A music puzzle game application for engaging in active listening

Kjetil Falkenberg Hansen[1,a)] Zheng Li[1,2] Hua Wang[1,2]

Abstract: In this paper we introduce an application for tablet devices with Android operating system called *The Music Puzzle*. This work is part of an ongoing projectcalled The Soundpark—Using modern smartphones to create interactive listening experiences for hearing impaired. In the Soundpark, we intend to provide different experimental applications for interacting with sound.
The aim of the presented study was to create interactive and game-inspired listening experiences for persons with hearing impairments (and possibly using hearing aids or having cochlea implants). Audio-based programs constitute a significant part of the Android Market, but the scopes of existing applications are limited. Modern smart devices open up new possibilities both in terms of using external information as input and providing real-time audio feedback to the user, and the Music Puzzle has a novel approach that explores the new possibilities.

1. Introduction

Hearing-impaired persons have different prerequisites from those with normal hearing for using audio-based products, entertainment and services, and in general enjoying music [2], [6]. Modern hearing aids meet many requirements to ensure an active participation in society, but these technologies are usually optimized for speech intelligibility. The Soundpark project (Ljudparken/The Soundpark–Using modern smartphones to create interactive listening experiences for hearing impaired, http://bit.ly/TaiH0I) was initiated to create interactive listening experiences through new sensor-based musical instruments and audio interfaces. The aim of that project is to provide better conditions to support persons with explorative applications for music and sound in mobile phones.

With the built-in sensors of smartphones and tablets, it is possible for users to manipulate recordings and explore sounds, as well as experiencing their auditory environment in new ways. Interactive listening applications will provide better conditions for taking part of commonly available audio-based entertainment; for instance, everyone now expect to have constant access to music through their devices. Music has been shown to have direct links with a person's well-being, social development, language, and several other aspects [11]. The challenges are to include hearing impaired with different prerequisites and help this group to engage in sound environments and services.

From other studies [3], we have seen that a common reflection from professionals working with hearing loss patients is that many choose to deactivate their hearing aids and only use these for limited amounts of time. Thus, there is a risk that developing "hearing skills" necessary for interpreting the signals from the aid is affected [5], [7].

Computer games is commonly known to have great potential as educational and training tools (see for instance [10]). Can an easily accessible and attractive audio-based game increase the amount of time of active listening among the hearing impaired? With the Music Puzzle, we contribute to the area of HCI targeting the *serious game* category.

2. Method

The development work follows a classic prototyping model described in [1] combined with target user group interviews and questionnaires. These inputs inspired, together with relevant literature on music therapy and new devices used in such circumstances [4], [8], [9], a game-like design based on audio. The interaction is mainly through physical interaction with the device, but to solve the task, focused listening is required.

In the early phase of the work, the interaction was evaluated for gameplay purposes, and therefore, test persons' degree of hearing impairment was not taken into account. In this paper, we report on this first development phase consisting of the original design concept, the first version for informal testing, and then the second version made for formal evaluation.

The original concept for the gameplay of the Music Puzzle can be summarized by the following steps:

(1) The user shakes the tablet to split a file into a number of fragments
(2) Some sound parameters are randomly adjusted for each fragment
(3) Sound fragments are visualized and distributed randomly around on the screen
(4) The user listens, judges and rearranges these sound frag-

1 KTH Royal Institute of Technology, School of Computer Science and Communication, Sound and Music Computing, Stockholm, Sweden
2 Blekinge Institute of Technology, Sweden
a) kjetil@kth.se

ments in correct order

(5) The user adjusts the sound parameter values to match the sounds perfectly

(6) The user must finish arranging all fragments correctly to win the game

Metaphorically, the tablet screen holds a "sound ball" that can be shattered to randomly disformed smaller pieces when shook hard: the immediate response from the player should be to rebuild it again. The sound parameters used for changing the sound fragments are equalization (filtering effects) and a combination of tempo and pitch changes.

Since the main purpose of the game is to activate listening, the Music Puzzle accommodates this kind of training in several ways. First, the task to recombine the sound pieces requires the user to listen closely to match the ending of one sound with the start of another. Second, changing the pitch parameter requires the user to match in the frequency domain. Third, adding filtering to the sound requires the user to match in timbre. Finally, all these tasks are performed on different sound stimuli.

3. The application

The Music Puzzle was primarily tested on a Samsung tablet, GT-P7300 Galaxy Tab 8.9 inches, but the application also works with other devices running Android 2.2 or above. The programming was done in Pure Data (Pd, http://puredata.info), Android 3.1, and for the visual representation on the screen, Android View components (see examples in **Fig. 1**). In the following, the components are described briefly.

3.1 Pure Data and libpd

We are employing the libpd, a thin layer on top of Pd that turns Pd into an embeddable audio library, as an audio engine to support and develop mobile applications. With libpd, the Android interface can interact with the Pd core and Pd patches to carry out a series of functions.

Various sender and receiver objects and messages are used for controlling the global communication between the Pd part and the Android interface. The development was first done with following objects from the standard Pd library (commonly called Pd Vanilla):

adc~ Recording using the soundcard input object.

soundfiler Read recorded sound into the main array.

Tables Assign the sound segments to subarrays using the `tabread4~` and `tabwrite~` objects.

phasor~ Randomly assign tempo and frequency values for each subarray, adjustable by reading the table at varying speed.

Filters Randomly assign filtering effect values for each subarray, filtered using the `bp~`, `lop~` and `hip~` objects.

In the subsequent implementations, some changes were made to the above scheme.

3.2 Android user interface and sensors

One of the reasons for choosing a mobile device is the access to various sensors. In this application, mainly the microphone, accelerometer and multi-touch screen features are used, but there are many possibilities for extending the functionality and experience by employing more of the sensors.

3.2.1 Accelerometer

The built-in 3D accelerometer in the mobile device can detect the acceleration when the user is shaking the device in any direction. In this way, depending on the acceleration (the shaking force applied by the user), the number of audio segments is determined accordingly. This value will be sent to the Pd patch. A hard shake corresponds to a finer fragmentation of the audio, and consequently a more difficult task for the player.

3.2.2 Touchscreen

The screen interface is mainly provided so that users can control the sound segments, which are visualized as colored balls on the screen, see Fig. 1. These do not provide information about the sound and its properties, as the game is meant to be kept audio-based as much as possible.

3.2.3 User interface

In the interface, there is a "record" button for capturing audio directly from the surroundings through the built-in microphone in order to employ this sound in the game. The actual recording is done from Pd. After the recording, Pd will return the sound to the user interface for the user to listen in entirety, and subsequently, to split the sound by shaking the tablet and then reassemble the pieces.

4. Implementations and testing

Three versions are described here, and the differences in design regarding gameplay, audio and interface will be discussed:

V1 The original concept. This was not implemented fully, but presented to a reference group.

V2 The first test version. This was tested by developers and colleagues.

V3 The evaluation version. This was used and evaluated by a test group.

The basic idea has remained the same between each version. The user shakes the tablet, and the shake force is used to determine how many pieces to split the whole sound into, typically between 3–10 pieces. The sound has a duration between around 5–15 seconds. These pieces are represented as balls and placed randomly around the screen, and the sound pieces assigned to the balls are changed randomly in pitch and frequency filtering. Some design modifications were done for each new version, however, and these will be described in the following.

4.1 Audio design

Initially, simple sampling rate conversions were done for changing in pitch, in combination with speed alterations. This gave too many artifacts in the sound to be very useful. For V3, the pitch shift was done in the frequency domain instead using the "vocoder" examples in Pd. Pitches are randomized in 5 steps: in two steps below the original, the original, and in two steps above the original.

Audio filtering is done using the three common methods for this in Pd: the lowpass (`lop~`), highpass (`hip~`) and bandpass (`bp~`) filters. Randomization is done in four timbres: lowpass below a threshold, highpass above a threshold, bandpass between

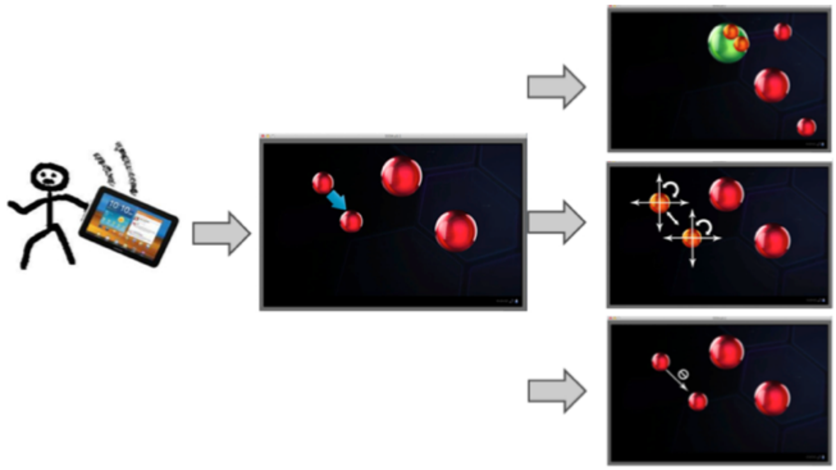

Fig. 1 An overview of the Music Puzzle gameplay in the first concept design. The user shakes the tablet to "smash" a sound file into pieces. In this process, some sound quality parameters are changed. Then, by listening to each fragment, represented visually by small balls, the user tries to match pieces that fit together in time, but also to adjust the altered sound parameters.

these values, and the unfiltered version.

4.2 Gameplay design

In V1, the balls could be tested for match in order by dragging them close together: they would either attach or bounce away. If they matched in order but not in sound parameters they would bounce but sliders for changing the parameters would appear. This design was judged to be too visually dependent and therefore revised. In subsequent versions a ball is not affected by the presence of another. Only the horizontal orientation is in use, and the balls must be ordered correctly from left to right (see **Fig. 2**).

Finding an intuitive way to change the rather abstract sound parameters was challenging. First, in V1, the parameters could be adjusted in a 2D space. In V2 this was changed to showing sliders when the balls matched in sequence but not parameters. In V3 the sliders appear after a long press on a ball. Several suggestions from users to take advantage of the vertical position on the screen for changing parameters were investigated, but were all found to be too visually dependent.

For finishing the game, V2 offered no other option than a silent evaluation of the balls' order. In V3 a playback functionality was added, where the sounds were played in order from left to right with their current sound parameters. From user's feedback, this functionality will be further developed to get a faster gameplay: by clicking balls in sequence, this series will be played back accordingly. In all versions, the game will only be successfully finished if all sound parameters have been adjusted right.

4.3 Interface design

The start-up screen has been updated since V1 with new functionality. The user could in V2 only start the game, while in V3 choose between "speech" and "music", record a sound, and listen to the whole piece. Based on the evaluation of V3, a new start-up screen will be implemented to include level of difficulty, and also more sound categories to choose from.

Only one change was made to the use of sensors. In V1 and V2 it was possible to use multitouch to move balls around simultaneously, but from V3 this is not necessary.

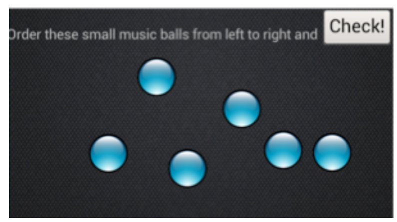

Fig. 2 The second implementation of the Music Puzzle game. Sound "pieces" are represented by balls randomly placed on the screen. The task is to order them correctly from left to right to rebuild the original, complete sound.

5. Results and discussion

User tests have been conducted after each implementation. With V1, a series of interviews with a reference group and more informal evaluations were done. The main outcome of these were design suggestions implemented for V2 as described above. With V2, the evaluation addressed usability, including the developer team and colleagues. This process further improved the application design and functionality.

The evaluation for V3 was performed with 9 subjects with normal hearing, between 21 and 27 years of age. Three sound models were rated by the gameplayers by preference, and by the level of difficulty using a 5-point scale (from very easy to very hard). All sounds had a duration of five seconds: the speech model consisted of a news reporter reading, the music model was a well-known Christmas carol, and the recording was made by the experiment leader to provide consistent conditions.

Half of the subjects found the speech model most preferable, while the recording model was only liked by one subject. This can be explained by the results plotted in **Fig. 3**, where speech had the highest portion of "medium" ratings. The interpretation of this is that the difficulty of solving the task was appropriately hard. Music had a higher rating of being easy, but also a high portion of being very hard. The recording had an even distribution between easy, medium and very hard. None of the testers found

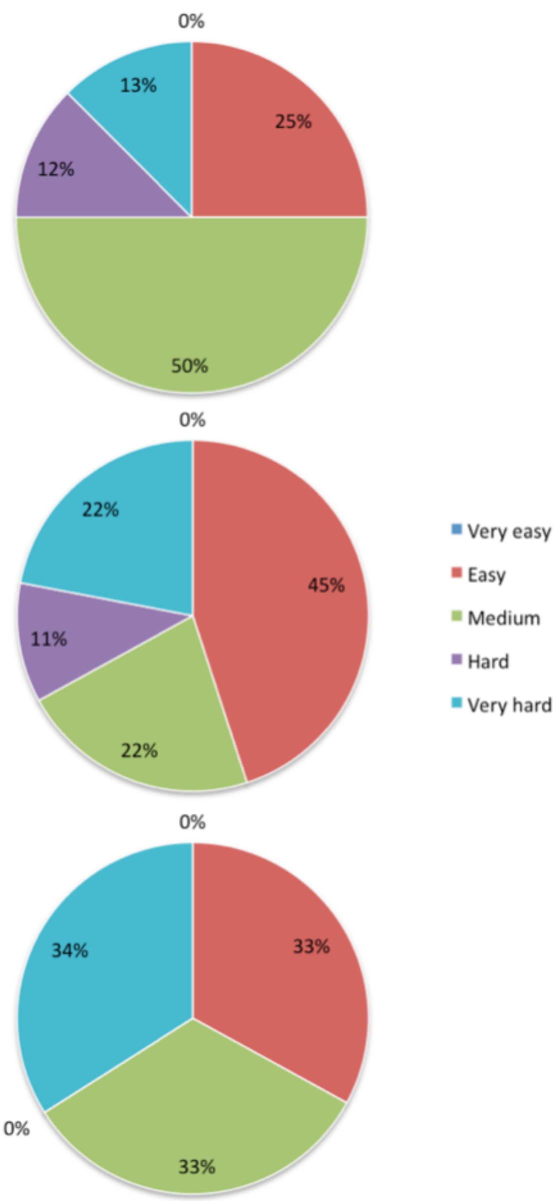

Fig. 3 Rating of difficulty level on a 5-point scale for five seconds of speech (upper chart), five seconds of music (middle chart), and using their own recording (lower chart). The difficulty ratings go clockwise from top from very easy (0% in all cases), easy, medium, hard, and very hard.

the task to be very easy.

There were a couple of test-procedure choices that can serve as further explanations of the results from the evaluation. Firstly, the experienced difficulty is expected to be very closely correlated to the number of fragments generated by the initial shaking gesture, regardless of which sound model was played. Secondly, the "recording" module was very sensitive to gaps or pauses in the audio, for instance between spoken words. Gaps may occur in any audio file and it poses a problem that needs to be addressed further.

6. Conclusions and future work

A novel sound-based game called Music Puzzle was implemented on an Samsung Galaxy Tab using Pure Data, libpd, Android and a set of the device's sensors. In the game the player has to match and align a number of short audio fragments to reassemble a complete sound file. The difficulty of mastering the game was found to be sufficient for maintaining the player's interest and engagement. For accommodating personal preferences, skills and ambitions, several difficulty levels are considered necessary for subsequent versions.

The evaluation version, V3, has been developed further for testing with the target user group. With this implementation of the Music Puzzle, the next step is to test more thoroughly with regards to audio aspects, and less with interface and design aspects. However, any immediate training results or health benefits coming from engaging hearing impaired users in active listening are not likely to be revealed: it is also outside the scope of this study to investigate the question. In the next user tests, the focus will therefore be on the user experience in combination with possible impact on listening habits.

The game is definitely engaging, and the authors believe that it has a broader potential than merely the initially envisaged one. For sight impaired persons, there are only a very limited number of games to play. Also for persons without sensory impairments, audio-based games offer fun, challenging, and enriching moments just like other games—or new possibilities where other game interfaces may be less desirable.

7. Acknowledgments

The Soundpark project is funded by the Swedish *Post och Telestyrelsen*, PTS.

References

[1] D. Benyon. *Designing Interactive Systems: A Comprehensive Guide to HCI and Interaction Design.* Addison Wesley Publishing Company, 2010.

[2] K. Gfeller, A. Christ, J. F. Knutson, S. Witt, K. T. Murray, and R. S. Tyler. Musical backgrounds, listening habits, and aesthetic enjoyment of adult cochlear implant recipients. *J Am Acad Audiol*, 11(7):390–406, 2000.

[3] Kjetil Falkenberg Hansen, Christina Dravins, and Roberto Bresin. Active listening and expressive communication for children with hearing loss using getatable environments for creativity. *Journal of New Music Research*, 2012. In press.

[4] Andy Hunt, Ross Kirk, and Matt Neighbour. Multiple media interfaces for music therapy. *IEEE Multimedia*, 11(3):50–58, 2004.

[5] Nina Kraus, Erika Skoe, Alexandra Parbery-Clark, and Richard Ashley. Experience-induced malleability in neural encoding of pitch, timbre, and timing. *Annals of the New York Academy of Sciences*, 1169:543–557, Jul 2009.

[6] Luis Lassaletta, Alejandro Castro, Marta Bastarrica, Rosa Prez-Mora, Rosario Madero, Josefa De Sarri, and Javier Gaviln. Does music perception have an impact on quality of life following cochlear implantation? *Acta Otolaryngol*, 127(7):682–686, Jul 2007.

[7] Valerie Looi and Jennifer She. Music perception of cochlear implant users: A questionnaire, and its implications for a music training program. *International Journal of Audiology*, 49(2):116–128, 2010.

[8] Wendy L. Magee. Electronic technologies in clinical music therapy: A survey of practice and attitudes. *Technology and Disability*, 18(3):139–146, 2006.

[9] Wendy L. Magee and Karen Burland. An exploratory study of the use of electronic music technologies in clinical music therapy. *Nordic Journal of Music Therapy*, 17(2):124–141, 2008.

[10] A. Mitchell and C. Saville-Smith. *The use of computer and video games for learning. A review of the literature*. The Learning and Skills Development Agency, 2004.

[11] UK Department for Education. Annex 3: The benefits of music (academic literature review). In Department for Culture, Media and Sport, editor, *The Importance of Music. A National plan for music education*, pages 42–44. Crown Copyright, 2011.