Virtual Environments and Semantics

Jesús Ibáñez-Martínez and Carlos Delgado-Mata

This article outlines different aspects and applications of virtual environments. The focus here is the relation and integration of these virtual environments with semantic information. In that sense, we show how virtual worlds can benefit from an additional layer of semantic information that allows for improvement of the interaction of the users with these worlds. Furthermore, we demonstrate how the worlds themselves are a powerful tool to show the semantics associated with the information represented from other domains. These aspects are illustrated via concrete examples that show the direct application and usefulness of the ideas presented.

Keywords: Interaction, Metadata, Semantic Information, Virtual Environments.

1 Introduction

Nowadays Virtual Environments (VEs) are a commonly used piece of technology. The actual environments show, on relatively cheap platforms, 3D (three-dimensional) worlds which were unthinkable just a few years back. However, now that we are starting to move in real time in impressively realistic virtual environments, we are coming across other problems and limitations that appear when we interact with the environments for a certain time. Even though the worlds are visually impressive, they lack other aspects of reality that rely on a certain degree of intelligence and, above all, on a different type of representation of the information. These problems need to be solved in order to achieve more interactive and user-friendly environments.

We, as human beings, perceive the environment surrounding us at a physical level. However, we conceptualise it at a more abstract level and we are fooled into believing that we really think and reason at this abstract level. For example, when looking at a tree, we visually perceive colours, textures, shapes, etc. However, we abstract all these features when thinking and reasoning about facts involving the concept of tree.

Current virtual worlds' representation models describe the worlds in such a manner that browsers can efficiently visualise the geometry of the worlds and, in some cases, can support low level interactivity. There is a gap between this low level representation of the worlds and the way we conceptualise them (and therefore the way we think and talk about them). Thus, a high level representation model (including semantics descriptions of the objects in the worlds) is desirable in order to support much richer user interactions at a more abstract level (for example querying for contents) and to deploy agents reasoning about the environments they inhabit.

This paper presents work which add this semantic level to virtual worlds, showing concrete examples where the worlds benefit from this level. Moreover, the paper exposes that the virtual environments themselves are a powerful tool to show the semantics related to the information (from other domains) they represent.

The next sections overview work related to several issues concerning the relationship between semantics and virtual environments. First we study current alternatives to encode metadata for virtual environments. Then, we describe several approaches and concrete examples for the addition of a semantic level to virtual environments. We divide these examples into three groups, depending on whether the semantic level is defined while, before or after the virtual environment is constructed. Finally, we show examples which employ virtual environments as a medium to represent semantics.

Authors

Jesús Ibáñez-Martínez is a Visiting Professor in the Dept. of Technology at the *Universitat Pompeu Fabra*, Barcelona, Spain. He received a PhD degree in Computer Science from the *Universidad de Murcia*, Spain. Amongst others, he was granted a Marie Curie fellowship at the Centre for Virtual Environments at the University of Salford, Manchester, UK. His research interests include the application of intelligent systems to improve user interaction, especially with virtual worlds. He has published extensively in this area. He is co-founder and co-chair of the biennial International Conference on Intelligent Virtual Environments and Virtual Agents (IVEVA). <jesus.ibanez@upf.edu>

Carlos Delgado-Mata is Computer Science Engineer from *Universidad Bonaterra*, Aguascalientes, Mexico (1990-1995). He graduated with honours and received a medal as the best student of his generation. He holds a Master in Computation Engineering from UNAM (*Universidad Nacional Autónoma de México*), (1996-1997). He received his PhD from University of Salford, UK. He has been programming leader in several projects, for example REMOTE. He is co-chair of the International Conference Intelligent Virtual Environments and Virtual Agents (IVEVA) 2006 and he was co-chair of the International workshop IVEVA 2004. He is author of more than 20 peer refereed articles on international conferences and journals. At the moment he is in charge of CINAVI (Centre for Research in Intelligent Virtual Environments) and of the Computation Academy in *Universidad Bonaterra*. He received a diploma on behalf of ACM (Association for Computing Machinery) because he was the pioneer in the ACM Programming contest in the region. He is partner of the newly form computer games developer Nibbo Studios. <cdelgado@bonaterra.edu.mx>

2 Encoding Metadata for Virtual Environments

The annotation of metadata related to the elements contained in a virtual environment is, clearly, a useful step towards establishing a level of semantic information about the environment. The set of metadata annotated in a world, along with ontologies and knowledge bases of the appropriate domain, may constitute a semantic level which improve the ability to interact with the world. This section briefly describes three current ways to encode metadata in virtual environments and 3D contents. In particular we explore mechanisms to encode metadata in VRML (Virtual Reality Modeling Language) worlds, GeoVRML and MPEG-7 (Moving Picture Experts Group).

2.1 Encoding Metadata in VRML Worlds

There are two ways to encode a kind of metadata in VRML worlds: the *WorldInfo* node and the parameterisation of VRML prototypes. Note that the latter was not explicitly conceived as a mechanism to encode metadata, though it can fulfil that task.

The *WorldInfo* node is a VRML node which contains general information about the world. This node is strictly for documentation purposes and has no effect on the visual appearance or behaviour of the world. The WorldInfo node contains two fields: *title* and *info*. The *title* field is intended to store the name or title of the world so that browsers can present this to the user (perhaps in the window border). Any other information about the world can be stored in the *info* field, such as author information, copyright and usage instructions. There is no convention about how to encode this information. Therefore, this field is mainly used for human management of the worlds.

The parameterisation of VRML prototypes can be employed to encode metadata related to a particular object in the world, even though it was not conceived as a way to encode semantics. The prototyping mechanism provides the capability to have semantic descriptions of objects which abstract out the geometric implementation, e.g. you could define a prototype called *House* with fields such as *numfloors*, *rooftype*, *numwindows*, *location*, etc. (all semantic information). To specify an implementation of this prototype, we can use something like "House {numfloors 2 rooftype "flat" ...}". This specification encodes semantic information about a VRML object (in particular a house). This information will be most likely employed to construct a geometric representation of the house on the fly. In fact, the parameterisation of VRML is not usually employed to encode metadata and there is no convention on how to encode it.

2.2 Encoding Metadata in GeoVRML

GeoVRML is an official Working Group of the Web3D Consortium. It was formed in 1998 with the goal of developing tools and recommended practice for the representation of geographical data using VRML. The desire is to enable geo-referenced data, such as maps and 3-D terrain models, to be viewed over the web by a user with a standard VRML plugin for their web browser. The group has produced the GeoVRML 1.1 specification, providing a number of extensions to VRML for supporting geographic applications. There is also an accompanying Open Source Java sample implementation of these nodes. Finally, these nodes are part of Amendment 1 to the VRML97 ISO standard.

GeoVRML 1.1 includes the *GeoMetadata* node, which aims to specify metadata describing any number of GeoVRML nodes. There are a number of organisations that are already working on standards and representations for geospatial metadata, such as the FGDC, ISO TC211, CEN TC287, OpenGIS Consortium, and others. Rather than adopting any particular standard, the purpose of the GeoMetadata node is to provide links to any of these complete metadata descriptions, with the option to also supply a short, human-readable summary. It is not the purpose of the GeoMetadata node to introduce a new metadata standard.

The GeoMetadata node contains three fields: *url*, *summary* and *data*. The *url* field is used to specify a hypertext link to an external, complete metadata description. Multiple URL (Universal Resource Locator) strings can be specified in order to provide alternate locations for the same metadata file. The *summary* string array contains a set of keyword/value pairs, with each keyword and its subsequent value contained in a separate string. Keywords include *title, description, coordinateSystem, horizontalDatum, verticalDatum, ellipsoid, extent, resolution, originator, copyright, date, metadataFormat, dataUrl* and *dataFormat*. This provides a simple, extensible mechanism to include metadata elements that are human-readable and easy to parse. The *data* field is used to list all of the nodes that implement the data described in the GeoMetadata node. If the data field is not specified, it is assumed that the GeoMetadata node pertains to the entire scene.

2.3 MPEG-7

MPEG-7, formally known as Multimedia Content Description Interface, is a standard to describe multimedia content. MPEG-7 will provide a suite of standardised tools that describe multimedia content that will be consumed by both humans and automated computer systems. MPEG-7 is expected to form the basis of a new generation of multimedia applications such as media search and retrieval systems, digital libraries, multimedia editing systems and multimedia directory systems. MPEG-7, for example, can be used to describe MPEG-4 content in ways that are not possible with MPEG-4 alone, dramatically increasing the likelihood that end users will be able to locate a specific piece of content when they want it. Likewise, MPEG-7 can be used to describe content stored in a variety of multimedia formats, such as RealAudio and RealVideo, Microsoft Media, Apple QuickTime and others. The MPEG-7 standard specifies:

- *Descriptors* which describe basic characteristics of audiovisual content that define the syntax and the semantics of each feature representation.
- *Description Schemes* that specify the structure and semantics of the relationships between components which may be both Descriptors and Description Schemes.
- *Description Definition Language* which allows new Descriptors and Description Schemes to be defined.
- *Systems layer* that allows synchronisation and ac-

cess of MPEG-7 description metadata with or without the corresponding multimedia data.

MPEG-7 does not specify how to extract descriptions, how to use descriptions nor how to evaluate the similarity of different content. Several visual descriptors have been defined, including color, texture, shape and motion for video. The shape descriptors include region-based shape, contour-based shape, 2D/3D shape and 3D shape. Thus, 3D shape information can be described in MPEG-7 through the 3D shape descriptor. Most of the time, 3D information is represented by polygonal meshes. As pointed out in [1], the 3D shape descriptor provides an intrinsic shape description of 3D mesh models. The main applications targeted by this descriptor are search, retrieval and browsing of 3D model databases based on shape similarities, i.e. at a low-level.

3 Semantic Level

This section describes several approaches and concrete examples for the addition of a semantic level to a virtual world. These examples are grouped into three types: examples that construct the world along with its semantic level, those that build the world on a pre-existent semantic level, and those which add the semantic level to a pre-existing world.

3.1 Virtual Environments Built along with A Semantic Level

Since a few years ago, the necessity for a semantic level to improve the **interaction** with virtual environments has been apparent.. In this sense, Cavazza discusses, in [2], requirements for a knowledge representation layer for virtual environments. He recognises that in some virtual world applications there is a need for simultaneous access to both concrete and abstract information. Concrete information is related to a direct visual representation of the simulation to be controlled. Abstract information is required in order to master the complexity of the scene and its dynamic evolution. Cavazza states that domain concepts are described in abstract terms, and have a nontrivial correspondence with real-world object configuration. Therefore a challenge for VE is to be able to apply abstract procedures to the VE as a whole, rather than to its constituent objects. Cavazza particularly bases his discussion on the kind of AI (Artificial Intelligence) techniques that have been developed for the simultaneous interpretation of natural language and dynamic scenes, or for the creation of animation sequences from natural language.

There is no standard (not even a de facto standard) to model the semantics of virtual worlds. As a result, some groups define a particular modelling approach. Usually each approach is application dependant. For example, in [3], the authors present an Informed Environment that creates a database dedicated to urban life simulation. They introduce a method for building virtual scenes with associated semantic information as well as for the exploitation of such scenes. The three-dimensional scene provided by the designer is divided into two parts, one for visualisation and another for database construction. The database contains geometrical and semantic information for mobile entity simulation. Their approach is to define some structured areas. The areas are either subdivided into sub-areas or grouped, depending on the level of information. Thus, by analogy to a geographical map, they decompose a large area into sub-areas with information inherent to the level of description. The semantic information is later used to simulate more realistic behaviours of virtual humans (for example humans walking on a sidewalk). This Informed Environment database model including semantic information has also been integrated with a Web application. This integration is described in [4].

NOVAE (Networked Open VirtuAl Environment) was a project aimed at the definition of components for an open architecture supporting interoperability of virtual worlds. In the context of NOVAE, in [5], the authors propose a semantic approach to increase interoperability between networked virtual worlds. They propose the ontology paradigm to define common knowledge in virtual worlds, and the Influence/Reaction model for both describing the behaviours

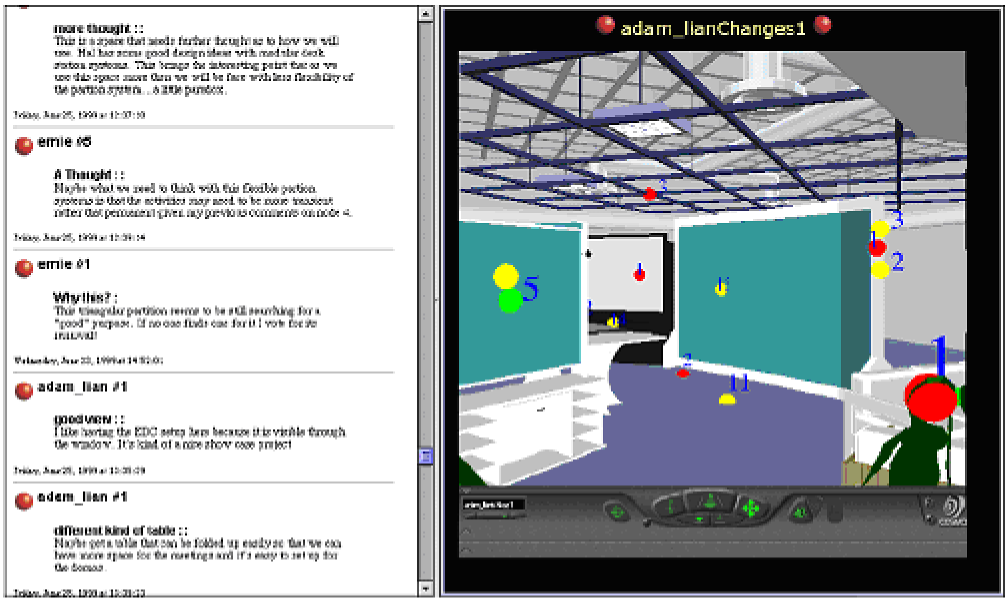

Figure 1: The User Interface of Redliner.

of the virtual entities and the reactions of the virtual worlds. In fact they extend the Influence/Reaction model (developed in multi-agents systems) to the concept of world resource. A resource represents a strong semantic feature of the virtual world and may be used by the designer to express a requirement or a property which is common to numerous virtual entities.

Some examples propose semantics as a fundamental issue for **modelling** virtual environments. In this sense, in [6], the authors determine the necessity of having tools which help the designer to model the VR application much more from the domain expert point of view. This lets them participate in the specification of the VR application in a more intuitive way using their own domain expertise and terminology. To achieve that, the authors introduce a new approach to design and develop a VR application where the domain expertise is used to generate it more easily. This approach uses ontologies as a way of grasping the knowledge of a domain, expressing the virtual world in terms of the end-user domain, generating the virtual world more easily and performing intelligent reasoning.

3.2 Virtual Environments Built on A Semantic Level

Some virtual environments are built on a preexistent semantic level. For example, some virtual environments represent real spaces by taking advantage of GIS (Geographical Information System) data. Perhaps the clearest examples of environments built on a semantic level are *digital cities*. They integrate urban information and create public spaces for people living in real cities. Digital cities, which are being developed all over the world, will provide the infrastructure for networking local communities.

Each digital city has a different goal. For example AOL digital cities aim to explore vertical markets. In fact, each AOL digital city [7] collects tourist and shopping information for the corresponding city. On the other hand, Digital City Amsterdam [8] was built as a platform for several community networks and thus specially focuses on social interaction among citizens.

According to [9] the design of digital cities follows a general three layer model. Its layers are as follows:

- Information layer, where digital archives, web pages and realtime sensory data from physical cities are integrated and reorganised using the city metaphor.
- Interface layer, where graphical representations provide a view of the city.
- Interaction layer, where resident and tourist interact with each other.

Usually a semantic level (mostly a geographical database) is employed to facilitate the integration of different kinds of data in the information layer in a coherent manner. In the interface layer 2D maps have been used extensively, though recent works on digital cities frequently include 3D technologies. Some examples of 3D spaces for digital cities are those of Helsinki [10], Los Angeles [11], and Kyoto [12].

3.3 Adding A Semantic Level to Virtual Environments

Some examples research the annotation of virtual environments adding a semantic level which can be perceived later by other users. In this sense, voice annotation systems are proposed in the Virtual Book project [13] or the V Anno project [14]. There users leave annotations markers in the 3D model and record voice annotations for future retrieval.

The post-it note metaphor was introduced in [15] to annotate 3D fluid models. Markers associated with text are embedded in a virtual environment and reviewed later by other users. A magic lens filter allows hiding or highlighting markers in a selected area. In [16] a prototype system is proposed allowing users to place grocery shopping list on a VRML refrigerator door (post-it style) so that relationships between annotation and places can be more easily recognised, based on the objects they are attached to.

IDT (Immersive Discussion Tool) [17] is a system to annotate architectural 3D models. IDT allows users to leave arrows to designate specific points of the design, circles to annotate larger areas or geometric constructions made of connected cylinders and spheres for representing 3D artifacts or gestures. Text can be attached to any mark left in the model.

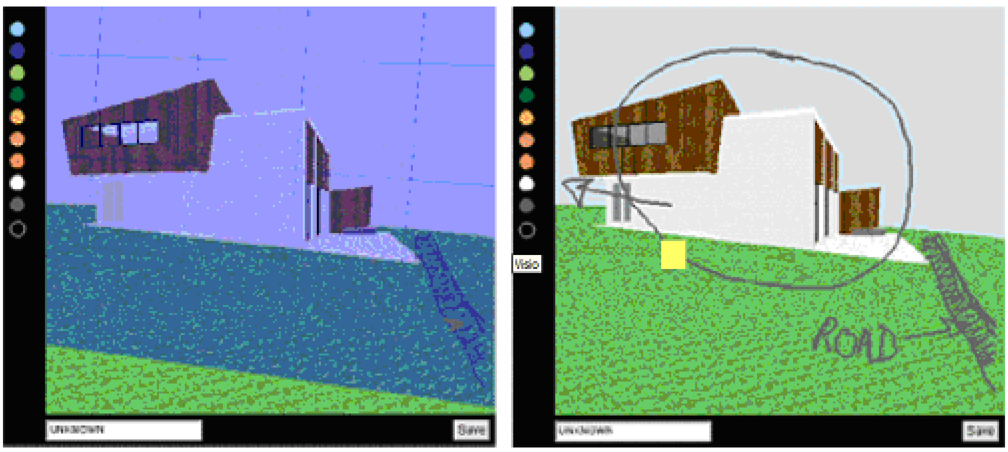

Figure 2: Space Pen.

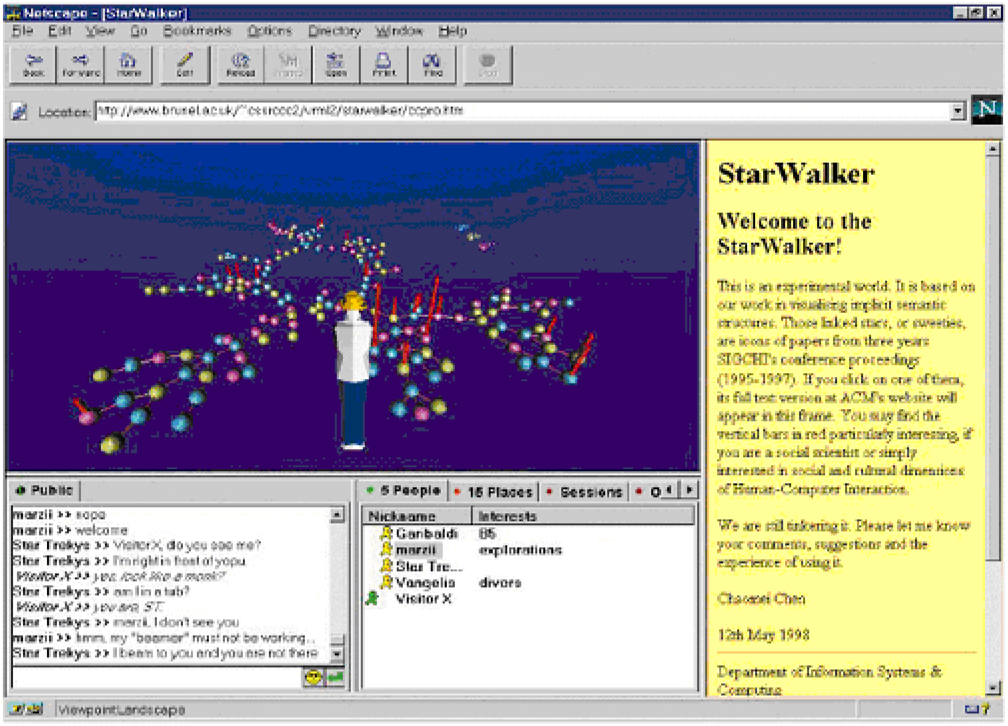

Figure 3: The User Interface of Starwalker Consists of A Virtual World Window, A Chat Window and A Session Window.

Redliner [18] is a system for collaborating designers which allow them to annotate 3D models by leaving annotation marks (represented as colored spheres) directly in the 3D representation of a building. Any member of the design team can log on to the Redliner Web site, identify a login name and choose an annotation colour. The Redliner interface (see Figure 1) contains a window with two frames: one containing user text annotations and the other is for 3D interaction with the design model. Comments listed on the left side of the Redliner window are sorted by time and by author. The right side displays the 3D model.

In Space Pen [19], as in Redliner, users can tag the model with notes associated with text comments. In addition, Space Pen users can also mark up the surfaces of the 3D model (see Figure 2) and save those drawing annotations for others to review. The drawing mark can also be associated with an explanatory text comment.

4 Virtual Environments to Represent Semantics

Some examples study the employment of virtual environments to represent semantics related to data. In this sense, in [20], semantically organised virtual spaces are examined, focusing on how such spatial organisation shapes users' cognition, interpretation and interaction. In particular, the authors describe the design of StarWalker, a multiuser virtual environment which illustrates authors' design principles for unifying spatial models, semantic structures, and social interaction patterns within multiuser environments.

The semantic structure of the virtual world in StarWalker's prototype uses three years of proceedings from the ACM Computer Human Interaction conference (CHI, 1995-1997). This spatial model was automatically generated by using a unique integration of latent semantic indexing (LSI) and pathfinder network-scaling techniques. The authors used LSI to generate interdocument similarities, and they used the Pathfinder network-scalling techniques to extract salient structural patterns. The semantic space obtained was rendered into a VRML model and made available on the Internet through Blaxxun's Online Community client-server architecture (Figure 3 shows the user interface). Then the authors studied how visitors used StarWalker.

The DocuDrama Timetunnel [21] (see Figure 4) tells the history of folders and documents in a project's lifecycle. It visualises interaction on folders and documents, shows interdependencies and coherence between activities. The aim of the DocuDrama Timetunnel is to provide an abstract view on project activity and to offer functionality to manage project data. It provides team-members with a generative tool to visualise project events history in various configurations. Moving through the tunnel the user can embark on a virtual journey through the project's lifetime, in which the tunnel symbolises the time axis of the project.

The user selects a time period of a project's lifetime for visualisation, for example the complete lifetime of a project. In that case, the entry of the tunnel represents the start of the selected time period. The closer the user gets to the end of the tunnel, the closer he gets to the end of the project's lifetime. The time tunnel consists of different slices. Each slice represents a step in the lifetime of the project. Depending on the selected time period a slice might represent a year, a month, a day or even only an hour. Small boxes placed on the wall of a time-slice denote interaction with the project's folders and documents. Each box represents a document. The position of the box inside the time-slice indicates the form of interaction that has been opened for writing. Boxes on the ceiling show documents which have been opened for reading. The user can retrieve the document's name and the date of the action by clicking on a box. The colouring of the walls supports the meaning of the position. The user is now able to move along the time axis, going back and forth in time, and following the project's activities over a time period. Many boxes piled up on the walls denote time periods with high activity. Empty time slices denote time periods at which no action has been performed.

The National Laboratory for Applied Network Research (NLANR) has a number of network measurement projects that produce large volumes of data. Altogether, these projects produce gigabytes of data and thousands of web pages of graphs and summaries each day. A large volume of data is necessary because it is not a priori known which parts of the data will contain interesting features and because different users are interested in different parts of the measured systems. As a consequence of the systems scale, it is not humanly possible for users to scan all of the data or web pages for interesting artifacts, but in not doing so, important network events may be overlooked. As an approach to address this need, NLANR developed the Cichlid [22] tool to visualise and animate data sets in a three-dimensional virtual environment. The data can be animated to show changes over time. Often a physical analogy of the data is developed in order to show related semantics; for example, network delays between a collection of sites, visualised as an interpolated surface and animated in time, creates a display similar to the view of flying over mountainous terrain. Different landscapes are related to different network conditions.

Pajek [23] is a program for analysing and visualising large networks (e.g. genealogies, flow graphs of programs, computer networks, transportation networks, social networks, etc.). Pajek contains many different tools to visualise networks, as well as several tools to export visualisations of networks to different formats including VRML. Therefore, as long as a network can capture the semantics of a system, a VRML virtual world can represent this semantics.

5 Conclusions

This article has outlined different aspects and applications of virtual environments. The focus here has been the relation and integration of these virtual environments with semantic information. In that sense, we have shown how virtual worlds can benefit from an additional layer of semantic information that allows improving the interaction of the users with these worlds. Furthermore, we have demonstrated how the worlds themselves are a powerful tool to show the semantics associated with the information represented from other domains. These aspects have been deployed via concrete examples that have shown the direct application of the presented ideas and their utility.

Translation by the authors

Figure 4: DocuDrama Timetunnel.

References

[1] Philippe Salembier. Overview of the MPEG-7 Standard and of Future Challenges for Visual Information Analysis, EURASIP Journal on Applied Signal Processing, vol 4, pp. 1-11, 2002.

[2] Marc Cavazza. High-Level Interpretation in Dynamic Virtual Environments, Working Notes: Intelligent Virtual Environments, Workshop at the 13th Biennial European Conference on Artificial Intelligence (ECAI-98), Brighton, UK, 1998.

[3] Daniel Thalmann, Nathalie Farenc, and Ronan Boulic. Virtual Human Life Simulation and Database: Why and How, International Symposium on Database Applications in Non-Traditional Environments, DANTE'99, 1999.

[4] Nathalie Farenc, Christian Babski, Fabien Garat, and Daniel Thalmann. Database and Virtual Human Representation on the WEB, Advances in Databases and Multimedia for the New Century -A Swiss-Japanese Perspective-,WorldScientific, 2000.

[5] Michel Soto and Sebastien Allongue. A semantic approach of virtual worlds interoperability, IEEE WET-ICE '97, Cambridge, MA, IEEE Press, 1997.

[6] Wesley Bille, Bram Pellens, Frederic Kleinermann, and Olga De Troyer. Intelligent Modelling of Virtual Worlds Using Domain Ontologies, Workshop on Intelligent Virtual Environments and Virtual Agents, IVEVA, Mexico City, Mexico, 2004.

[7] AOL digital cities, <http://www.digitalcity.com/>.

[8] Peter van den Besselaar and Dennis Beckers. Demographics and Sociographics of the Digital City, Community Computing and Support Systems, Lecture Notes in Computer Science, vol 1519, pp. 109-125, Springer-Verlag, 1998.

[9] Toru Ishida. Understanding Digital Cities, Digital Cities: Experiences, Technologies and Future Perspectives, Lecture Note in Computer Science, vol 1765, Springer-Verlag, 2000.

[10] Risto Linturi, Marja-Riitta Koivunen, and Jari Sulkanen. Helsinki Arena 2000 - Augmenting a Real City to a Virtual One, Digital Cities: Experiences, Technologies and Future Perspectives, Lecture Note in Computer Science, vol 1765, Springer-Verlag, 2000.

[11] William Jepson and Scott Friedman. The Virtual World Data Server & The Virtual Los Angeles Project.

[12] Toru Ishida, Junichi Akahani, Kaoru Hiramatsu, Katherine Isbister, Stefan Lisowski, Hideyuku Nakanishi, Masayuki Okamoto, Yasuhiko Miyazaki, and Ken Tsutsuguchi. Digital City Kyoto: Towards A Social Information Infrastructure, International Workshop on Cooperative Information Agents (CIA-99), Lecture Notes in Artificial Intelligence, vol 1652, 1999.

[13] Jouke C. Verlinden, Jay David Bolter, and Charles van der Mast. Virtual Annotation: Verbal Communication in Virtual Reality, Proceedings of the European Simulation Symposium, ESS'93, 1993.

[14] Reid Harmon, Walter Patterson, William Ribarsky, and Jay Bolter. The Virtual Annotation System, IEEE Virtual Reality Annual International Symposium, pp. 239-245, 1996.

[15] Maria M. Loughlin and John F. Hughes. An Annotation System for 3D Fluid Flow Visualisation, Proceedings of the IEEE Conference on Visualisation '94, pp. 273-279, 1994.

[16] C. M. Adriano, A. B. Raposo, I. L. M. Ricarte, and L. P. Magalhaes. Changing interaction paradigms in annotation environments, EDMEDIA 2000, Montreal, Canada, 2000.

[17] David Latch Craig and Craig Zimring. Support for Collaborative Design Reasoning In Shared Virtual Spaces, Media and Design Process, Proceedings of ACADIA '99, pp. 126-137, 1999.

[18] Thomas Jung, Ellen Yi-Luen Do, and Mark D. Gross. Immersive Redlining and Annotation of 3D Design Models on the Web, Proceedings of the 8th International Conference on Computer Aided Architectural Design Futures, pp. 81-98, 1999.

[19] Thomas Jung, Mark D. Gross, and Ellen Yi-Luen Do. Space Pen: Annotation and Sketching on 3D Models on the Internet, Proceedings of the 9th International Conference on Computer Aided Architectural Design Futures, pp. 257-270, 2001.

[20] Chaomei Chen, Linda Thomas, Janet Cole and Chiladda Chennawasin. Representing the semantics of virtual spaces, IEEE Multimedia, April/June, 1999.

[21] Leonie Schäfer, Uta Pankoke-Babatz, Wolfgang Prinz, Ava Fatah Schieck, Amanda Oldroyd. DucuDrama, Journal of Virtual Reality, Springer, vol 7, num 1, pp. 43-53, 2003.

[22] Jeffery A. Brown, A. J. McGregor, and H-W. Braun.

