Curating Ethnomusicology in Cyberworlds for Ethnomusicological Research: World Music in Wonderland

Michael Frishkopf, Michael Cohen et Rasika Ranaweera

Exposer les paysages sonores
Volume 37, numéro 1, 2015

URI : id.erudit.org/iderudit/1039658ar
https://doi.org/10.7202/1039658ar

Résumé de l'article

Nous décrivons ici un cybermonde musical qui est à la fois un espace virtuel de conservation ethnomusicologique et un espace de recherche : l'ethnomusicologie des cyberespaces musicaux contrôlés. Notre cybermonde diffère de la plupart des collections musicales en ligne en ce qu'il permet une expérience sociale d'immersion. Si l'on considère qu'une telle exposition numérique de recherche ethnomusicologique est en elle-même une forme de pratique sociale et musicale, cela exige également de nous que nous fassions l'ethnomusicologie de telles expositions. En général, la recherche en ethnomusicologie s'est effectuée sous la forme de travaux qualitatifs en situations non contrôlées. En contraste, nous concevons ce cybermonde musical d'usage comme un laboratoire ethnomusicologique virtuel, une plateforme de recherche ayant pour finalité de meilleures façons de concevoir les expositions de musique en ligne pour la découverte, l'apprentissage et la contemplation esthétique, ainsi que pour la contribution à notre compréhension d'ensemble du rôle de la musique dans l'interaction humaine et la formation des communautés.
CURATING ETHNOMUSICOLOGY IN CYBERWORLDS FOR ETHNOMUSICOLOGICAL RESEARCH

World Music in Wonderland

Michael Frishkopf
University of Alberta
Michael Cohen
University of Aizu
Rasika Ranaweera
Aizu Laboratory Inc.

Introduction: Ethnomusicology of, and through, cyberworlds

Ethnomusicology can be defined as the branch of the human sciences that studies music in its social-cultural contexts, especially the ways in which people interact through shared musical experience and discourse about music, and how music thereby facilitates the emergence of social groups and communities (Nettl, 2005). Methodologically, ethnomusicology centers on qualitative research, mainly ethnographic fieldwork relying upon participant-observation and informal interview techniques (Fine, 2001; Barz and Cooley, 2008). Variables typically cannot be controlled.

Cyberworlds open new avenues for ethnomusicological research. A cyberworld is an online social space, with implications for real-world social interaction and culture-formation. Cyberworlds can model the real world, but are also embedded within it. Cyberworlds are thus of tremendous interest to many scholars working in the social sciences and the humanities (Kong, 2001; Taylor, 1997). As cyberworlds incorporating music become increasingly prominent (especially in multiuser videogames), the task of studying them falls to ethnomusicology. The ethnomusicologist seeks to comprehend social dimensions of musical cyberworlds, to enhance their musical functions, and to further understand music in social-cultural contexts more generally, since cyberworlds are closely related to the real world, and impact it strongly.
Indeed, this task has already begun, with several ethnomusicological studies of online communities and virtual gaming (Lysloff, 2003; Miller, 2007), as well as reflections on the virtual fieldwork enterprise (Cooley, Meizel and Syed, 2008). However, until now ethnomusicologists have studied “naturally occurring” cyberworlds, rather than constructing cyberworld laboratories expressly for research. For the most part, ethnomusicology has not relied on controlled experimentation at all. Subject matter, methodology, and technological limitations have largely precluded ethnomusicologists (like historians) from this sort of scientific research, by which variables may be manipulated and their relationships examined.

Now it is not only possible to build a cyberworld as the focus for ethnomusicological research, but necessary as well, since cyberworlds now comprise a significant component of contemporary sociomusical reality, including musical social media and multiplayer videogames. Musical cyberworlds can enable a new paradigm for ethnomusicology. Instead of observing musical interactions in the world-as-encountered, one can study a virtual world whose parameters are, to a great extent, under the researcher’s control. Such a cyberworld becomes a laboratory for ethnomusicological research, and a means of better understanding other musical cyberworlds, providing, for the first time, a controlled environment for ethnomusicology.

“World Music in Wonderland”¹ (hereafter “WMiW”) is a virtual reality groupware environment – a cyberworld in which each user is represented as an avatar, capable of walking/teleporting, conducting voice/text chats with other users, and listening to spatialized audio/music. Each track, positioned by a visual marker, broadcasts looped sonic content within an audio sphere (its “nimbus”). Within this virtual space (as shown in Figures 1, 2, and 3), each real-world user appears as (one or more) avatars. As in the familiar video game paradigm, avatars are capable of moving (walking, flying, or teleporting), communicating (via speech or text) with other users, listening to spatialized audio, and browsing metadata; real-world users receive sensory inputs corresponding to the immersive binaural experience of their corresponding avatars. WMiW is built upon “Open Wonderland,”² a pure Java framework (originally developed as “Project Wonderland” by Sun Microsystems, now supported by an independent foundation) for creating collaborative 3D virtual worlds (Kaplan and Yankelovich, 2011)

like “Second Life.” WMiW builds upon technology and experience resulting from an earlier Wonderland-based research project, Folkways in Wonderland (FiW), which enables virtual explorations of Smithsonian Folkways albums, positioned on a giant cylindrical map.

To enter the WMiW cyberworld, a user connects to a public server hosted over the Internet (currently deployed in Canada and Japan) using a web browser, and downloads our extended Wonderland client. After authentication, the user can explore music in multiple ways, including visually (dereferencing placemarks and bookmarks or browsing a map), auditorily (entering a track’s “nimbus” or sonic sphere, as described in Greenhalgh and Benford, 1995), and socially (through discussions with other users). The system is collaborative: multiple avatars can enter a space, audition track samples, and contribute their own sounds (typically speech) to the mix via voice chat. By default, avatars can directionally hear within a space all sound sources (musical tracks and sounds produced by other avatars), attenuated for distance and mixed according to a spatial sound engine that emulates binaural hearing. Avatar-represented users are free to explore the cyberworld, using keyboard and mouse/trackball/trackpad controls to navigate through the surrounding virtual environment.
(including galleries, streets, and nature, as seen in Figure 2), while interacting with one another and listening to music. When tracks are near each other, overlapping nimbus projections create a dense mix, which is appropriate when exploring an entire collection by moving one’s avatar among distributed songs. However, in order to listen to a particular track, an auditory focus function is available which causes other musical streams to be blocked. When the audition of music is disturbed by cacophony from nearby tracks, such “narrowcasting” operations can be invoked to refine one’s soundscape (Alam, Cohen, Villegas and Ahmed, 2009; Fernando, Adachi, Duminduwardena, Kawaguchi and Cohen, 2006). An exotic multipresence feature (Cohen, 2000) allows the user to be simultaneously represented in the cyberworld by more than one avatar, for radically flexible avatar deployment.

Related Research

Our system integrates various functionalities that are typically offered only separately by more specialized programs. In this section, we consider several classes of such focused applications.

Music Information Retrieval

Finding a particular recording is generally supported by traditional search interfaces via metadata (Hughes and Kamat, 2005), but there is a growing need for improving search techniques via different information retrieval strategies. Damm, Fremerey, Kurth, Müller, and Clausen (2008) introduced a novel user interface for multimodal (audiovisual) music
Curating Ethnomusicology in Cyberworlds

presentation as well as intuitive browsing and navigation. Many music search engines exist. For instance, Musipedia\(^7\) offers melody search functions. Similarly, the Music Ngram Viewer\(^8\) encodes songs for look-up. The Folktune Finder\(^9\) also has melody and contour search. MusicSim (Chen and Butz, 2009) uses audio analysis techniques and user feedback for browsing and organizing large music collections. Although most such applications and interfaces facilitate locating music and visualizing collections, it is also important to take into account what information is desired and how that information will be used after retrieval (Downie, 2002). The mobile music player by Kuhn, Wattenhofer, and Welten (2010) incorporates several smart interfaces to access larger personal music collections and visualize content using similarity maps.

**Social (Distributed) Music Audition**

Many research systems have been developed for music consumption, both stand-alone and distributed, of which work by Frank, Lidy, Peiszer, Genswaider, and Rauber (2008) is representative. Boustead and Safaei (2004) compare various architectures for delivery of streamed audio, including techniques for optimization based on similarity of distribution of avatars in a virtual space with that of human players in the real world. Such groupware systems are instances of collaboration technology for synchronous but distributed (not collocated) sessions.

The major commercial labels haven’t fully capitalized on the way many people really consume, share, and experience digital music. Napster anticipated distributed music sharing, but presented an asynchronous experience. Many people, especially younger listeners, enjoy music through networked music audition services. Such systems often offer social media features, generalized as “groupware” among human-computer interaction researchers and scientists. For instance, Last.fm promotes “scrobbling,” publishing one’s music-listening habits to the Internet, to monitor when and how often certain songs are played, but such journaling is an asynchronous practice. SongPop\(^10\) is a social multiplayer online music identification game, in which players compete against others in real time to identify song snippets. (In 2012 it was the highest-rated game on Facebook.\(^11\)) Both

\(^7\) http://www.musipedia.org
\(^8\) http://www.peachnote.com
\(^9\) http://www.folktunefinder.com
\(^10\) http://www.songpop.fm
\(^11\) http://en.wikipedia.org/wiki/SongPop
Shazam\textsuperscript{12} and SoundHound\textsuperscript{13} feature real-time maps of music neighbors and other users are listening to, as “My Music” and “Explore,” respectively.

In the future, online communities, currently used primarily for interactive 3D social interaction and online video games, will be increasingly used for browsing media, listening to live performances, or even performing together. The primary example of such a not-quite-mainstream environment is Second Life, which allows virtual concerts and runs from a distributed network of 40,000 servers (but might eventually be eclipsed by its founder’s subsequent immersive environment venture, High Fidelity\textsuperscript{14}). Although network and processing latency precludes a totally satisfying real-time experience for globally distributed online musicians, prerecorded tracks (such as those served by WMiW) can be streamed for a “concert-like” experience. Boustead, Safaei, and Dowlatshahi (2005) considers server-side optimization of compiled soundscapes, including accommodation of limited bandwidth and soundscape compilation distribution to clients for load-sharing. For a perfect network, running at the speed of light, packets would take about 100 ms to get halfway around the world (“worst best case”), which delay would be fine for conversations, but probably distractingly audible for distributed performance.

WMiW music search features are limited to text-based search on its tracks’ metadata tags. What distinguishes WMiW from the above-described applications is its multimedia, social character: collaborative music audition, integrated text chat, voice chat, spatial music rendering, and figurative presence and natural spatial navigation, for real-time, interactive, dynamic consultation and an immersive experience.

Our system is an instance of social music browsing, or distributed music audition, allowing collaborative music exploration and ethnomusicological journeys, realizing some of Alan Lomax’s vision of a “Global Jukebox” (Lomax, 1997). Crossing groupware social audition with music information retrieval yields collaborative music information seeking, which is what WMiW is intended to foster.

**WMiW Architecture**

Our music browser is implemented as a module in Open Wonderland. The Wonderland framework consists of the “Darkstar” game server, which

\begin{itemize}
  \item \textsuperscript{12} http://www.shazam.com
  \item \textsuperscript{13} http://www.soundhound.com
  \item \textsuperscript{14} https://highfidelity.io
\end{itemize}
provides a platform for Wonderland to track the frequently updated states of objects in the world, and ‘VoiceBridge,’ a pure Java open source audio mixing application, which communicates directly with the Darkstar server, providing server-side mixing of high-fidelity, immersive audio (Kaplan and Yankelovich, 2011).

Artistic, geographic, audio-related, and generic information describing the “Diversity Cape Breton” music collection is curated in XML (Extensible Markup Language) format, an open standard maintained by the W3C (World Wide Web Consortium) for interoperable unicode documents (Lam, Ding and Liu, 2008), conforming to MX: IEEE 1599 (Baggi and Haus, 2009), a comprehensive, multilayered music description standard. Music has been notated and annotated for centuries in many cultures with symbols, but modern attempts have been made to create standards based on XML. MML (Music Markup Language) is a syntax for encoding different kinds of music-related information, whereas MUSICXML is designed for the exchange of

15. https://www.w3.org
17. http://www.musicxml.com
scores.

Figure 4. Metadata display window: The ‘Track Details’ tab shows song information, the ‘Playlist’ tab displays the entire collection as an outline, and the ‘History’ tab lists tracks visited by the user. Other operations, invoked by buttons at the bottom, allow (left to right, respectively) auditory focus by auditioning a single track, browsing selected track information at the web portal,18 teleporting to the origin of a track, muting a track, and ‘opening map window’ with an OpenStreetMap (OSM) to provide detailed, zoomable, topographic information. Narrowcasting state of a track (mute and solo) can be overridden by another narrowcasting operation. Narrowcasting state can be checked by clicking on a track or selecting a track from the “history” tab in the metadata window. When a mute operation is conducted, the icon changes to a slashed speaker and the album frame alters its color to red (when the track is clicked again). Subsequent clicking the button toggles the mute state. Similarly, the respective album frame becomes green and the headphone icon changes when the solo operation is invoked.

Mx, standing for musical application using XML, inherits all the features of XML — including inherent human-readability, extensibility, and durability (Ludovico, 2009) (Baratè, Haus, Ludovico and Perlasca, 2016) — and unifies features of MML and MusicXML with some additional features, including the concept of layers. The six Mx layers, which allow integrated representation of several aspects of music, are: General, Logic, Structural, Notational, Performance, and Audio, described as follows:

- General – music-related metadata, including title, author, date, genre, performance, and recording information (as shown in Figure 4, left section)

• Logic – music description from a symbolic point of view

• Structural – identification of music objects and their mutual relationships

• Notational – graphical representations of a score

• Performance – parameters of notes played and sounds synthesized, specified by performance languages (Russo, 2008) such as MIDI

• Audio – digital or digitized recordings of the piece.

Even though the “Diversity Cape Breton” curation has no information corresponding to the Mx logical, structural, notational, or performance layers, the schema allows empty layers (Ludovico, 2008), and there are no restrictions preventing browsing of other music collections when such information is available (as seen in Figure 5). Note that layers may contain URLs as well as directly accessed data, for extra flexibility.

Discussion: Curation as research platform

Using World Music in Wonderland, a cyberworld for curating ethnomusicology, as a virtual laboratory, we pose the following question: How do social actors, represented by avatars, interact in such an immersive cyberworld when presented with a specific collaborative task? A laboratory environment enables us to control variables and thus answer — at least within this restricted domain — questions about such dependencies with rigor that cannot be achieved in the real world, through data gathering either from the panoptic perspective of system administrator, or from the narrower but deeper immersive perspective of embedded fieldworker: the participant-observer qua avatar. In particular, we are concerned with understanding the relations between two primary clusters of independent variables known by ethnomusicologists to shape the emergence of musical community: the social and the musical. Here, social variables include the number and demographic profiles of participants whose avatars inhabit the cyberworld, while musical variables include the number and kinds of music tracks that populate it. Variables within either cluster can be manipulated: the former through participant selection, the latter by loading different collections of music tracks into WMiW.
Conclusion & Future Research

We have presented a novel application for listening to world music inside a virtual space. Rather than finding tracks using traditional interfaces, an avatar- or avatars-represented user can explore music immersively while adjusting their soundscape with narrowcasting. Users can invoke mute or solo functions to listen only to particular songs when cacophony might distract.

Research will, at the outset, be exploratory, but we anticipate that the present phase of cyberworld building and observation will lead to the formulation of hypotheses and, subsequently, more focused experimentation designed to test them. We believe that this process will produce results suggesting better ways of designing musical cyberworlds as a means of ethnomusicological curation, as well as a site for ethnomusicological research, a laboratory where broader principles underlying the role of music in human interaction and community formation can be studied. Controlled research in and about a custom-built musical cyberworld can
usefully supplement, though never supplant, traditional real-world fieldwork in ethnomusicology.

Acknowledgment

Development of WMiW was funded primarily by the Social Sciences and Humanities Research Council of Canada, with additional support from folkwaysAlive! at the University of Alberta.
References


