

Journal of • Virtual Worlds Research

jvwresearch.org ISSN: 1941-8477

Volume 2, Number 3
Technology, Economy, and Standards.

**Community
Creation
Commerce**

Artwork by Anshe Chung Studios



Volume 2, Number 3 Technology, Economy, and Standards October 2009

Editor

Jeremiah Spence

Guest Editors

Yesha Sivan
J.H.A. (Jean) Gelissen
Robert Bloomfield

Reviewers

Aki Harma
Esko Dijk
Ger van den Broek
Mark Bell
Mauro Barbieri
Mia Consalvo
Ren Reynolds
Roland LeGrand
Vili Lehdonvirta

Technical Staff

Andrea Muñoz
Kelly Jensen
Roque Planas
Amy Reed

Sponsored in part by:



virtual worlds research consortium

The Journal of Virtual Worlds Research is owned and published by:

The JVWR is an academic journal. As such, it is dedicated to the open exchange of information. For this reason, JVWR is freely available to individuals and institutions. Copies of this journal or articles in this journal may be distributed for research or educational purposes only free of charge and without permission. However, the JVWR does not grant permission for use of any content in advertisements or advertising supplements or in any manner that would imply an endorsement of any product or service. All uses beyond research or educational purposes require the written permission of the JVWR.

Authors who publish in the Journal of Virtual Worlds Research will release their articles under the Creative Commons Attribution No Derivative Works 3.0 United States (cc-by-nd) license.

The Journal of Virtual Worlds Research is funded by its sponsors and contributions from readers. If this material is useful to you, please consider making a contribution. To make a contribution online, visit: <http://jvwresearch.org/donate.html>



Journal of • Virtual Worlds Research

jvwresearch.org ISSN: 1941-8477

**Volume 2, Number 3
Technology, Economy, and Standards
October 2009**

Supporting Soundscape Design in Virtual Environments with Content-based Audio Retrieval

By Jordi Janer, Nathaniel Finney, Gerard Roma, Stefan Kersten, Xavier Serra
Universitat Pompeu Fabra, Barcelona

Abstract

The computer-assisted design of soundscapes for virtual environments has received far less attention than the creation of graphical content. In this “think piece” we briefly introduce the principal characteristics of a framework under development that aims towards the creation of an automatic sonification of virtual worlds. As a starting point, the proposed system is based on an on-line collaborative sound repository that, together with content-based audio retrieval tools, assists the search of sounds to be associated with 3D models or scenes.

Keywords: content-based; audio retrieval; freesound; virtual worlds; soundscape.

Supporting Soundscape Design in Virtual Environments with Content-based Audio Retrieval

By Jordi Janer, Nathaniel Finney, Gerard Roma, Stefan Kersten, Xavier Serra
Universitat Pompeu Fabra, Barcelona

Virtual worlds are primarily populated with 3D models of real world objects and spaces. While the graphical representation of virtual objects has been extensively addressed, the representation of the sounds they produce is less well supported in currently popular virtual worlds. One example of the imbalance between graphical and sonic content is Google's 3D Warehouse initiative, which serves as a repository of 3D models that can be integrated in virtual worlds. This situation may lead to ending up with visually appealing but sonically poor virtual worlds.

Generating the soundscape of a virtual environment is still a tedious manual process. To add sound to a virtual object, the designer needs either to find an appropriate sample from a sound effects database, or to adjust a large number of synthesis parameters in the case of physical modelling.

Instead, we propose to use a large on-line collaborative sound repository that, together with content-based audio retrieval tools, can automate the sonification of virtual worlds. Our framework, currently under development, assists the search of sounds associated with 3D models and scenes, partly by relating text queries to social tags in the sound database, and partly by ranking search results using concepts borrowed from ecological acoustics.

Characterization of soundscapes

The design of sound in virtual environments (VE's) relies on the techniques and traditions of sound design for film and video games (Chion M., 1991). Sound effects are typically created by foley artists or obtained from commercial sound effects databases. With the popularization of internet-based and socially oriented virtual environments, sound design faces new challenges and opportunities. Users generate their own objects and sounds are produced in their interaction with the virtual environment and other users. For this process to be automatic we need to automatically characterize a given soundscape and search for sounds that best fit that characterization.

Soundscape classification can be addressed from different perspectives. A classification scheme based on the physical characteristics of the produced sound was proposed by Pierre Schaeffer (1966), which categorizes sounds using three pairs of criteria: (1) *Masse*, which is a 'fuzzier' generalization of pitch; (2) *Facture*, which is an energy envelope; (3) *Durée/Variation*, or duration and variability; and finally, the more subjective *Équilibre/Originalité*, which is related to the complexity of the signal. Originally published in 1977, R. Murray Schafer (1994) distinguished three types of sounds within a soundscape: keynote sounds, signals and soundmarks. Schafer also proposed a classification of sounds based on the reference to the source: Natural, Human, Sounds and Society, Mechanical, Quiet and Silence, and Sounds as Indicators. More recently, Gaver (1993a, 1993b) has contributed to create a solid framework for ecological acoustics. He proposed a taxonomy of environmental sound, providing specific categories for sounds considering whether they are generated by solids, liquids or aerodynamics.

Some systems have already addressed the automatic generation of soundscapes by using existing sound classifications. Using a lexical database, a system presented by Cano et al. (2004) generated a complete ambiance combining sound snippets related to a high-level concept (e.g. “beach”). This system used a structured commercial sound FX database, but an application to virtual worlds could not be done from that since there is a lack of correspondence to the actual objects and the generated soundscape. A recent approach to sound retrieval by Chechik, G. et al. (2008), proposes ranking the results of text queries using content-based audio retrieval techniques. While useful for general audio search in structured and unstructured databases, this method doesn't take into account the specifics of sound design. Therefore, it is still limited for the purpose of creating virtual world soundscapes.

Use of collaborative sound repositories

The principal contribution of the proposed system is the use of content-based audio retrieval from online collaborative sound repositories, employing concepts from ecological acoustics. Given appropriate interfaces, users, or the actual system, could rapidly find the appropriate sounds for 3D models through web-based search, which would facilitate the creation of soundscapes for virtual worlds.

In terms of technology, the proposed system benefits from content-based audio retrieval algorithms. Repository sounds are labelled with user-generated tags called folksonomies (Martínez, E. et al., 2009), which result in an unstructured database. For all sound in the database, a number of acoustic descriptors are automatically extracted. Searching for a sound associated with a virtual object starts with a text query. The system uses the Wordnet lexical database (Fellbaum, C., 1998) to semantically relate the query with the tags of the sound repository. Search results are ranked according to an ecological acoustics taxonomy (e.g. solid, liquid, gas). Ranks for each sound in each of the concepts in the taxonomy are obtained using automatic audio analysis and machine learning classification. In initial experiments, we used the state of the art Support Vector Machine (SVM) classifier LIBSVM by Chang C. and Lin C. (2001). These experiments show that given a sufficient number of examples, a few descriptors suffice to produce reasonable results using this approach.

User-generated media might represent an important factor in the expansion of virtual environments. Most popular virtual environments allow users to create and furnish their own spaces. We argue that closed commercial Sound FX databases do not fit into this model—on the one hand because of the prices and licenses associated with their use, and on the other, because they cannot be augmented by users. Therefore, our system uses Freesound.org (2005) as a collaborative sound repository, which currently offers over 70,000 sound snippets under a Creative Commons license.

Bibliography

- Cano, P. et al, (2004). Semi-automatic ambiance generation, in Proceedings of the Conference on Digital Audio Effects, Naples, pp. 319–323.
- Chang C. and Lin C. (2001). A library for support vector machines. Retrieved June 2009, from LIBSVM Web Site: <http://www.csie.ntu.edu.tw/~cjlin/libsvm> .
- Chechik, G. et al. (2008). Large-Scale Content-Based Audio Retrieval from Text Queries, in Proceedings of MIR'08, Vancouver.
- Chion. M.(1991). L'audio-vision (son et image au cinema), English translation: Audio-vision, Sound on Screen. Armand-Colin.
- Fellbaum, C. (Ed.) (1998). WordNet: An Electronic Lexical Database Cambridge, MA: The MIT Press (Language, speech, and communication series). <http://wordnet.princeton.edu/>
- Freesound.org (2005)... Retrieved June 2009, from Universitat Pompeu Fabra Web Site: <http://www.freesound.org>
- Gaver, W. W. (1993a). What in the world do we hear? An ecological approach to auditory event perception, in *Ecological Psychology*, vol. 5, no. 1, pp. 1–29.
- Gaver, W. W. (1993b), How do we hear in the world? Explorations of ecological acoustics, in *Ecological Psychology*, vol. 5, no. 4, pp. 285–313, .
- Martínez, E. et al. (2009). Extending the folksonomies of freesound.org using content-based audio analysis, in Proceedings of the Sound and Music Computing Conference, Porto.
- Schaeffer, P. (1966). *Traité des objets musicaux*. Paris: Editions du Seuil.
- Schafer, R.M. (1994). *Our sonic environment and the soundscape: the turning of the world*. Rochester, VT: Destiny Books.