Audio Engineering Society

Convention Paper

Presented at the 130th Convention
2011 May 13–16 London, UK

Interactive Audio Realities: An Augmented / Mixed Reality Audio Game prototype

Nikos Moustakas[1], Andreas Floros[2], and Nicolas Grigoriou[3]

[1] Dept. of Audiovisual Arts, Ionian University, Corfu, 49100 Greece
nikos.av06@gmail.com

[2] Dept. of Audiovisual Arts, Ionian University, Corfu, 49100 Greece
floros@ionio.gr

[3] Dept. of Audiovisual Arts, Ionian University, Corfu, 49100 Greece
nikolas@ionio.gr

ABSTRACT

Audio-games represent a game alternative based on audible feedback rather than on visual. They may benefit from parametric sound synthesis and advanced audio technologies (i.e. augmented reality audio), in order to effectively realize complex scenarios. In this work, a multiplayer game prototype is introduced which employs the concept of controlled mixed reality in order to augment the sound environment of each player. The prototype is realized as multiple user audiovisual installations, which are interconnected in order to communicate the status of the selected control parameters in real-time. The prototype reveals significant relevance to the-well known on-line multiplayer games, with its novelty originating from the fact that user interaction is realized in augmented reality audio environments.

1. INTRODUCTION

Video games represent a major field of nowadays-new media production. The consecutive evolutions on audiovisual technologies and reproduction areas, combined with the significant merits obtained by the employment of advanced artificial intelligence modeling have offered the opportunity to developers to design highly realistic and immersive game environments. These environments are suitable for realizing complicated interactive scenarios, not only for entertainment, but also for educational purposes (i.e. in terms of Serious Games) [1]. Under the above perspective, sound represents an important component of a video game authentic realization, since its parallel reproduction with the visual scenery allows the establishment of the desired interaction paths with the player.

Based on the needs of visually impaired humans who were a priori excluded from the video games target audience [2], a new kind of games has evolved: the audio games [3]. In such applications, the feedback to the player is strictly provided and perceived by audible means. Modern sonic interaction design techniques allow the implementation of complex and experimental gameplay scenarios, supported by different interaction and audio playback frameworks, such as the Augmented Reality Audio (ARA) technology [4]. ARA involves the enrichment of the real world sound field with an overlay of a "virtually" synthesized sonic environment and has been previously used in various application domains, such as distant communications, information services and location-aware environments [5].

Recently, an augmented reality audio game prototype was introduced by the authors [6]. It consisted of a) the real world visual environment physically represented by the game-play physical space b) the real world sound component which was present within the game physical space and c) the virtual audio information which was synthesized in real-time and reproduced in the 3D space. During the game-play, the real and virtual audio components were mixed for synthesizing the overall, virtually enriched (i.e. augmented) sound environment. Hence, the game-play evolution was realized in an augmented reality audio environment that maps the dynamic information related with the game scenario to a number of sound objects.

This work aims to go a step further by extending the above augmented reality approach to the more general mixed audio reality interaction concept. By exclusively amplifying the well-established general concept of mixed reality [7] - [8] for audio only, it can be considered that within an interactive mixed audio reality environment, the augmented audio reality of a user is partially (or fully) defined by the sonic reality of another one. Hence, a strong interconnection / interaction between the two remote users/players can be defined, very similar to the well-known multiplayer game rules and scenarios. We hereby demonstrate and evaluate the proposed interactive mixed/augmented audio framework in terms of an audio only game prototype (termed as "eidola multiplayer", see Fig. 1), in order to assess it's potential advances in terms of playability, realistic scenario realization and user immersion.

Figure 1 The "eidola multiplayer" game prototype logo

The rest of the paper is organized as following: Section 2 overviews the fundamental concepts of augmented and mixed audio interactive realities, followed by the analytic description of the introduced game prototype provided in Section 3. In Section 4, some game implementation issues are discussed and the results obtained after demonstrating it to typical human audience are presented. Finally, Section 5 concludes the work and outlines further research enhancements that should be considered in the future.

2. AUGMENTED AND INTERACTIVE MIXED REALITIES OVERVIEW

In a typical ARA environment, the real world sound field is enriched with synthetic sounds produced by virtual sound sources. Obviously, this sonic enrichment must be realistic and authentic, providing a natural feeling of the two concurrently overlaid sound layers to the acoustic receiver [9]. Provided the human ability to localize the instantaneous spatial positions of the active sound sources, it is necessary to reproduce the virtual sound objects effect in the three dimensional space. Towards this aim, binaural technology [10] is very frequently used, due to the required simple sound reproduction through a pair of audio channels, as well as the increased playback portability that can be achieved through headphones.

More specifically, the binaural technology employment purpose is twofold, as: a) binaural recording of the natural sound environment is required. This recording is performed using in-ear microphones, usually embedded on the stereo ARA headset employed for binaural signal reproduction [11] and b) binaural processing using the correct set of Head Related Transfer Functions (HRTFs) is applied for rendering the perceptually correct spatial position of the virtual sound source. The output signals derived by the above recording and processing tasks are driven to an ARA mixer [12], a device that overlays them and produces the desired, virtually enriched binaural soundfield.

For effective ARA realizations, some specific implementation issues should be considered. For example, equalization must be employed in order to correct the potential sound coloration that may be imposed by the leakage through the ARA headset and the changed resonances in the closed ear-canal. Moreover, head-tracking [13] represents a demanding extension to the complete ARA system, since it allows the free movement of the user, without affecting the relative position of the virtual sound sources surrounding him. Finally, room acoustics modeling may significantly impact the sound localization efficiency, mainly by allowing the "capture" of small unconscious head movements. Alternatively, in such small movement cases, high-quality, time-varying HRTF interpolation techniques can be employed, that also perceptually improve the effect of moving sound sources [14]. However, this approach introduce significantly increased binaural processing computational load.

In existing augmented reality audio applications [4], [15], the virtual audio objects' instantaneous positions are algorithmically controlled. However, if we consider multiple, ideally alike ARA environments, then a virtual audio object can be assigned to other users / players that concurrently interact within these setups. Obviously, in such a case, the complete sonic environment of a specific user consists of: a) the real-world audio data tracked by the in-ear microphones of his ARA headset b) the augmented sound component produced by the algorithmically controlled virtual audio objects and c) the augmented sound component that corresponds to a remote user that concurrently participates within his own augmented reality audio context.

The latter component introduces a strong interconnection between the participating users' augmented audio realities, based not on algorithmic or artificial intelligence rules, but on human interaction principles. This additional interaction layer is transparent to the local users, who perceive it as a part of their overall virtual audio scenery. However, in practice, the dynamic adaptation of this layer obeys the remote user reactions, which in turn are defined by the ARA environment scenario. Hence, complex interaction scenarios can be realized, similar to those found in today multiplayer game applications.

3. PROPOSED GAME DESCRIPTION

In this Section, the "eidola multiplayer" game employed for demonstrating the proposed mixed audio reality interaction concept is presented. More particularly, the game scenario is initially described, focusing on the game-play details that correspond to the proposed novel interaction path, followed by an analysis of the overall game architecture.

3.1. The interactive mixed reality scenario

The "eidola multiplayer" scenario was designed as an extension of the "eidola" ARA game presented in [6]. According to its scenario, a number of invisible creatures (named "eidola") were continuously moving following non-linear traces within the game 3D installation space. Each creature was producing a specific sound, used for identifying its position. Aim of the game was "eidola" elimination, a task that was performed when the user was approaching them within a specific distance limit and pulled a trigger of a virtual gun. "Eidola" scenario was designed to be as simple as possible, as the main focus of the game realization was the assessment of the ARA employment in audio games platforms.

In this work, the above simple game scenario was enriched, by introducing multiple (two) players chasing "eidola" creatures within two absolutely alike rooms. Again, "eidola" elimination is performed when the local player's position coincides with their current position. However, one moving being is not controlled by the game-engine (i.e. the 3D trace scheduler, see next Section for details). Instead, it corresponds to the remote user, thus forming the mixed audio reality interaction concept. Additionally, some objects (such as radios, TV sets, etc.) are located at specific positions within both rooms; these objects produce natural sounds comprising the ambient, real-world sound environment. Ambient sounds are additionally reproduced through the virtual audio component, using pre-recorded sounds and adaptive playback gain adjustment based on the instantaneous position of the local user.

The sound assigned to the remote player was also "virtual" and adjustable (in terms of real-time playback gain variations). It consisted of human footsteps recordings, appropriately pre-processed to match the physical acoustic environment (i.e. reverberation) of both rooms. This selection allowed the easy discrimination of the moving being that corresponded to the remote user and realized in practice the required augmented acoustic interaction path. The objective of each player was to eliminate as many "eidola" as possible, including the remote user. Fig. 2 illustrates the graphical representation of the described "eidola multiplayer" scenario. It should be also noted that the number of the remote players was here limited to two for keeping the game implementation complexity as low as possible; however, in general, there is no other restriction for extending the scenario to more users.

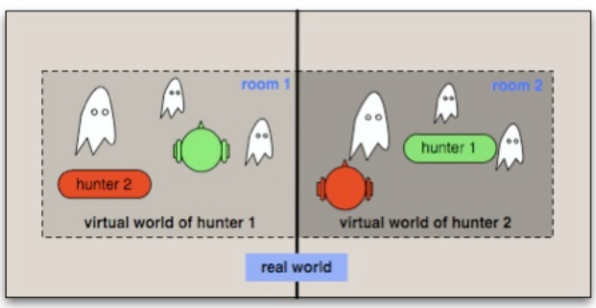

Figure 2 The game scenario graphical representation

3.2. Game architecture

By summarizing the above scenario details, each player's augmented audio reality is constructed using a) a real-time, gain-adjustable playback engine for reproducing the sound (i.e. footsteps) that corresponds to the sound objects of the remote users b) the real-world audio component recorded by the in-ear microphones c) the virtual world audio component that includes the sounds assigned to both the non-human "eidola" creatures and the virtual surrounding objects of the game world.

In order to obtain the previous audio information, a random 3D trace scheduler is responsible for controlling the movement of the non-human "eidola" sound objects within the installation space. Additionally, a binaural renderer-processing module converts all the virtual objects' audio outputs to a binaural signal, taking into account the room geometry and absorption characteristics, as well as the local user position and head direction. Finally, an ARA mixer mixes the real sound environment captured using in-ear microphones with the binaural renderer output, while a general interaction controller collects and appropriately processes the signals obtained from the user position and head direction identifier/tracker, as well as the virtual gun trigger. All these interaction sub-systems were implemented following the description provided in detail in [6].

Fig. 3 illustrates the general game architecture, focusing on how the previously described system modules are distributed among the two discrete installations' environments. The architecture follows the well-known master/slave communication model, as this provides the less complex means for synchronizing the two independent, but interactive, audio environments. The necessary communication between the master and slave nodes is performed through the Open Sound Control (OSC) protocol [16].

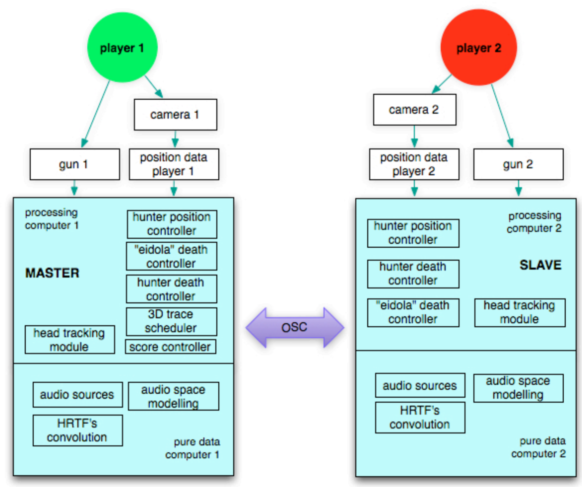

Figure 3 The "eidola multiplayer" master/slave architecture

More specifically, each node controls a specific installation room and consists of a) the interaction controller and b) the augmented audio mixer and reproduction controller. Both controllers were implemented in software, using the Processing development platform [17] and a Pure Data optimized patch for carrying out the binaural sound reproduction in real-time [18] respectively. The node defined as "master", additionally hosts the 3D trace scheduler and transmits the relative spatial information to the local and the "slave" eidola death controller through the OSC protocol. The latter controller is responsible for locally identifying and keeping track of the current status of "eidola" (dead or alive). Moreover, the master host is responsible for counting and storing the current game score using the score controller. The rest game functionality is realized at both computer hosts, following exactly the same architecture.

In addition to the system architecture described before, the overall ARA Mixer functionality is also illustrated in Fig. 4. Multiple wireless audio technologies operating in different frequency ranges were selected for providing the input to and delivering the output from the ARA mixer, in order to minimize potential wireless channel interferences.

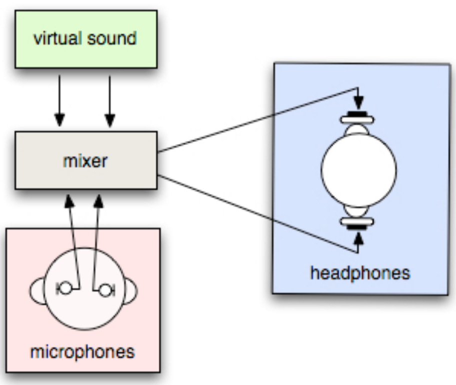

Figure 4 The ARA mixer functionality

3.3. The user/player equipment

The user equipment represents an inextricable part of the game environment. Due to its required portable nature, the overall design had to carefully consider power feed issues based on re-chargeable batteries, as well as electromagnetic interference imposed by the multiple wireless protocols operated concurrently.

Fig. 5 illustrates the architecture of the portable user equipment. It incorporates the necessary headphones / microphones stereo pre-amplifiers, the required audio signal wireless transmitter and receiver, as well as a headphone equalizer for correcting the potential sound coloration mentioned in Section 2. Moreover, the head-tracking system is also integrated. It was developed using a modified Bluetooth-enabled computer joystick construction appropriately applied on the back part of the human head, achieving accuracy in the range of +/-5 degrees. Head tracking data were transmitted to the local hunter position controller through Bluetooth ACL links. The same links were also employed for delivering the status of the gun trigger of each player, provided that it was pressed.

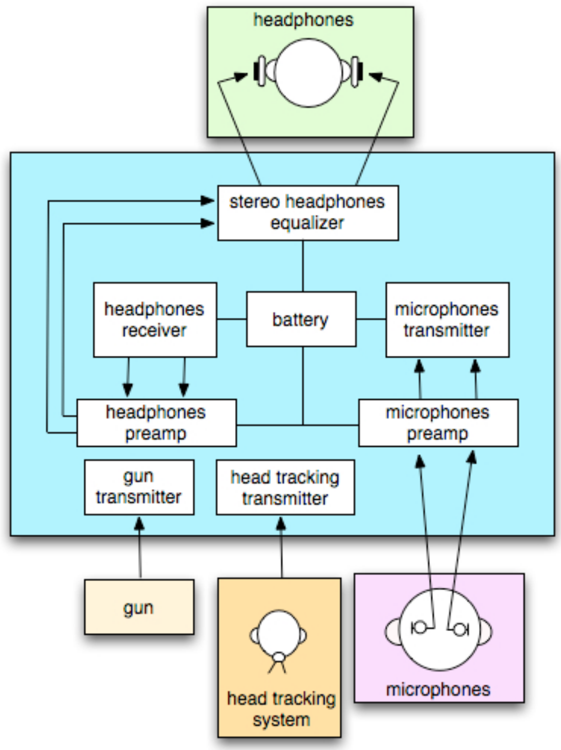

Figure 5 The portable user equipment architecture

4. GAME IMPLEMENTATION AND RESULTS

The "eidola multiplayer" interactive installation setup was performed in two enclosures with dimensions 3.2m (width) x 3.2m (length) x 3.5m (height). In each room a camera was placed on the celling in order to achieve the desired effective detection surface for the local user localization. The wall color additionally provided the necessary contrast for the player's position detection sub-system.

4.1. Subjective test methodology

The assessment of an audio game experience from the player point of view using objective measurements is a rather difficult task [19]. Hence, the evaluation of the proposed interactive mixed reality framework was based on the subjective reports obtained in the form of questionnaires filled by the participated game – prototype users. All users were graduate or post-graduate/post-doc university students and researchers. The basic assessment criteria were defined in the form of a total number of 13 appropriately designed questions. These criteria employed in this work can be summarized as following: a) the level of the perceived user immersion, a parameter that is strongly related with the overall experience achieved by the augmented reality audio environment b) the weighted, perceptual significance of the virtual and real audio components, mainly in terms of the game-play realization c) the effectiveness of interaction between the two remote players as a key factor for supporting the game-scenario and d) the comparison between the expected and the achieved user experience.

All the questions were statements, allowing answers of type "yes" or "no", or between a 5-item scale ranging from strongly negative to strongly positive, or between pre-defined short answers. Typical statements included are: a) "Is this your first time playing an audio game", b) "How difficult was to locate the spatial sound position", c) "Was the game scenario clearly realized", d) "Did you feel an active/real part of the game" and e) "Which audio component was dominant in the game-play". Additionally, the questionnaires included a number of participant demographics, such as male and female, ages discrimination ranging from 18 to 29 and the graduate level and type of studies.

15 pairs of players evaluated the previously described game prototype over a period of two days. All players were given the user equipment described in Section 3.3. Additionally, prior to entering the game physical space, the same brief summary of the game scenario and scope was provided to all users through electronic audiovisual means. However, no information was provided in terms of the evaluation goals.

4.2. Summary of the results

Table 1 summarizes the participants' responses at statements answered using the 5-item scale, where A – D corresponds to the following statements: A: How do you rate the novelty of the game (1: poor, 5: absolutely novel) B: What was the required time for getting used to the game environment (1: very long time, 5: very short time) C: Were the instructions provided prior to entering the game environment enough for understanding the scenario (1: useless 5: absolutely adequate) D: How competitive was the game (1: not competitive, 5: absolutely competitive) E: How easy was to locate the sound sources positions (1: very hard, 5: very easy) F: How easy was to distinguish the different sounds assigned to the virtual objects (1: very hard, 5: very easy) and G: How do you rate the overall immersion level and experience of the game (1: poor, 5: excellent).

Table 1 Summary of the 5-item scale results

	1	**2**	**3**	**4**	**5**
A	-	12,5%	25%	37,5%	25%
B	-	25%	62,5%	12,5%	-
C	-	-	12,5%	12,5%	75%
D	-	12,5%	25%	37,5	25%
E	25%	25%	-	25%	25%
F	12,5%	37,5%	37,5%	12,5%	-
G	-	-	-	25%	75%

From the above results it became clear that the game prototype was considered to have high novelty, while it additionally achieved a very positive overall experience and perceived immersion assessment. Focusing on particular design factors, spatial sound source localization and the selected sound design was found to significantly vary based on the listening experience of the corresponding players. For example, participants with study profiles related to sound and music better evaluated sound source localization and separation. The same users were also the majority of the winners in all test cases. According to this result, both the sound design employed as well as the specific ARA implementation issues discussed in Section 2 extend the relative requirements met in common audiovisual productions (i.e. cinema films and legacy multimedia productions) and should be very carefully considered when realizing ARA applications.

It should be noted here that nearly 90% of the participants answered that this was their first time that they played an audio game. This fact can explain the relatively long times required by most users for getting used to the game scenery. However, it is the authors' belief that this also strengthens the overall experience results' significance, since a) the lack of any prior acquaintance with audio games did not negatively affect the user experience, while b) it seems that the demonstrated interactive, mixed reality audio game was clearly characterized as an acceptable, different game type that all the participated players wish to play again in the future.

An additional, significant outcome is related to the mixed audio reality interaction concept: 80% of the players provided positive feedback, as they considered significant the presence of the remote player within the scope of their own augmented audio environment. Additionally, 75% of them mentioned that they would prefer to have more that one remote enemy to compete with. Finally, 65% of the participated audience felt that the game scenario was equally supported by both the virtual and real world audio components, while the rest answered that the virtual component was dominant. This represents an evidence that the augmented nature of the game environment predominates against a legacy virtual reality approach in the context of the requirements imposed by the particular game scenario.

Apart from the above processed subjective results, some general player observations and comments were obtained. In particular, a number of players requested larger room spaces, while some other suggested that the scenario should be extended, allowing the moving beings to kill the player as well. All the registered user comments will be considered in future game versions.

5. CONCLUSIONS

This work introduces a novel framework that extends the Augmented Reality Audio approach employed in previously published works focused on entertainment / audio games applications. Based on the definition of an additional type of virtual sound objects that are mapped to remote users that concurrently act in alike augmented audio sceneries, an additional interaction path is established within the augmented audio environment. This interaction path allows the dynamic adaptation of the virtual audio environment of a specific user, based on human-based reactions and not on algorithmically controlled feedback.

For the purposes of this work, the proposed mixed reality interaction framework was demonstrated using an ARA game prototype. The game scenario involved all the parameters required for assessing the additional interaction path described previously. A sequence of tests carried out using subjective means (i.e. appropriately designed questionnaires) has shown that the proposed approach clearly achieves high levels of perceived user immersion, while it significantly enhances the observed user interaction, allowing the realization of more complicated and adaptive game scenarios. The demonstration of the above game-prototype has also shown that the sound design and ARA implementation requirements are more demanding in the context of the proposed interaction-enriched framework. Hence, significant attention should be paid during the design phase of future applications.

It is the authors' near future intention to further exploit the derived results and conclusions for developing more sophisticated game scenarios and supporting them through advanced and novel means of interaction.

6. ACKNOWLEDGEMENTS

The authors wish to acknowledge Prof. George Papadelis from Aristotle University of Thessaloniki for his valuable contribution in carrying out the subjective measurements of this work.

7. REFERENCES

[1] A. Stapleton, "Serious Games: Serious Opportunities", In Proc. Of the Australian Game Developers' Conference, Melbourne, VIC, 2004.

[2] K. Bierre, J. Chetwynd, B. Ellis, D. M. Hinn, S Ludi and T. Westin, "Game Not Over: Accessibility Issues in Video Games", In Proc. of the 3rd International Conference on Universal Access in Human-Computer Interaction, 2005.

[3] J. Friberg and D. Gärdenfors, "Audio games: new perspectives on game audio", In Proc. of the 2004 International Conference on Advances in computer entertainment technology, pp. 148 - 154. doi:10.1145/1067343.1067361, 2004.

[4] A. Härmä et al., "Augmented reality audio for mobile and wearable appliances", Journal of the Audio Eng. Soc., Vol. 52, No 6, June 2004, pp. 618-639.

[5] T. Lokki, H. Nironen, S. Vesa, L. Savioja, A. Härmä, and M. Karjalainen, "Application Scenarios of Wearable and Mobile Augmented Reality Audio," Audio Engineering Society 116th Convention, Berlin, May 2004 (preprint 6026).

[6] N. Moustakas, A. Floros and N. Kanellopoulos, "Eidola: An Interactive Augmented Reality Audio-Game Prototype", Audio Engineering Society 127th Convention, New York, October 2009 (preprint 7872).

[7] Y. Ohta and H. Tamura (editors), "Mixed Reality: Merging Real and Virtual Worlds", Springer, ISBN: 978-3540656234, 1999.

[8] Remote: a mixed reality installation (http://www.haque.co.uk/remote.php, last accessed: March 4th, 2011).

[9] M. Tikander, "Sound Quality of an Augmented Reality Audio Headset", In Proc. of the 8th Int. Conf. on Digital Audio Effects (DAFx-05), Madrid, Spain, Sept. 2005.

[10] J. Blauert, "Spatial Hearing" (revised edition), The MIT Press, Cambridge, Massachusetts, 1997.

[11] M. Tikander, M. Karjalainen and V. Riikonen, "An Augmented Reality Audio Headset," in Proc. of the 11th Int. Conf. on Digital Audio Effects (DAFx-08), Espoo, Finland, Sept. 1–4, 2008.

[12] V. Riikonen, M. Tikander and M. Karjalainen, "An Augmented Reality Audio Mixer and Equalizer", Audio Eng. Soc. 124th Convention, Amsterdam, May, 2008 (preprint 7372).

[13] D. Begault and M. Wenzel, "Direct Comparison of the Impact of Head Tracking, Reverberation and Individualized Head-Related Transfer Functions on the Spatial Perception of a Virtual Sound Source", Journal of the Audio Eng. Soc., Vol. 49, No 10, Oct. 2001.

[14] K. Hartung, J. Braasch and S. Sterbing, "Comparison of Different Methods for the Interpolation of Head-Related Transfer Functions", In Proc. of the AES 16th Int. Conf. on Spatial Sound Reproduction, 1999, pp. 319-329.

[15] A. Walker, S. Brewster, D. McGookin, and A. Ng, "Diary in the sky: A spatial audio display for a mobile calendar", In Proc. of IHM-HCI 2001, Lille, France, pp. 531–540, 2001.

[16] Open Sound Control (http://opensoundcontrol.org/ , last accessed: March 9th, 2011).

[17] Processing (http://processing.org/ , last accessed: March 9th, 2011).

[18] M. Noisterning, T. Musil, A. Sontacchi and R. Höldrich, "3D Binaural Sound Reproduction using a Virtual Ambisonic Approach", In. Proc. of International Symposium on Virtual Environments, Human-Computer Interfaces and Measurement Systems (VECIMS 2003), Switzerland, July 2003.

[19] F. Conway and N. Paterson, "Spatial Audio and Reverberation in an Augmented Reality Game Sound Design", In Proc. of the AES 40th International Conference, Tokyo, Japan, October 8–10, 2010.