



Université 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

<b>1</b>	<b>Introduction</b>	<b>4</b>
<b>2</b>	<b>Subject and definitions</b>	<b>5</b>
2.1	Subject presentation . . . . .	5
2.2	What is a 3D musical instrument? . . . . .	6
2.3	Immersion . . . . .	7
2.4	Control . . . . .	7
<b>3</b>	<b>Three-dimensional displays</b>	<b>9</b>
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	Criteria . . . . .	10
3.3	In-depth presentation of some 3D display methods . . . . .	11
3.3.1	Two-view 3D displays . . . . .	11
3.3.2	Horizontal parallax multiview 3D displays . . . . .	12
<b>4</b>	<b>Presentation of 3D musical instruments</b>	<b>16</b>
4.1	History of the 3D musical instruments . . . . .	16
4.2	The DRILE . . . . .	17
4.2.1	Hierarchical Live-Looping . . . . .	17
4.2.2	Piivert, a new controller . . . . .	18
4.2.3	Drile world . . . . .	18
4.3	Aerial Percussion . . . . .	21
<b>5</b>	<b>Implementation</b>	<b>22</b>
5.1	Objectives . . . . .	22
5.1.1	Finding a display . . . . .	22
5.1.2	Implementing suitable renderings . . . . .	22
5.1.3	Customization . . . . .	22
5.2	Implementation . . . . .	22
5.2.1	Chosen display techniques . . . . .	23
5.2.2	Aerial Percussion implementation . . . . .	24
<b>6</b>	<b>Conclusion</b>	<b>26</b>



# List of Figures

2.1	Picture of a musician using DRILE . . . . .	6
3.1	Common visual cues . . . . .	10
3.2	Red/cyan color filter 3D glasses[Vie] . . . . .	11
3.3	Polarized 3D system[Pol] . . . . .	12
3.4	A lens array[gla] . . . . .	13
3.5	A parallax barrier[gla] . . . . .	14
3.6	3D multiviewpoint display[3Dm] . . . . .	14
3.7	Pepper's ghost . . . . .	15
4.1	Piivert with pressure buttons on the bottom . . . . .	19
4.2	Tunnels . . . . .	19
4.3	Tree nodes containing colored worms. Appendages represent tree edges . . . . .	20
4.4	Scene selection for a worm translation . . . . .	20
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
5.1	A screenshot of the Aerial Percussion renderer . . . . .	25

# 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 setup our own 3D musical instrument.

## Chapter 2

# Subject and definitions

### 2.1 Subject presentation

#### 3D Musical instruments

At the **SCRIME**<sup>1</sup> and **LABRI**<sup>2</sup>, three-dimensional musical instruments have been implemented within the context of research in interactive virtual reality and music computing.

The **DRILE** [BDCH<sup>+</sup>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 positionned 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.

---

<sup>1</sup>Studio de Création et de Recherche en Informatique et Musique Électroacoustique

<sup>2</sup>Laboratoire Bordelais de Recherche en Informatique

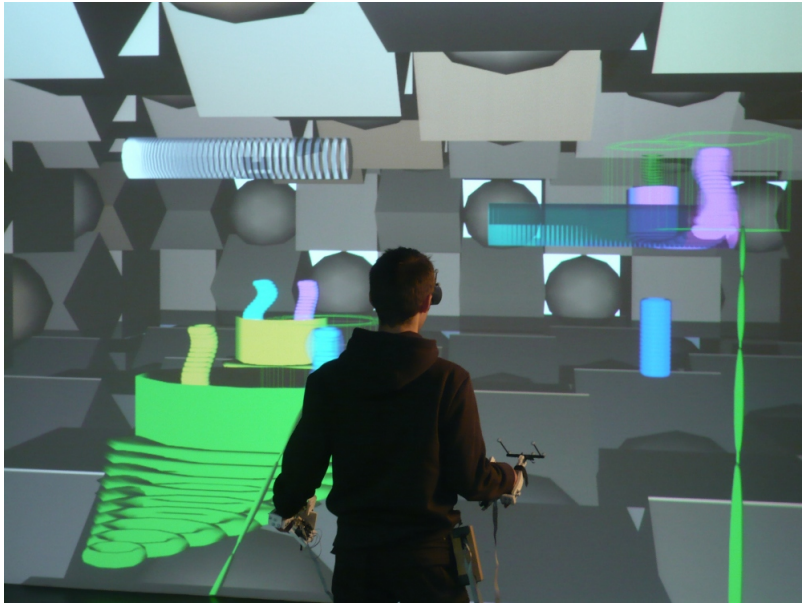


Figure 2.1: Picture of a musician using DRILE

## 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 psychologic 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 vibrate. 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 violonist 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 **HMD**<sup>1</sup>, a **CAVE**<sup>2</sup> or other virtual reality devices and methods like Fishtank 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 problematics while conceiving **PIIVERT** [BHDC10], the control interface to the **DRILE**.

---

<sup>1</sup>Head-mounted Display

<sup>2</sup>CAVE 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 litterature, 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<sup>+</sup>06].

## 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

Since a decade, a new generation of screens appeared: 3D screens. 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].

##### Wavelength selective displays

To view in 3D, it is necessary to provide our brain two images of the same scene taken from two different points of view. This is the principle of stereoscopy. The distance between the two points of view is used to compute the depth of field to achieve the 3D effect.

Different methods exist so that each eye receives the intended image.

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



Figure 3.2: Red/cyan color filter 3D glasses[Vie]

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

This is a passive method.

##### Time-sequential two-view 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.

### Passive method

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



Figure 3.3: Polarized 3D system[Pol]

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 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.

### Time-parallel two-view displays

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

### Simple stereoscopy

There are multiple designs available. One would be a variation on the polarized glasses : instead of having the images one after the other, they are put together, for instance by using a dual projector. However, this can lead to crosstalk.

Another 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.

### 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].

## 3.3.2 Horizontal parallax multiview 3D diplays

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

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.

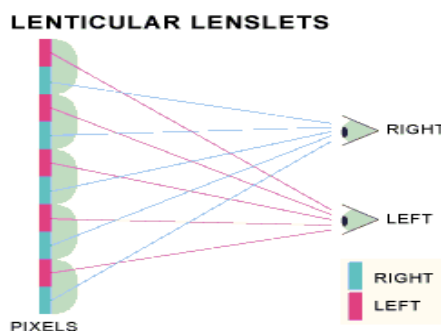


Figure 3.4: A lens array[[gla](#)]

### 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 3.7; 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.

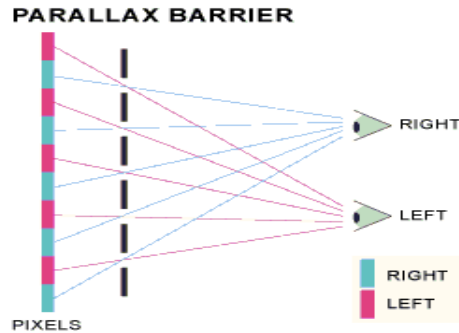


Figure 3.5: A parallax barrier[[gla](#)]

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 this 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 masks. However, a recent amelioration on this technique greatly improves brightness[[LZWW14](#)].

### Multi-Projector Displays

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.



Figure 3.6: 3D multiviewpoint display[[3Dm](#)]

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.

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.

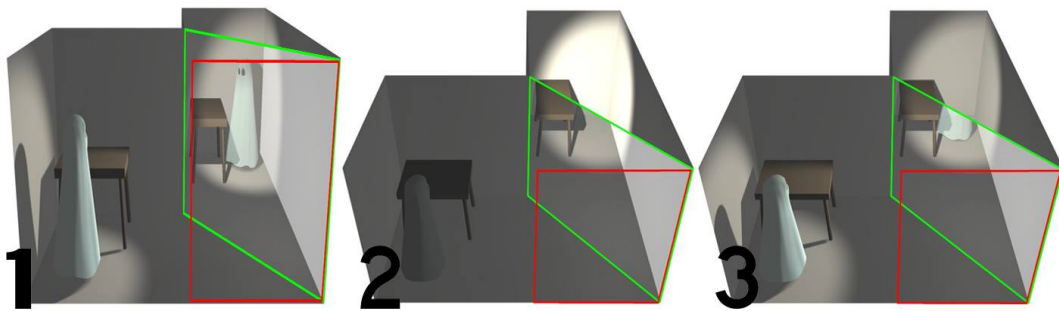


Figure 3.7: Pepper's ghost

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  $45^\circ$  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

De nombreux instruments de musiques immersifs se concentrent sur la navigation dans un environnements 3D virtuelle. Tout d'abord le projet Phase [RLC<sup>+</sup>05] explore la génération, la prise en main et le controle de son ou de musique à l'aide d'un capteur haptique et d'une représentation visuelle pouvant guidée l'utilisateur. Un second projet, Plumage [JAC<sup>+</sup>07], est une interface pour le contrôle interactif de la composition audio spatialisées. Des plumes dispersées dans une scène 3D représente des grains sonores, génèrent du son lorsque des têtes de lectures les parcours. Les têtes de lectures sont contrôlées directement par l'utilisateur. Néanmoins ces deux projets ne permettent pas de manipuler directement la structure de la synthèse sonore, mais seulement de la manipuler.

Une autre gamme d'instrument 3D se concentre sur une unique synthèse sonore. Dans ce cas nous pouvons trouver par exemple le *Virtual Xylophone*, le *Virutal Membrane* ou encore la *Virtual Air Guitar* [MPLKT05]. Un autre exemple d'interaction 3D avec une synthèse sonore unique est celle de Mike Wozniowski [WSC06]. Son application permet a un utilisateur de naviguer dans une scène 3D comportant à certain point précis des générations de son. L'utilisateur entend les sons en fonction de sa position et de son orientation dans la scène 3D.

La percussion aérienne est un intstrument 3D que nous pourrions mettre dans cette classe d'instrument.

Le DRILE propose une nouvelle utilisation de la 3D. Le DRILE utilise l'interaction 3D pour pouvoir manipuler plus aisément la structure même d'une musique.

Le DRILE et la percussion aérienne ont été conçu pour la performance musicale.

## 4.2 The DRILE

### 4.2.1 Hierarchical Live-Looping

// ANCIENNE VERSION MARIE AMELIORE PAR DAMIEN ET JEAN The DRILE musical process is based on the hierarchical live-looping technique.

#### What is Live-Looping?

Live-looping is a composition technique which came to light owing to famous modern composers such as Steve Reich. The concept is relatively simple : composing music using only samples that are played in loop. 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 guitarists and electronic musicians.

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

#### Hierarchical Live-Looping

Inspired from Marczak [MEP07] works, hierarchical live-looping organizes live-looping in a tree. Leafs are composed of samples (i.e audio extract or synthetized sound) and a musical effects list. Then, each node are made up of a node children list, a musical effects list, and a musical content. This 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. //

// NOUVELLE VERSION MARIE L'instrument Drile qui constitue une partie importante de notre étude, est basé sur la technique live-looping hiérarchique, il est donc bon de parler de cette technique.

Tout d'abord le live-looping classique est une technique qui consiste à enregistrer des échantillons audio ou de contrôle, et de lire ces enregistrements en boucle, en temps réel . L'idée de base de cette technique est de prendre (capturer) et de mettre en boucle, une partie d'une performance ou d'une chanson en temps réel puis de pouvoir ajouter d'autres morceaux ou arrangements par dessus. Cette technique est très utilisée chez les chanteurs et les instrumentistes car elle leur permet de se créer facilement un accompagnement avec très peu d'instrument. La technique du live-looping classique est limitée car elle ne permet que d'empiler des boucles musicales ou rythmiques.

Le live-looping hiérarchique permet d'obtenir des structures musicales complexes en ajoutant au live-looping classique une structure arborescente.

Fonctionnement de cette structure arborescente :

On peut voir une composition musicale comme un ensemble d'arbres avec des nœuds et des feuilles.

Dans cette structure arborescente, les feuilles contiennent des échantillons sonores bruts et les nœuds les informations ainsi que les commandes permettant de lire ou de modifier les échantillons sonores (mode de lecture, suite d'effets, etc.) On distingue deux sortes de nœuds (parents et enfants), une action sur les nœuds parents interagit sur les nœuds enfants qui peu-

vent se trouve trouver sur le même arbre ou sur un arbre voisin.

Il existe un ensemble d'opérations qui permet de manipuler ces arbres. La possibilité de fusionner ces différentes opérations, facilite la construction de structures musicales variées et complexes, contrairement à la technique live-looping classique. Des logiciels ont été développés pour utiliser le live-looping hiérarchique basés notamment sur la représentation 3D, ce qui permet aux musiciens à l'aide d'interfaces utilisateurs d'agir directement sur la structure musicale dans un environnement immersif ou pas.

→ TRADUCTION MARIE The Drile instrument stands for an important part of our study. It is based on the hierarchical Live-looping technique. A technique that needs to be seen in details. First of all, the classical Live-looping technique consists of recording audio or control samples, then reading this recording in loop and in real time. The core idea of this technique is to catch and loop a given performance or song sample and add several other samples.

This technique is very used by singers and musicians cause it's handle to add musical arrangement with very few instruments. Classical Live-looping technique is limited: it can only stack music loops or rhythmic loops.

Hierarchic Live-looping allows to obtain complex music structures by adding a tree structure. How does this tree structure works ? Let's consider this musical composition as a set of trees with nodes and leaves.

In this tree structure, leaves contain pure song sample and nodes contain informations and commands as read or modify the said samples.

There is two kinds of node, parent and children. An action on a parent node interacts with the children nodes on the same or on a neighbour tree.

Contrary to the classical live-looping, the hierarchical technique, through a whole set of operations , allows the manipulation of these trees; namely the possibility of merging these different operations as well as facilitating the construction of the varied and complex musical structures. Softwares has been developed in order to use hierarchic Live-Looping. Especially in 3D representations, that way, musicians can directly interact with the musical structure in a immersive environment or not.

//

#### **4.2.2 Piivert, a new controller**

To interact with the environment, the author developped an input device called Piivert (shown in 4.1). This device is separated in two controllers that are properly tied to each hand. That improves the musical control of the user. It combines a 6DOF tracking -that allows ray casting and some gestures detection- and high-sensitivity pressure sensing.

The author approach regarding the choice of musical gestures was thought following 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 have an important part in the instrument thanks to its pressure sensing and vibrating feedback (i.e excitation can be used to trigger sequences).

#### **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.



Figure 4.1: Piivert with pressure buttons on the bottom

### 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.

Those 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.

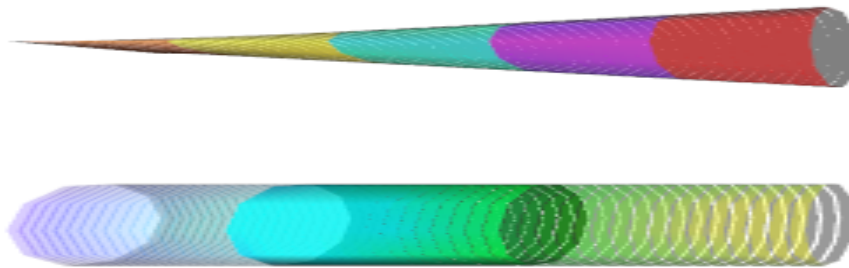


Figure 4.2: Tunnels

### 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 we use an *interaction plane* which is at a fixed distance and allows manipulation of Drile objects. Depth-manipulations are maintained for less accurate uses.

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.

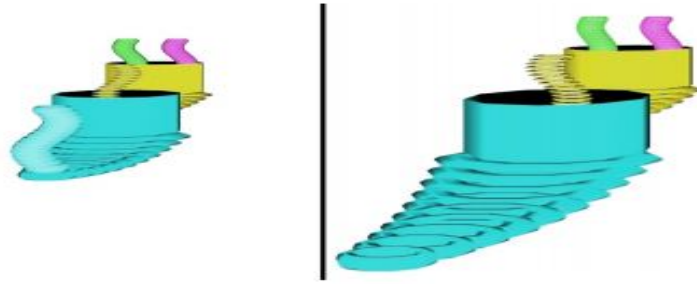


Figure 4.3: Tree nodes containing colored worms. Appendages represent tree edges

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 simulate the fact of moving aside in front of a real scene and then it becomes easier to access objects in the back of the scene. Common tree operations (merging, duplication, ..) are mapped to Piivert gestures that look natural to improve immersion. For example merging gesture is colliding two worms.

## Scenes

During a performance, the user might want to use pre-built musical patterns, those ones are stored into *scenes* (4.4). He might also want to translate playing nodes from one to another scene in the aim of doing a nice musical transition. The 3D environnement creates a good visual feedback of the scene content which can help the user to choose the right scene.

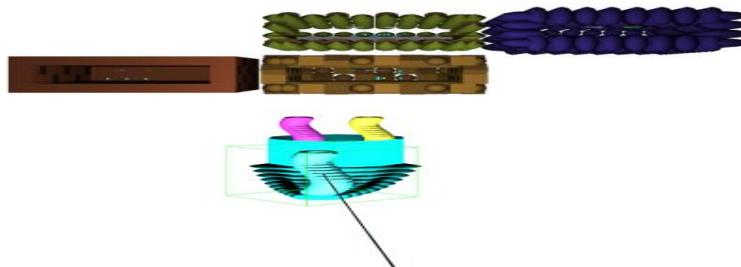


Figure 4.4: Scene selection for a worm translation

## Collaboration

The Drile environnement is a multiple users interface as long as each user have one controller. The large display is also an advantage.

## Evaluation of the instrument

The author used a dimension space representation (??) to evaluate his instrument according to [BFMW05]. We can see that the Drile environnement "grants" two important sides in this

representation : first the Feedback Modalities axe especially with the 3D environnement and the Piivert controller, finally the Inter-Actors axe thanks to its multiple user interface and its large display.

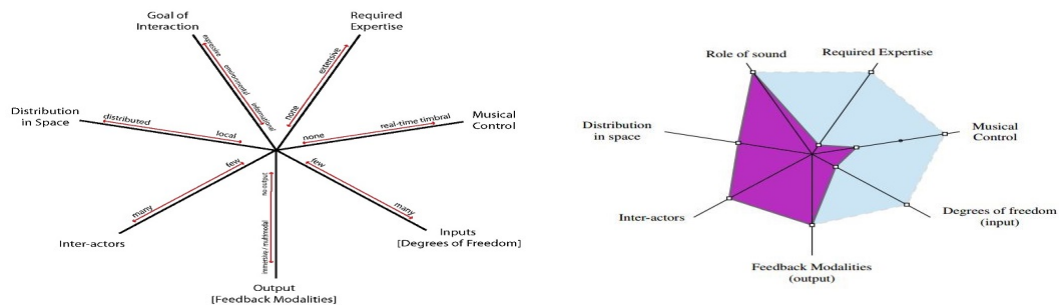


Figure 4.5: On the left the dimension space representation, on the right the Drile representation (in purple for beginner interface, in blue for complete control)

### 4.3 Aerial Percussion

# 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 setup 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. autostereoscopic 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.

#### Setup time

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

#### 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.

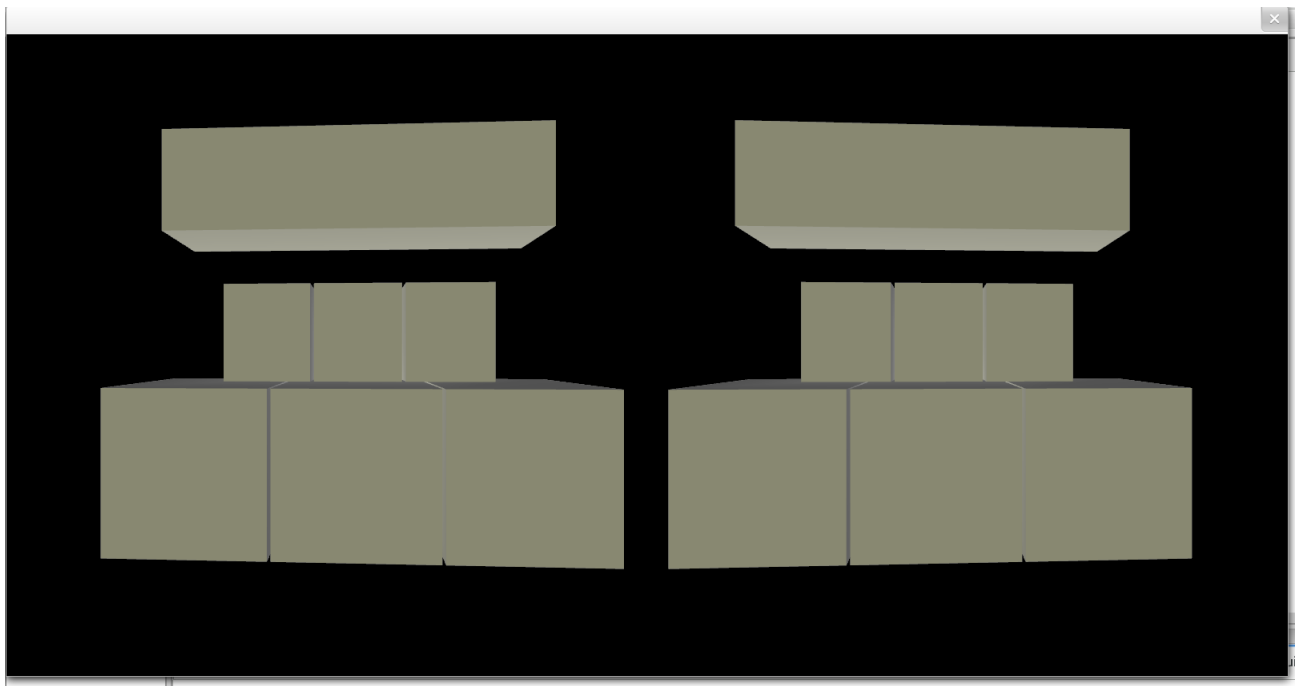


Figure 5.1: A screenshot of the Aerial Percussion renderer

## 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  $1920 \times 1080$ . Hence, the viewport is  $3840 \times 1080$ .

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 [3.3.1](#).

### **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 when the musician puts a stick inside a virtual drum element.

## **Chapter 6**

# **Conclusion**

Penser à mettre perspectives

# Glossary

**accomodation** . [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** . [9](#)

**display** a visual output device for the presentation of images [[PS12](#)].. [9](#)

**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. [4](#)

**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<sup>+</sup>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<sup>+</sup>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<sup>+</sup>06] LLOYD KAUFMAN<sup>12</sup>, 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](http://en.wikipedia.org/wiki/Pepper's_ghost).
- [Pol] [http://en.wikipedia.org/wiki/Active\\_shutter\\_3D\\_system](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<sup>+</sup>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 Wozniowski, Zack Settler, 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.