Shaping 3D Multimedia Environments: The MediaSquare

Michael Dittenbach,
Helmut Berger,
Ronald Genswaider,
Andreas Pesenhofer
E-Commerce Competence Center–EC3
iSpaces Research Group
Donau-City-Straße 1,
A–1220 Wien, Austria
{firstname.lastname}@ec3.at

Andreas Rauber,
Thomas Lidy,
Dieter Merkl
Intitut für Softwaretechnik und
Interaktive Systeme
Technische Universität Wien
Favoritenstraße 9–11/188
A–1040 Wien, Austria
{rauber,lidy}@ifs.tuwien.ac.at
dieter.merkl@ec.tuwien.ac.at

ABSTRACT

In this paper we describe *The MediaSquare*, a 3D Multimedia Environment we are currently developing, where users are impersonated as avatars enabling them to browse and experience multimedia content by literally walking through it. Users may engage in conversations with others, exchange experiences as well as collaboratively explore and enjoy the featured content. The combination of algorithms from the area of artificial intelligence with state-of-the-art 3D virtual environments creates an intuitive interface that provides access to automatically structured multimedia data taking advantage of spatial metaphors.

Categories and Subject Descriptors

H.5.2 [**Information Interfaces and Presentation**]: User Interfaces; H.4.3 [**Information Systems**]: Communications Applications; I.2.6 [**Artificial Intelligence**]: Learning

General Terms

Human Factors, Design

Keywords

3D collaborative environments, self-organizing map, multimedia, game engine

1. INTRODUCTION

Millions of users interact, collaborate, socialize and form relationships with each other through avatars in online environments such as Massively Multi-User Online Role-Playing Games (MMORPGs) [1, 5, 6]. While the predominant motivation to participate in MMORPGs is still "playing", an increasing number of users spend a significant amount of time in 3D virtual worlds without necessarily following a predefined quest. Generating and publishing content in 3D virtual spaces is an emerging trend on the Internet with *Second Life*[1] being the most prominent representative at the time of writing. On the one hand, such 3D virtual worlds address the aspect of social interaction by providing instruments to interact and to exchange experiences with other users that go beyond the possibilities of conventional text-based chat rooms. Especially one's inherent presence in space and the awareness of others facilitates the initiation of social contacts. On the other hand, using 3D virtual worlds has the advantage of communicating via commonly accepted spatial metaphors [2]. Similarity of objects can be expressed by spatial relations, i.e. the more similar two objects are, the closer they are placed together. Furthermore, users can interpret each other's interests by how close they are to one another and the objects in space. Having a common point of reference and orientation within the virtual space as well as being aware that other users can see their actions and objects in the same way, are important features regarding communication between users about particular locations. Consequently, users are supported in building a mental model of the information space, to understand its characteristics and to grasp which information is present and how the respective items relate to each other.

The MediaSquare, a 3D multimedia environment, takes advantage of these spatial metaphors and allows users to explore multimedia information that is automatically structured and organized within space (see Figure 1). The information is either organized based on the actual content or by transforming predefined structures into a room structure. Currently, *The MediaSquare* implements the following scenarios. The first scenario is a 3D Music Showroom that enables users to browse and listen to songs within the collaborative virtual environment. To this end, acoustic characteristics are extracted from music tracks by applying methods from digital signal processing and psycho-acoustics. The features describe the stylistic content of the music, e.g. beat, presence of voice, timbre, etc. and act as features for the training of a self-organizing map (SOM) to arrange similar

[1] `http://secondlife.com`

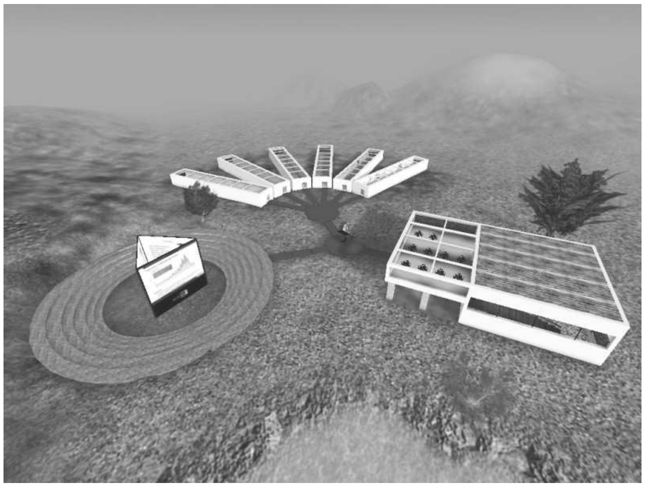

Figure 1: The MediaSquare.

music tracks in spatially adjacent regions. More specifically, the self-organizing map is an an unsupervised neural network model that provides a topology-preserving mapping from a high-dimensional input space onto a two-dimensional output space [3].

In a second scenario, a 3D scientific library has been implemented. This library enables users to explore scientific documents such as posters or papers in this immersive 3D environment. In this case, a directory structure is used to create a room structure in which the content is presented. In its final version it is envisioned to integrate a 3D Video and Image Showroom whose content is automatically arranged in analogy to the Music Showroom.

In a nutshell, *The MediaSquare* presents an impressive showcase for combining state-of-the-art multimedia feature extraction approaches and unsupervised neural networks assembled an immersive 3D multimedia content presentation environment. This allows geographically separated individuals to immerse into a collaborative virtual environment, interact with each other and collectively experience the featured content.

2. THE SYSTEM ARCHITECTURE

The MediaSquare is a three-dimensional virtual environment that allows multiple users to explore large multimedia repositories such as music or text collections as well as video or image galleries. The system architecture is depicted in Figure 2. It is based on the Torque Game Engine[2], whereof the server is the core of the system. The server, running in dedicated mode, is responsible for the creation of the virtual world, the coordination of the avatars' positions as well as the media and communication handling. In Torque, virtual environments are described in terms of mission files, which define the terrain, sky textures, buildings, interiors, water areas, etc. Torque clients connect to the server and are responsible for the user interface, graphics rendering and sound playback. On connecting, it checks for the most recent version of the environment and automatically downloads updates files, if necessary.

[2] `http://www.garagegames.com`

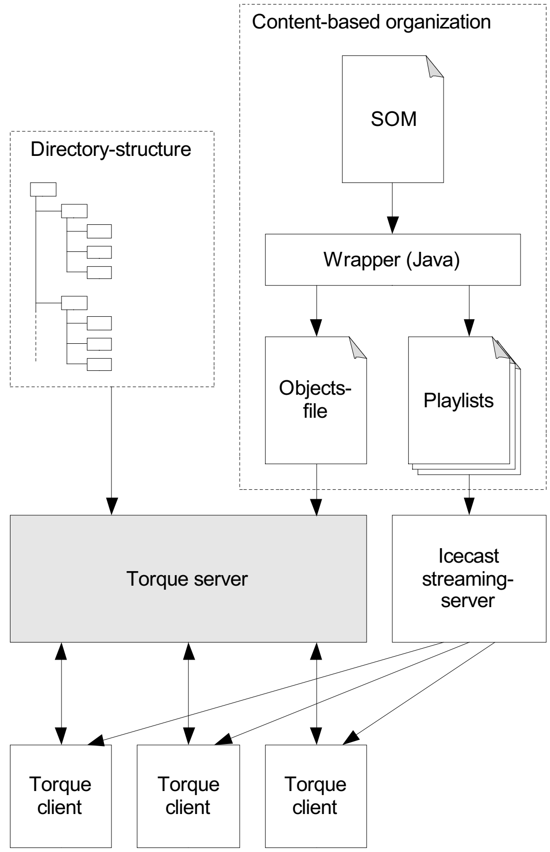

Figure 2: System architecture.

In order to populate the environment with data, the system offers two modes. The first mode allows automatic content-based organization of data using a self-organizing map. Prior to integrating the mappings generated by a SOM, it is necessary to define designated areas, i.e. marker areas, in the virtual world that specify the positions of the media data. This is done with the built-in game editor of the Torque Game Engine. The marker areas are rectangles that are invisible during runtime. These contain properties that specify which SOM, or specific parts thereof, are selected for a particular area in the environment. So, several SOMs can be represented simultaneously and each map can be split into multiple parts. For example, this can be used to distribute them across several rooms of a building.

A further property defines the type of interiors that are placed in the area. Template files, created with the Torque editor, contain sets of objects that are grouped together. An interior set represents one unit of the SOM in the virtual environment. A Java-based wrapper application translates the SOM files into a configuration file, imports the maps, reads the marker areas and the template files for the virtual environment. Then the wrapper creates an object file containing information about interiors, media objects and their respective positioning.

The second mode is based on a predefined directory tree which is mapped onto a corresponding room structure. So, arbitrary data can be integrated into the environment. The top-level elements in the hierarchy are folders that structurally group related data. Every folder contains a file that provides metadata such as title, author, date and the number of group elements. It also contains an image that is used as a texture for labeling the objects in the virtual world. The actual data are stored in sub-folders. Consider, for example, a slide show consisting of a number of slides. In this case, the slides are stored as separate images to be used as textures on a presentation screen.

On startup, the server reads the configuration file and loads a mission file. Additional objects, i.e. the interior sets and audio sources as specified by self-organizing maps, are added via scripting. Moreover, the server scans the directory structure and maps the data onto the corresponding visual representation. Then, the virtual environment is created and may be distributed to the clients. When a client connects to the server, an avatar model and the user name must be chosen in order to get access to *The MediaSquare*. With their avatars, users can explore the world and communicate with each other.

3. THE MEDIASQUARE

3.1 The 3D Music Showroom

The music library used in *The MediaSquare* is the collection from Magnatune[3], which is distributed under the creative commons license for non-commercial use. It contains about 1,500 MP3 files and is subdivided into the genres classical, electronic, jazz, blues, rock, pop, metal, punk and world music.

In this particular scenario, the wrapper application inspects the mission file for area markers, loads the corresponding SOMs into the internal object structure and reads the template files for the interior. For every unit of the SOM a location for an audio source and its respective interior is determined. An interior template set for the music showroom consists of a table, chairs, a playlist and a speaker. The texture of the playlist object is created dynamically from the song metadata, which is extracted from the ID3 tags of the music files. Finally, the information about the objects is written to the objects file including the name, position, type (interior, playlist, audio emitter object), rotation (according to the rotation of the marker-area), 3D shape file, the URI of the audio stream and the texture for playlists. For audio streaming, the program creates playlist files for every represented SOM unit and a configuration file for the Icecast[4] streaming server. It broadcasts audio like a radio station, thus, it is ensured that all users are listening to the same music when at the same location. Depending on the user's position relative to the audio sources, one ore more spatialized audio streams are audible. When the user's avatar is close to an audio source a head-up display shows the currently playing track as well as the corresponding playlist (see Figure 3). On clicking the left mouse button the audio stream of the respective source skips to the next track and

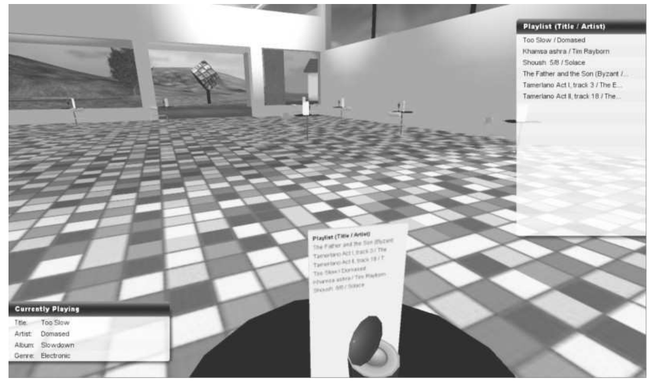

Figure 3: Music showroom with GUI elements.

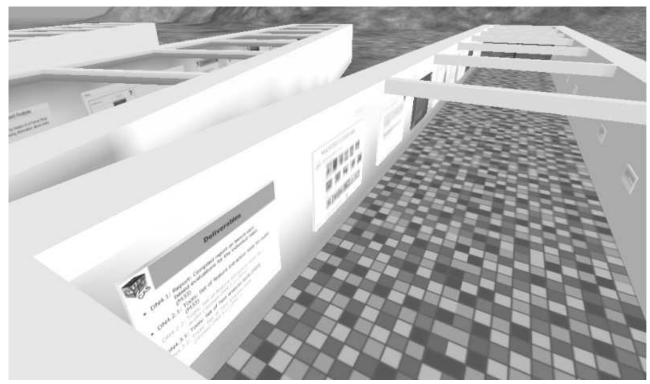

Figure 4: Directory structure mapping.

all users close to this particular audio source will notice the change.

3.2 The Scientific Library

The second scenario implements a 3D scientific library. This library enables users to explore scientific documents such as presentations, posters or papers in this immersive 3D environment. *The MediaSquare* prototype provides access to the scientific results of the MUSCLE Network of Excellence. In this case, a directory structure is used to create a room structure in which the content is presented. The presentations are grouped according to different scientific meetings and each directory contains several presentations that have been given there. The metadata file associated with a meeting describes the location and the date it was held. In case of a presentation the metadata file contains the title and the author names.

The 3D virtual world contains several buildings whereof each represents a particular meeting as shown in Figure 4. These buildings feature display screens that are used for visualization of the data. Labels describing the meetings' locations and dates are placed at the corresponding entrances. Each building features presentation screens that are attached to the walls. Labels containing the metadata are attached on the exact opposite. The textures of the screens change in predefined time intervals.

[3] `http://magnatune.com`

[4] `http://www.icecast.org`

4. CONCLUSIONS

In this paper, we have presented the current state of our 3D multimedia environment, which allows multiple users to collectively explore multimedia data and interact with each other. The data is automatically organized within the 3D virtual world either based on content similarity, or by mapping a given structure (e.g. a branch of a file system hierarchy) into a room structure. With this system it is possible to take advantage of the features of spatial metaphors such as relations between items in space, proximity and action, common reference and orientation, as well as reciprocity.

Future work includes improved user interface capabilities, a tighter integration of the single components of the system as well as the integration of additional feature extraction modules for other file types. An extension of the SOM overcoming the limitation of a rectangular map structure is the Mnemonic SOM [4], which will be used in the future. In this model, the map nodes are assigned according to standard size parameters such as rows and columns of nodes, or approximate number of nodes, with the required map nodes being projected onto an arbitrary map shape. Originally intended to better explain a trained SOM by using recognizable map shapes (e.g. outline of a country). The advantage of this model for our approach is to better fit the information space spanned by the SOM into more complex, non-rectangular room structures on the virtual world.

5. ACKNOWLEDGMENTS

This work was partially funded by the Austrian Federal Ministry of Economics and Labour under the k-ind research program and the MUSCLE Network of Excellence (project reference: 507752).

6. REFERENCES

[1] E. Castronova. *Synthetic Worlds: The business and culture of online games.* University of Chicago Press, Chicago, IL, 2005.

[2] S. Greenberg and M. Roseman. Using a room metaphor to ease transitions in groupware. In M. Ackerman, V. Pipek, and V. Wulf, editors, *Sharing Expertise: Beyond Knowledge Management*, pages 203–256. MIT Press, Cambridge, MA, January 2003.

[3] T. Kohonen. *Self-organizing maps.* Springer-Verlag, Berlin, 1995.

[4] R. Mayer, D. Merkl, and A. Rauber. Mnemonic SOMs: Recognizable shapes for self-organizing maps. In M. Cottrell, editor, *Proceedings of the Fifth Workshop on Self-Organizing Maps (WSOM'05)*, pages 131–138, Paris, France, September 5–8 2005.

[5] B. Woodcock. An analysis of MMOG subscription growth. `http://www.mmogchart.com/`.

[6] N. Yee. The psychology of massively multi-user online role-playing games: Emotional investment, motivations, relationship formation, and problematic usage. In R. Schroeder and A. Axelsson, editors, *Avatars at Work and Play: Collaboration and Interaction in Shared Virtual Environments*, volume 34 of *Computer Supported Cooperative Work.* Springer-Verlag, Heidelberg, Germany, 2005.