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 C

January 16, 2014

Contents

List of Figures

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 suitables 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**[1] and **LABRI!**[2], 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 positionned around the instrument and the musician. According to the position, the orientation, and the speed of the sensors, sounds are generated.

Required work
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.

[1] Studio de Création et de Recherche en Informatique et Musique Électroacoustique
[2] **LABRI!**

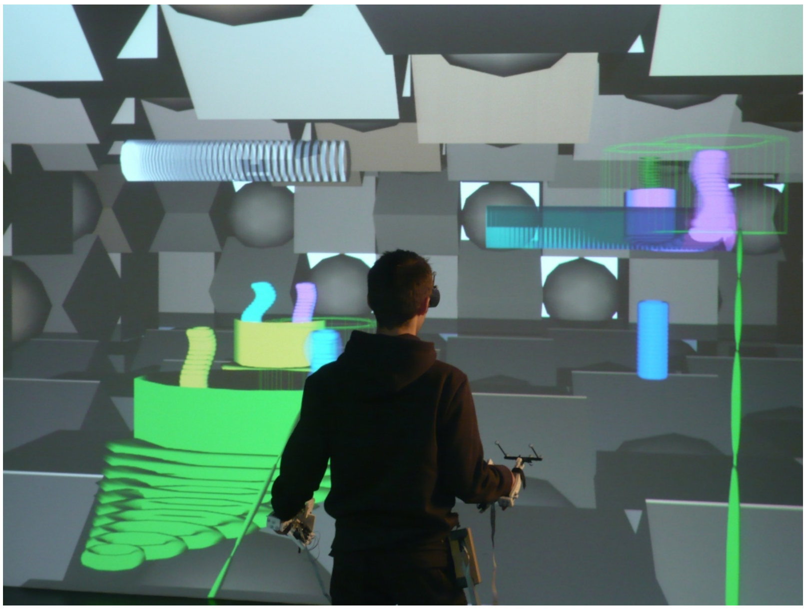

Figure 2.1: Picture of a musician using DRILE

2.2 What is a 3D musical instrument?

TODO a reformuler : bof le wikipédia

A 3D musical instrument, or an immersive virtual musical instrument, represents sound processes and their parameters as 3D entities of a virtual reality so that they can be perceived not only through auditory feedback but also visually in 3D and possibly through touch as well as haptic feedback, using 3D interface metaphors consisting of interaction techniques such as navigation, selection and manipulation.

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 environement 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 musican 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 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**[1], a **CAVE**[2] 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**.

[1] Head-mounted Display

[2] 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.

Hence, we have to qualify what is 3D and what it is not.

3.1.2 Parameters

In the litterature ([Oko76], [PS12]), the main idea is to relate to the human brain and body capabilities to define 3D vision. 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

Psychological cues
Occlusion Linear perspective Atmospheric perspective Shading

Figure 3.1: Psychological cues

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 attemps 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 :

- 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 (red and cyan glasses). However, the result is of poor quality.

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. The eyes can only see the image intended for them, 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.

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

Figure 3.3: Polarized 3D system[Pol]

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
These displays use glasses and generate two simultaneous images. They include double projection systems that combine two projectors and polarizing filters matched with appropriate glasses.

Advantages

- The glasses are easy to build, and therefore widely available from suppliers.
- The quality of the polarizer and the nature of the polarization retention screens are important.

Auto-stereoscopy
This technique provides a terrain vision without wearing glasses adjusted, the observer is placed at the correct distance from the screen, with each eye sees the image that corresponds to it (an image will be sent to the right eye and another to the left eye and the brain will add recreating the effect of relief).
The disadvantages:

- He must be placed precisely over the screen, if the observer is not the precise position relative to the screen he will see superimposed images and is unreadable.
- The system is also tiring for the eyes.
- He does not allow the visualization of the stereoscopic image to multiple viewers at the same time.

3.3.2 Horizontal paralax multiview 3D diplays

There are two main types of technology for Multiview parallax displays: Or 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 late images giving an impression of relief (3D).
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.
Thus the principle according to the imaging angle of view we see the red image or the blue image. An observer looking at a lens 3D screen does not see the same screen pixel flat between his left eye and his right eye, although both eyes set the same lens.
On the figure below, the left eye sees a lens emitting red color (from a red pixel) while the right eye sees simultaneously the same lens blue.
With horizontal lenses, projecting changing images, the image is seen depending on the viewing angle of the lens.

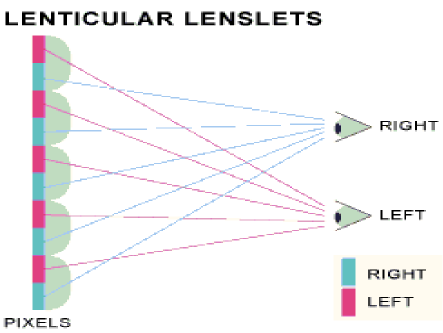

Figure 3.4: A lens array[gla]

Parallax Barrier Displays

Each cache is precisely placed so that the observer, itself placed at the correct distance from the screen, with each eye sees the image that corresponds to it. The images are divided into columns of a pixel wide. Thanks to the cache and inserting columns of each image, an image is sent to the right eye and one in the left eye and the brain will add recreating the relief effect.
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 flat between his left eye and his right eye: this is because the barrier parallax is not contiguous to the flat screen, a distance

of a few millimeters between the two panels.
On the figure below, the left eye sees through a hole in a red color (from a red pixel) while the right eye sees a blue simultaneously through the same hole.

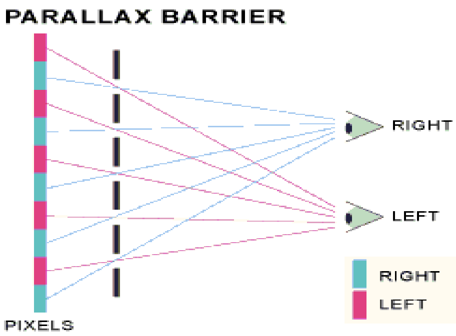

Figure 3.5: A parallax barrier[gla]

This technique provides a terrain vision without wearing glasses.

The disadvantages:

- It must be placed precisely over the screen, if the observer is not the precise position relative to the screen he will see superimposed images and is unreadable.
- We are not machines and we move without us realizing what is not compatible with the system.
- It does not allow visualization of the stereoscopic image to multiple viewers at the same time.

Multi-Projector Displays:

This technique consists of positioning in a circle several projectors displaying all an angle different image after these images are projected onto a special screen, double lenticular lens example (spherical lens), the easiest way to project these images on a cylinder fog.

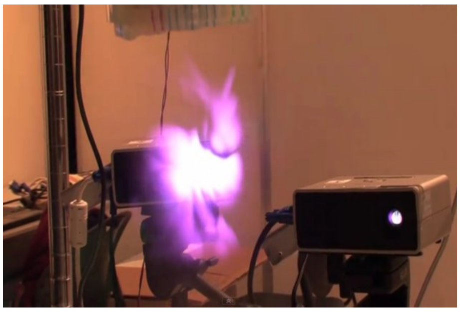

Figure 3.6: 3D multiviewpoint display[3Dm]

It is possible to turn around the rabbit in 3 dimensions and view it from all angles.
The advantage:

- One advantage of this technique, the size of the 3D image may be much larger, there is currently no limit.

The disadvantages:

- Multiple projectors are needed (for projector).
- Headlights must be precisely aligned.

Pepper's Ghost

Glasses

Head-mounted displays

Hologram

Autostereoscopic screen

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+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+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 Wozniewski [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 livelooping

4.3 Aerial Percussion

Chapter 5

Implementation

5.1 Objectives

Apart from the research work, the application of our research to two musical instruments (the DRILE and the Aerial Percussion) is required.

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, antoher 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.

5.2.2 DRILE Implementation

Technologies used

Pictures

5.2.3 Aerial Percussion implementation

Technologies used

Displaying the data

Truc sur les angles

Pictures

Conclusion

Glossary

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.. 8

display a visual output device for the presentation of images [PS12].. 8

linear perspective the way our visual perception of objects are affected by their position and dimension. 8

live looping A REMPLIR. 4

occlusion when an object is hidden by another.. 8

- L'article [HDFP11]
- L'article [PS12]
- L'article [MPWL13]
- L'article [KHY+12]
- L'article [Cad99]
- L'article [BHDC10]
- Le livre [Oko76]
- L'article [RLC+05]
- L'article [JAC+07]
- L'article [MPLKT05]
- L'article [WSC06]

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.

[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.

[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.

[KHY+12] Youngmin Kim, Keehoon Hong, Jiwoon Yeom, Jisoo Hong, Jae-Hyun Jung, Yong Wook Lee, Jae-Hyeung Park, and Byoungho Lee. A frontal projection-type three-dimensional display. *Opt. Express*, 18(20):20130–20138, 2012.

[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.

[Pol] http://en.wikipedia.org/wiki/Active_shutter_3D_system.

[PS12] Waldir Pimenta and Luıs 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.

[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.