Design and evaluation of Sonic Tennis: an audio-only iPhone game

Stefano Baldan
University of Milan
via comelico 39
Milan, Italy
stefano.baldan@unimi.it

Stefania Serafin
Aalborg University Copenhagen
A.C. Meyers Vaenge 15
Copenhagen, Denmark
sts@create.aau.dk

ABSTRACT
Sonic tennis is an audio-based game for mobile devices, which relies on rhythmic interaction as the main element of the gameplay. Its aim is to be entertaining but also useful for ear training purposes, and accessible to both visually impaired and sighted users. The first prototype of the game has recently been qualitatively tested and publicly demonstrated at two different international conferences on human computer interaction. This paper focuses on the collected feedback from the results of the test as well as from the informal observations of the product "on the field", and on its use as a starting point for a second iteration of the design process.

Categories and Subject Descriptors
H.5.2 [**User Interfaces**]: Auditory (non-speech) feedback

General Terms
Design, Experimentation

Keywords
Qualitative testing, audio game, mobile devices, sound design, rhythmic interaction, motion-based

1. INTRODUCTION
Traditionally, audio only games have been seen mostly as a niche targeted mostly to the narrow audience of visually impaired people [5], and unsuccessful from the commercial point of view [8].

However, relatively few academic studies have been investigating the role of audio only games and design principles related to them. As an example, in [4] an auditory version of space invader is presented. In [3] the authors present the challenges encountered when designing and implementing Beowulf, an audio only games. Although the game was described as a failure, boneless by designing it the authors gained some important knowledge on design guidelines for audio only games. Previous research also shows that audio only games are not only challenging and enjoyable to sighted users, but also have a great potential in terms of ear training and enhancement of one's listening capabilities [9][2]. Another advantage of audio-based games is their relatively low requirements in terms of processing power compared to their visual-based counterparts [7][6]. This element, together with the very limited or sometimes absent use of the screen, makes audio-based games well suited for mobile devices or other embedded computing applications lacking a computer display.

In this paper we present an evaluation of Sonic Tennis, an audio only game first introduced in [1]. The evaluation is based on qualitative feedback provided by the players as well as observations by the authors.

2. SONIC TENNIS
Sonic Tennis is an audio only game that simulates a tennis match. The game can be played in single mode or multiplayer mode. For the purpose of this evaluation, we consider the two players mode. The goal of the game is to win a tennis match by merely listening to the auditory cues provided by the phone. Specifically, auditory cues regarding the timing of a hit and the location (right or left side) are provided to each player. The players must react to the cues by hitting a left strike or a right strike according to which direction the opponent's ball is coming from. Moreover, player need to listen to the timing of the opponent's strike, and react to it.

3. TESTING
The game was qualitatively tested by twenty participants, of which ten were males and ten females, aged 20 to 41 (mean age 25, standard deviation 5), sighted and without hearing deficits. Participants were organized in pairs, and to each of them it was asked to play the game in two different conditions. One of the two matches had to be played with the participants facing each other, thus seeing the physical actions of the opponent, while the other match had to be played with the participants placed out of their respective line of sight. The order in which the two different matches had to be played was randomized among the different pairs in order to reduce bias in the test. After having played the game in both conditions, for as long as they wished, participants were asked to fill an online questionnaire.

Our testing group was composed mostly by casual gamers. Out of the twenty participants, only six never play videogames

in a regular week. Other six participants usually play videogames for less than one hour per week, five for less than one hour per day, three for one to three hours per day and nobody for more than three hours per day. On the other side, participants with musical training were a minority. Thirteen of them never play musical instruments during the week, while other five play for less than one hour per week and the remaining two for less than one hour per day.

Participants were asked to tell if they had some experience with particular kinds of computer games and means of interaction. five out of twenty were not familiar with games for mobile devices, while twelve played them occasionally and the remaining three played them often. Twelve participants had no experience with motion-based devices like Nintendo Wii remote, Microsoft Kinect or Playstation Move, eight used them occasionally and nobody used them often. Moreover, just four out of the twenty participants had previous experience with an audio-based game before trying Sonic Tennis.

Most of the users experienced problems in playing the game, to various extent. The most common difficulty, pointed out by eight participants, was encountered in following the rhythmic interaction, namely recognizing the right timing to hit the ball and acting accordingly. Five users experienced networking issues during the game, such as sudden disconnections and unwanted delays in the delivery of the messages, causing them to miss the ball even if their timing was right. Two participants experienced some problems in the gesture recognition, especially for the overhead hit, used to serve the ball into play at the beginning of the match or after a point is scored. Two other players asked for improvement in the sound design, differentiating more the sounds coming from the two different sides of the court and making the timing cues clearer, louder and sharper. One player was particularly disturbed by seeng the actions of the opponent when playing in front of her, because she perceived them as inconsistent with respect to the timing and distance of the audio feedback she was hearing. Finally, one participant noticed that the racquet strikes were triggered also when she swang the device backwards, namely towards her back.

Users were asked to state if the fact of seeing or not seeing their opponent affected somehow their way of playing the game, and if yes to indicate in which condition they felt more at ease. Fourteen participants answered affrmatively to the first questions, and nine of them preferred to see their opponent rather than not. The questionnaire ended with an open request for general comments. One participant suggested the use of wireless headphones, as the awinging cable can get annoying when rapidly moving the device. Others suggested to raise the volume of the sound samples or porting the game also on the Android platform. Finally, a trained sound designer suggested to add some continuous audio feedback as an additional timing cue.

Comments and informal observations were also collected during two public demonstrations of the game at international conferences on human computer interaction, namely Computer Human Interaction (CHI) 2013 in Paris, France [1] and New Intefaces for Musical Expression (NIME) 2013 in Seoul, South Korea. In both occasions the problems related to the network became even more evident, due to the poor quality of the public wi-fi link and to the high traffic handled by the access points. The more skilled players also asked to have more control over the ball, using the nuances in their gestures to alter the direction and speed of the ball.

4. IMPROVING SONIC TENNIS

Answers from the participants collected during the test and user feedback "on the field" collected during public demonstrations of the work were therefore used to drive a second iteration of the design and development of Sonic Tennis. Analyzing the various comments, four main areas of improvement could be identified.

4.1 Timing

As already stated in section 3, the main difficulty encountered by the majority of the users was trying to recognize the right moment to hit the ball following just the audio cues delivered by the application. From the designers' perspective this is a somewhat desired result, because the acoustically induced rhythmic interaction is exactly where the main challenge of the game lies. Responding consistently and precisely to the sounds produced by the virtual ball, without seeing it moving, is something which is meant to be difficult at first and which should need practice to be improved. Nevertheless, the game should also not be too difficult during the first matches, or it might convey a sense of frustration which would cause the player to leave the game.

To address this issue, three different levels of difficulty have been introduced when playing a match against the computer. Easier levels feature slower balls (i.e. more time passes between the hit of the racquet and the bounce on the ground), higher failure rates from the computer and a longer span of time in which the ball can be hit, viceversa harder level present faster balls, lower failure rates and a shorter span of time in which the ball can be hit. Beginner players can choose the easy bot to practice safely, and move on the harder ones when they need more challenge. The introduction of continuous sound feedback after the ball bounces might also help in better recognizing the sweet spot in which the ball should be hit, but we will write about it more extensively in section 4.4.

4.2 Network

Issues related to the network were the second most observed problem by test participants, and the most frequent complaint during public demonstrations. They seem to be the greatest source of annoyance and frustration while playing the game, because a delayed message often result in a lost ball, even if the timing of the player was correct. In a few cases the quality of the connection was so poor that users experienced sudden disconnections, thus having to restart the match from the beginning. Network delays also ruin the sense of immersion in the game, as they cause delays in the playback of the audio samples as well, thus desynchronizing them from the actions performed by the user.

As with many other applications requiring real-time communication over a shared network infrastructure, these problems cannot be fully resolved without having complete control over the network infrastucture itself. In the case of

messages sent over the Internet, we have to rely on its *best effort* model and just hope that they will arrive in time to be processed. Many issues can still be mitigated though, simply trying to avoid the exchange of messages over the Internet whenever possible.

In the current version of Sonic Tennis all the game logic is implemented on the server and each single event in the game, namely hits, bounces and points scored, is communicated to the "dumb" clients via message passing. This architecture allows rapid prototyping and implementation of new features in the game, without having to recompile and resubmit the client to the Apple Store, thus avoiding the long and tedious review process. For the same reason even the single player mode was implemented through bots running remotely on the server, which therefore require connection to the Internet and suffer from the same network delays and failures of the multiplayer mode. Moving the game logic on the clients and using the remote server just for matchmaking would result in a more cumbersome deploying process, but would allow a local implementation for the single player mode and a multiplayer mode based on bluetooth when playing the game with both opponents in presence. Network traffic for remote matches over the Internet could also be reduced, sending messages just for the hits and locally calculating the right timing for ball bounces and points scored.

4.3 Gesture recognition

Some users experienced difficulties in having their swings recognized as correct forehands or backhands by the application. The most difficult gesture to perform seemed to be the overhead swing, used to serve the ball at the beginning of the match or after a point is scored.

In the current implementation, a swing is detected when the absolute value of the accelerometer reading on the z axis, namely the one coming out of the display, goes above a fixed activation threshold. No other hits can be detected until that value goes below a fixed disactivation threshold, usually lower than the previous one. A positive reading means that the device has been moved in the direction of the display, therefore triggering a forehand, while a negative value means that the device is moving towards its back side, therefore a backhand is triggered. As soon as the hit is detected, the attitude of the device is also read. If the *roll* of the device, namely its rotation angle around the vertical (y) axis, is negative, the left side of the phone is pointed towards the ground. The player is therefore most likely right-handed, so the forehand will correspond to a hit coming from the right and the backhand to a hit coming from the left. If the roll is positive, the right side of the phone is facing the ground. The player is therefore most likely left-handed, so the forehand will be a hit coming from the left and the backhand a hit coming from the right. To detect an overhead hit, the device must be held in an upright position. When this happens, the *pitch* of the device, namely its rotation angle around the horizontal (x) axis, is close to $\pi/2$.

The inaccuracy of this approach in detecting some particular gestures is likely due to the fact that it uses just one sample to describe the whole action. Instead of sampling the device attitude just after having reached the activation threshold, readings can be accumulated until reaching the disactivation threshold and then used to extract a more statistically representative value of what happened (e.g. maximum, minimum, average, median).

4.4 Sound design

A few observations regarded the choice of the sounds used to create the soundscape of the game and to give cues on when and how to hit the ball. Most of them pointed out excessive similarity between the sounds coming from the two different sides of the court, lack of realism and difficulty in hearing ball bounces in a noisy environment. Quite unexpectedly nobody complained about the very simple stereo panning used to discriminate between balls coming from the left or from the right, which proved very effective and was intuitively understood by all the users, sometimes even without giving previous explanations.

The current soundscape of Sonic Tennis is mostly created using publicly available sound samples, released under Creative Commons [**?**] licenses, altered in pitch, amplitude and reverb to give an impression of further distance for the sounds coming from the opponent's side of the court. The use of samples coming from dedicated field recordings, binaural filtering and/or real-time sound synthesis techniques with an approrpiate engine such as libpd [**?**] could improve the realism of the soundscape and the effectiveness of the timing audio cues even in noisy environments. Being such an important aspect in an audio-based game, specific listening sessions might be considered in order to guide the choice of the most effective approach in designing the acoustic environment.

A broadband and continuous sound feedback, increasing in volume after the ball bounces and slowly decreasing after having reached the sweet spot, could also be used as an additional timing cue. Initial work in this direction has already been done, recording some noise-making toys with a dummy head in the anechoic chamber of the Multisensory Experience Lab at Aalborg University Copenhagen. At a first listening the results look promising, but a proper evaluation must be done to understand if and to what extent such a sound effect could improve or rather degrade the gaming experience.

4.5 Game rules

Although playing Sonic Tennis proved to be a quite challenging task for most of the people who tried it, some users felt particularly at ease with its style of interaction and quickly became skilled players. Once having mastered how to consistently hit the ball, the game started to become boring and repetitive for them. Some users explicitly asked to have more control over the ball, using their gestures to change its speed and direction.

The gesture recognition improvements proposed in section 4.3 could be exploited for this purpose. Averaging or taking the maximum reading of the accelerometer, the strength of the hit can be inferred, and the speed of the ball can be changed. Timing can also be taken into account to tweak this parameter: a very strong and perfectly synchronized hit might result in a very fast ball, whereas a weak and not so well timed hit might result in a slower ball, easier to catch for the opponent. The strength of the hit might also influence

the tolerance interval in which the hit must be performed. A very strong hit might be more effective but have more chances to miss, while a weaker hit would result in a slow but safe reply. Following the same principle, rotation on the x axis might be used to influence the direction of the ball: a high rotation would result in a diagonal hit, whereas a low rotation would send the ball down the line.

To add more variety to the game, some special events might appear at random during the match. For example, a very slow ball might give the opportunity to the opponent to perform a very powerful overhead smash, or at some point a different sound might warn that the ball has bounced on the net, requiring an underhand hit to rescue it.

5. CONCLUSIONS AND FUTURE WORK

Sonic tennis provided a nice proof of concept and case study to explore a novel approach to the design of audio-based games, which relies on rhythmic interaction rather than verbal storytelling or soundscape navigation. The first qualitative tests and public demonstrations of the application unleashed the potential of the idea, offering invaluable hints and feedback on how to improve the initial concept. Several guidelines have been proposed based on this feedback, which might be used as a starting point for a second iteration of the design and development process of the game. Our further work on the project will follow these guidelines, trying to get closer to the goal of making an engaging, enternaining and educational game for both visually impaired and sighted users.

6. REFERENCES

[1] Stefano Baldan, Amalia de Götzen, and Stefania Serafin. Mobile rhythmic interaction in a sonic tennis game. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, pages 2903–2906. ACM, 2013.

[2] J. Friberg and D. Gärdenfors. Audio games: new perspectives on game audio. In *Proceedings of the 2004 ACM SIGCHI International Conference on Advances in computer entertainment technology*, pages 148–154. ACM, 2004.

[3] M. Liljedahl, N. Papworth, and S. Lindberg. Beowulf: an audio mostly game. In *Proceedings of the international conference on Advances in computer entertainment technology*, pages 200–203. ACM, 2007.

[4] R McCrindle and David Symons. Audio space invaders. In *Proceedings of the Third International Conference on Disability, Virtual Reality and Associated Technologies*, pages 59–65. Citeseer, 2000.

[5] Nigel Papworth. ispooks: an audio focused game design. In *Proceedings of the 5th Audio Mostly Conference: A Conference on Interaction with Sound*, page 11. ACM, 2010.

[6] N. Röber. Playing audio-only games: A compendium of interacting with virtual, auditory worlds. 2005.

[7] N. Röber and M. Masuch. Leaving the screen: New perspectives in audio-only gaming. In *In: 11th Int. Conf. on Auditory Display*. Citeseer, 2005.

[8] Emmanouel Rovithis. A classification of audio-based games in terms of sonic gameplay and the introduction of the audio-role-playing-game: Kronos. In *Proceedings of the 7th Audio Mostly Conference: A Conference on Interaction with Sound*, pages 160–164. ACM, 2012.

[9] S. Targett and M. Fernström. Audio games: Fun for all? all for fun. *ICAD2003*, 2003.