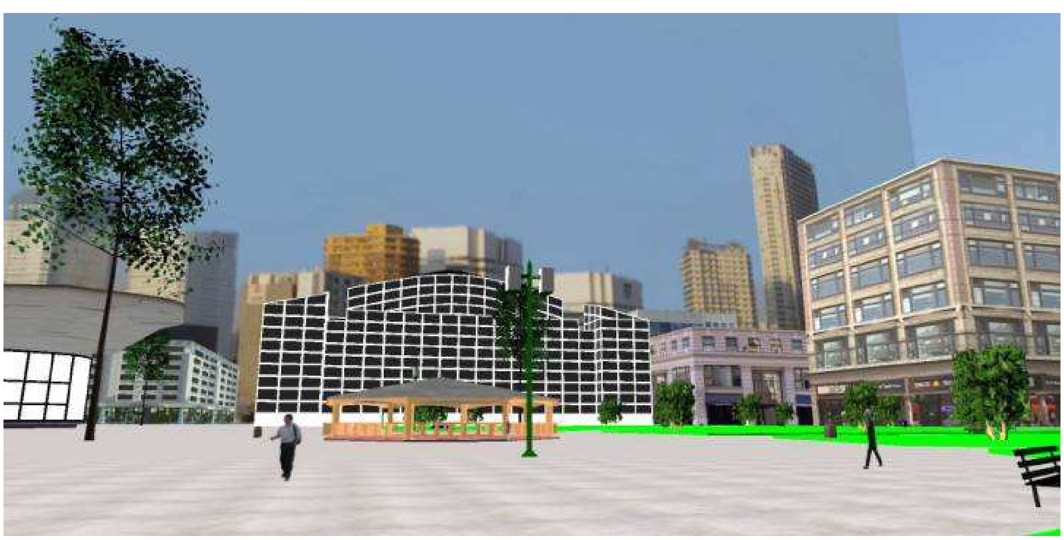
2.2 Web3D Standards

Web3D is a term referring to open, international media standards or recognized practices of the Web3D Consortium such as Virtual Reality Modeling Language (VRML), GeoVRML, Humanoid Animation (H-Anim), and Extensible 3D (X3D) [4]. These standards are developed by the Web3D Consortium in the context of WWW technologies and eventually ratified as International Standards Organization (ISO) standards. The standards are royalty-free and freely published.

Web-based standards for interactive 3D content provide data durability, interoperability, integration. As hardware speeds have improved, interactive virtual environments written in VRML back in 1997 are running faster and smoother than ever. Few, if any, non-standard media types have demonstrated this kind of longevity.

Because Web3D standards describe scenegraph behavior in an abstract way (with no assumptions about the underlying implementation), these worlds can run with





Figures 1 and 2: VRML scenegraphs used across CAVE (w/ stereo + headtracking + wand) and desktop (w/ mouse) displays consistent semantics across multiple platforms while marshalling and connecting data and media resources from across the web. As more commodity input and output devices become available to the mass audience, the standards specifications must break from the desktop legacy.

3 Requirements and Value

Integration of existing information and reuse of virtual world content can provide significant savings in the development and deployment of VR applications. Consider a typical authoring scenario: Arthur is building a cultural heritage application for his home city, which includes geospatial, architectural, historical, and commercial information. All this information can be collected and linked with or within the context of a virtual environment.

Some textured models such as buildings around the main plaza may already exist from a previous project; some may have to be modeled. Some geodata and facts exist on the web. In building the VR application, Arthur wants to reuse and much virtual world content as possible… but how many formats and translations are required? How much information is lost or inaccessible?

Whether this virtual environment is viewed on the desktop or in some immersive venue, the model of the city and buildings and their appearances should be the same. It is not only the objects of the virtual world, but their behaviors as well. For example, when a user 'touches', 'selects' or 'clicks' a doorknob, the door should open, animating around its hinges; a light switch in the model should turn on and off a light.

The value in a standard representation is in maintaining and deploying virtual environment information without translation or recompiling for each platform. In addition, the integration with existing media on the WWW is of great value. Ideally, media resources such as other objects, worlds, textures, movies, audio, and links can be combined and presented in conjunction with a platform-independent scengraph.

One could imagine many similar scenarios where reuse and integration of content is a value. For example, more efficiency in the production and management of virtual environment assets would benefit simulation, visual analysis, training, education, games, and HCI research just to name a few.

Let's consider another authoring scenario briefly. This time a grad student needs to put together a virtual environment to test some wayfinding aids. The usability evaluation will be presented on an HMD and in the CAVE. During the development process, the models and the environment must be inspected by collaborators at another university where they only have desktops. To collaborate on the environment's design, the environment must be transportable across these platforms. Using freely available models on the web, the grad student builds two, rich, standards-based environments in the space of a few weeks and posts them for the collaborators. They are later run in the immersive venues (e.g. Figures 1 and 2).

I consider the common requirements between Immersive VR and Web3D to be:

- Real-time 3D scene graph
- Meshes, lights, materials, textures, shaders
- Integrated video, audio
- Animation
- Interaction
- Behaviors
- Scripts
- Application programming interfaces

While its detractors are many, if not ill-informed, VRML can be considered a successful international standard. Indeed it is currently used in industry and academia all around the world. People continue to innovate on top of the language and the 'old' content still runs. Still, VRML has limitations that are very real. It is a large, monolithic standard without some important functionality and with some underspecified behaviors.

In order to improve on VRML's shortcomings, some major re-thinking had to be done. X3D's design goals (circa 2000) capture the new philosophy behind the standard:

- Support new graphical, behavioral, and interactive types
- Separate runtime architecture from data encoding
- Provide robust application programmer interfaces (APIs) into the 3D scene
- Modularize the architecture (Lightweight core + add-on components)
- Define subsets of the specification ("profiles") that can meet different market needs
- Allow for the specification to be implemented at varying levels of service
- Allow for the specification to be extended more easily
- Eliminate, where possible, unspecified or underspecified behaviors

4 Technology Pathways

Years of development by Web3D members has resulted in a number of ISO-approved specifications for interactive 3D graphics. In this section, I describe the scope of the standards: their functionality, architecture, and encodings.

4.1 Functionality

The following functionality is covered within the scope of the standard. By understanding the capabilities of standards representations, developers can assess how well it meets their requirements. This list is adapted from a tutorial by Virtual Realms [5]:

- **Navigation** - many control options - walk, fly, examine or none (engines may also support proprietary modes)

- **Viewpoints** - pre-defined camera positions
- **Models** - primitives (box, sphere, cone, cylinder), extrusions, indexed face set (mesh), line set, point set, elevation grid and text (engines may also support proprietary spline and NURBS geometry)
- **Materials** - diffuse colour, specular, emissive, ambient, shininess, transparency, colour per vertex
- **Metadata** – each node in the scengraph can carry metadata information
- **Audio** - fully spatialized 3D audio in WAVE or MIDI format
- **Textures** - support for JPEG, GIF, PNG and MPEG1 video. (engines may also have proprietary support for Flash, RealMedia, AVI, multi-texturing and environment mapping)
- **Lighting** - directional, point or spotlight
- **Environments** - background, switches, hyperlinks (anchor), billboards and fog
- **Performance** - LODs (levels of detail), visibility distance culling
- **Animation** - animate position, rotation, scale, points, colour and much more. Scope for many separate animations in one world all with different time lines and triggered by different events. Almost every attribute can be animated.
- **Sensors** - sense user activity such as touch, drag, keypress (plane, cylinder, sphere, and key sensors). Environmental sensors include time, proximity, and visibility.
- **Scripting** - Interfaces directly with ECMAScript; also with Java, the web browser (DOM) and any programming language residing on the client/server
- **Routes** - scripts, animations and object properties can be "wired" together in an infinite number of ways to create any effect
- **Compact** - extremely small file size with gzip compression
- **Modular** - references external textures, models, scenes and scripts
- **Extensible** - if the core nodes aren't enough you can create your own nodes using Prototypes

Most recent efforts in the standards pipeline include expanding this functionality via: a CAD component, a Physics component, a Layering component, Binary Compression plus XML encryption / authentication, and a Medical/ Volumetric component.

4.2 Architecture

The architecture of X3D is broken up into Components and Profiles. By componentizing the node set and providing various levels of support, X3D is much more flexible in serving numerous application needs. New field types and nodes have been added to the spec and rationalized into components with a coherent node inheritance model. Collections of components are called Profiles.

The **Profiles** architecture allow a broad range of descriptive functionality:

- Interchange (basic models, appearances, lights)
- Interactive (sensors & animation + ; currently used in MPEG4)
- Immersive (VRML 2 ++)
- Full (any additional components)

4.3 Encodings

Beyond the new nodes and the profile archicteure, X3D also introduces additional encoding beyond the 'VRML-Classic' UTF-8 text encoding. Equivalent scenegraphs may be encoded as XML documents or as Binary files. Like VRML, Gzip compression may also be applied to the files. These additional encodings provide developers the ability to generate, store, and transmit the scenegraph with a multitude of databases, tools, and networks. The X3D MIME types and file extensions specified for a webserver are as follows:

```
AddType model/x3d+xml .x3d
AddType model/x3d+vrml .x3dv
AddType model/x3d+binary .x3db
AddEncoding gzip .x3dvz
AddEncoding gzip .x3dbz
```

The addition of the Binary encoding is step forward on a number of fronts. While Gzip compression can provide a significant reduction in files size, it does not use the semantics of the scenegraph to compress the information. In contrast, the X3D Binary Encoding applies special routines for things like geometry and normal information, providing enhanced compression and parsing performance. In addition, the integration with the World Wide Web Consortium's (W3C) Authentication and Encryption provides the means to sign digital assets.

The XML encoding of X3D casts interactive 3D content as a first-class citizen of the information enterprise. Not only can X3D scenes be validated through a DTD or Schema, but they can also be integrated with other XML essentials such as the DOM, SVG, SMIL, and the Semantic Web. In addition, the XML encoding has enabled the use of other standard tools such as Transformational Stylesheets (XSLT) for generating and distributing interactive visualizations of XML data [e.g. 6].

5 Roadmap

5.1 Development

Currently active Web3D Working Groups include:

- X3D User Interface Working Group
- X3D Earth
- Medical X3D
- X3D Networking
- H-Anim

At recent ACM Web3D, IEEE VR, and HCI/VR International conferences, members of the Web3D User Interface Working Group have run or participated in workshops addressing how to relate standards scenegraphs to the variety of input and output devices on the market.

One of these position papers [7] in last year's MRUI Workshop at IEEE VR [8] proposes an 'economy of specification' approach for future standards. By this reasoning, it should be sufficient to standardize scenegraph 'hooks' where navigation, and pointer events (i.e. for selection, and manipulation) can be sent. Applications that load X3D content would be responsible for binding specific devices and mapping interaction techniques and interface logic according to their tasks and platform. The Web3D UI Working Group is continuing to develop and specify the core specificantion changes to provide a truly device-independent standard.

5.2 Conformance

Tools in many domains export VRML and now a growing number export X3D. With over 31 commercial and open source X3D implementations at the time of this writing, the obvious question of conformance arises. While the US National Institute of Standards Technology (NIST) had produced a set of VRML conformance tests in the late 1990s, they were voluntary. In the first half of 2007, the Web3D Consortium began a new official conformance program with updated tests, and rules for use of the Trademark 'X3D'. The new program launched with the official certification of the open source Xj3D tookit as an Interchange Conformant Browser. The conformance test suite is continuting to be developed for other profiles and will continue to enhance interoperability and longevity.

4 Summary

The historical legacy of VRML's design goals left many in the VR community considering it as a 'toy' technology. With the increased power in commodity graphics hardware, it is time for another look. Consider VRML's descriptive functionality for rich interactive scenegraphs rivals or exceeds many established VR tookits. The durability of VRML standards information, such as models and simulation results, has also exceeded the staying power of many VR tools and toolkits.

Most important for the purposes of this workshop, is to recognize the new face of standards scenegraphs: fully-featured graphics, componentized architecture, multiple language encodings, and robust API bindings. These new capabilities present developers and authors with a rich pallette to build virtual worlds and applications.

The question remains, will your pet feature be in the next Standard? Join the Web3D Consortium in opening up new horizons on the WWW!

Acknowledgment

The author wishes to thank the membership of the Web3D Consortium whose work has built the foundations for interactive 3D content as a first-class media type in the networked information ecology. All Trademarks are property of their respective owners.

References

[1] Bowman, D. and Raja, D. A Method for Quantifying the Benefits of Immersion Using the CAVE. *Presence-Connect* (online journal), vol. 4, no. 2, 2004. Available at: http://www.presence-connect.com

[2] Bowman, D., Raja, D., Lucas, J., and Datey, A. Exploring the Benefits of Immersion for Information Visualization. *Proceedings of HCI International*, 2005.

[3] Narayan, M., Waugh, L., Zhang, X., Bafna, P., and Bowman, D. Quantifying the Benefits of Immersion for Collaboration in Virtual Environments. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, 2005, pp. 78-81.

[4] Web3D Consortium, Web3D Specifications http://www.web3d.org (accessed 2/2007)

[5] Virtual Realms Web3D tutorial http://www.virtualrealms.com.au/vrml/tute01/tutorial.htm (accessed 2/2007)

[6] Polys, N. F. Stylesheet Transformations for Interactive Visualization: Towards Web3D Chemistry Curricula. Paper presented at: Web3D Symposium (St. Malo, France, ACM Press). 2003.

[7] Polys, Nicholas F., & Ray, Andrew, "Supporting Mixed-Reality Interfaces through X3D Specification". Workshop on Mixed-Reality Interface Specification IEEE VR 2006, http://w3-mmt.inf.tu-dresden.de/vr2006/slides/MRUI-2006-WorkshopProceedings.pdf (accessed 2/2007)

[8] Workshop on Mixed-Reality Interface Specification IEEE VR 2006 http://w3-mmt.inf.tu-dresden.de/vr2006 (accessed 2/2007)

Dr. Nicholas F. Polys has developed interactive 3D graphic content and systems for over 9 years. His research interests lie at the center of Human Computer Interaction: the intersection of visualization, virtual environments, and perception. After his undergraduate research in Cognitive Science at Vassar College (1996), he jumped into the networked information space of the WWW developing audio, visual, and 3D assets and software. As a co-author of the international standard Extensible 3D (X3D) and a Director of the Web3D Consortium, he has presented papers, tutorials, and workshops at numerous conferences. His doctoral work at Virginia Tech (2006) examined perceptual cues and layout behaviors for Information-Rich Virtual Environments for desktop to immersive platforms. Currently, Nicholas is a researcher for Virginia Tech Research Computing building information architectures and user interfaces for computational scientists.