Music Mood Wheel – Auditive Interfaces for Mobile Music Devices

Andreja Andric
State University of Milan
Via Comelico 39/41
20135 Milan, Italy
aandreja@dico.unimi.it

Pierre-Louis Xech
Microsoft Research Cambridge
7 J J Thomson Avenue
Cambridge, CB3 0FB, UK
pierre-louis.xech@microsoft.com

ABSTRACT

Digital mobile devices for listening to music nowadays contain thousands of musical titles. Searching for music content is becoming difficult and time-consuming. Many outdoors listening situations do not allow the user to distract the attention from the current activity and spend too much time in finding the song that fits best with the current mood and context. We are proposing a simple auditive interaction model and prototype, in order to make the selection of the desired music contents on a mobile device more comfortable and versatile in everyday life.

1. VISION, PROJECT AND CHALLENGES

We started from the idea that a stimulus for listening to music rarely comes up with a concrete name or song title: but rather with a wish to create a specific ambient, to reinforce a pleasant emotion, to fill up a space or to suppress boredom. Besides, the size of private music collections overwhelmingly exceeds a person's ability to recall which compositions comply best with the current mood [1]. In our project we challenge the "recall and search by title" procedure exhibited by most mobile music devices (like for instance the iPod) with the "radio wheel" metaphor [2]. When listening to a "vintage radio device", we simply turn the frequency wheel around until we hear something that catches the attention. As soon as we lose interest in the contents currently broadcasted, we choose another station, by moving the wheel again. Nevertheless, our approach does not follow the music similarity design pattern and it ignores textual metadata as well. One of the most exciting challenges was to check if the recent advances in extracting high level descriptors from the audio signal, like tonal strength, BPM or mean audio loudness estimation can contribute in achieving our vision. We were inspired by the work in automatic playlist generation [1], new interfaces for browsing music collections [2], music recommendation systems [3] and non-visual interaction with mobile devices [4].

In our project we plan to implement the "Music Mood Wheel" vision through the delivery of a working prototype targeting devices running Windows Mobile 5.0 (e.g. new generation of smartphones with hard drive). We developed first a rough prototype on a Tablet PC running Windows XP, and are now making comparative user experiments between this device, a Samsung Portable Media Center and the Apple iPod. We will therefore use this experimental corpus for designing and implementing the final prototype, now scheduled for being released mid 2006.

The Music Mood Wheel project started mid 2005 as a one year collaboration project between the State University of Milan, Computer Science and Communications department, and Microsoft Research Cambridge, External Research Office, Intelligent Environments Group [5].

2. OUR DEVICE AND ITS EXPERIMENTAL EVALUATION

Our rationale was to first approximate the new device with off the shelves PC technology (e.g an ultra light laptop) and then to target the "modern mobile device par excellence", i.e. the Smartphone. We developed a kind of a rough prototype that consisted of a tablet PC, a trackball and a pair of headphones. The user in the experiment carried the tablet PC in the sack on their back. Our software ran on the tablet PC, while the trackball and the headphones were connected to it with cables. In this way, the trackball played the role of a mobile device without screen. The trackball button was used for play/pause, while the ball was used for music search.

Instead of being organized as a tree of options, the songs were placed in an imaginary 2D space. The device was developed in three somewhat different variants with regards to what the axes of the 2D space represent. In each of the three variants, each axis represents a feature automatically extracted from the audio. The features regarded the rhythmical, harmonic and loudness-related aspects of music. Having this in mind, we chose the features that we felt might be relevant for the perception of music (like rhythm velocity) and that are transparent, in other words, easy to understand by an ordinary user.

The goal of the experiment was to establish whether the proposed device that offers search by audio content lets the user find the desired songs more quickly, and whether it lets the user find better songs.

We made a large joint collection from the songs collected from the users themselves. In this way, the collection

generally reflects what the social group of our participants listens to.

There were two groups of users. First group did the experiment using either Apple iPod or Samsung Portable Media Center (PMC). The second group worked with our trackball device.

As the idea was not to use song titles, or any other textual metadata, we decided to give the listeners a list of verbal descriptions of the music to look for. The descriptions loosely described certain emotions, sensations, or broad descriptions of music content, the kind of one might use in real life situation, for example: "I'd like something more joyful". These were: joyful, energetic, danceable, relaxing, funny, rhythmic, intimate, exciting, sensual, and melancholic. These ten descriptions were obtained from another group of participants by means of a questionnaire.

The experiment scenario was organized around a walk in a park during which the user tries to find ten songs out of a large collection (3500 songs) according to the ten different descriptions.

Each participant went through the list of descriptions in a different order. The orders of the descriptions for each participant were chosen at random. Each participant was told to find a song that (a) he/she likes, and (b) corresponds to the given description. After each search, the participant was asked to rate the songs on the scale 1-5. The participant could also give up the search if frustrated or feeling unable to find the desired song.

We had 40 participants that took part in the experiment. They were students of the University (19-27 years old). Their private music collections contain 4000 songs at average. The 77% of them play an instrument or participate in a musical activity.

From the obtained results, the Trackball devices perform better than traditional taxonomically based search methods (Figure 1).

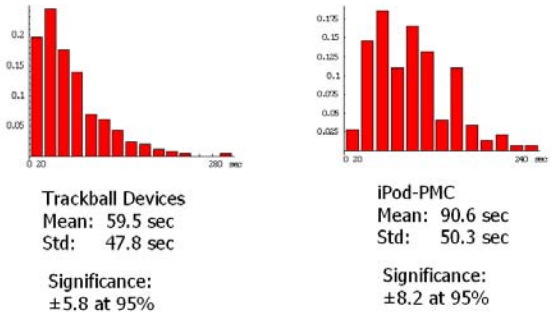

Figure 1. The distributions of mean times

The search times obtained with the Trackball device shows typical exponential decay with the main centred as normal distributions while iPod-PMC couple shows a random distribution with apparent linear decay.

The difference exists because of the different attitudes in the two types of search. Searching taxonomically (iPod-PMC) implies a typically linear search based on multiple random starting points. These starting points are based on previous personal knowledge and they are responsible for the randomness of distributions. On the other side searching a song by moving in a space ordered by a scale (Trackball device) is like cutting the search space into two parts, each time we listen to a song: one part where we know that we are not likely to find the song we are looking for, and the other part where we probably will find it.

On the contrary, we did not obtain any significant differences in mean ratings and in the number of situations where the user eventually gave up the search.

3. CONCLUSION AND FUTURE WORK

As a conclusion, we are encouraged in our approach and intuition - the eyes free, pure audio browsing approach is a promising path to follow.

The evaluation studies that include some kind of music preference measurement are rare; hence the value of our evaluation studies lies as well in establishing ground and gaining experience in measuring phenomena of such fugitive nature

The main objective of our next stage of experimental research will be to answer the following question: what is the relationship between a scale obtained by human rating of melody and rhythm, and the scale determined by the values of the parameters we used in the above described experiment?

In parallel, we are working on an advanced prototype on a smartphone running the Windows Mobile 5.0 Operating System.

REFERENCES

[1.] S. Pauws, B. Eggen, "PATS: Realization and user evaluation of an automatic playlist generator", in Proc. of the 3rd International Conference on Music Information Retrieval (ISMIR 2002), IRCAM – Centre Pompidou, Paris, France, October 13-17, 2002.

[2.] M. Podlaseck, R. Hoch, E. Schoenberg, "This Is Not a Catalog: Philip Glass's Work Online", IEEE MultiMedia, 2003. pp 4.-7.

[3.] P. Cano, M. Koppenberger, N. Wack, An Industrial Strength Content based Music Recommendation System, Proceedings of the 28th ACM SIGIR, 2005, August 15-19, Salvador, Brasil

[4.] A. Pirhonen, S. Brewster, C. Holguin, "Gestural and Audio Metaphors as a Means of Control for Mobile Devices", Proceedings of CHI 2002, April 20-25, 2002, Minneapolis, Minnesota, USA

[5.] Microsoft Research Cambridge Call for Proposals: "Create, Play and learn": Innovative technologies with social and cultural value. http://research.microsoft.com/ero/iep/CfProposals_awarded.aspx