Gestural Control Of Wavefield synthesis

Grani, Francesco; Di Carlo, Diego; Portillo, Jorge Madrid; Girardi, Matteo; Paisa, Razvan; Banas, Jan Stian; Vogiatzoglou, Iakovos; Overholt, Daniel; Serafin, Stefania

Published in:
Sound And Music Computing Conference Proceedings 2016

Publication date:
2016

Gestural Control Of Wavefield synthesis

Francesco Grani[1], Diego Di Carlo[2], Jorge Madrid Portillo[3], Matteo Girardi[4], Razvan Paisa[5], Jian Stian Banas[6], Iakovos Vogiatzoglou[7], Dan Overholt[8], Stefania Serafin[9]
SMC, Aalborg University Copenhagen
{[1]fg, [8]dano, [9]sts}@create.aau.dk
{[2]ddicar16, [3]jmadri15, [4]mgirar15, [5]rpaisa11, [6]jbanas14, [7]ivogia00}@student.aau.dk

ABSTRACT

We present a report covering our preliminary research on the control of spatial sound sources in wavefield synthesis through gesture based interfaces.

After a short general introduction on spatial sound and few basic concepts on wavefield synthesis, we presents a graphical application called *spAAce* which let users to control real-time movements of sound sources by drawing trajectories on a screen. The first prototype of this application has been developed bound to WFSCollider, an open-source software based on Supercollider which let users control wavefield synthesis. The *spAAce* application has been implemented using Processing, a programming language for sketches and prototypes within the context of visual arts, and communicates with WFSCollider through the Open Sound Control protocol. This application aims to create a new way of interaction for live performance of spatial composition and live electronics.

In a subsequent section we present an auditory game in which players can walk freely inside a virtual acoustic environment (a room in a commercial ship) while being exposed to the presence of several "enemies", which the player needs to localise and eliminate by using a Nintendo WiiMote game controller to "throw" sounding objects towards them. Aim of this project was to create a gestural interface for a game based on auditory cues only, and to investigate how convolution reverberation can affects people's perception of distance in a wavefield synthesis setup environment.

1. INTRODUCTION

The evolution of audio technology allowed for new listening setups to be experimented and evaluated. Long gone are the days of Thomas Edison's phonograph in 1877. Without doubt a milestone in the history of audio engineering, Edison's invention was able to both record and playback sound, however spatial fidelity was rather underwhelming, as the entire process was monophonic. Notably, not long after phonograph introduction, in 1881, a stereophonic playback device called the théâtrophone has been proposed by Clement Ader. The principle was simple - two microphones were placed across the opera stage and the signal collected by them was output to a pair of telephone receivers, placed in the opera house's foyer [1]. Later extensive research in this field slowly lead towards the commercial use of stereophony [2]. For some purposes it has been enhanced with an addition of a central speaker - mainly in cinemas, due to large dimensions of the screens. In consumer grade applications, stereophony has started to become widespread in the late 1950s with the invention of methods to engrave two channels onto a vinyl disc. At the same time, spatialisation of sound sources is an expressive tool that music composers had put into use since centuries. Dozens are the compositions of the 16th century Italian composer Giovanni Luigi da Palestrina that make use of spatial distribution of musicians. With the rise of the era of electronic music during the second half of the 20th century, the number of composers who pushed the boundaries of the available techniques in order to pursue their creative needs in terms of spatial sound just increased, often leaving commercial solutions behind the "brute force" *ad-hoc* methods adopted by composers and their sound technicians (just to mention few cases: Karlheinz Stockhausen's *Gesang der J́unglinge* (1955), Varese's *Poeme Electronique* (1958)). In particular cases artistic needs ended up in the construction of dedicated venues such as the Acousmonium, designed in 1974 by Francois Bayle to host spatialised sound concerts [3]. In most cases however the bridge between science and art has been very short, leading to various experiments in the field of recording and mixing techniques (e.g., [4]) that quickly brought us to a series of available techniques for the recording and reproduction of almost any desided sound field. Results achieved in sound spatialization techniques for systems of loudpseakers span today from stereo panning to more extended multichannel configurations, such as ITU 5.1 Surround [5], VBAP [6] and [7], DBAP [8], ViMiC [9], first and higer orders of Ambisonics [10], [11], as well as this project focus, wavefield synthesis (WFS) [12], [13].

Of all the above mentioned techniques, however, only WFS let in principle the listener to perceive the same designed soundfield in the same way and with the same auditory perspective from any point in space. This peculiar characteristic makes WFS a privileged technique to be adopted in situations like the ones described in this work.

2. BACKGROUND

The use of wavefield synthesis to recreate 3D sounds is not something new, having been pioneered in the late 80s at the TU Delft University, Netherlands [12]. In recent years, with the increase in availability of computational power the technology has gained a commercial interest, supported by hardware and software solutions such as the ones proposed by IOSONO [14] and Sonic Emotion [15], [16] as well as several open source engines available to control, simulate and render WFS [17], [18], [19], [20].

2.1 Basics of Wavefield Synthesis

The foundation of WFS lies on the theory concept of Christiaan Huygens: *Each point on a wave front can be regarded as the origin of a point source. The superposition of all the secondary sources form a waveform which is physically indistinguishable from the shape of the original wave front* [21].

The principle has been originally used to describe water and optical waves, and was first formulated for acoustics in 1988 at the TU Delft after being pioneering described in the 50s by Snow et al. [2]. A WFS system does require a large number of loudspeakers, placed as close as possible to the next one in order to create an array with as few discontinuity as possible. Each loudspeaker of the array corresponds this way to a secondary sound source and needs to be driven by a dedicated/independent signal thus requiring a large number of audio channels, equal to the number of loudspeakers; the signal for each channel is calculated by means of algorithms based on the Kirchhoff-Helmholtz integrals and Rayleigh's representation theorems [22], [23]. Due to the physical and software limitation, WFS systems are enduring several approximations, which introduce certain limitations and artifacts.

A first approximation needed to minimize complexity is to reduce the control of the sound field from a 3D to a 2D space (an horizontal -unlimited- plane). A second approximation consists into limiting the amount of secondary sources to a finite number (a finite set of loudspeakers); this approximation leads towards the consequence that the frequency range whereas a WFS system provides artifacts-free sounds gets reduced to the portion of the acoustic spectrum that is located below a threshold frequency, named "spatial aliasing frequency"; above this frequency artifacts will occur in the form of "ghost sound images". To cope with this limitation it is desirable to place the loudspeakers at the minimum possible distance -in our case 16.4cm, thus introducing a spatial aliasing threshold of 1048 Hz- and to design a sonic content which is not unbalanced towards hi frequencies. Another approximation consists on the fact that linear arrays of loudspeakers have a limited physical length and this generates what is called "truncation error", a phenomenon that limits the angles of incidence of sound sources in which a good result of WFS can be achieved. Further interferences can be introduced by the loudspeaker construction itself, as well as by the acoustics of the room in which the system is installed. An exhaustive description of WFS limits can be found in [24].

In wavefield synthesis technique it is usual to distinguish between three categories of sound sources that can be reproduced.

- Point sources: virtual sources that are placed anywhere outside the inner area of the loudspeakers array.
- Plane waves: sources that are ideally placed at an infinite distance, thus their incident wavefront can be described as plane.
- Focused sources: sound sources that are located inside the area covered by of the loudspeakers array.

3. SPAACE

3.1 Introduction

The spatial characteristic of a composition has been an important topic for the avant-garde musicians in the past decades [25] and it still is a relevant quality of a musical piece [18]. From an artistic point of view, conveying spatial musical ideas and thoughts could underestimate the technical issues that must be faced during the development and implementation process of a software for musical purposes. Hence, contemporary composers and sound engineers have to find a trade-off during such process of composing new musical material. Moreover, learning new technologies or softwares for spatial music could be time consuming for composers that do not have a deep knowledge in the computer music field. Therefore, the *spAAce* application attempts to provide the following advantages:

1. a quick way to sketch and test movements of sound sources.
2. improvisation with spatial sound sources during live performances.

The latter point could lead to new approaches of performing live concerts in a live electronics scenario. Indeed, WFSCollider is employed just as engine render while the composers can focus on creating trajectories for sound sources and expanding musical expression.

3.2 State of the art

As introduced in 2, several different softwares for spatial sound movements have been developed in the recent years [26] [27] [28] and several spatial techniques have been implemented such as VBAP, DBAP, Ambisonics and Wavefield synthesis (WFS). Most of the spatial rendering engines come with a Graphical User Interface, such as SoundScape Render [28], Spat [26], WFSCollider [27] itself and others.

With the spreading of user-friendly GUI development environments for mobile and web app, some applications have been developed for these rendering engines, which allow real-time finger-based interaction. Some of them are more are mixing-oriented, providing a real-time positioning of the sound sources in the space, while others, such as Trajectoires [29] and Spatium [30], allow the users to move the sound sources and create complex paths both in time and in space. However, these new applications are still in the embryonic stage and there is a lot of work still to be done in order to design the most suitable interface that can capture the actual intentions of the performers. The key concept of *spAAce* is the combination of several modes of interaction that should encourage and enable artistic sound spatialization.

3.3 Design process

Knowing what kind of technologies are usually employed, the development path has led us to employ Processing [31] as development platform, since it has many libraries and a strong community of developers. The next step has been asking to composers and sound engineers for interviews to test the concept of the *spAAce* application and to receive general feedback, hints and suggestions. The composers and sound engineers reached are highly involved in contemporary music production and live electronics performances. These interviews gave a solid starting point to implement the main core of the application. An iterative procedure of *implementation - testing - bugfixing* has been then employed for the development of the application.

3.4 Software architecture

The graphical user interface has been developed to be controlled by a multi-touch screen like a tablet or similar. The interface allows the user to create, control and delete trajectories for the sound sources. There are three types of trajectories:

- line trajectory
- circular trajectory
- free hand drawing

These trajectories are displayed as buttons on the left upper corner of the screen. Sound sources can be dragged with a finger and when placed on the top of one trajectory, getting automatically to follow the trajectory's path with a default speed. This speed can be changed with a knob on the bottom right side of the screen. Additionally, there is a control panel on the bottom left side of the screen where users can select each sound source to create control groups and perform mass editing. Lastly, since the OSC protocol is employed in order to communicate with WFSCollider, the users can set the proper IP Address and OSC port by clicking on the "Network" button located on the right upper side of the screen.

3.5 Device and Controller

The controller is not constrained to a particular hardware since the development has been done in Processing, which is a multi-platform application. In our testing sessions, a Wacom 22" multi-touch screen has been used, and a Leap Motion has been added in order to track hand movements and perform specific actions, such as moving sound sources towards a particular direction, applying spatial effects or even to control more than one sound source at the same time. Overall, the controller is designed so that the expression in terms of spatialization can be improved as much as possible.

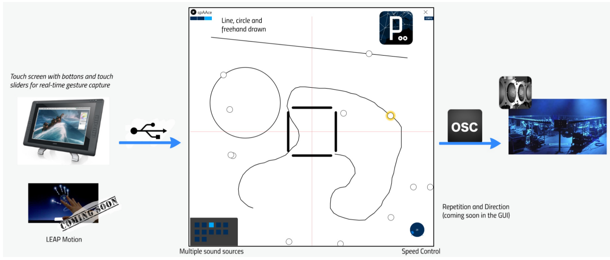

Figure 1. The *spAAce* architecture.

3.6 Physical Setup

Since *spAAce* is a brand new project, it is continuously tested in the Multi-Sensory Experience Laboratory of Aalborg University in Copenhagen. All the experiments have been conducted with the following setup: A WFS system of 64 loudspeakers - presented with more details in 4.3; a computer running WFSCollider server for sound rendering (the OSC server); a laptop running the Processing sketch *spAAce*, placed near the center of the WFS system and connected to the server via Ethernet (the OSC client); a Leap Motion and a Wacom Cintiq 22" touch standing in the center of the system as a input/output devices for the client.

3.7 Evaluation

To understand if *spAAce* is correctly abiding to our user's needs, we have tested the prototype of our system in two consecutive phases: an iterative phase followed by a final one. The goals of the iterative testing were to gage product usability and evaluate the features we implemented, so that we could converge and focus only on the main ones. The goal of the final test was to provide a general and more complete evaluation of our system and to show some concrete results. We focussed on investigating subjective qualities inherent to the musical experience, such as enjoyment, expressiveness and perceived affordances, both from the *performer* and the *audience* side. In this paper we focus on the performer's side.

In total we had nine participants for the *performer* testing (all students of Aalborg University). This participant number does not provide statically significant evidence, however it is a reasonable number of participants for evaluating the overall system prototype and retrieve some useful feedback and comments. Seven of the subjects were male, and 6 of them had previous musical experience. Four of them did not have any previous experience with similar softwares, and the rest reported to have some degree of experience. The age of the students was between **19 and 29** years old. Sixteen graduated students from Aalborg University in Copenhagen voluntarily participated in the *audience* test which, however, is not presented in this report. The experiments have been conducted the in the *Multi-Sensory* Laboratory of the Aalborg University in Copenhagen, and we used the same setup described in 3.6. To collect feedback from the participants a survey with qualitative questions has been used and we applied a 7-point Likert scale.

Testing the performers

To evaluate all the performer-related parameters, we designed a two-parts test. The first part aims to evaluate the system and the interface exploration and learnability. For this test the participants were divided randomly in two groups, one with a small training, the other without. The groups were asked to perform some basic tasks to explore the main function of the system. Thus, assuming that all participants achieve the same basic knowledge after completing the first tasks, participants were asked to perform their own creative spatial composition as second part of the test. There were no time constraints for the testing and after the completion of the every tasks, users were asked to fill out a *user evaluation* survey.

3.7.1 Results

According to the data measured, the GUI has revealed to be easy an intuitive to use (data collected show an optimal score with a mean of 5.88 and a low deviation of 1.05). However this result might not be strongly reliable since not enough participants were tested and some of them knew the application on beforehand. Figure 2 shows the overall system experience regarding user usability.

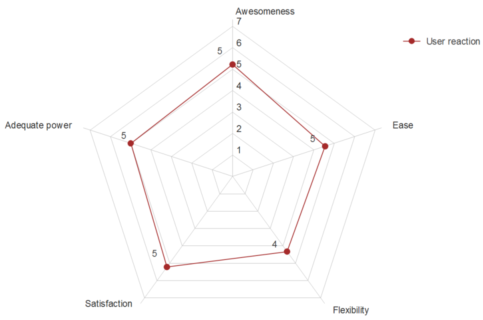

Figure 2. Radar chart that visualises the mean of the 1-7 Likert scale results.

Evaluating the performer experience

Here we tested if the user-friendly interface and the natural approach of the Wacom tablet provides an easy and simple interaction. The answer is basically yes, but on the other hand, Leap Motion was found by all users difficult to manage and would probably need practice in order to be an effective control interface. Figures 3 and 4 show how the users rated the learning curve. The main conclusion is that the system has almost no entry-fee thanks to the already well known tablet interaction, designed and developed with a user-centered framework. Of course the issues with the Leap Motion remain, but research shows that training and improvement of gestures and mapping could lead to a good mastering of the system. Also, comments from the test subjects show that people found Leap Motion quite useful and natural in order to control sound sources, even if the experience was harsh and frustrating at first.

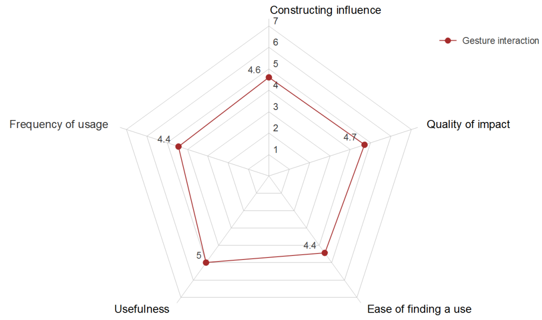

Figure 3. Mean of the Likert scores for evaluating the gesture interaction between the performer and the system.

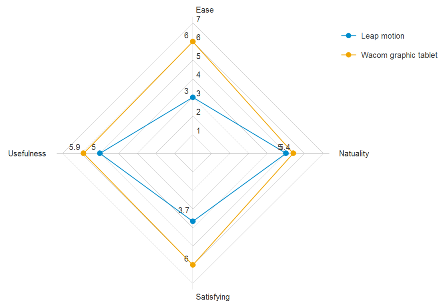

Figure 4. Mean of the Likert scores for evaluating the input devices.

Conducting vs Playing

According to the t-test *p*-value (0.36, alpha = 0.05) there is a significant difference between spatial music trained people and the others. The results show how participants thought that they were mixing rather than playing with the sound sources. We believe the title of the question was probably misunderstood. The ”mixing” rated values belonged to participants highly trained in WFS techniques and/or amateur composers, while the participants who rated ”playing” were less trained in these regards. There is perhaps an expectation bias and also a problem of terminology which could be improved in further tests.

Satisfaction and Enjoyability

As show in figures 3, 4 and 2 the satisfaction was very high, and a very positive feedback, ofter even passionate, was felt with respect of the subjects enjoyability and satisfaction. This shows that *spAAce* can be very enjoyable, even if the sound sources movements are not fully understood. However we think that the novelty of the system and the astonishment of watching the system in the laboratory space, has influenced greatly the subjects.

4. WFS GAME

4.1 Introduction

One goal of this project is to investigate how convolution reverb affects people’s perception of distance in a wavefield synthesis setup environment. In order to achieve this, an auditory game prototype has been developed and to keep the focus on auditory perception, players do play the game blindfolded. The style of the game is horror/survival and the user is exposed to several “enemies”, which he/she needs to localize and eliminate by using a Nintendo Wi-iMote game controller to “throw” sounding objects towards them. There are three types of enemies with different mechanics and sonic characteristics that will be described in a next section. They all are created by using point sources and focused sources, and they are wither static or moving around or towards the player after they appeared in the virtual space around the player.

4.2 Impulse Response Reverb

The environment of the game resembles a commercial ship, thus a background ambience soundscape was designed containing sounds such as an air fan, water drops from a broken pipe, wind sound coming from outside the ship and rat squeaks. Acquiring the impulse response from a ship was essential, since this project relies on investigating the role of convolution reverb in distance perception for WFS. The Impulse Response was captured using the ESS (Exponential Sine Sweep) method [32], in a big metal ship owned by the Illutron Collaborative Interactive Art Studio [1] . The following equipment was used in the process: a MacBook Air Laptop, a Dynaudio BM5 MK I speaker, a Rode NT2 omnidirectional microphone and a Focusrite Scarlett 8i6 audio interface. The recording and deconvolution was handled via the Apple Logic Pro X internal Impulse Response Utility.

4.3 Hardware and Software

Since the game was designed to allow the player to move freely in the area inside the array of loudspeakers, a shooting system has been implemented coupling a WiiMote with two motion capture markers captured by an array of 16 OptiTrack Flex 3 infra-red cameras. Of the two MoCap markers one had been placed on the player’s shoulder and another one on the WiiMote.

Unity3D was running as a local debug software as well as an interpreter from VPRN to OSC, since NaturalPoint cannot send OSC data. The Unity5 game engine, the WiiMote OSCulator 2.13.3 receiver and the WFSCollider sound engine software were all running on a separate Mac Pro computer (dual Intel Xenon 12 core processor, 64 GB DDR3 RAM).

The WFS audio stream is delivered from an RME MADIface USB interface via two DirectOut ANDIAMO 2 MADI converters, each connected to 32 M-Audio BX5 D2 loudspeakers. In total the WFS system delivers sound trough 64 loudspeakers aligned one to the other and calibrated in their output level. The WFS system consists of 4 arrays of 16 loudspeakers each, displaced to form a square of 4 by 4 meters inside which users can freely move.

The wavefield synthesis had to happen in realtime after the user input and according to the enemies positions, to maintain the desired playability. Also for this project the choice of WFS engine fell on WFSCollider, the audio spatialization engine for Super Collider developed by Wouter Snoei at The Game of Life Foundation [2] . Beside the capability of rendering wavefield sound, WFSCollider also serves as an intuitive digital audio workstation (DAW) offering functionalities such as multi-track mixing, effect chains, auxiliary buses, featuring also an easy OSC control on every parameter, thus making it very suitable for the desired setup of this work. In WFSCollider sound sources are triggered and controlled in position and properties by control messages coming from Unity5 and OSCulator.

4.4 Sound Design

Three types of enemies were designed for the game, which will be described as Enemy 1, Enemy 2 and Enemy 3. The numbering represents the order of apparition. All these enemies spawn randomly at different locations, from three different “rings” or levels of distance. Enemy number one will always appear from the further area, while enemy number two will be appearing from the closest one, leaving enemy number three to appear from the mid one. See Figure 5 where E1, E2, E3 represents the three enemies; the rings represent the three different areas of distance where enemies are coming from; the square represents the physical space enclosed by the WFS array, and P represents a player.

The lifetime of each enemy is 1 minute. This limit is implemented to compensate for an issue encountered in the pilot tests: sometimes, the user cannot hit an enemy.

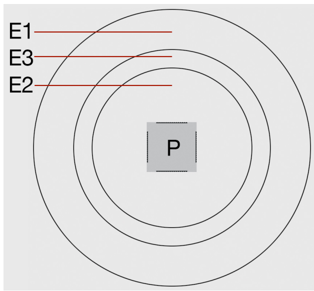

Figure 5. Map of the virtual space.

1. Enemy 1 slowly moves on a linear trajectory towards the player and tries to “hit” him/her, emitting a continuos flow of sound while it moves. E1 depicts a human dragging a heavy metal object and the most evident sound characteristic of this enemy is its slow footstep movement. Several sounds are used to create it: a pair of foot sounds alternating, a recorded heavy breathing sound, as well as the sound of a metal object dragged on a metal surface. All sounds are grouped together; as long as their virtual position is located outside the ring of speakers they are rendered as point sources, and when they get close to the player and “enter” the loudspeaker area, they become focused sources.

2. Enemy 2 position is static. E2 represents a woman who is breathing fast and sobbing while spinning a

[1] http://illutron.dk

[2] http://gameoflife.nl

chain, its sound characteristics are then female screams and a swinging flail weapon sound. This enemy is immobile and it alternates short silences and sounds. Three sounds were used to create her, a chain links clinker, a recorded sobbing/ breathing sound and a vocal sound. Just as Enemy 1, these sounds are grouped together and if the enemy appears in the area behind the loudspeakers they are rendered as point sources, otherwise they are rendered as focused sources if E2 appears in the area inside the loudspeakers array.

3. Enemy 3 combines together some of the mechanics and sound characteristics of E1 and E2. Its position is static but every 20 seconds it spawns a series of moving distractive sounds, which travel around the virtual space where the player is, making it harder of the player to locate and eliminate him/her. E3 symbolises a ward drum player, with a twist, and only one sound source is used to create it: a rhythmical uninterrupted drum loop. For the distraction sound, several male exhale sounds were used, being processed to sound like a wrath. These two sounds (E3 and its "distractors") combine together one continuous sound cue with a series of short sounds that appear and go. Just as the other two enemies, all the sounds are point sources when their virtual location is located outside the speaker area, and focused source otherwise.

4.5 Test Design

Each subject is introduced to the game mechanics by going through a training phase which consists of three stages, each lasting one and a half minutes and dedicated to set the player familiar with the relation between the gesture he/she has to perform (direction and force of the gesture) to "throw" a sound against an enemy, and the distance at which the sound is thrown. In this phase the subject is already blind-folded and is requested to locate and hit the virtual sound by "shooting" another sound with the WiiMote towards it, according to the subject's perception of how far the target sound is located.

The training sound to hit resembles a synthetic metronome beat and is located into one of the three circular areas visible in Figure 5; the player receives a sound feedback to understand if the shoot was good (the gesture was performed with the exact force needed to launch the sound into the desired area) or not. The sound to hit remains the same on all three stages but the distance increases from the inner to the outer circular areas as the stages progress. Once a participant has been familiarised with how the game works, the actual testing starts.

The real game/test comprises also of three stages, one for each of the three enemies. In each of the three stages the participant is exposed to eight instances of every enemy, four of these are presented with impulse response reverberation and four without, randomly assigned. The participants actions are tracked throughout the test and logged to files. The log entries include player position, collisions coordinates and timing, number of shots fired during each of the phases and number of enemies spawned and hit.

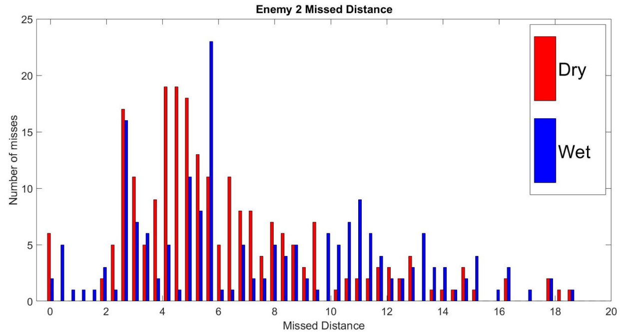

Figure 6. "Spatial precision": missed projectile distance.

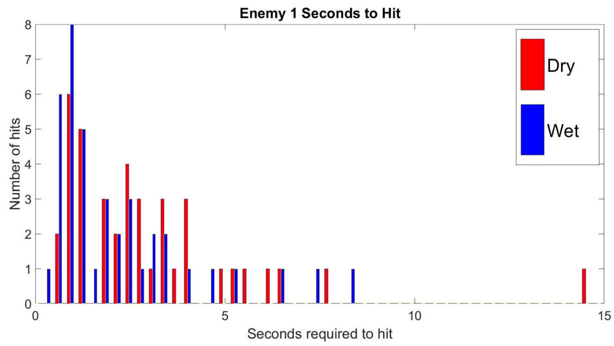

Figure 7. "Game performance": time required to hit an enemy.

4.6 Experiment Results

The test was performed on 9 males and 1 female participants aged between 21 and 27, all of them reporting to have musical training and suffer no hearing loss. The results have been gathered in three main categories. These include a distance between the projectile impact position and a target enemy (namely "spatial precision"), the time required to correctly shoot an enemy and the total accuracy of shots (number of good shots versus bad shots). In each of these categories a paired t-test has been performed to verify whether or not the presence of reverberation (wet/dry parameter) had an influence on the participants' achieved scores. This procedure was carried out three times, once for each of the enemies that the test participants were exposed to, thus getting a total of 9 tests. The paired t-test revealed that results provide no statistical significance required to determine whether or not entries related to the reverberant environment condition differ from those logged in the dry condition (= with no impulse response based reverberation - only the natural dry reverberation provided by the lab room in which the WFS system is placed). Among all the nine tests, only one yielded significant difference between two conditions - the one ran on missed projectile distance entries during an Enemy 2 phase (p=0,046). However, even in this case, the difference between mean values is equal to 20.1949-16.5106=3.6843, a relatively low number -Figure 6.

Statistical analysis of data does not bring a solid answer to the hypothesis that a difference is in place between the performance achieved in shooting at the correct distance in wet or dry reverberation conditions. Also the analysis performed on the number of seconds required to hit an enemy, shows no difference in all nine cases, so only one plot is here presented as an example of the results (Figure 7.), leaving further reflections to the discussion part.

4.7 WFS Game Discussion

The analysis of data shows no significant difference in the results performed with and without convolution reverberation, nevertheless it is worth mentioning that both the system used as a tool to perform the test, and the experiment design itself have possibly affected the outcome substantially. First of all, the gestural interface was commonly reported by subjects to be counter-intuitive and non-reliable and hence it can be partially blamed for an overall poor performance of the users (in terms of accuracy, time required to aim and average missing distance); moreover, this aspect raised frustration and distraction from the task. Consequently, participants tended to become tired towards the end of the test, which led to further deterioration of their score. The main reason behind this issue has been addressed as the delay between the motion capture system and the WiiMote input data flows. The stream of data from the MoCap computer, to the computer receiving the WiiMote data, is affected by a small lag, that causes incorrect reading on the users hand position in the moment when they trigger the WiiMote button to "throw" their sonic weapon. This small lag sometimes causes a wrong reading of the relative position of the two markers (the one placed on the player's shoulder and on the WiiMote), which in the end can generate a wrong shooting angle. This error is more pronounced in users who perform a very fast and energetic movement with the WiiMote. This problem could be overcame by changing the shooting mechanism. Another solution to overcome the lag would be to redesign the data flow either making use of a single computer, or relying only on the WiiMote internal sensor data fusion to generate an accurate shooting direction.

The evaluation aspect of this project was revolving about the impact of convolution reverb in a WFS system, but this is not the only way to create artificial room simulations; different techniques could be adopted instead of a direct convolution of the sound sources: for future studies another option could be to model reverberation as four planar waves representing physical walls and fed with all the signals to be convolved. Also incorporating completely different approaches, such as Schroeder reverberators might be worth investigating. It is in the end worth mentioning that this project completely omits the proprioception aspect of the experience. Early tests suggest that the perception of the shooting hand might influence the shooting performance from player to player. A further experiment investigating this aspect could provide useful information in understanding the analysed data, as well as provide useful knowledge for designing interactions and interfaces for alike systems. At last, also sound design aspects could be affecting the results and be worth investigating more, since besides comparing moving sounds and static sounds, the sounds themselves embed different temporal and spectral contents which might affect subject's perception.

5. CONCLUSIONS

We presented two preliminary studies of gesture control of Wavefield Synthesis performed in our research group: a graphical application called *spAAce* which let users to control real-time movements of sound source by drawing trajectories, and an acoustic-based game which aimed to investigate the impact of convolution reverberation over the perception of distance in a wavefield synthesis scenario.

5.1 spAAce

The first prototype of this application has been developed bound with WFSCollider, an open-source software based on Supercollider. In order to communicate with the software, Open Sound Control protocol has been employed. The *spAAce* application has been implemented using Processing, a programming language for sketches and prototypes within the context of visual arts. This application aims to create a new way of interaction for live performance of spatial composition and live electronics. Promising results have been found in the small test performed, encouraging the authors to further implement the system. Future work will focus on a more extended test (to eliminate possible biasing caused to the users by the novelty of the "experience") and on further development of the available tools used to create and control sound trajectories. Also, a more general version of the *spAAce* software is desired to be developed in a later stage, to run the software on iOS and Android tablets.

5.2 WFS GAME

While the test results for this project were mainly inconclusive, the miss distance for Enemy 2 was statistically proven to be influenced by the reverb status, accepting the null hypothesis, indicating that a dry sound sources were slightly easier for the participants to hit. Nevertheless, the platform used for this research is worth further development as it provides more possibilities for examining embodied interaction in a virtual auditory environment. Also, besides the considerations on further possibilities of study on how to implement a more effective setup for the experiment purposes, another interesting way of exploiting it would be to include interaction between people, so that more users can interact with the environment.

6. REFERENCES

[1] F. Rumsey, *Spatial audio*. CRC Press, 2012.

[2] W. Snow, "Basic principles of stereophonic sound," *Audio, IRE Transactions on*, no. 2, pp. 42–53, 1955.

[3] B. François, "Pour une musique invisible: un acousmonium," *Festival International du Son Haute Fidélité Stéréophonique*, pp. 125–134, 1975.

[4] J. M. Chowning, "The simulation of moving sound sources," *Journal of the Audio Engineering Society*, vol. 19, no. 1, pp. 2–6, 1971.

[5] I. Recommendation, "775-1, multichannel stereophonic sound system with and without accompanying picture," *International Telecommunication Union, Geneva, Switzerland*, 1994.

[6] V. Pulkki, "Virtual sound source positioning using vector base amplitude panning," *Journal of the Audio Engineering Society*, vol. 45, no. 6, pp. 456–466, 1997.

[7] ——, "Generic panning tools for max/msp," in *Proceedings of International Computer Music Conference*, 2000, pp. 304–307.

[8] T. Lossius, P. Baltazar, and T. de la Hogue, *DBAP–distance-based amplitude panning*. Ann Arbor, MI: MPublishing, University of Michigan Library, 2009.

[9] N. Peters, T. Matthews, J. Braasch, and S. McAdams, "Spatial sound rendering in max/msp with vimic," in *Proceedings of the 2008 International Computer Music Conference*, 2008.

[10] M. A. Gerzon, "Ambisonics in multichannel broadcasting and video," in *Audio Engineering Society Convention 74*. Audio Engineering Society, 1983.

[11] M. Gerzon, "Surround-sound psychoacoustics: Criteria for the design of matrix and descrete surround-sound systems." *Wireless World. Reprinted in An anthology of articles on spatial sound techniques, part 2: Multichannel audio technologies. Edited by F. Rumsey. Audio engineering society, Inc, 2006.*, 1974.

[12] A. J. Berkhout, "A holographic approach to acoustic control," *Journal of the audio engineering society*, vol. 36, no. 12, pp. 977–995, 1988.

[13] A. J. Berkhout, D. de Vries, and P. Vogel, "Acoustic control by wave field synthesis," *The Journal of the Acoustical Society of America*, vol. 93, no. 5, pp. 2764–2778, 1993.

[14] S. Brix, F. Melchior, T. Roder, S. Wabnik, and C. Riegel, "Authoring systems for wave field synthesis content production," in *Audio Engineering Society Convention 115*. Audio Engineering Society, 2003.

[15] E. Corteel and T. Caulkins, "Sound scene creation and manipulation using wave field synthesis," *Rapport technique, IRCAM*, 2004.

[16] R. Pellegrini and C. Kuhn, "Wave field synthesis: Mixing and mastering tools for digital audio workstations," in *Audio Engineering Society Convention 116*. Audio Engineering Society, 2004.

[17] J. Ahrens, M. Geier, and S. Spors, "The soundscape renderer: A unified spatial audio reproduction framework for arbitrary rendering methods," in *Audio Engineering Society Convention 124*. Audio Engineering Society, 2008.

[18] M. Baalman, "Application of wave field synthesis in the composition of electronic music," in *International Computer Music Conference, Singapore*, 2003, pp. 1–4.

[19] M. A. Baalman, "Updates of the wonder software interface for using wave field synthesis," *LAC2005 Proceedings*, p. 69, 2005.

[20] W. Snoei. Wfscollider 2.2.1. [Online]. Available: https://github.com/GameOfLife/WFSCollider

[21] T. Sporer, "Wave field synthesis-generation and reproduction of natural sound environments," in *7th International Conference on Digital Audio Effects (DAFx-04), Naples, Italy*, 2004.

[22] S. Spors, R. Rabenstein, and J. Ahrens, "The theory of wave field synthesis revisited," in *124th AES Convention*, 2008, pp. 17–20.

[23] E. N. G. Verheijen, "Sound reproduction by wave field synthesis," Ph.D. dissertation, TU Delft, Delft University of Technology, 1998.

[24] S. Spors and R. Rabenstein, "Spatial aliasing artifacts produced by linear and circular loudspeaker arrays used for wave field synthesis," in *120th AES Convention*. Citeseer, 2006.

[25] K. Stockhausen, "Musik im raum," in *La Rassegna Musicale: Problemi della musica, oggi (Numero Speciale)*, G. M. Gatti, Ed. Giulio Einaudi editore, 1961, vol. 31, IV, pp. 397–417.

[26] J.-M. Jot and O. Warusfel, "Spat: a spatial processor for musicians and sound engineers," in *CIARM: International Conference on Acoustics and Musical Research*, 1995.

[27] "CodeOfLife wfscollider," https://github.com/GameOfLife/WFSCollider.

[28] M. Geier, T. Hohn, and S. Spors, "An Open Source C++ Framework for Multithreaded Realtime Multichannel Audio Applications," in *Linux Audio Conference*, Stanford, USA, Apr. 2012. [Online]. Available: http://lac.linuxaudio.org/2012/download/lac2012_proceedings.pdf

[29] X. Favory, J. Garcia, and J. Bresson, "Trajectoires: une application mobile pour le contrôle et l'écriture de la spatialisation sonore," in *27ème conférence francophone sur l'Interaction Homme-Machine.* ACM, 2015, p. a5.

[30] R. Penha and J. Oliveira, "Spatium, tools for sound spatialization," in *Proceedings of the Sound and Music Computing Conference*, 2013.

[31] "Processing," https://processing.org/.

[32] A. Farina, "Advancements in impulse response measurements by sine sweeps," in *Audio Engineering Society Convention 122*. Audio Engineering Society, 2007.