



University of Bordeaux
2013-2014

Projet d'Étude et de Recherche

Three-dimensional musical instrument

Supervisor: Aurélie BUGEAU

Clients: Myriam DESAINTE-CATHERINE, Joseph LARRALDE, and Florent BERTHAUT

Mohamed BOURARA, Jean BUI-QUANG, Omar OURHI, Jean-Michaël CELERIER

Marie IMMACULA OMISCAR and Damien CLERGEAUD

January 23, 2014

Contents

1	Introduction	4
2	Subject and definitions	5
2.1	Subject presentation	5
2.2	What is a 3D musical instrument?	6
2.3	Immersion	7
2.4	Control	7
3	Three-dimensional displays	9
3.1	Definition of a 3D display	9
3.1.1	The problem	9
3.1.2	Parameters	9
3.1.3	Presentation of common visual cues	9
3.1.4	Definition of a 3D display	10
3.2	Classification of the 3D displays	10
3.2.1	Criterions	10
3.3	In-depth presentation of some 3D display methods	11
3.3.1	Two-view 3D displays	11
3.3.2	Presentation of the methods	11
3.3.3	Multi-view displays	13
3.3.4	Omni-view Displays	14
4	Presentation of 3D musical instruments	17
4.1	History of the 3D musical instruments	17
4.1.1	3D Navigation instruments	17
4.1.2	Single sound synthesis instruments	17
4.1.3	The Drile	17
4.2	The Drile	17
4.2.1	Hierarchical Live-Looping	18
4.2.2	Piivert, a new controller	18
4.2.3	Drile world	19
4.3	Aerial Percussion	21
4.3.1	Hardware: Polhemus Liberty	22
4.3.2	Software: SetKreator and Flock of Birds	23
5	Implementation	24
5.1	Objectives	24
5.1.1	Finding a display	24
5.1.2	Implementing suitable renderings	24

5.1.3	Customization	24
5.2	Implementation	24
5.2.1	Chosen display techniques	25
5.2.2	Aerial Percussion implementation	26
5.2.3	Physical set-up of the system	27
6	Conclusion	29
Glossary		30

List of Figures

2.1	Picture of a musician using Drile	6
3.1	Common visual cues	10
3.2	Red/cyan colour filter 3D glasses	12
3.3	Polarized 3D system	12
3.4	A lens array	13
3.5	A parallax barrier	14
3.6	3D multiviewpoint display	15
3.7	Pepper's ghost	15
4.1	Piivert with pressure buttons on the bottom	19
4.2	Tunnels	20
4.3	Tree nodes containing colored worms. Appendages represent tree edges	20
4.4	Scene selection for a worm translation	21
4.5	On the left the dimension space representation, on the right the Drile representation (in purple for beginner interface, in blue for complete control)	21
4.6	Polhemus Liberty and sticks.	22
4.7	SetKreator	23
5.1	A screenshot of the Aerial Percussion renderer	26
5.2	Our installation	27
5.3	Our installation. 1 and 3 are the projection screens, 2 is the Plexiglas screen.	28

Chapter 1

Introduction

A three-dimensional musical instrument might sound quite abstract for the bystander. One can think of it as a musical instrument taking place in the virtual reality or augmented reality domain. While an exact definition might be hard to settle because every instrument will be different in core features to others, a general definition might be an instrument which can have either :

- A visual representation in a three-dimensional space
- Interactions in a three-dimensional space

The two points are generally shared, however it is harder to display the instrument in 3D than to interact with it.

The display can have two goals : [BHDC10]

- Giving visual cues to the spectators of the musician's actions.
- Helping the musician to perform.

One of the main focuses of this exposé will be to assess the different display techniques suitable to a 3D instrument, and the other will explain how it can improve existing 3D instruments.

The last part of this report will be about the choices we had to make in order to set-up our own 3D musical instrument.

Chapter 2

Subject and definitions

2.1 Subject presentation

3D Musical instruments

At the SCRIME¹ and LABRI², three-dimensional musical instruments have been implemented within the context of research in interactive virtual reality and music computing.

The DRILE [BDCH⁺10] is a 3D musical instrument which allows manipulation of the structure of a song using [live looping](#), in an immersive virtual reality scene.

The AERIAL PERCUSSION is a 3D musical instrument which generates sounds using the position in space of sensors which are put at the end of drumsticks. Virtual 3D shapes like cubes, cylinders, are positioned around the instrument and the musician. According to the position, the orientation, and the speed of the sensors, sounds are generated.

Objectives

We were asked to implement a prototype of a 3D render and display device, for musical performance. It is necessary to take into account the constraints inherent to a musical performance environment, as well as the constraints of the instruments.

Here are some constraints for the performance :

- The musician has to be in front of the audience.
- The musician requires visual cues inherent to the utilisation of the instrument, and the audience must see the instrument to understand the gestures and actions of the musician.

To enact this implementation, a precise explanation of the nature of the 3D musical instruments is required.

Plan for this section

We will first make a short presentation of 3D musical instruments, present and will then define the concepts of immersivity and interactivity.

¹Studio de Création et de Recherche en Informatique et Musique Électroacoustique

²Laboratoire Bordelais de Recherche en Informatique

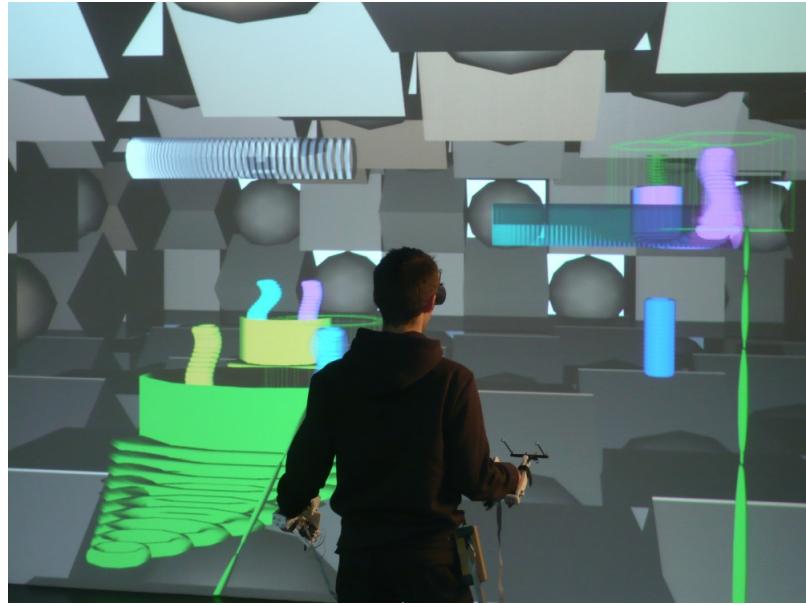


Figure 2.1: Picture of a musician using Drile

2.2 What is a 3D musical instrument?

A 3D musical instrument is an instrument which is designed thanks to the progress in virtual reality and computer music. Virtual reality provides display and control methods while computer music technologies provides sound and music synthesis. The goal is to use these tools to design an instrument which would be precise, easy to handle, and have an important artistic quality.

Example for the Drile

For instance, the picture 2.1 shows a musician with special glasses, as well as joysticks with force-feedback haptic sensors (Piivert [BHDC10]). The user handles 3D shapes in a 3D environment to influence the music generation. A specific part of this report will be dedicated to a precise study of the **DRILE**.

2.3 Immersion

Definition

Immersion is a psychological state where the subject stops about taking care of its own physical state. The immersion is quite important in virtual reality. For instance, for the Aerial Percussion, it would come down to the state the performer is when he stops thinking consciously of the disposition of the shapes he interacts with. The musician will then be immersed in the virtual 3D environment which consists in the shapes disposition.

Hence, to immerse an user, multiple parameters are accessible to the 3D instrument designer. They are mostly linked to the senses of the human body. In our project, we will mainly focus on vision, and more precisely on 3D display devices.

Interactivity & reactivity

Interactivity is an important aspect of the immersion capacities of a musical system. The user is more likely to get an immersive feeling in a virtual world, if the world instantly reacts to his actions and gestures [BD95]. This leads us to an important part of the 3D musical instrument : the interface, the controls that the performer requires to operate the instrument.

2.4 Control

In order to understand and define what is the control of a 3D musical instrument, it is necessary to think of the instrument in two different ways :

1. First of all, it is a musical instrument, which implies multiples constraints :
 - It has to be adapted to the human body shape so that the musician can manipulate it.
 - It has to be precise enough for the performer to be able to learn how to play the instrument.
2. But it is also an interactive immersive system, which means that it requires :
 - A lot of interactivity.
 - The usage of senses for a feedback : for instance, immersive visual and haptic feedback (as well as auditive, since it is a musical instrument).

Musical instrument gestures

In order to conceive a 3D musical instrument, it is necessary to understand the movements that the musician does while playing. For instance, the Cadoz gesture segmentation [Cad99] is an attempt to differentiate different families of gestures while playing.

Cadoz defines three kinds of gestures that all come to play when playing music:

- Selection gestures : The musician selects a component of the instrument that he will play on. For instance, for stringed musical instruments, where the same tone can be achieved on different strings, but with a different timber, it is the choice of the string on which the note will be played.
- Modification gestures : It is an action that modifies the physical state of the instrument. For instance, it would be the case of the guitar player who presses its hand on the guitar strings against the wood.

- Excitation gestures : These actions are the ones generating the actual sound of the instrument, by making the air vibe. On a guitar, it would be picking a string, and on a violin, it would be moving the bow against a string. This gesture is the one the artist can put expression inside : for instance, a violinist can press his bow softly or hardly, in order to change the nuance.

Control and immersion

In order to correctly play an instrument, the user requires some kind of manipulation comfort : it must not be painful or too tiring to play. Hence the requirement for immersion : being immersed means that the user does not need to put effort into playing the 3D musical instrument, he becomes part of it. An easy way to improve immersion is to make the environment react to the performer's movements.

This can be achieved by using head tracking, with a **KINECT** for instance, or a **HMD¹**, a **CAVE²** or other virtual reality devices and methods like Fish tank VR [RCVD97]. This allows to adapt the scene's projection to the movements of the head of the user : for instance, if he turns his head to the right, the display will adapt by showing him what he would see if it was real.

Haptic feedback buttons also increase the consciousness of the user's actions, which implies an increased precision.

Florent Berthaut thought about most of these problems while conceiving **PIIVERT** [BHDC10], the control interface to the **DRILE**.

¹Head-mounted Display

²CAVE Automatic Virtual Environment

Chapter 3

Three-dimensional displays

3.1 Definition of a 3D display

While it is commonplace to hear about 3D display in television or smartphone advertisement nowadays, the distinction between 2D and 3D might be more difficult to settle.

3.1.1 The problem

If we take the simple definition : a 3D display is a display that can show 3D images, it is really ambiguous, because of what is supposed to be "3D". For instance, for years, video games have been advertising 3D engines and spectacular 3D graphics, even without the depth provided by what is now called 3D displays or 3D movies in cinema. Another literal but limited definition for a 3D display would be a display that really exists in three-dimensions; one could think for instance of programmable matter and claytronics, or at least of a display that would be able to show a scene or an object from any point of view.

Hence, we have to qualify what is 3D and what it is not, in order to build a real definition.

3.1.2 Parameters

In the literature, the main idea is to relate to the human brain and body capabilities to define 3D vision ([Oko76], [PS12]). For instance, a big part of the "3D" feel is due to the fact of having two eyes that looks in the same direction, but from a slightly different angle, but it is not the only parameter.

The visual cues of 3D vision are separated in two families:

- Physiological cues. They will relate to the capabilities of the human body.
- Psychological cues. They will relate to the information inference capabilities of the human brain.

3.1.3 Presentation of common visual cues

Most of the visual cues on 3.1 are explained in-depth in [PS12], however a short explanation is provided in the glossary for many of them.

A complete classification of the current state of the art in 3D displays using these visual cues is present in [MPWL13] on Table 2.

Psychological cues	Physiological cues
Occlusion	Stereoscopy
Linear perspective	Convergence
Atmospheric perspective	Accommodation
Shading	Retinal image size [MPWL13]
Motion parallax	Texture gradient [HR12]
Kinetic depth	

Figure 3.1: Common visual cues

3.1.4 Definition of a 3D display

The definition retained in [PS12] is the following : a 3D display is a display that uses at least one of the physiological cues. We will see that most of the displays use the stereoscopy cue, because it is the one that provides the most convincing 3D experience[KKN⁺06] thanks to **stereopsis**.

3.2 Classification of the 3D displays

One of the main problem while trying to find a proper **display** for a given application is to choose a relevant classification for the displays, that allows a choice with criterions relevant to the application.

3.2.1 Criterions

There was a lack of proper nomenclature in the literature for a long time [PS12]. However, some attempts have been made to find relevant criterions that would be general enough to cover the current display techniques, but also the ones that are not yet thought of.

Different classifications

The first classification was in [Oko76], and it was really based upon the different kinds of displays existing at the time:

- Lens-sheet three dimensional pictures.
- Projection-type three dimensional displays.
- Holography.

However, it did not hold well against the emergence of new techniques, like volumetric displays for instance.

Other classifications would limit themselves to only a subset of 3D displays.

Hence the need for a classification that would not base itself on the different technologies, but on criterions that would be inherent to the idea of display and human vision.

Chosen classification

In [PS12], the main idea is to classify the displays according to two axes :

- The display depth (flat or deep).

- The number of points of view from which the image can be seen (duoscopic, multiscopic, or omniscopic).

3.3 In-depth presentation of some 3D display methods

This section describes the different technologies used by manufacturers of 3D displays, and also explains some 3D visualization systems in detail. But first we will define what are stereoscopic and auto-stereoscopic displays.

Stereoscopy is the set of techniques used to reproduce a depth perception from two planar images.

Autostereoscopy is a method of image representation, either in three-dimension or stereoscopic, which requires no additional device to render the 3D effect.

3.3.1 Two-view 3D displays

This is the category of most commercial 3D display technologies, like the common 3D TV screens or 3D movie theater technologies. The main feature of these screens, which distinguishes them from conventional screens, is their ability to display stereoscopic images : each eye of the observer will receive a different point of view. There are multiple technologies that are able to power such screens; we will explain some. But an exhaustive list is present in [MPWL13].

About the time

There are two major technologies that are in use to create a 3D effect from two image streams.

Time-sequential displays

The idea is that the images are shown one after the other, but it has two requirements :

- The switching between images must be fast : a refresh rate of 48 Hz is theoretically minimal for a 24 fps movie, however research [HDFP11] shown that 58 Hz is a practical minimum.
- Each eye must see only the images directed to it. There are multiple technologies to achieve this, passive and active.

Time-parallel displays

This is the opposite of time-sequential two-view displays. Both eyes receive a stream of images at the same time.

3.3.2 Presentation of the methods

Wavelength selective displays

The most famous are the anaglyph glasses, pictured on figure 3.2. However, the result is of poor quality and unpleasant after some time.

In this case, images are filtered by colour. The two images are superimposed and the glasses incorporate a colour filter. Each eye can only see the image intended for it, so this can be used to create a stereoscopic effect.

This is generally used as a time-parallel method.



Figure 3.2: Red/cyan colour filter 3D glasses

Polarized method

This technique uses polarized glasses. The right lens is polarized in one direction while the left lens is polarized in the other direction, as in figure 3.3.



Figure 3.3: Polarized 3D system

This is not an expensive method, however the image is often less bright. It is generally used in movie theaters.

Active method

This technique uses active glasses, with LCD shutters that needs to be perfectly synchronized with the screen refresh rate : one of the glass goes black while the other allow the light to pass, which allows the eyes to receive only the wanted image at the wanted time.

This method is more expensive, but reviews often said it offers a better experience.

It is a time-sequential method.

Multi-screen display

This design uses multiple LCD displays, one seen in transmission and the other in reflection in a mirror, in order to change the polarization of the reflected screen.

This way, with polarized glasses, it is again possible to see with stereoscopy.

Auto-stereoscopy

This technique is interesting, because no special headgear is required : the correct light information is sent directly by the screen to the eyes.

However, the viewer has to be at a precise position in front of the screen for it to work. Hence, it only works with a single person except in the case of a very large and expensive screen.

Some recent designs can for instance use reflective screens [SSR13].

3.3.3 Multi-view displays

Multiview or multiscopic displays are displays that can be seen from a finite, discrete, number of point of view and show a different picture corresponding to this point of view.

The focus on this part will be on horizontal parallax displays : the different point of view are on the horizontal axis.

There are two main types of technology for multiview parallax displays: either by applying a parallax barrier, or by application of a lenticular panel.

A parallax barrier is a mask of parallel black stripes that reveals the different parts of the underlying image depending on the direction of observation.

The same effect can be obtained with the lenticular sheets, which are linear networks of narrow cylindrical lenses. Both techniques are used to show a unique view of each position in the viewing area. Thus, a viewer feels the motion parallax and binocular stereo without the use of special glasses.

Lenticular Displays

The application of a lenticular panel to deflect the rays from the pixels of the screen in order to reproduce a stereoscopic display giving the images an impression of depth.

A lenticular panel is composed of a regular grid of small spherical lenses. Each lens covers an area of several pixels of the flat screen and will divert the direction of the light emitted by the pixels covered.

As we can see on figure 3.4, each eye only sees the right or left pixels. An observer looking at a lenticular 3D screen does not see the same screen pixel between his left eye and his right eye, although both eyes see the same lenses.

With horizontal lenses, projecting changing images, the perceived image depends on the angle between the eye and the lenses.

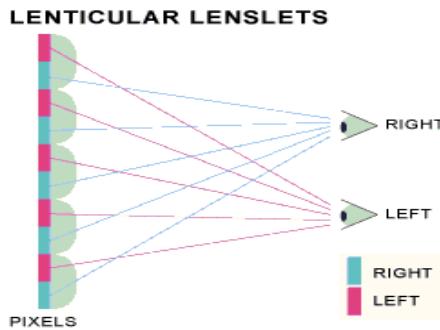


Figure 3.4: A lens array

Parallax Barrier Displays

Instead of lenses, multiple small masks are put in front of the pixels. The images are divided into columns of a width of one pixel. Each mask is precisely placed so that the observer, himself set at the correct distance from the screen, views the image that corresponds to each eye.

Thanks to the cache, an image is sent to the right eye and another in the left eye, as in figure 3.5; the brain recreates the 3D picture by stereopsis.

In comparison with the lens array, we can see each spherical lens replaced by an opaque surface and a small hole in its center. Therefore, as in the case of the lens array, an observer looking at a small hole in the 3D screen does not see the same screen pixel between his left eye and his right eye: this is because the parallax barrier is not contiguous to the flat screen, but there is a distance of a few millimeters between the two panels.

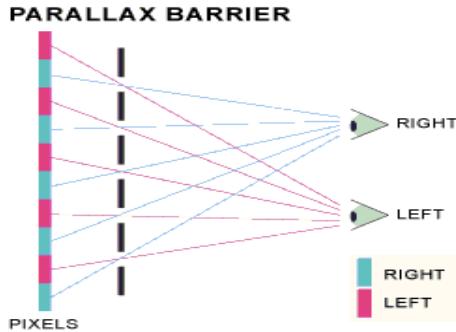


Figure 3.5: A parallax barrier

This technique provides a 3D vision without wearing glasses.
However, there are some disadvantages :

- It must be placed precisely over the screen, if the observer is not in a precise position, he will see superimposed images which will make the image look garbled.
- Hence, movement is not compatible with this system, but it is very difficult for human beings to stay still for the length of a movie, for instance.
- It does not allow visualization of stereoscopic image for multiple viewers at the same time.
- Brightness is generally lower due to most of the light being occluded by the mask. However, a recent amelioration on this technique greatly improves brightness [LZWW14].

3.3.4 Omni-view Displays

These displays can be seen from more than a point of view and would show a correct image from any accessible point of view.

Multi-projector display

This method consists in positioning several projectors in a circle. They all display the image under a different angle. The images are projected onto a special screen. For instance, a double lenticular lens example (a spherical lens) works well, but the easiest way is to project these images on a cylindrical fog, as in figure 3.6.

Lasers can also be used : they would provide more contrast, but shapes cannot be very detailed.

In this example, it is possible to turn around the rabbit in 3 dimensions and view it from every angle.

This technique has some advantages :

- There is no limit of size for the projection, which makes it good for live performance.



Figure 3.6: 3D multiviewpoint display

However, it has important requirements :

- Multiple projectors are needed.
- Headlights must be precisely aligned.

Pepper's Ghost

Pepper's ghost is an optical illusion technique used during plays, in circuses and with some magic tricks.

This artifice was invented in 1862 by Henry Dircks. He took advantage of an optical effect which gives the impression of a ghost appearing and vanishing in the scene.

In the same year, John Henry Pepper, decided to start working on illusions after going to an Henry Dircks' show. He realized that with some technique improvements, this illusion would be more cost-effective for theater renters.

The name of this method comes from this man because he made this effect popular.

The illusion is produced by glass sheets and lighting effects. The actor who plays the ghost is placed in a dark room, invisible from the audience (an hidden scene), in such a way that the public cannot directly see him[Pep]. The figure 3.7 explains how it works.

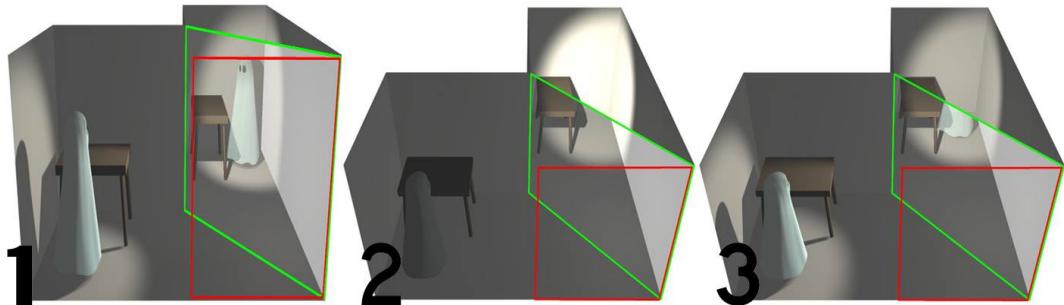


Figure 3.7: Pepper's ghost

1. A viewer looking through the red rectangle sees a ghost floating next to the table. The illusion is created by a large piece of glass situated at a 45° angle between the viewer and the scene (green outline). The glass reflects a room hidden from the viewer (on the left), sometimes called a "Blue Room" that is built as a mirror-image of the scene.
2. If the mirror-image room (left) is darkened, it does not reflect well in the glass. The empty room (top) is brightly lit, making it very visible to the viewer.

3. When the lights in the mirror-image room are raised (with the empty room being dimmed slightly to compensate), the ghost appears out of nowhere.

Nowadays, the hidden scene has been replaced by a video projector. This enables the director of the spectacle to display anything on the screen. Pepper's ghost used to be a theater-specific trick, but today digital devices revive this technique.

Chapter 4

Presentation of 3D musical instruments

4.1 History of the 3D musical instruments

4.1.1 3D Navigation instruments

Many immersive musical instruments tend to focus on navigation in a 3D virtual environment. First of all, there is the Phase project [RLC⁺05] which explores generation, handling and control of sound (or music), through a haptic sensor and a visual representation that guides the user. Secondly, there is Plumage [JAC⁺07] which is an interface for interactive control of spatialized audio composition. Feathers are scattered in a 3D scene which represents sound grains. They generate sound when they are crossed by reading heads. Those heads are controlled by the user. Nevertheless these two projects do not allow manipulation of direct sound synthesis structure, only of existing sounds.

4.1.2 Single sound synthesis instruments

Another range of 3D instruments focuses on a single kind of sound synthesis. For instance, the *Virtual Xylophone*, the *Virtual Membrane* or the *Virtual Air Guitar* [MPLKT05] are all recreations of existing instruments in a virtual world. Another example is Mike Wozniewski's instrument. In his program, the user can navigate in a 3D scene where some precise points generate sounds[WSC06]. The user hears the sounds according to his position and orientation in the scene.

The aerial percussion is a 3D instrument that could be in this instrument group.

4.1.3 The Drile

The Drile offers a new usage of 3D. It uses 3D interaction to manipulate more easily the internal structure of a music.

Both the Drile and the aerial percussion were built for musical performance.

4.2 The Drile

The Drile is a musical instrument which innovates in multiple domains. From a musical point of view, it uses an original structural approach of the music : hierarchical live-looping. In

the virtual reality domain, the Drile uses 3D and haptic feedback to improve the musician's immersion.

4.2.1 Hierarchical Live-Looping

The Drile musical process is based on the hierarchical live-looping technique.

Live-Looping

Live-looping is a composition technique which came to light thanks to famous modern composers such as Steve Reich. The concept is relatively simple : music is composed using only samples that are played in loops. Stacking those loops and adding multiple audio effects on them can easily produce complex musics. It has been used a lot for musical performance especially by singers, guitarists because they can build their own backing tracks easily. Many electronic musicians are using this technique to produce entire songs.

Nevertheless, one of its limits is that one cannot manipulate complex structures but only sequences of samples. So the structure remains quite linear.

Hierarchical Live-Looping

Inspired from Marczak [MEP07] works, hierarchical live-looping organizes live-looping in a tree structure which leads to a more complex musical structure.

Leafs are composed of raw samples (i.e audio extract or synthesized sound) and a musical effects list that is to be applied on the associated sample. Then, each node is made of a node children list, a musical effects list, and a musical content. This last content is the result of the children musical contents.

In that way, adding an effect to a leaf will only affects its own musical content. Whereas adding an effect to a node will apply the effects to the whole musical content produced from all children. It becomes possible with operations on tree (duplication, merging...) to manipulate a more sophisticated musical structure.

All these tree manipulations get simplified with the 3D representation of this structure.

4.2.2 Piivert, a new controller

To interact with the virtual environment, the author developed an input device called **Piivert** [BHDC10] shown in 4.1. This device is separated in two controllers that are properly tied to each hand. This improves the musical control of the user. It combines a 6-DOF¹ tracking and high-sensitivity pressure sensing. The 6-DOF¹ tracking allows ray casting and gestures detection.

The author approach regarding the choice of musical gestures was inspired by Cadoz's [Cad99] research. Indeed, Cadoz suggested three categories of gestures : selection, modification and excitation gestures. As graphical interaction does not fit to the last category of gestures, the Piivert has an important part in the instrument, thanks to its pressure sensing and vibrating feedback. For instance, excitation can be used to trigger sequences.

¹DOF!

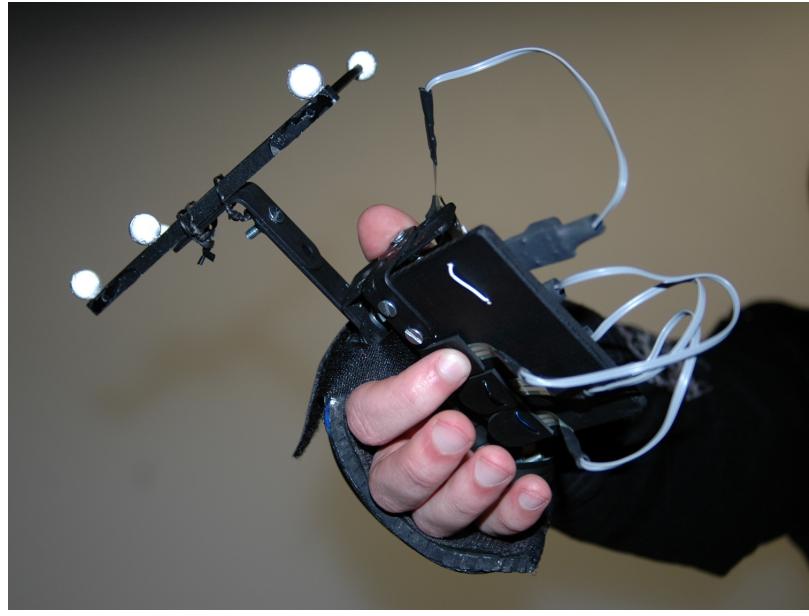


Figure 4.1: Piivert with pressure buttons on the bottom

4.2.3 Drile world

Objects

In order to use the hierarchical live looping technique, the Drile implements several objects. This world tries to take the best part of the graphical interface.

Worms

The first object is what we call a *worm*. It is the representation of the nodes of the live-looping trees. They look like their associated analysed audio spectrum in order to ease the identification of each sound during a manipulation.

These worms have graphical parameters such as color hue, size, transparency and so on. They are mapped to different audio effects parameters in such a way that modifying a worm's appearance would modify its audio effects. The current mapping are based on the result of a user study : size/volume, color hue/pitch, transparency/distortion... Moreover, the mapping is relative to each audio sample, so that each worms starts with the same appearance. To improve interaction, a graphical parameter also gives the current value of the effect. The rotation of a worm on the y-axis provides information about the reading position of the associated audio.

Tunnels

Worms audio parameters and their associated graphical parameters can be modified by grabbing and sliding them through *tunnels* as shown below 4.2. Each tunnel is used for one specific parameter modification.

Hierarchy manipulation

As the size of worms is used to control node volume, manipulating worms through depth-axis would create perception problems. To settle that an *interaction plane* is used, which is at a fixed distance and allows manipulation of Drile objects. Depth-manipulations are maintained for less accurate uses.

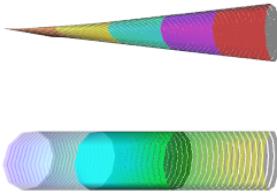


Figure 4.2: Tunnels

Live-looping trees are represented by connected worms (4.3) organized by depth. The closest level to the users is on the interaction plane. Knowing that children nodes are the furthest from the user, accessing the different levels of the tree is possible by pulling and pushing the tree using Piivert gestures.

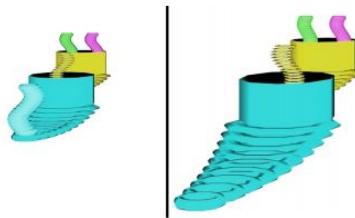


Figure 4.3: Tree nodes containing colored worms. Appendages represent tree edges

Also, the user can take advantage of the 3D world by grabbing one worm directly with the ray-casting and then manipulating it. This can be facilitated with the head-tracking technique which simulates the fact of moving aside in front of a real scene, making access to objects in the back of the scene easier. Common tree operations (merging, duplication...) are mapped to Piivert gestures that look natural to improve immersion. For example, a merge can be made by colliding two worms.

Scenes

During a performance, the user might want to use pre-built musical patterns. They are stored into *scenes* (4.4). He might also want to translate currently playing nodes from one scene to another in order to perform a nice musical transition. The 3D environment creates a good visual feedback of the scene's content: this can help the user to choose the right scene.

Collaboration and Learning

The Drile environment is a multi-user interface, but each user needs a controller. A large display is also an advantage. Because the Drile requires some expertise to have a proper handle on the music creation approach, two levels are implemented, one for beginners and one for experts. Beginners have access to higher level functions to learn the fundamental manipulations.

Evaluation of the instrument

The author used a dimension space representation (figure 4.5) to evaluate his instrument[BFW05].

The Drile environment fulfils two important aspects of this representation :

1. The Feedback Modalities axis, especially with the 3D environment and the Piivert controller,

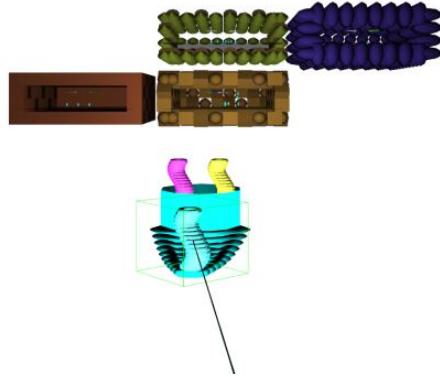


Figure 4.4: Scene selection for a worm translation

2. The Inter-Actors axis thanks to its multiple user interface and its large display.

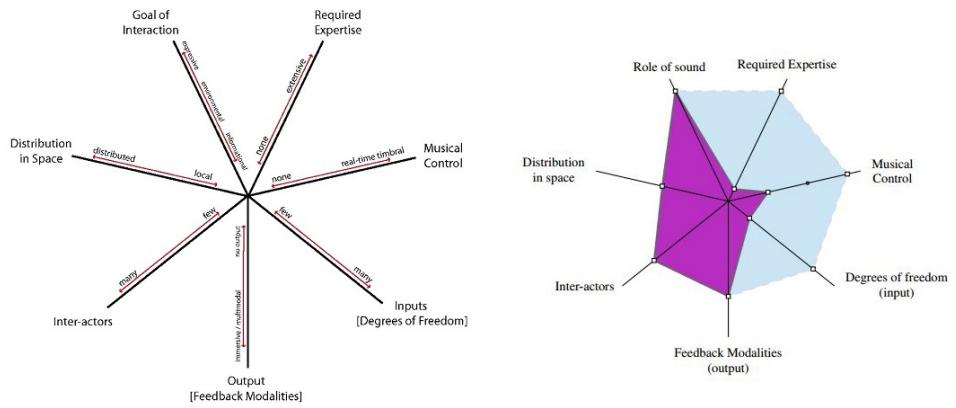


Figure 4.5: On the left the dimension space representation, on the right the Drile representation (in purple for beginner interface, in blue for complete control)

4.3 Aerial Percussion

The aerial percussion is an innovative musical instrument which differs from usual percussions because it does not use actual physical material to generate sounds. The area where the performer throws his sticks is virtual.

The position of the sticks in the space is measured permanently thanks to sensors. Input data is processed by a software which detects virtual impacts in real time and associates them to specific sounds.

The interest of the computer is the access of an unlimited sound range. This percussion also allows the musician to set aside the physical constraints of the instrument, since he can enter inside the drums with his sticks. Therefore, he can choreograph his movements and the visual performance becomes more important.

4.3.1 Hardware: Polhemus Liberty

The Polhemus Liberty is a device that returns the position and orientation and position of two sensors in real time.

Core

A transmitter generates a magnetic field with three fixed antennas placed orthogonally from one to each other in a ten centimeters side cube. This cube, the transmitting base, is linked with the central block with a cable and is situated at the center of the zone where the acquisition occurs. The diffused magnetic field is captured by a second antenna inside each sensor.

Sensors

They are small cubes of less than one cubic centimeter, located at the end of the cables connected to the core. The input signal coming from the sensors is then sampled, treated by a **DSP¹** in the central bloc and finally sent to the computer via USB.

All the components of this percussion can be found on figure 4.6. From left to right there are the sticks, the central bloc and the transmitting base.



Figure 4.6: Polhemus Liberty and sticks.

However the device is not sufficient on its own, there is also a software part.

¹Digital Signal Processor

4.3.2 Software: SetKreator and Flock of Birds

Flock of Birds

FoB¹ is a small software which receives Polhemus data through the local network and generates an **OSC**² network stream from it. **OSC** is then easier to manipulate using **MAX/MSP**, **PUREDATA**, or **OPENFRAMEWORKS**.

SetKreator

SetKreator has been developed at the SCRIME by Joseph Larralde and Sébastien Lebreton. It is a virtual instrument editor for the aerial percussion. Basic volumes (parallelepipeds, cylinders...) inside a sphere which indicates the volume where the Polhemus sensors can be used, like in figure 4.7. It is also able to associate each shape to a particular sound synthesis system. It receives position data from the Polhemus via an **OSC** stream.

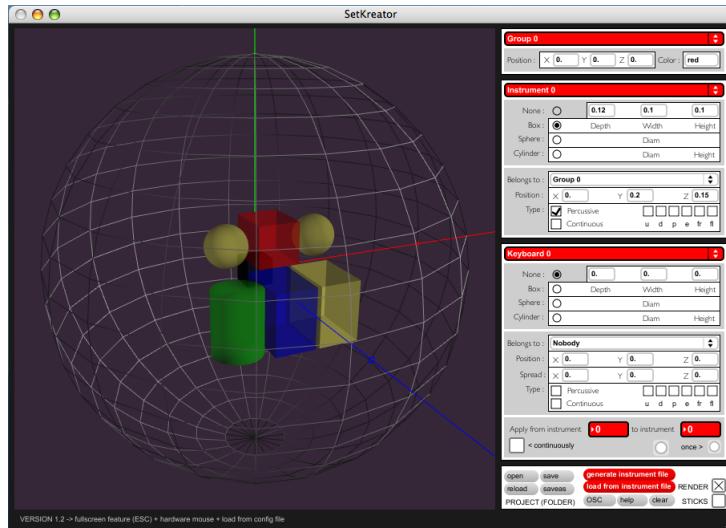


Figure 4.7: SetKreator

¹Flock of Birds

²Open Sound Control

Chapter 5

Implementation

5.1 Objectives

Apart from the research work, it is asked from us to apply our research to two digital musical instruments (the DRILE and the Aerial Percussion).

The goal of our work is to enact a live show with these two instruments, that allows for both the performer and the spectators to see the musical instrument in three dimensions.

5.1.1 Finding a display

The first task is to find a suitable display method that would allow :

- The performer to interact with the instrument
- The spectators to see the performer as if he was part of the 3D scene
- If possible, a stereoscopic feel.

5.1.2 Implementing suitable renderings

There is already some existing work for the rendering engine of the Drile, however there is nothing for the aerial percussion.

We have to make renders from two different viewpoints : one for the performer, another for the spectators.

5.1.3 Customization

If we have time left, we are to add some customizations to the aerial percussion rendering, in order to make it look like a real show, with special effects, flares, particles, textures...

5.2 Implementation

We will describe here the multiple choices that have been made during this project, and the reason behind these choices, as well as the result of our implementation.

5.2.1 Chosen display techniques

There are multiple factors to take into account :

1. The availability of the technology.
2. The potential price of the required materials.
3. The time to set-up the display.
4. The scaling for a medium-sized audience.
5. The compatibility with the double requirement : a view for the performer, and another for the spectators.

We are now going to study these requirements point by point.

Availability

This is the main problem : many of the display devices presented in 3 have only been the subject of research and not of a real implementation sold by a company (e.g. holograms). Also, the development state of some technologies might not be sufficient for what we are striving for (e.g. auto stereoscopic displays which are only present in very small screens like smartphones).

Price

Some technologies might be irrelevant only because of the amount of money needed to get a working implementation. For instance, an active 84" 3D HDTV generally costs more than ten thousand dollars, which is unsuitable to this project.

Set-up time

Some methods might require a very long time to set-up. While we don't have a required maximum time to set-up the show, we should try to keep it as low as possible. For instance, the Pepper's Ghost technique is quite long to set-up, because there is a lot of massive hardware, projectors, screens, to set-up.

Scaling

Since this is for a show, we need a system that will allow everybody in the room to enjoy the performance. The estimate is at about 40 persons : we need a display that provides big enough viewing angles and is big enough for everybody to be able to enjoy it. A square display with a side of two meters would be ideal to enable complete immersion.

Double-view requirement

This is one of the hardest requirements, because it can easily double the quantity of required hardware. For instance, if we were to use 3D TVs, we would need one TV for the viewers and one for the performer.

The choice

We chose to go with the Pepper's Ghost technique, with polarized stereo glasses. This was mainly because it was the only method efficiently available with all the tools at our disposal : no additional fees are required.

Due to computer limits, we could not test the two views, but it is implemented in our software. We mainly did focus on the Aerial Percussion, because we did not know how to use the Drile code.

5.2.2 Aerial Percussion implementation

We will described here the implementation of the Aerial Percussion 3D renderer.

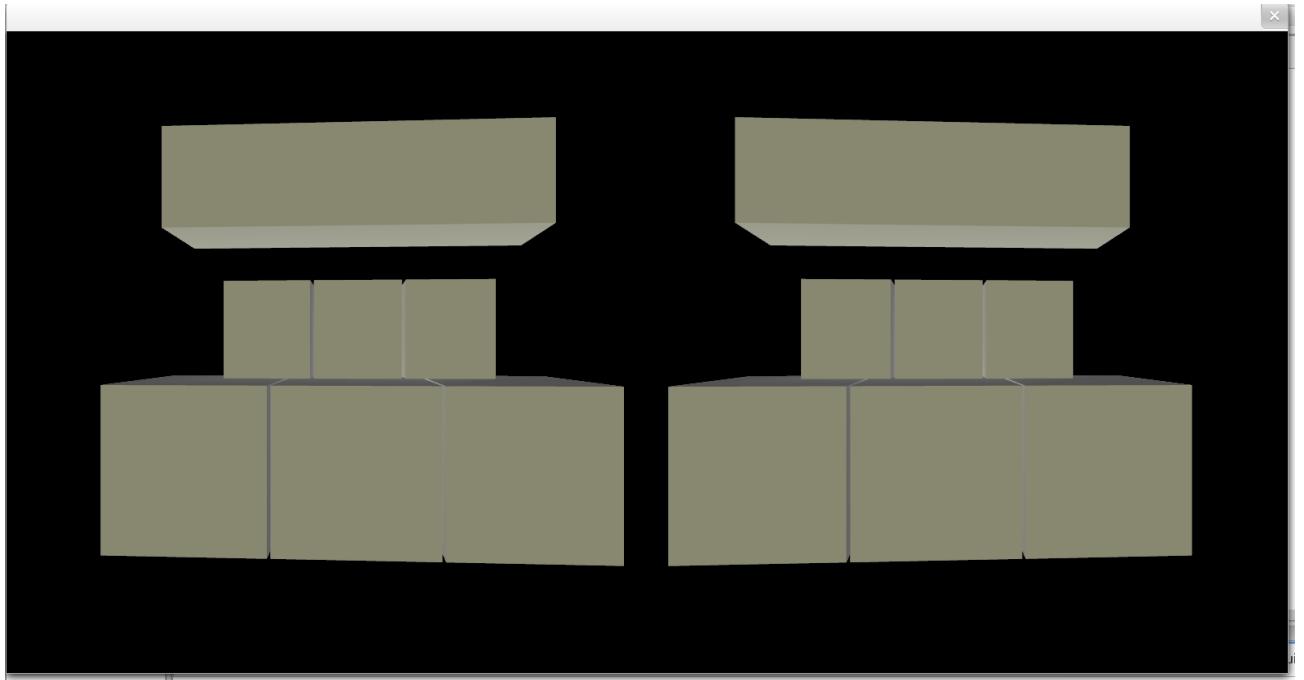


Figure 5.1: A screenshot of the Aerial Percussion renderer

Technologies used

We worked with **OPENFRAMEWORKS**, a **C++** abstraction over many different media libraries, like **OPENGL**, **OPENNI** and many others. It makes the implementation of creative software easy while retaining the power and speed efficiency of the **C++**.

Our software listens to **OSC** messages, which means that it would be easy to put it on two computers and have one computer display the images for the spectators, and the other for the musician.

Displaying the data

Pepper's Ghost technique

In order to have the Pepper's Ghost technique work efficiently, we have to display our shapes on a black background, so that only the bright parts reflects on the glass screen.

Stereoscopic display

The method uses two projectors, one on the top of another. Each projector has a polarizing filter and is connected to a different input of the computer.

The computer, running **DEBIAN GNU/LINUX**, is set up so that there is a single viewport of twice a projector's horizontal resolution. Each projector receives one half of the viewport.

For instance, the actual resolution of a projector is 1920×1080 . Hence, the viewport is 3840×1080 .

The software then runs at the same resolution, and the image is displayed from two point of views, that slightly differs on the horizontal axis, as shown on figure 5.1. We used frame-buffer objects to achieve this. This allows to generate a 3D effect when the images are projected on the same point in space by the dual projector system, because each eye will receive the correct image, as we could see on paragraph 3.3.2.

Additional features

We added the feature to change the distance between the eyes in order to find the most pleasant one, using the left and right keyboard keys. We can also zoom in and out using the up and down keys.

The space key will change the point of view from front to back.

The software also reacts by changing the colour of a box when the musician puts a drum-stick inside a virtual drum element.

5.2.3 Physical set-up of the system

Figure 5.2 presents the set-up for the show.

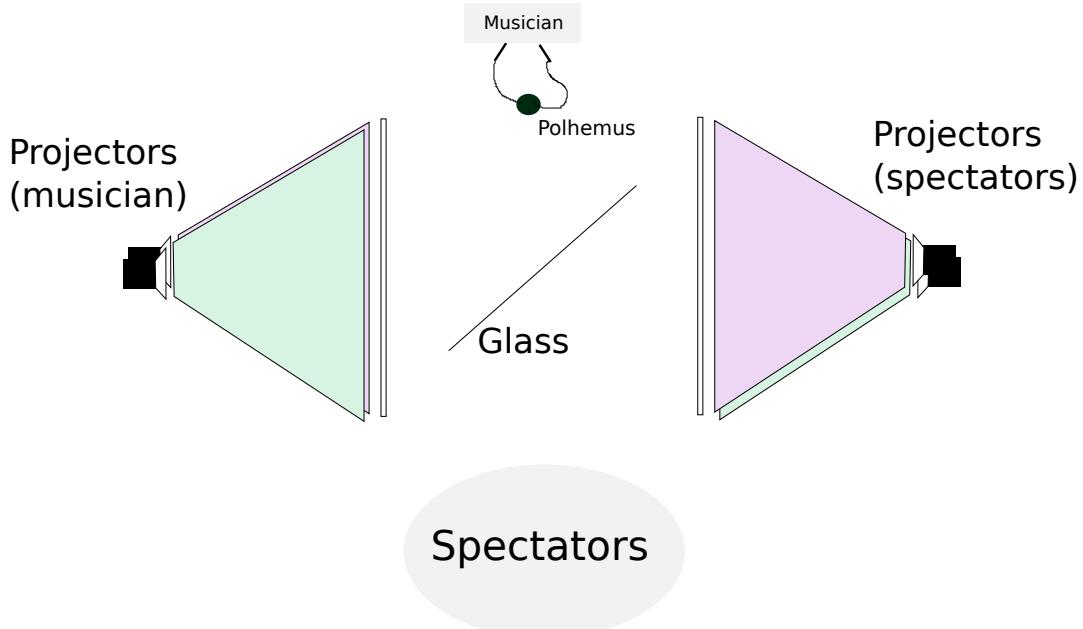


Figure 5.2: Our installation

Figure 5.3 shows the result.

The picture is taken from the spectator's point of view. Due to the small size of the glass screen, there are artefacts on the borders.

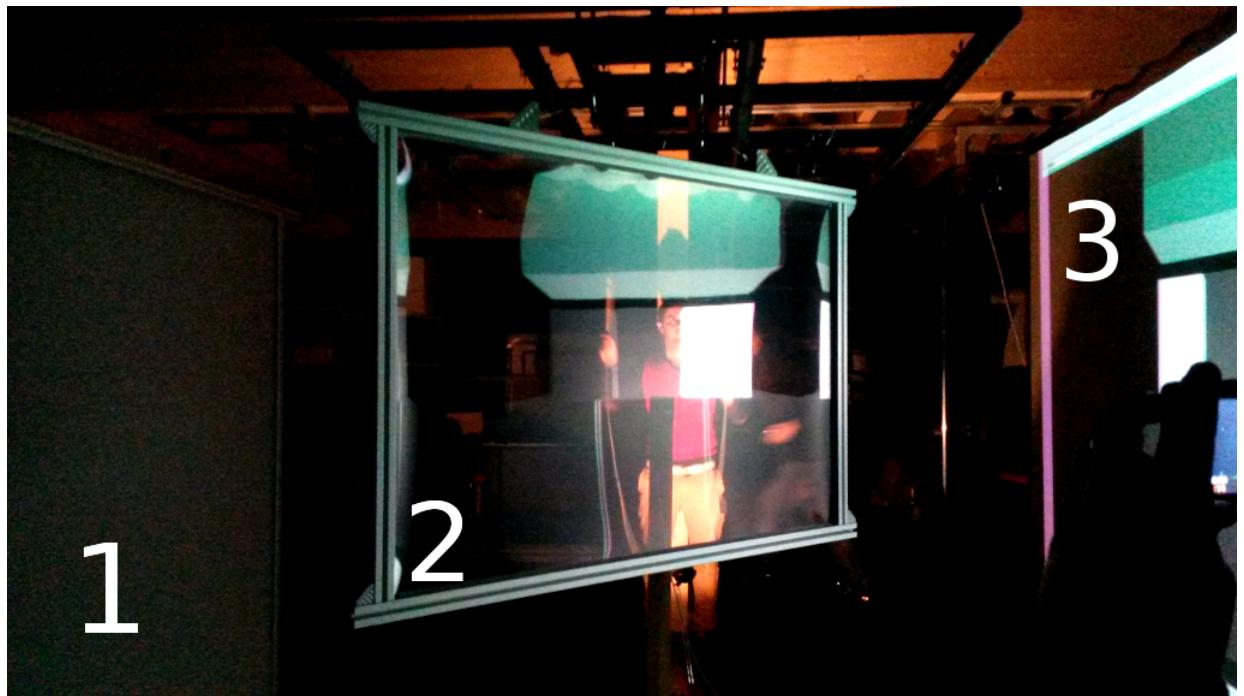


Figure 5.3: Our installation. 1 and 3 are the projection screens, 2 is the Plexiglas screen.

Since the computer only had two video outputs, the system is only able to drive one half of the projectors : here, the spectators half.

Chapter 6

Conclusion

This project was the occasion to discover a part of computer science that is cross-domain with both optics and electronics in one part, and with art and music on the other part.

We were not able to complete every required part of the implementation, mainly due to the complexity of the Drile code, however we still managed to have a working software for the aerial percussion.

There is quite a lot of room for improvement, however. First, on the artistic side, the software can be made to look more appealing to the eyes, for instance by using textures and particles. Also, for the Pepper's Ghost technique, the Plexiglas screen could be replaced with a transparent tarpaulin : it seems it gives better results and it would also allow for a bigger projection size. Also, with more time and budget, it could be possible to try to implement the render for other types of displays than two-view, for instance multi-view or omni-view. Lasers were proposed but due to the short amount of time, we couldn't try working with them.

Finally, it would be interesting to try with a two-computer setup to have spectator and performer view : this would be the best way to use the Drile.

Glossary

accomodation the change of focus of the eyes in order to perceive clearly what is looked at. [9](#)

atmospheric perspective the impression of depth given by the refraction of the air. For instance, we can say that mountains are far because they appear more blue than close mountains. [9](#)

convergence when the eyes rotate to aim on the point in space a person focuses on. [9](#)

display a visual output device for the presentation of images [PS12]. [10](#)

kinetic depth the visual cues we have of an object in motion. [9](#)

linear perspective the way our visual perception of objects are affected by their position and dimension. [9](#)

live looping A REMPLIR. [5](#)

motion parallax when two objects, one further from another, seem to translate at a different speed if the observer is moving. [9](#)

occlusion when an object is hidden by another. [9](#)

shading the gradient in color and shades that would appear due to the shape of objects and color, intensity, and direction of light. [9](#)

stereoscopy the result of the human eyes receiving two different images. [9](#)

Bibliography

- [3Dm] <http://www.ubergizmo.com/2011/03/fog-projectors-3d-multi-viewpoint-display>.
- [BD95] Frank Biocca and Ben Delaney. Immersive virtual reality technology. *Communication in the age of virtual reality*, pages 57–124, 1995.
- [BDCH⁺10] Florent Berthaut, Myriam Desainte-Catherine, Martin Hachet, et al. Drile: an immersive environment for hierarchical live-looping. *Proceedings of New Interface for Musical Expression 2010*, 2010.
- [BFMW05] David Birnbaum, Rebecca Fiebrink, Joseph Malloch, and Marcelo M Wanderley. Towards a dimension space for musical devices. In *Proceedings of the 2005 conference on New interfaces for musical expression*, pages 192–195. National University of Singapore, 2005.
- [BHDC10] Florent Berthaut, Martin Hachet, and Myriam Desainte-Catherine. Piivert: Percussion-based interaction for immersive virtual environments. In *3D User Interfaces (3DUI), 2010 IEEE Symposium on*, pages 15–18. IEEE, 2010.
- [Cad99] Claude Cadoz. Musique, geste, technologie. *Les nouveaux gestes de la musique*, pages 47–92, 1999.
- [gla] <http://www.agency1903.com/blog/2010/8/18/z-axis-power>.
- [HDFP11] Nicolas S Holliman, Neil A Dodgson, Gregg E Favalora, and Lachlan Pockett. Three-dimensional displays: a review and applications analysis. *Broadcasting, IEEE Transactions on*, 57(2):362–371, 2011.
- [HR12] Ian P Howard and Brian J Rogers. *Perceiving in Depth, Volume 2: Stereoscopic Vision*. Number 29. Oxford University Press, 2012.
- [JAC⁺07] Christian Jacquemin, Rami Ajaj, Roland Cahen, Yoan Olivier, and Diemo Schwarz. Plumage: Design d'une interface 3d pour le parcours d'échantillons sonores granularisés. In *Proceedings of the 19th International Conference of the Association Francophone D'Interaction Homme-Machine*, pages 71–74. ACM, 2007.
- [KKN⁺06] LLOYD KAUFMAN12, James H Kaufman, Richard Noble, Stefan Edlund, Sunhee Bai, and Teresa King. Perceptual distance and the constancy of size and stereoscopic depth. *Spatial vision*, 19(5):439–457, 2006.
- [LZWW14] Guo-Jiao Lv, Wu-Xiang Zhao, Jun Wang, and Qiong-Hua Wang. Shared pixel based parallax barrier 3d display with high brightness. *Optik-International Journal for Light and Electron Optics*, 2014.

- [MEP07] Raphael Marczak, Tuteur ENSEIRB, and Bernard PERROT. *Etude d'une représentation hiérarchique liant micro et macro-structures musicales*. PhD thesis, Master's thesis, University of Bordeaux, 2007.
- [MPLKT05] Teemu Mäki-Patola, Juha Laitinen, Aki Kanerva, and Tapio Takala. Experiments with virtual reality instruments. In *Proceedings of the 2005 Conference on New Interfaces for Musical Expression*, pages 11–16. National University of Singapore, 2005.
- [MPWL13] Mostafa Mehrabi, Edward M Peek, Burkhard C Wuensche, and Christof Lutteroth. Making 3d work: A classification of visual depth cues, 3d display technologies and their applications. 2013.
- [Oko76] Takanori Okoshi. *Three-dimensional imaging techniques*. Academic Press, 1976.
- [Pep] http://en.wikipedia.org/wiki/Pepper's_ghost.
- [Pol] http://en.wikipedia.org/wiki/Active_shutter_3D_system.
- [PS12] Waldir Pimenta and Luis Paulo Santos. A comprehensive taxonomy for three-dimensional displays. *WSCG 2012 Communication Proceedings*, pages 139–146, 2012.
- [RCVD97] George Robertson, Mary Czerwinski, and Maarten Van Dantzich. Immersion in desktop virtual reality. In *Proceedings of the 10th annual ACM symposium on User interface software and technology*, pages 11–19. ACM, 1997.
- [RLC⁺05] Xavier Rodet, Jean-Philippe Lambert, Roland Cahen, Thomas Gaudy, Fabrice Guedy, Florian Gosselin, and Pascal Mobuchon. Study of haptic and visual interaction for sound and music control in the phase project. In *Proceedings of the 2005 conference on New interfaces for musical expression*, pages 109–114. National University of Singapore, 2005.
- [SSR13] Quinn Smithwick, Lanny S Smoot, and Daniel Reetz. Autostereoscopic display system with one dimensional (1d) retroreflective screen, April 4 2013. US Patent 20,130,083,291.
- [Vie] <http://www.newegg.com/Product/Product.aspx?Item=N82E16800991111>.
- [WSC06] Mike Wozniewski, Zack Settel, and Jeremy R Cooperstock. A spatial interface for audio and music production. In *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, volume 2006, 2006.