

A Virtual World Prototype for Interacting with a Music Collection

Jukka Holm¹ and Arto Lehtiniemi²

¹ Tampere University of Technology, Human-Centered Technology, P.O. Box 589, 33101 Tampere, Finland

² Nokia Research Center, Visiokatu 1, FIN-33270, Tampere, Finland
jukka.holm@tut.fi, arto.lehtiniemi@nokia.com

Abstract. This paper studies the idea of using virtual worlds as an interface to access music collections or music recommendation services. In the evaluated 3D prototype, the user moves around a city and interacts with various characters, the looks of which reflect the musical style that they are recommending. In a user study with 41 participants, it was learned that the concept did not fit well to most users' music consumption habits, and using the prototype for only browsing music was considered to be too tedious and boring. The application should contain more activity and interactive elements, and there should be a clear goal to strive for. To develop the concept further, a better solution would be to modify an existing game or virtual world application by adding music recommendation capabilities.

Keywords: Music, visualization, character, avatar, virtual world, playlist, music recommendation, musical genre, user experience, AttrakDiff, Demola.

1 Introduction

A virtual world is a persistent and shared online environment, where the users can interact and socialize with each other in real-time [3]. Virtual worlds are used for various purposes including gaming, social networking, education, marketing, and advertisement. In these applications, the user is typically represented with a human-like avatar that (s)he can move around the world interacting with other characters and different objects.

The potential of virtual worlds has also been noticed in the music industry. Second Life [12] includes various clubs and radio stations playing different musical styles, and it has become a popular place for live concerts where the performance of a real-life artist is streamed to the virtual world and performed by representative avatars. Since 2006, major artists such as U2 and Duran Duran have performed in Second Life [1]. In the academia, avatars have been used to access music recommendation services and to summarize users' musical preferences [5, 9, 10].

In this paper, we propose the concept of using virtual worlds as an interface for accessing music collections or music recommendation services. The virtual world can be, e.g., a city consisting of numerous suburbs, buildings, vehicles, and characters.

Each city element works as an interface to a song or album in the music collection or some specific music player function. For example, different looking characters could recommend different type of music from the service, speeding a car could fast-forward a song, shooting a character could delete a song, and each bar or discotheque could represent a different musical genre. The world could also be dynamic; old buildings (representing rarely listened albums or genres) could be torn down, new bars open, and so on.

To study the idea in practice, a simplified off-line version of such an application was developed for the PC platform. In the prototype, the user moves around a city and interacts with various characters to receive new music recommendations. The look of each character reflects in some way the musical style that it is recommending. This paper discusses the prototype implementation in more detail, presents the results of a user study with 41 participants, and proposes a broad range of ideas for the future development of the concept.

2 Prototype

The virtual world prototype was implemented as a Demola [4] student project for the PC platform. As creating an extensive virtual world (such as the one described in the Introduction) would have required significant resources, our approach was to make a simplified off-line prototype to test the feasibility of the overall concept. The prototype was implemented as a Flash application to enable the use of different platforms including personal computers and mobile devices such as Nokia N900. As suitable mobile devices or Apple iPad were not available at the time of implementing the prototype, the main test platform was a HP Pavilion tx2500 touch-screen laptop.

In the prototype (Fig. 1), the user moves around a city using the arrow keys and interacts with various characters by pressing the enter key. While each character can move freely within the city limits, they spend most of their time in their natural surroundings (e.g. the heavy metal character is typically located close to the black and



Fig. 1. Screenshot from the prototype

metallic buildings). When interacted with, each character recommends new music from a genre that is somehow related to the looks of the character. The user is able to listen to the songs and add them to his/her playlist by pressing the tab key. The user can also enter any house in the city, but due to project's time limitations all houses contain the same room (a game hall) and the user cannot do anything in the room.

To simplify the prototype implementation, the number of characters was limited to eleven. From our research's point view, the biggest problem with the original Demola prototype was that several characters had visual characteristics that represented other musical genres than the playlist linked to the character. Thus, we decided to modify some of the characters and change most of the music. Due to implementation details, only minor graphical modifications such as changing colors and adding symbols on top of the characters were possible. Unfortunately, we were not able to change the body shapes, clothing, and other such features selected by the students.

Fig. 2 shows the eleven modified characters. These included (from top left to bottom right): 1. female with a cyan dress (representing dance music), 2. female with a red and black dress (grunge music), 3. female with a pink dress (pop music), 4. female with black leather clothes, strong make-up and skull decorations (heavy metal music), 5. "Raiden" – male character with Eastern and ninja elements (world music), 6. main character from the Blade movie (soundtrack music), 7. Homer Simpson (soundtrack music), 8. Megaman (game music), 9. Astroboy (cartoon music), 10. "Princess Li" – female character decorated with Eastern symbols (world music), and 11. "Question mark blade" – Blade with different clothes and question marks on the chest and the forehead (random or miscellaneous music).



Fig. 2. Characters used in the prototype

3 Research Method

To study how well the prototype works in practice, we arranged a user study with 41 participants. The selected research method was a combination of observation, semi-structured interview, and questionnaires. The prototype was tested in various locations such as participants' own homes, authors' homes, as well as the premises of Tampere University of Technology and Nokia Research Center. Each session lasted

for about 30 minutes. All participants tested the software for the first time, and they were not informed about how the prototype works or what they should do with it.

In the beginning of a session, the users had to fill in a short background information questionnaire. After that, the users were able to use the software freely, move within the city, and interact with the characters. When meeting a character, the users were asked to describe it with a couple of words and describe what type of music comes to their mind when looking at the character. After the user had listened to the associated playlist, (s)he was asked to rate how well the music corresponded with the looks of the character on a seven-point Likert scale. Due to the time limitations of the user study, only four characters per user were studied.

After rating the characters, a short, semi-structured interview was conducted. The participants were asked to describe their first impressions on the prototype, and various quality aspects of the prototype were studied by asking questions and filling in a questionnaire. Both hedonic and pragmatic aspects [6] were studied, and nine users also filled in the AttrakDiff questionnaire [1].

The Finnish participants were selected using convenience sampling from varying age groups and education levels. 78% were male and 22% female. 7% (3 participants) were between 12 and 17 years old, 12% (5) were 18-30, 56% (23) were 31-40, and 25% (10) were 41-55. All participants from the 12-17 age group were male, and the average age was 34.5. 61% were university graduates and the rest were split between various types of education such as elementary school, college, and university students. 7% were professional and 49% hobby musicians.

51% of the participants considered listening to music as their hobby. All except two listened to music at least once a day, and the average time per day was 2.1 hours. During listening, 20% did not usually do anything else but concentrate on the music. Popular listening contexts included home (55%), car (45%), school, work, sports, and the like. The most popular genres included pop (liked by 85% of the participants), rock (81%), metal (63%), soul, rnb & funk (61%), and alternative (56%), while the least popular genre was gospel (5%). 43% of the participants searched actively for new music to listen to. The most common sources included the Internet (65%), friends (63%), magazines (33%), radio (33%), and Spotify (20%). 88% had used YouTube for listening to music, 51% had used Spotify, 32% Last.fm, and 29% iTunes Genius.

4 Results

Based on the user study, we were able to answer our initial research questions and also received several ideas for the future development of the concept. In the following, the results are discussed in a detailed manner.

4.1 Quality Aspects

In the end of the interview, various quality aspects of the prototype were studied by asking questions and filling in a questionnaire. In the questionnaire, the participants had to rate nine different aspects of the software on a seven-point Likert scale (1=totally disagree, 7=totally agree). Fig. 3 shows the results using boxplots. Unless otherwise noted, the ratings did not correlate with participants' age, gender, or level of experience with music recommendation systems. We also experimented with the Chi-Square test, but the conditions for the test were not met in any of the cases.

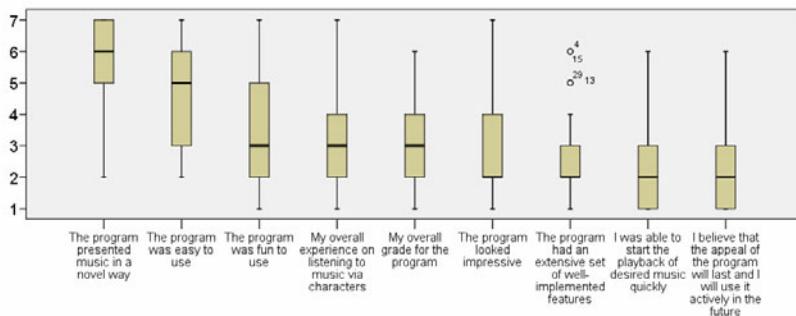


Fig. 3. Boxplots for the quality aspects of the prototype

On the positive side, the participants considered the concept to be a novel way to present music (median 6.0). During the interview, comments such as “*Great idea*” and “*A promising concept but very draft implementation*” were common.

Most participants also felt that the prototype was easy to use (median 5.0). The features were easy to find, and in general help files or tutorials were not needed. However, some participants did not fully understand the basic idea before it was explained to them. Two persons thought that the application was a first-person shooter game and were wondering where all the action was.

44% of the participants considered the prototype to be more fun than the traditional music players. However, for the statement “The program was fun to use”, the median fell on the “negative” side (3.0). There was moderate correlation with age (Spearman’s $\rho=0.44$, significant at the $p<0.01$ level (2-tailed)), i.e., older participants did not consider the prototype to be as fun as the younger participants.

The graphics were considered to be old-fashioned (“*retro feeling*”) or simply bad. Many participants also felt that the female characters were too sexual. As a result, the median for “The program looked impressive” was as low as 2.0.

For the statement “The program had an extensive set of well-implemented features”, the median was also 2.0. When asked about the negative sides of the prototype and/or the idea, the following issues came up: difficulties and required time in finding characters representing desired music, “nothing to do” while moving in the city, and the lack of textual search function (requested by 44% of the participants). In addition, 88% felt that this type of application does not fit their current music consumption habits. When listening to music, they either just want to concentrate on the music and do nothing else, or the music plays in the background while something else (e.g. reading or driving) is being done. Several participants commented that they would not use this type of music player even if the implementation was much better.

Finding music using characters was considered to be too tedious (“*It takes a lot of time to listen music with this approach: too much running, too little to do, too few people*”). As a result, the median rating for the statement “I was able to start the playback of desired music quickly” was only 2.0. During the interview, it became clear that people were not willing to spend extra time looking for music. Instead, the process should be as fast and easy as possible.

While the median for lasting appeal was as low as 2.0, the results correlated clearly with age (Spearman's $\rho=0.55$, significant at the $p<0.001$ level (2-tailed), i.e., younger participants were more interested in using the prototype longer than the older ones. Only 22% mentioned that they would like to use the prototype for a longer time.

When asked the question "Would you like to use this type of application rather than traditional music player applications?", only one participant gave a positive reply. 37% answered that they could use a virtual world to complement other applications. As this is in conflict with the 22% of lasting appeal, some participants may have referred to an imaginary future product instead of our prototype. The interest in using virtual worlds to complement traditional music players decreased with age (Spearman's $\rho=-0.55$, significant at the $p<0.001$ level (2-tailed)), and some participants even stated that they were too old for using this type of software.

The medians for "My overall grade for the program" and "My overall experience on using characters to listen to music" were both 3.0. However, the mean for the latter was 0.4 higher as some participants considered the idea to be slightly better than the actual implementation. The overall grade correlated clearly with age (Spearman's $\rho=-0.52$, significant at the $p<0.01$ level (2-tailed)) and somewhat with gender ($\rho=-0.32$ with $p<0.05$). In other words, older participants and women did not like the prototype as much as younger participants or men. Moderate correlation between the overall experience and age was also found ($\rho=-0.45$ with $p<0.05$).

In addition to our own questionnaire, nine users also filled in the AttrakDiff questionnaire [1]. AttrakDiff is a tool for evaluating the pragmatic and hedonic quality of products, and it consists of 28 evaluations on a seven-point semantic differential scale. According to the used online tool, our prototype was considered to be simple, inventive, creative, and innovative. In terms of pragmatic and hedonic quality, the user interface was rated to be neutral and it just met ordinary standards. The results also suggested that "product's attractiveness value is located in the average region" and "the overall impression of the product is moderately attractive."

The results seemed to be well in line with the results derived from the interview and own questionnaire. As the use of the AttrakDiff method did not seem to bring any extra benefit to our study, other participants did not fill in the questionnaire. Also, we found it quite tedious and unpractical to translate the terms used in the tool to those participants who were not fluent with English.

4.2 Characters

After the users had listened to some of the songs linked to a character, they were asked to rate how well the looks of the character matched with the music using a seven-point Likert scale (1=no match, 4= neutral, 7=perfect match). As mentioned earlier, we had very limited means to affect the looks of the characters in the development phase. This was also reflected in the results (Fig. 4), which were not as promising as in our two avatar related studies [9, 10].

Before listening to the music, the users were also asked to describe the character with a couple of words and guess what type of music it could possibly represent. These first impressions revealed many reasons for the poor ratings of some characters.

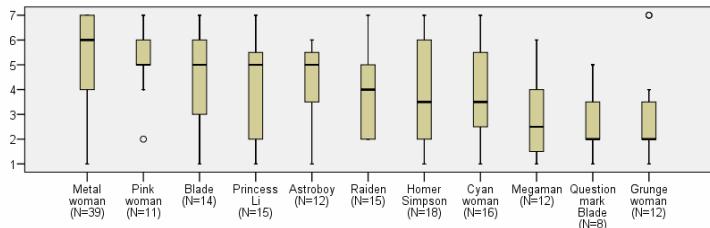


Fig. 4. Boxplots for how well the music and the characters match. In the case of each character, N refers to the number of the participants who met and rated the character.

The winner of the study was the heavy metal character, which relied on findings from our earlier research [7, 8]. Most participants (39/41) met the character immediately after starting to use the prototype, and felt that the looks matched well with the songs that the character was mapped to (median 6.0). Ratings less than four were explained by some participants who considered the character to be quite futuristic (or even a cyborg), and thus associated it with industrial, electro, or gothic music. Artist Marilyn Manson was mentioned quite a few times. Several participants would have preferred a male character to represent metal. To improve the design, the character should have looked more down-to-earth or like a male black metal fan. In fact, the latter approach was taken in [9] and [10] and proved to be very successful.

Another successful character was the pink woman representing pop music (median 5.0). One reason for this might have been the color selection, as in [7] it was found out that 48% of the participants associated pink with pop. The single outlier value is explained by one person who associated the character with clubs and trance music.

For unknown reasons, the cyan female character was often associated with pop music and the grunge woman with dance, techno, and hip-hop. The single outlier value for grunge is explained by a 12-year participant who recognized that grunge bands have often worn such shirts. The short dress was also associated with supermodels, movie stars, and RnB videos.

A major complaint for the pink, cyan, and grunge female characters was the body shape, which gained undesired attention from both genders. Many participants felt that the characters were too sexual (“*Why does she have such big tits?*” female 38), had silicone implants, and/or looked like prostitutes. Due to the exaggerated sexuality, many participants did not pay enough attention on the dress colors and other visual cues, and some even thought that the same character had just changed her dress. To improve the recognition rates, the body shapes should be less extreme and differ clearly from each other. Also, the characters should stay closer to their natural surroundings (e.g. pop fans outside a pop club) instead of moving around the city.

The main problem with the movie, cartoon, and game characters was that many participants did not associate them with soundtrack music. Instead, the answers were based on the looks of the character, or the participants imagined what type of music the character could listen to. Thus, Homer Simpsons was associated with reggae, funk, adult rock, 70's hits, children's music, country, and “funny” music, Astroboy was mapped to techno, punk, and funk, and Megaman to techno, pop, and metal. Blade was associated with genres such as metal, gangsta rap, hip-hop, and punk. The

main reasons for the metal association were the leather clothes and the black color, which is typically associated with metal music [7].

Raiden was commonly described with terms such as martial arts, ninjas, kung-fu, East, Asia, Japan, and Vietnam. However, as Raiden's world music playlist included also African songs, some participants decided to give it a negative rating. One participant (male, 33) commented that "*I understand what you are after here, but I have to give a bad rating as the playlist includes African music.*" On the other hand, many participants considered Princess Li to be quite confusing because the looks were a mixture of various cultures. In addition to world music, Li was associated with techno, female rock, hip-hop, and blaxploitation soundtracks. As a conclusion, one should not try to represent world music with a single character but there should be a different character for each country or region.

The "Question mark Blade" character was a complete failure (median 2.0). While the question mark is often associated with unknown music [8], it did not work well in this context. Participants focused more on the other properties of the character, and many did not even notice the question mark. The character was associated with various music styles such country, blues, rap, and funk, and those two participants who recognized the character just thought that it represented Blade as an old man.

5 Ideas for Further Development

During the interview, we also received several ideas for the future development of the concept. To start with, most participants felt that there should be more interactive content, characters, and things to do. Examples included fighting, destroying, driving, entering buildings, customizing one's own avatar, hidden features, and playing mini-games. Many participants were a little confused regarding to what to do in the city, and it was said that here should be a clear goal to be strived for while listening to the music. The virtual world could also be extended to include areas outside the city.

The city could be enhanced with different types of ambient sounds (e.g. speech and traffic sounds) and background music. These sounds could change when the user moves around the city and comes closer to the points of interest. Music played inside a club could get louder as the user walks closer, and there could be background music in shopping centers, elevators, and other places where such music is typically played.

The relationship between the characters and the music should be clearer, and the characters could show more hints on which musical style they represent. There could also be a setting for adjusting which musical genres are visible in the city. By default, different suburbs could represent different musical styles and characters could hang around outside the clubs.

The desired characters should be easier and faster to find. One way to accomplish this would be a map indicating the current location of the user and the movement of selected characters. Quick shortcuts were also seen as important for jumping to the next character in the city and for finding certain characters. While moving in the city, the user could toggle different view modes such as FPS or birds-eye.

Character playlists should be dynamic and change (for example daily) to provide different songs from the same genre. Certain areas of the city (such as the central square) could provide dynamically changing top lists of music listened to in the city.

By mixing different characters, users could create cross-genre playlists. The characters could also represent real people such as known artists with their own music recommendations.

The interaction with the characters should be improved. There could be a possibility to talk with the characters, and speech bubbles could appear on the top of them. Some participants were also interested in chatting with the other users of the system. It was proposed that the interaction could resemble the interaction between the characters in Habbo Hotel and Sims.

Some participants felt that the prototype did not include enough information on the songs. Thus, when a song/artist has been selected and is being played, the system could retrieve additional information from Wikipedia or other web page. The virtual world could also give access to exclusive content such as remixes of popular songs.

6 Conclusions and Future Work

In this paper, we studied the idea of using a virtual world as an interface for accessing music collections or music recommendations services. In the simplified prototype implementation, different characters acted as an interface to access the music. The looks of each character reflected the musical genre that they were mapped to.

The prototype was evaluated by 41 persons. Due to the poor implementation of the prototype and the use of only Finnish participants, the results are only indicative. Still, based on our understanding, it is not worth the time and money to start implementing game or virtual world –like features on top of an existing music application.

Firstly, exploring music collections using the prototype did not fit well with the music consumption habits of the participants. Most people prefer to concentrate only on the music, or the music plays in the background while something else is being done. Finding new music should be as fast and easy as possible, and thus using the prototype was considered to be too tedious.

Secondly, providing only music recommendations in a virtual world environment does not provide enough entertainment and value for the users. To make a music-related virtual world application more appealing, it should also include a lot of action (e.g. fighting and driving) and interactivity in addition to the basic music player features. There should also be a clear goal to be strived for while listening to the music.

During the user study, we received several ideas for the further development of the concept. Should they be implemented, the prototype would start to resemble more and more games such as Grand Theft Auto, virtual worlds such as Second Life, or social networking services such as Habbo Hotel. Thus, rather than building the virtual world from scratch on top of a music player application, a better solution would be to enhance an existing game or virtual world with music recommendation and searching capabilities. This solution would fulfill the users' need for more action and interactivity, and minimize the financial risks of developing a new type of service. The resulting application could also evoke the interest of gamers and people that are not active music listeners.

In addition to the music related features, the user experience of such an application would depend on the genre and implementation details of the original application. As

one cannot speculate how such a combination would be perceived by the users, this would be an interesting topic for a future user study that should preferably be done in collaboration with the developer of the original application.

References

1. Andrews, R.: Second Life Rocks (Literally),
<http://www.wired.com/techbiz/it/news/2006/08/71593>
2. AttrakDiff, <http://www.attrakdiff.de>
3. Book, B.: Moving Beyond the Game: Social Virtual Worlds. In: State of Play 2 Conference (2004)
4. Demola, <http://www.demola.fi>
5. Haro, M., Xambo, A., Fuhrmann, F., Bogdanov, D., Gomez, E., Herrera, P.: The Musical Avatar – A Visualization of Musical Preferences by Means of Audio Content Description. In: Audio Mostly. ACM Press, New York (2010)
6. Hassenzahl, M., Platz, A., Burmester, M., Lehner, K.: Hedonic and Ergonomic Quality Aspects Determine a Software's Appeal. In: CHI 2000 Conference on Human Factors in Computing Systems, pp. 201–208. ACM Press, New York (2000)
7. Holm, J., Aaltonen, A., Siirtola, H.: Associating Colors with Musical Genres. Journal of New Music Research 38, 87–100 (2009)
8. Holm, J., Holm, H.: Associating Icons with Musical Genres. In: International Conference on Internet and Multimedia Systems and Applications. ACTA Press, Calgary (2008)
9. Holm, J., Lehtiniemi, A.: Evaluating an Avatar-Based User Interface for Discovering New Music. In: 9th International Conference on Mobile and Ubiquitous Multimedia. ACM Press, New York (2010)
10. Holm, J., Siirtola, H., Laaksonen, L.: Associating Avatars with Musical Genres. In: 14th International Conference on Information Visualisation. IEEE, Los Alamitos (2010)
11. Lehtiniemi, A.: Evaluating SuperMusic: Streaming Context-Aware Mobile Music Service. In: International Conference on Advances in Computer Entertainment Technology, pp. 314–321. ACM Press, New York (2008)
12. Second Life, <http://secondlife.com>