

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MASTER THESIS

Ensuring Self-Haptic Consistency for Immersive Amplified Embodiment

Author:
Sidney BOVET

Supervisor:
Dr. Ronan BOULIC

*Submitted in fulfillment of the requirements
for the degree of MSc in Computer Science*

in the
Immersive Interaction Group
School of Computer and Communication Sciences

January 19, 2017

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Abstract

Ensuring Self-Haptic Consistency for Immersive Amplified Embodiment

With the rise of consumer-grade Virtual Reality (VR) technologies, high quality VR equipment now no longer is only accessible to research institutes or universities but to the general public as well. This opens up a lot of possibilities.

One of these opportunity appears to be in the field of stroke rehabilitation. Due to the temporary lack of oxygen in parts of the brain, a stroke may result in partial paralysis of a limb. In such a case the rehabilitation process is long, tedious, and often demotivating. The way we treat this condition is through exercises and movements of the affected limb, hence making use of neuroplasticity in order to fully recover the original range of motion. This process however is really demotivating, because the patients are forced to face their handicap. We however believe that there is a simple, yet effective solution that will help stroke patients stay motivated, and that is by providing them which a VR game that transposes and amplifies their movements in a virtual environment.

In the past few years, huge improvements have been achieved in the area of motion capture. New techniques allow for instance to reflect skin-to-skin contacts more precisely than ever, such as the Egocentric Coordinate formalism. We propose a modification of this coordinate system in order to apply a distortion to the movements of a tracked performer, while preserving self-haptic consistency. In other words, we introduce a way to create an application in which, when patients touch their skin, the virtual avatar does so as well, but when they lift their hand 10 cm above the surface of the skin, the avatar may show a 20 cm gap if a doubling distortion is applied.

We also propose an experiment in order to understand how well we accept a movement distortion, which is key here because we do not want patients to detect the distortion, or at least not too heavily. Indeed, the goal is to keep them motivated at a reaching task of some sort, but if their eyes constantly remind them that what they see is not what they achieve, we might loose the motivating effect we are looking for.

This work therefore lays the foundation for further research in the field of amplified embodied interactions, which surely will result in dramatic improvement in terms of treatment time and life quality for stroke patients.

Keywords: Virtual reality, embodiment, agency, retargeting, motion capture, movement amplification, distortion

Acknowledgements

The acknowledgments and the people to thank go here...

Contents

Abstract	iii
Acknowledgements	v
1 Introduction	1
1.1 Stroke Rehabilitation	1
1.2 Virtual Reality	1
1.2.1 Immersion, Embodiment, and Presence	2
Immersion	2
Embodiment	2
Presence	3
1.2.2 Haptics and Self-haptics	3
1.3 Motion Capture	3
1.4 Amplified Embodiment	3
1.5 Serious Games	4
2 Egocentric Coordinates	5
2.1 Motion Capture and Inverse Kinematics	5
2.2 Egocentric Coordinates	5
3 Distortion Model	7
3.1 Egocentric Coordinates Distortion	7
3.2 Linear Function	8
3.3 Other Functions	9
3.4 Egocentric Normalization Factor	10
3.5 Reachable Sphere	11
3.6 Testing Application	12
4 Experiment	15
4.1 Just Noticeable Difference	15
4.2 Equipment and Software	15
4.3 Experiment design	16
4.3.1 Task	16
4.4 Hypotheses	18
4.5 Procedure	18
4.6 Subjects	19
5 Results and Discussion	21
5.1 Distortion	21
5.2 Experiment	22
5.3 Software Issues	23
6 Conclusion	25
6.1 Further Work	25

A Questionnaires	27
A.1 Characterization Questionnaire	27
A.2 Detection Question	27
Bibliography	29

1 Introduction

In order to properly introduce this project we need to briefly cover multiple subjects, ranging from stroke rehabilitation to motion capture techniques. All these seemingly different subjects revolve around the idea that we can do something to help stroke patients recover faster by having them participate to serious Virtual Reality (VR) games in which their movements are distorted to keep them motivated.

1.1 Stroke Rehabilitation

As explained by the NHLBI¹, a stroke is a medical condition occurring when parts of the brain stop being provided with proper blood flow. Without such oxygenation, brain cells quickly die. The resulting brain damage may induce various symptoms, such as loss of vision to one side or paralysis. It is the latter that is of interest to us, and more precisely the motor recovery process involved after the stroke itself has been identified and treated. As exposed by [1], there were more than 10 million stroke cases in 2013, which represents an increase of around 60-75% with respect to 1990.

The motor recovery process involves so-called constraint-induced movement therapy, a technique that is at least around 100 years old as proposed by Oden [2] in 1918. In his experiment he remarked that monkeys that were forced to use their almost-paralyzed limb through binding of the other one recovered much faster than the ones with unbound arms. This rather archaic procedure has been improved over the years, but it still exploits the same idea: it takes advantage of neuroplasticity and involves movement exercises of the paralyzed limb. The more the movements are exercised the better the motor recovery will be.

The process as a whole is a long one, involving checkups and exercises at both the hospital and home. Many factors play a role in how successful the recovery process will be, one of them and maybe the most important one being motivation. As argued by [3], the more a patient participates to a rehabilitation task the greater the motor recovery will be. Keeping participants motivated during such tasks is thus essential.

1.2 Virtual Reality

We now jump to a completely different topic, but a careful reader will quickly understand how both are closely linked and of interest to us.

Virtual Reality can be defined as “The computer-generated simulation of a three-dimensional image or environment that can be interacted with in a seemingly real or physical way by a person using special electronic equipment, such as a helmet with a screen inside or gloves fitted with sensors.” [4] The most common type of VR device used nowadays are Head-Mounted

¹National Heart, Lung, and Blood Institute: <https://www.nhlbi.nih.gov/health/health-topics/topics/stroke> [Last visited: 08.01.2017]

Displays (HMD), which are constructed as a screen in front of which two lenses are fixated, allowing the device to be held in front of the eyes while focusing the screen's content at infinity. Coupled with inertial sensors, such device allows one to look around at a virtual environment. Other VR displays also exist, such as the CAVE: a cube with screen-faces surrounding a user proposed by [5]. In the recent days, companies such as Oculus VR and HTC have begun commercializing HMD and VR becomes more widespread than ever.

1.2.1 Immersion, Embodiment, and Presence

As VR becomes more and more accessible to the general public and broadly used in the industry, the accepted meaning of the following terms sometimes diverged from their original definitions. As an effort to clarify these, here is a summary of the concepts these words describe.

Immersion

This first word unfortunately is the least understood by the general public. Immersion has a clear definition offered among others by [6], [7] that we think is preferable to be recalled here: it refers to the capability of a system to deliver a convincing set of sensory stimuli. It is an objective measurement of parameters such as screen resolution, audio equipment, and sensors used.

Embodiment

Embodiment is defined in the field of cognitive neuroscience and philosophy of the mind by [8], [9]. It encompasses the relevance of sensorimotor skills and the role the body has in shaping the mind, as well as the subjective experience of using and 'having' a body. It is formally defined by [10] as follows: "E is embodied if some properties of E are processed in the same way as the properties of one's body".

One may from that perspective embody a tool, such as a pen or a hammer, even though that tool is not considered as being part of one's body. As described by [10], the *sense* of embodiment (SoE) refers to the fact that one *feels* such phenomena as opposed to only *knowing* that it exists. As an example, learning that an organ is part of our body makes us embody that organ even though we cannot feel it, whereas feeling like the tip of our pen actually is part of our hand creates a SoE towards that pen.

One key factor of the SoE is the **Sense of Agency**. It is defined by Tsakiris et al. [11] as "the sense of intending and executing actions, including the feeling of controlling one's own body movements, and, through them, events in the external environment." To continue with our pen example, the fact that one feels like controlling the position and pressure of the pen on a sheet of paper contributes to the SoE towards that pen. The Sense of Agency is therefore closely linked to our research topic, given that we aim at changing the way one interacts with a virtual environment by distorting the movements of the controlled virtual body. We are therefore acting directly on the foundations of the Sense of Agency.

Although of lesser interest to us, it should be noted here that the SoE rises from two other components, alongside Agency: the **Sense of Body ownership** and the **Sense of Self-Location**. The two following definitions are the ones proposed by Kilteni et al. [12]. Body ownership denotes

the attribution of a body as our own. The sense arising from it comes twofold: receiving sensory information from that body, and the cognitive action of processing such information. The Sense of Self-Location on the other hand describes the feeling that our body occupies a given volume in space, that we *are inside* a body. We are used to all of these senses because they are a given in the real world—unless one is injured or hindered in some way or another—but they may be manipulated using VR, in the same way we affect the Sense of Agency in our work.

Presence

What most people tend to call immersion actually is the sense of presence, described by [13], [14] as the "sense of being there". As explained by [9], the central concept of the state of presence is that despite *knowing* that it is a simulation, the user *acts in* and *reacts to* the virtual environment as if it were real.

1.2.2 Haptics and Self-haptics

'Haptic' relates to the sense of touch, and more specifically to the one associated to manipulating objects. Feeling the pressure of a sphere between one's fingers when grasping a ball is a haptic feedback. Self-haptic similarly denotes the haptic feedback of touching one's own body part. The feedback thus is dual: by touching one's own arm the brain will receive both the information of the hand touching something and the arm being touched by something. Moreover, before the contact even occurs the brain predicts such self-contact using proprioceptive information about the position of both body parts in space, and is thus very sensitive to such feedback inconsistency.

1.3 Motion Capture

Motion capture can be defined as the action of capturing one's movements in order to reproduce such a motion in one context or another. Video game animations for instance are often performed by actors, whose walking or fighting movements are captured and then played back on the game's characters. The captured movement is often slightly altered, e.g. in order to fit it onto an avatar of different morphology, as proposed by [15].

1.4 Amplified Embodiment

The concept of amplified embodiment can easily be understood as the merging of both Sections 1.2.1 and 1.3. By capturing one's movements and transposing those on an avatar, we may create a SoE towards that avatar. The goal of amplified embodiment is to distort that avatar's movements by amplifying them while preserving the SoE, which might be partially or completely lost if the distortion is too heavy or non-continuous for instance.

1.5 Serious Games

To conclude this chapter we take a quick look at another distinct topic, one that closes the loop and creates a connection between stroke rehabilitation, VR, motion capture, and amplified embodiment.

As defined by [16], [17], a serious game is a game that neither has entertainment, enjoyment, nor fun as its primary purpose. Instances of serious games include educational ones such as Blupi at Home², a 1988 video game where young children could learn the alphabet by exploring a house with a yellow, egg-shaped character and playing various mini games. A more recent example is the “Zombies, Run!”³ application, which is a running mobile application using storytelling coupled with GPS localization to keep people motivated at running outside.

We hope that by now the thread linking all of the above subjects is clear enough: we are interested in discovering by how much we can distort one’s movements in the context of a VR serious game aimed at motor recovery while altering the sense of embodiment as little as possible, with the goal of keeping the patients motivated at performing the rehabilitation task.

²<http://www.ceebot.com/blupi/maison-e.php> [Last visited: 08.01.2017]

³<https://zombiesrungame.com/> [Last visited: 08.01.2017]

2 Egocentric Coordinates

In this chapter we briefly cover the Inverse Kinematics problem and how its parameters are acquired, and we then dive more deeply in the description of a special coordinate system based on the relative positions of different body parts around a point.

2.1 Motion Capture and Inverse Kinematics

As stated by Paul [18], Inverse Kinematics (IK) refers to the use of kinematic equations of a chain of rigid bodies in order to produce the desired pose of that chain so that it satisfies a position—or goal—for that chain’s end effector. Such techniques are useful for instance to compute the joint position of an industrial robot whose task is to tighten a screw at a given location. Another, more interesting application to us of such algorithms is to compute the pose of a virtual body. Such avatar is described as a kinematic chain, and its feet, elbows, knees or head positions are considered as IK goals. An IK solver may be numeric [19], [20], or it may be analytical, such as the one proposed by Molla et al. [21]. In this case, a closed-form expression is provided that takes one or several IK goals as input, and gives a vector of joint positions and orientations to be set in order to satisfy the goals.

These desired positions may be defined by some game engine so that a character points a tool at a desired location, but they are also often obtained using motion capture techniques, as mentioned in Section 1.3. Such techniques usually take advantage of special suits equipped with markers, whose individual positions are acquired either by the marker itself, or using tracking devices. As explained by [22], a suit may for instance be fitted with magnetic sensors and used in a room where a coil emits a given magnetic field, so that each sensor knows its position and orientation relative to the emitting base, acting as an absolute reference. Another way to acquire a position is to place a blinking LED at that location and tracking it using several cameras. If the position of each tracker is known beforehand (e.g. through space calibration) then the marker’s position can be recovered using triangulation.

All the above techniques focus on getting the absolute position (and sometimes orientation) of an IK goal. This is a necessary step for any animation application, but an end effector’s position need not be expressed as an absolute value throughout the application’s pipeline. The semantic information conveyed by a performer may in fact be better preserved by using an alternative, relative coordinate system.

2.2 Egocentric Coordinates

The idea of using reference points others than the origin of the world’s axes in order to describe an IK goal is already a few years old [23], but Molla et al. [15], [24] took it a step further by using one’s own body parts as reference points. This allows to correctly map the semantics of performed motions onto avatars of different proportions and sizes. Such body-centered coordinate system is called Egocentric.

A crude body representation is computed from markers placed on the performer. The limbs are represented as capsules while the trunk and head are sampled at multiple locations and then translated into a series of triangles forming a crude mesh. Given such body representation with n body parts, the position \mathbf{p}_j of an IK goal j is then represented as shown in Equation 2.1.

$$\mathbf{p}_j = \sum_{i=1}^n \hat{\lambda}(\mathbf{x}_i + \mathbf{v}_i) \quad (2.1)$$

The closest point on surface i is denoted \mathbf{x}_i , while \mathbf{v}_i is the relative displacement vector going from \mathbf{x}_i to \mathbf{p}_i , and λ is a normalization factor. This last value is computed as $\lambda = \lambda_p \cdot \lambda_\perp$, whose proximity and orthogonality factors are defined in Equations 2.2 and 2.3. $\hat{\lambda}$ is obtained by further linearly normalized λ such that $\sum \hat{\lambda} = 1$.

$$\text{Proximity: } \lambda_p = \begin{cases} \frac{1}{\epsilon} & \text{if } \|\mathbf{v}\| \leq \epsilon \\ \frac{1}{\|\mathbf{v}\|} & \text{otherwise} \end{cases} \quad (2.2)$$

$$\text{Orthogonality: } \lambda_\perp = \begin{cases} \cos(\epsilon) & \text{if } \cos(\alpha) \leq \epsilon \\ \cos(\alpha) & \text{otherwise} \end{cases} \quad (2.3)$$

In Equation 2.3, α denotes the angle between the surface normal \mathbf{n} at the closes point \mathbf{x} of a body part, and the relative displacement vector \mathbf{v} . The presence of ϵ , a tiny constant, helps prevent instability due to floating point precision numbers. The justification for measuring orthogonality is that (1) if a surface normal points at a joint, chances are they are interacting in some way, and (2) orthogonality holds semantic information about gestures, such as holding a hand in front of the heart.

The movement deformation we introduce in this work relies heavily on this Egocentric representation of the IK goals, and more precisely acts on the relative displacement vectors \mathbf{v} described above. The next chapter describes our distortion model, as well as exactly how we adapted the Egocentric formalism in order to apply this distortion to an avatar's movements.

3 Distortion Model

In this chapter we begin by describing how we modified the preexisting motion retargeting software in order to obtain the desired distorted behavior, and then we detail our distortion model and what function we used. The end of the chapter is dedicated to additional adaptations we propose to the work of Molla et al. [15].

3.1 Egocentric Coordinates Distortion

As briefly mentioned in Chapter 2, we are taking advantage of the Egocentric Coordinate formalism in order to introduce our distortion model. We modify each relative displacement vectors \mathbf{v}_i according to some value γ . A distorted position \mathbf{p}_j is thus obtained using the equation 3.1 below, which has been obtained by modifying Equation 2.1 using a function that we are going to detail in the next few lines.

$$\mathbf{p}_j = \sum_{i=1}^n \hat{\lambda}(\mathbf{x}_i + D(\mathbf{v}_i, \gamma)) \quad (3.1)$$

Figure 3.1 shows an example of a distorted position obtained by changing the length of all vectors \mathbf{v}_i . This arm model was implemented from the ground up using the FABRIK algorithm as an IK solver [25]. More considerations on this particular implementation can be found in Section 3.6.

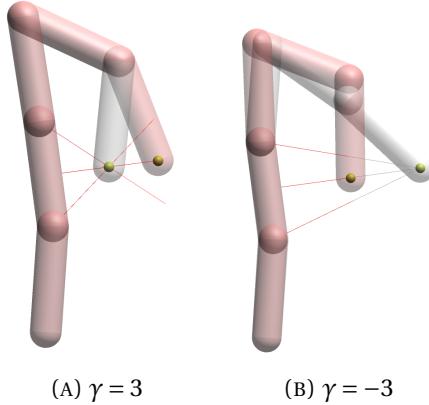


FIGURE 3.1: Two examples of our linear distortion applied to a simple IK arm with multiple segments. The gray lines are the relative displacement vectors and the red ones are their distorted counterparts. Similarly, the gray arms represent the pose the real arm would take whereas the red ones show the two resulting distorted poses.

We will now describe the function we will be using in our subsequent experiment, and then propose a distortion expression that may be better-suited for real-world applications.

3.2 Linear Function

We are considering candidate functions D to use in Equation 3.1. For ease of experimentation and clarity, we are looking for a linear function $f(x) = ax + b$. Figure 3.1 gives an example of what we aim to achieve, while Figure 3.2 below gives a more mathematical point of view of the distortion we are looking for, especially in terms of a , the slope of the function. This plot, as well as all of the other plots of this report, were obtained using the Plotly API [26].

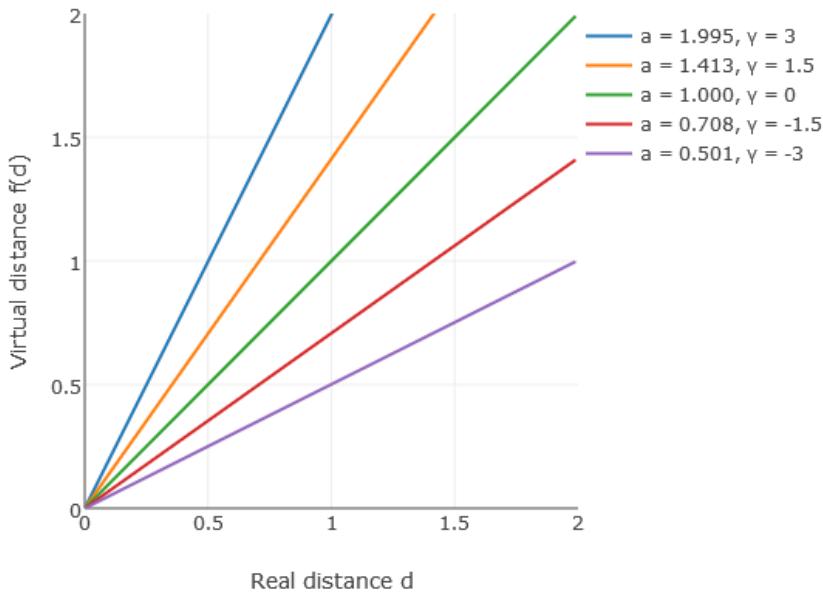


FIGURE 3.2: An example of a few distortion functions for various values of slope a and gain γ .

First of all we want to preserve self-haptic contacts. Such contact happens when a relative displacement vector satisfies $\|\mathbf{v}\| = 0$, which means that we need $f(0) = 0$, and thus $b = 0$. Moreover, we also rejected the possibility of distorting such skin-touching positions orthogonally to the normal at the contact point because it would imply a much more complex distortion model taking into account the local topology of the body. We also feel that it would not improve the distortion behavior significantly.

Going back to our linear function, one may intuitively see that the slopes should be arranged around $a = 1$, which results in no distortion at all and that we want to correspond to $\gamma = 0$. It can also be figured out that there is a correspondence between slopes above and below the line $f(x) = x$. For instance, for a given virtual distance to cover, a slope of 0.5 makes the traveling distance twice as long, whereas a slope of 2 halves the required movement.

Formally, we are modifying each relative displacement vector as specified in Equation 3.1, with γ representing a gain, measured in dB, and f defined by Equation 3.2 below.

$$D(\mathbf{v}, \gamma) = \hat{\mathbf{v}} \cdot \|\mathbf{v}\| \cdot 10^{\frac{\gamma}{10}} \quad (3.2)$$

In this equation, \mathbf{v} is a vector, $\hat{\mathbf{v}}$ is its normalized counterpart, and $\gamma \in \mathbb{R}$. The last factor, $10^{\frac{\gamma}{10}}$, comes from the definition of a gain, in dB [27], based on two values P_1 and P_2 of a single, yet undefined unit:

$$\begin{aligned}\text{gain} &= \gamma = 10 \cdot \log_{10}\left(\frac{P_1}{P_2}\right) \\ \frac{\gamma}{10} &= \log_{10}(\text{slope}) \\ \text{slope} &= 10^{\frac{\gamma}{10}}.\end{aligned}$$

A value of $\gamma = 3$ thus indicates that the virtual movement will roughly be twice the amplitude of the registered one ($1.995 \approx 2$), while a gain of $\gamma = -3$ means one will have to travel twice as big a distance (0.501) as perceived in order to cover it. Figure 3.1 shows two examples of distortion, and Figure 3.2 gives a few instances of this linear distortion functions with varying values of γ and the corresponding slope a .

3.3 Other Functions

Before deciding to use a simpler, thus easier to quantify, linear function for our experimentation process, we tried out different functions that we think are of interest for further applications. Two of these functions are described here as a reference for further investigation.

As for our linear function, we require the distortion to be null around $\|\mathbf{v}\| = 0$. Similarly, we introduce an action range a_r after which the distortion should be null again. The general form of the function applied to our relative displacement vectors then becomes the one described in Equation 3.3.

$$D(\mathbf{v}, s, a_r) = \begin{cases} \hat{\mathbf{v}} f(\|\mathbf{v}\|, s, a_r) & \text{if } \|\mathbf{v}\| \leq a_r \\ \mathbf{v} & \text{otherwise} \end{cases} \quad (3.3)$$

Note that we changed the ‘ γ ’ parameter for ‘ s ’, which is due to this parameter no longer denoting a cleanly defined gain, but a vaguer concept of strength. Two notable instances of the function f were implemented. They are shown in Figure 3.3 and are described hereafter.

Exponential

Based on an expression proposed by Khouri [28], this function has the following form:

$$f(d, a_r, s) = a_r \cdot \left(\frac{d}{a_r}\right)^{2-s}$$

As one can observe on Figure 3.3, this function has the advantage of smoothly transitioning from a non-distorted position to a maximum discrepancy, and then back to an undistorted position near $d = a_r$. This behavior however comes at a cost. At high strength values (e.g. $s = 2.4$) we observe a severe jump in position when the hand departs from the skin: in a few centimeters in the real world, the virtual hand already jumps to almost two thirds of the action range. For equally high negative values the behavior is similar in that the hand seems to

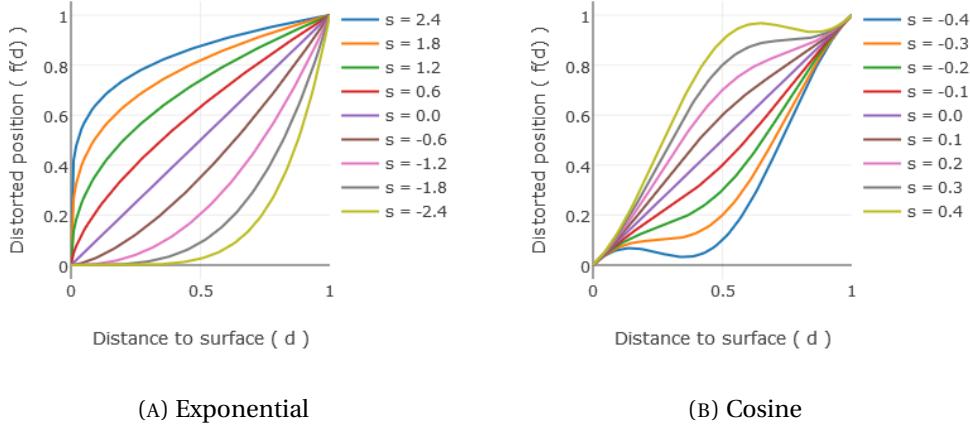


FIGURE 3.3: Multiple instances of the two alternative distortion functions we propose in Section 3.3. Both have been plotted using $a_r = 1$, only varying the strength parameter s .

‘wait’ on the skin for a long while before speeding up to rejoin the hand and approaching the other end of the action range. This high velocity discrepancy gets easily noticed and further investigations should be made as per what strength is acceptable.

Cosine

Observations made on the previous function led us to designing this second function, that is smoother in terms of velocity discrepancy near the ends of the range of the function. It is defined as:

$$f(d, a_r, s) = \frac{s}{2} \left(1 - \cos \left(\frac{2\pi}{a_r} \cdot d \right) \right) + d$$

Obviously, being smoother in terms of velocity near both ends of the distorted area leads to trade-offs. In our case the function introduces higher velocity discrepancies at other locations, as one can observe on Figure 3.3b. At high strengths (e.g. $s = 0.4$) the function even forces the virtual hand to go backwards in order to rejoin the real position before approaching a_r .

Again, both of these functions are an attempt at fixing the behavior of the distortion when going further away from the skin, but further experimentation is needed in order to determine how they affect both our acceptance of a distorted movement as our own, and the SoE through their respective effects on the Sense of Agency.

3.4 Egocentric Normalization Factor

We added one more modification to the definition of the position proposed by [15] which we modified to obtain Equation 3.1, and more precisely the way λ is defined. As originally explained by [24] as well as in Section 2.2, it is computed as the product of two importance factors, proximity and orthogonality, respectively denoted λ_p and λ_{\perp} .

The former importance factor was initially defined as $\lambda_p = \frac{1}{\|v\|}$. In practice we find that this formula does not give enough importance to nearby body parts, and we decided to change it slightly as $\lambda_p = \frac{1}{\|v\|^2}$, thus becoming proportional to the inverse square of the distance between the surface and the point. We find this definition elegant since it matches the form of an inverse square law (such as Coulomb's law, or Newton's law of universal gravitation), of the form:

$$\text{intensity} \propto \frac{1}{\text{distance}^2}$$

This indeed has a physical justification: a point-source geometrically dilutes its radiated energy, force, or any other conserved quantity in three-dimensional space, hence the influence of the inverse square of the distance between that point-source and a given surface. In our case we can consider the “amount of distortion” received by any point in space in a similar way. Moreover, this gives more consistent practical results in the sense that when moving the hand horizontally in front of the torso, that hand is less affected by other body parts such as the legs or the head.

3.5 Reachable Sphere

Molla et al. [15] proposed a special case for handling the position of the feet: they use the orthogonal projection of the ankle on the floor as an additional reference point, hence adding one Egocentric component to the coordinates of that joint.

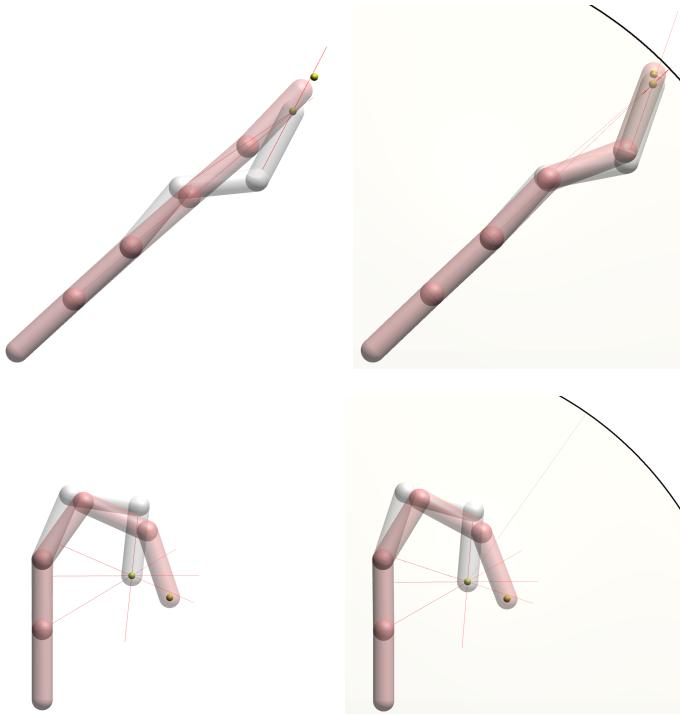


FIGURE 3.4: An example of undesired pose behavior near full arm extension (top left) and how we propose to solve it (top right) using the reachable sphere, in light yellow and outlined on the top right corner for readability. A pose far from the surface of the sphere is almost not affected (bottom left and right). The captured arms are in gray, the distorted ones in red, and the red lines represent the distorted coordinates.

We think that interactions close to the far end of the reachable volume of each arm may benefit from such an additional reference point as well, especially when considering distortions made to the wrist position. To do so, we propose the addition of a sphere of radius equal to the arm's length and centered on the shoulder of that arm, that we call “reachable sphere”. When approaching full arm extension, the position of the wrist relative to that reachable sphere becomes prominent and forces to distortion to act as if the hand was approaching one's skin: it would tend towards a null distortion. Figure 3.4 illustrate the type of pose behavior we attempt to avoid by introducing this reachable sphere, as well as shows that it does almost not affect movements closer to other body parts.

In the experiment described in Chapter 4 we however did not implement this reachable sphere for two reasons. First, it greatly alters the linear nature of the distortion function we proposed in Section 3.2 and thus is not desired in such an experimental context. The second reason is more practical: the retargeting software is quite complex and unfortunately poorly commented, which makes such a modification very hard to achieve within the timespan of this project.

3.6 Testing Application

The simple kinematic chain we used for preliminary testing and to obtain for instance Figure 3.1 was built using the description of the Egocentric Formalism provided by [24]. For ease of use, we integrated a preexisting IK solver to this implementation, namely the FABRIK algorithm proposed by Aristidou and Lasenby [25].

We encountered one issue when using this testing application once the distortion was implemented: depending on the sequence of movements of the IK goal, the actual kinematic chain and the distorted one began to have diverging intermediary joint positions, such as shown in Figure 3.5.

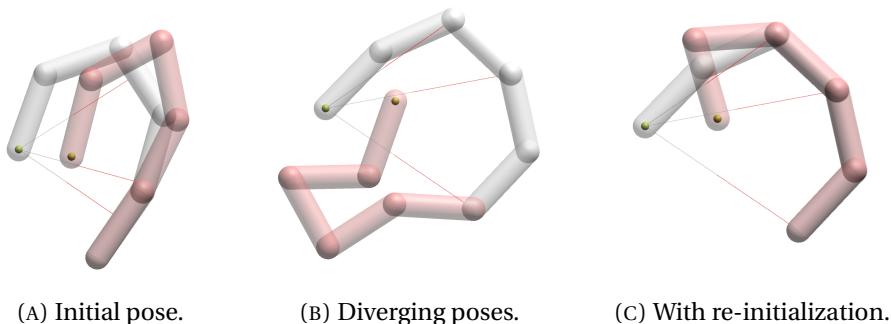


FIGURE 3.5: In gray are the real arms whose movements are supposedly captured through their respective IK goals (yellow spheres inside the last capsule), and in red are the distorted poses, which are the ones a end user would see in the virtual environment. An example of initial IK chains (left) and the diverging poses after a certain sequence of movements (center). The pose can then be corrected by initializing the distorted arm at the real arm's position (right).

Two options exist in order to solve this issue. The first one is to introduce secondary IK goals at what can be considered as the ‘ellbows’ of the arm. The second option is to re-initialize the distorted arm at the real arm pose before running the IK iterations.

While we chose the latter for our testing application, the former is probably better-suited for real world applications, where markers are available all over the body and may serve to determine such secondary IK goal.

4 Experiment

We are trying to estimate the limits of self attribution of a distorted movement. We will do so by estimating the just noticeable difference (JND) in visual stimuli discrepancy. This means estimating the just noticeable *distortion* made to the movement and hence the visual stimuli. The JND is estimated by using the adaptive staircase method introduced by Meese [29].

This method tries to estimate the JND by finding the point at which the subjects are uncertain whether a given visual stimuli corresponds to the actual movement they performed. These are found by changing the intensity of the distortion, based on whether the subject judged the last trial as distorted or not, and the JND is computed as the mean of the last few staircase turns (i.e. going from an increasing trend to a decreasing one or vice-versa). The detection judgment is gathered using a Yes/No prompt called the detection question : “Did the movements you saw exactly correspond to the movements you performed?”

4.1 Just Noticeable Difference

The JND will be measured in term of γ , the gain of the distortion function introduced in Equation 3.2. In general, if $\gamma = -3\text{ dB}$ the subjects are hindered by having to travel two times the distance between the targets, whereas if $\gamma = 3\text{ dB}$ the movement will be amplified and the required motion will be reduced by 50%.

Due to the nature of the Egocentric Coordinates and how the distortion is applied (respectively detailed in Chapters 2 and 3), this will not exactly be a metric of the difference in the distance that the subjects have to cover in order to reach the target, such as the metric used by [9]. It however gives a good understanding of the strength and the effect of the applied distortion.

4.2 Equipment and Software

The HMD used for this experiment will be the Oculus Rift in its first consumer version, with a resolution of 1080×1200 pixels per eye and a refresh rate of 90 Hz. Head orientation is acquired using the internal sensors of the HMD in order to minimize latency, and its orientation drift around the vertical axis is corrected using a set of markers placed on it. These markers also serve to determine the position of the camera in the virtual environment.

Audio is delivered using a pair of Bose® QuietComfort 35 wireless headphones. They feature high quality environmental noise cancellation and are additionally used to stream unlocalized white noise, in order to fully isolate the subjects. If needed, they are also used as a communication means between the subject and experimenter, the latter making use of a microphone.

Motion capture is performed using a PhaseSpace Impulse X2 optical tracking system. The setup uses 18 infrared cameras and a suit equipped with 38 LED markers, which are arranged as described by Molla [24] in order to track limb rigid bodies.

The virtual environment was developed using the Unity game engine. The room setup is derived from [9] with a chair, a carpet, and a few point light sources. The carpet is here to better match the real world, in which the tracking space has a carpet on its floor. The physical chair has been measured so that it would match the virtual seat we are using. Additionally, the wall in front of the subject features textual information on the next target to be hit or on resting times.

The avatar is animated using the IK and retargeting algorithms proposed by Molla et al. [15], [21], the latter of which is modified as described in Chapter 3.

4.3 Experiment design

We manipulate three factors: the sign of the distortion (positive or negative), respectively yielding a helped and hindered movement, the starting value of that distortion (0 or 3 dB, and the target sequence (see below). Chapter 3 gives a complete overview of the concept of distortion and its sign, as well as how it is implemented.

4.3.1 Task

While the whole set of IK goals will be distorted during the experiment, we will be focusing on the dominant hand movement. The task is performed in a seated position in order to avoid any unnecessary movement of the lower limbs, and has the subjects reach three targets. One of them may be in the air in front of them, while the starting and finishing location is the same relative to their skin. The reaching task is performed with the directing hand, and the subjects are instructed to keep their other hand at their side.

There are four different target orders: **Chest-Air-Chest** (O_1), **Leg-Air-Leg** (O_2), **Chest-Leg-Chest** (O_3), and **Leg-Chest-Leg** (O_4). Both chest and leg targets are located as depicted on Figure 4.1 and are picked according to the subject's handedness: left-handed subjects for instance will have to reach for their left thigh and right shoulder. The first target to be displayed, T_1 , is always one on the skin and requires the subjects to perform a self-contact in order to activate it. After a random time between 200 and 300 ms, the target activates. Then the subject moves the hand to the next target, which activates after 100 ms. Once this is done the subject goes back to T_1 , and the detection question is finally asked.

If the sequence requires an intermediary air target (O_1 and O_2), that target's position is computed such that the subjects have to move a predefined distance $d = 75\%$ of arm length between T_1 and T_{air} . Given that a whole sphere of positions is possible, and in order to disambiguate that position, we require that T_{air} is also at the same distance d from the third, unused target T_3 . The air target position therefore is on the intersection between the two spheres of radius d centered on T_1 and T_3 , and we chose the topmost position as a reasonable last disambiguation criterion. Figure 4.1 shows how such position is computed.

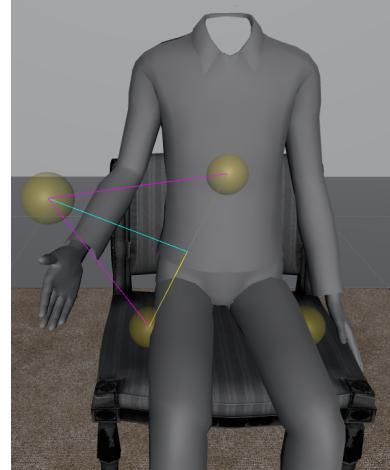


FIGURE 4.1: An example of target placement with lines showing how its position was reconstructed. The magenta lines are the one of desired length.

This target placement is useful in that it helps requiring movements of equal effort and does not change the ability to perform the trials across subjects.

One trial—or staircase run—consists of a reaching task, followed by the detection question. Based on the answer to this question, the experiment software modifies the distortion for the next staircase run as follows:

"Yes" The discrepancy is increased.

"No" and $\gamma \neq 0$ The discrepancy is decreased.

"No" and $\gamma = 0$ The parameter is not changed¹.

The amount of each increment or decrement is dynamic: it starts at 0.5 and is halved after the first staircase turn. That value is then kept for the rest of that staircase. This helps going quickly from the starting distortion to the point where subjects are unsure whether the movement they saw was theirs, and then explores the space around that point.

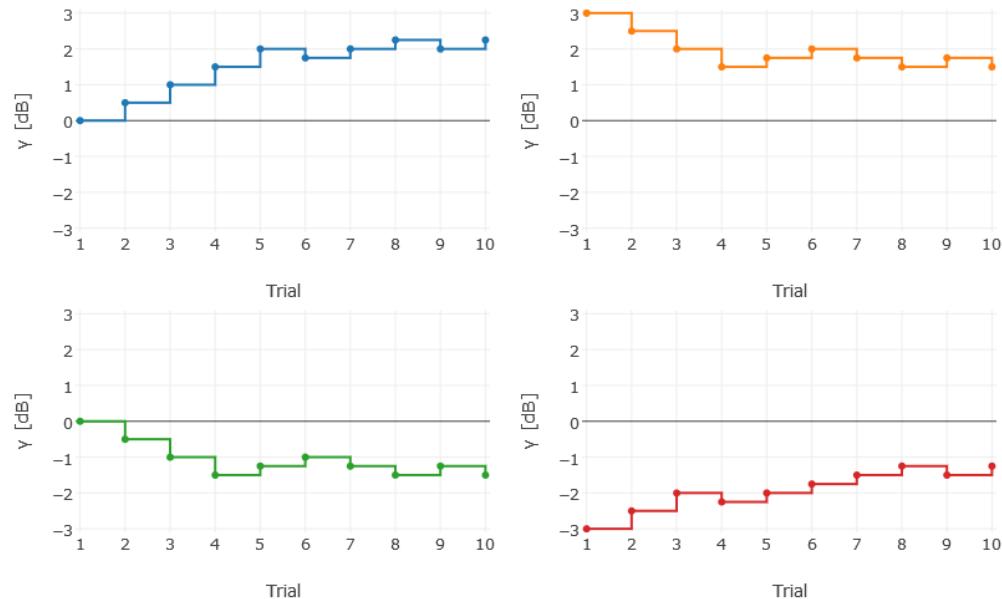


FIGURE 4.2: Four made up partial instances of staircases. Each one features ten trials and four staircase turns, including the one happening in the last trial. The sign of the staircase is the same for each row, while the starting distortion is the same for each column.

A staircase is completed either when the subjects change direction 7 times or when they performed 20 trials in that staircase. Figure 4.2 shows four instances of ten consecutive trial, depicting the four different initial staircases configuration used for each possible target order.

¹We do not change the distortion in this case because this would invert the sign of the distortion, thus introducing a distortion of different sign.

4.4 Hypotheses

Considering the target orders O_n we introduced in Section 4.3.1 and their resulting JNDs, denoted JND_n , the hypotheses we have for the experiment are the following:

- H1: The absolute value of the JND will be higher when the distortion is positive.
- H2: For a positive distortion, $JND_3 \approx JND_4$
- H3: Also for a positive distortion, $JND_1 < JND_2 < JND_3$
- H4: For a negative distortion, $JND_1 > JND_2$

H1 is based on the findings of [9], who found that a hindering movement distortion is more quickly detected than a helping one. The justifications for all other hypotheses are based on the type of movement asked of the subjects.

Considering the similarity of the tasks O_3 and O_4 , we believe they will exhibit similar JNDs, hence H2.

H3 translates our belief that the presence or not of body parts other than the starting one (the leg in the case of O_2 for instance) along the path of movement have an impact on the detection threshold. The symmetry of the repartition of the said body parts when performing O_1 (i.e. the legs below and the head above) will be beneficial to the resulting motion, whereas having the whole chest on one side of the task O_2 and nothing on the other one will lead to an increased JND. Similarly, both O_3 and O_4 will be negatively affected due to the nature of their respective paths. [30]

Also reasoning about trajectories, we believe that a negative distortion will have a more visible impact on O_1 than on O_2 , as stated by H4. The reason for this is that on a vertical path starting on the leg, subjects may travel a longer distance before reaching the joint limits of their arm, as opposed to during a horizontal trajectory starting at the chest. Indeed, the further one can go from the chest is roughly one arm's length, whereas starting from the leg allows one to reach much higher than that.

4.5 Procedure

The subjects are welcomed and introduced to the protocol described here, and then introduced to the tracking equipment. A consent form is signed and a characterization form is then filled in by the subjects. The questionnaire features background questions such as age and handedness, but also regarding any previous VR experiment participation, or experience with HMDs. They are then asked to remove their shoes and helped putting the motion capture suit on. A calibration is then performed as described by [15].

Before beginning the actual staircase trials, a familiarization phase takes place. The subjects go through two series of five trials with constant distortion. The first one has no distortion at all and the second has a big one of $\gamma = 3$ dB. The subjects are instructed to always answer "Yes" during the first staircase, and "No" during the second one, so that they become familiar with the whole procedure.

After these training trials, the real staircases begin. As described earlier, each staircase ends either after 20 trials, or when seven staircase turns are registered. At the end of each staircase there is a time for the subject to rest, during which they are encouraged to ask for a glass of water if needed. If they do, the experimenter helps them with the handling of the headphones

and HMD. When they are ready to continue, the experimenter presses a button in order to resume the experiment with the next series of trials.

4.6 Subjects

A few physical limitations will be applied to filter the subjects of this experiment. Their height will be required to be between smaller than 180cm and have a body mass index between 18 and 27. Those criteria are due to our motion capture equipment and especially the size of the suit on which the markers are placed. We also require that they have a normal or corrected to normal vision, and be fluent in both written and spoken English.

The subjects were paid 20 CHF per hour for their participation in the experiment.

5 Results and Discussion

We published a web site showing videos and interactive examples to complement this report. It is available at sbovet.ch/distortion.

As briefly mentioned towards the end of Section 3.5, the code base used in this project, and particularly the retargeting part implementing the whole Egocentric Coordinate formalism, is problematic. Although well formated and making use of carefully chosen variable names, the code is intrinsically complex and has sadly not been well documented nor commented.

This unfortunately forced us to reconsider the schedule for this project due to issues exposed in a video available on the aforementioned web page. This rescheduling prevented us from performing the experiment described in Chapter 4 at the time of writing this report. Luckily, we will be able to continue this project in the next six months thanks to a project grant by the Hassler Foundation, hence making it possible to run the experiment and publish its results in a subsequent report.

The remainder of this chapter therefore does not describe the results of the experiment itself, but what we have been able to achieve software-wise in terms of distortion and experiment implementation.

5.1 Distortion

A few examples of distortion and the resulting poses have already been shown on Figures 3.1 and 3.5. We propose here a more detailed summary of our results and a discussion of them.

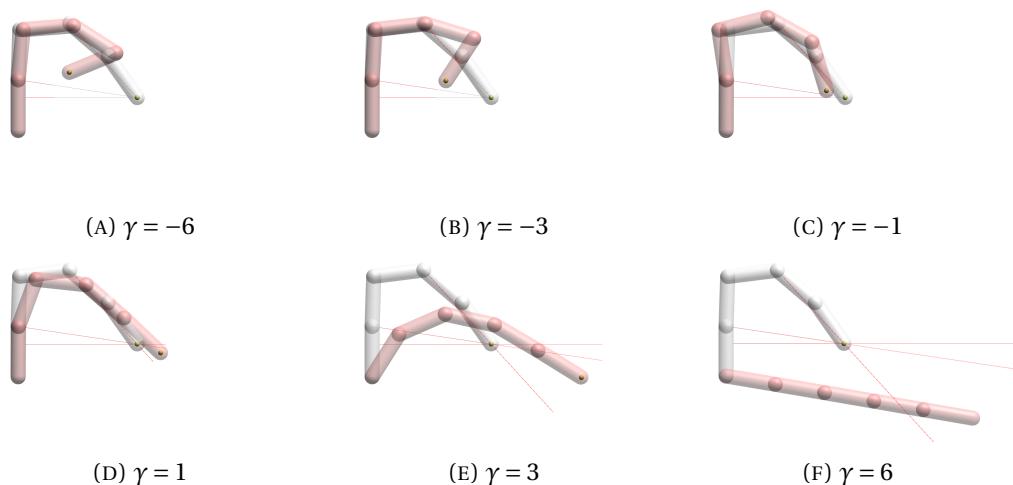


FIGURE 5.1: An example of distorted poses. As always, the captured arms are in gray, the distorted ones in red, and the red and gray lines respectively represent the distorted and original Egocentric coordinates.

As one can observe on Figure 5.1, the distortion works the way it is expected to. As a reminder, a gain of $\pm 6\text{dB}$ corresponds to a movement whose amplitude is respectively multiplied by

3.981 and 0.251, which means it is changed almost fourfold in each direction with respect to a non distorted one. A value of $\gamma = \pm 1$ dB similarly modifies a movement by 1.259 or 0.794.

The behavior of the bottom right arm in Figure 5.1 is precisely the one we proposed to avoid using the reachable sphere described in Section 3.5 but we do not report on it here because, as previously justified, we did not implement it in our experiment.

An interactive web application is available on the web site mentioned above. One may play there with both the real arm position and the gain of the distortion, as well as toggling the presence or not of the reachable sphere. The reader is encouraged to play with the IK chain in order to get a better feeling for how the distortion works, especially how it behaves near the skin.

In order to better understand how such a distortion applies to a captured motion, Figure 5.2 shows a subject performing a reaching motion towards the air target. The virtual hand is at the same position on both pictures, but as one may observe, the subject has his hand at different positions. This is due to the fact that in one case (higher hand position) no distortion was applied, while the other (lower hand position) sees a distortion of 3 dB.

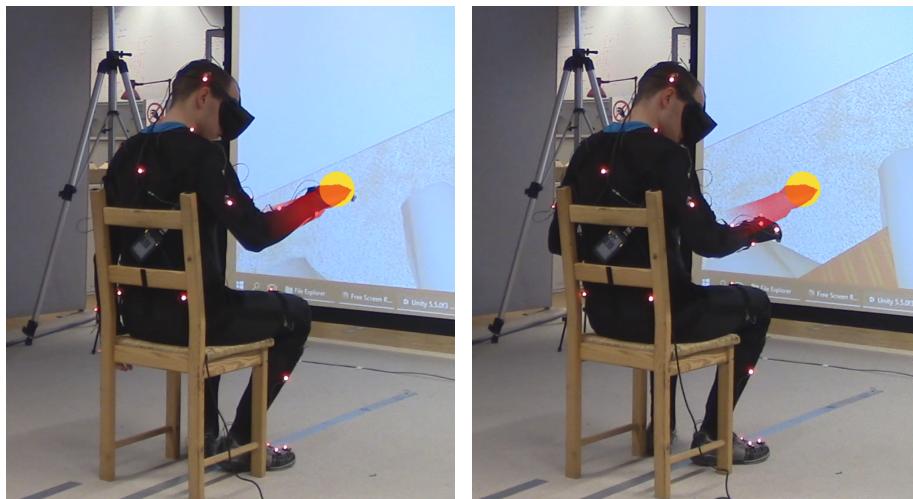


FIGURE 5.2: Pictures of a subject, taken from the same point of view, reaching for the same target but with two different distortions of 0 dB and 3 dB, with a visual representation of where the target and virtual hand are (yellow dot and red arm). The left image had no distortion and the right one a distortion of 3 dB.

This means that in order to perform the same perceived movement, the subject once had to cover the whole distance between his leg and the target, and could in a second time roughly travel two times less in order to reach the same target.

5.2 Experiment

The aforementioned web site also displays two videos, the first of which is summarizing the experiment process. As stated above, a second video shows some of the issues we encountered towards the end of the project, the resolution of which we will prioritize since these are the only reason we could not perform the experiment at the time of writing this report.

5.3 Software Issues

As stated at the beginning of this chapter, we encountered some issues during the end of this project. A video entitled ‘Motion Capture Issues’ is available on the website and summarizes said problems.

The first one is a strange torso movement that seem dependent on the angle of both elbows. The camera position is tracked through a different pipeline than the one causing the issue, which has two implications: it avoids any vestibular sensory discrepancy, which is known to cause cybersickness [31], but it also causes the torso to have an inconsistent position with respect to the subject’s point of view.

A second issue is that of calibration imprecision. Although much a much better calibration process than any previous method we used in the past, it still suffers some imprecisions. The radii of the capsules on each limb segment are for instance computed as the mean of the distance between the previously fitted bone of that limb and the two markers on that limb. For parts such as the thigh, which may sometimes have more of a conic shape, this means that there is slight interpenetrations near the knee and similarly tenuous hovering of the hand towards the hips. While not an issue at all in a typical use case, it becomes problematic when applying a positive distortion to the movements: such small imprecision becomes amplified and one sees much more clearly the error.

Additionally, we currently have an issue with the limbs performing a big jump at a few locations locations near the skin when a positive distortion is applied. For instance when moving outwards from the thigh with constant speed, the virtual hand first jumps a few centimeters before showing a constant speed as well. This issue will most likely cause the subjects to identify a distorted movement much earlier and thus biasing our experiment. We believe it is related to the overlapping of certain body parts, hence the narrow localization of the problem, and will need some further investigation in order to be solved.

6 Conclusion

We believe the content of this report is promising. The results obtained in terms of distortion behavior are satisfying and preliminary testing was very interesting and corroborates the findings of Debarba [9], stating that a helping distortion is more easily accepted than a hindering one.

While additional work is required before a complete understanding of this phenomenon occurs, we are confident that the second part of this project will lead to interesting findings and will eventually start a new era in stroke rehabilitation therapy.

We could for instance imagine a simplified calibration process involving only the upper limbs, along with a cheaper tracking solution such as the ones proposed by consumer-grade VR products. That would allow patients to train various movements at home and thus would dramatically reduce the cost of such therapy while allowing the clinicians to have remote access to detailed reports on their patients, with metrics like movement speed or distance covered with the affected limb. Such metrics may also be key in keeping the patients motivated, in that they can track the said measurements and see how they are improving.

6.1 Further Work

Before being able to issue a rehabilitation VR application, we must better understand to what extent the human perception allows such distortions to happen. This is going to be achieved through performing the experiment described in Chapter 4 as a pilot study, and then adapting the protocol to the results we will find. Points that may need refinement include the method and metric used (adaptive staircase and gain), the different target placements, and other experiment details such as the question asked.

A prerequisite to running the experiment however is to solve the issues exposed in Section 5.3. Priority will be given to fixing these. In parallel, a student's project will be focusing on the calibration process, which we hope to make faster and more accurate. If the project is successful, it will be included in our procedure.

On a wider scope, additional investigation will be required in order to fully understand the impact of this and other distortions on the Sense of Embodiment, and more precisely on its Agency component. More complex interactive experimentation should be conducted, and it might be interesting to assess the SoE through physiological measurements (e.g. galvanic skin response) while manipulating the strength of the distortion. Such measurement have already been carried out by Debarba et al. [32] in the context of perspective changes, and were found useful in assessing the differences between the conditions.

We hope that this subject will continue gaining interest and will involve more and more researchers, so that not only stroke patients but also VR in general can benefit from innovations in this field.

A Questionnaires

A.1 Characterization Questionnaire

TABLE A.1: The characterization questionnaire used at the begining of the experiment.

Question	Answer
Identifier	Numeric answer
Age	Numeric answer
Gender	Male / Female
How often do you participate on experiments using Virtual Reality equipments?	Never / A few times / Every month / week / day
How often do you use head mounted displays?	Never / A few times / Every month / week / day
How often do you play video games?	Never / A few times / Every month / week / day
How often do use the Microsoft Kinect, Nintendo Wii or Playstation move?	Never / A few times / Every month / week / day
Hand of preference?	Left / Right
Area(s) of expertise / study / work / interest?	Text answer
Are you a student?	Bachelor / Master / PhD / No

A.2 Detection Question

The detection question introduced in Chapter 4 that was asked multiple times to the subjects was the following : “Did the movements you saw exactly correspond to the movements you performed?”

Bibliography

- [1] T. Vos, R. M. Barber, B. Bell, A. Bertozzi-Villa, S. Biryukov, I. Bolliger, F. Charlson, A. Davis, L. Degenhardt, D. Dicker, *et al.*, “Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: A systematic analysis for the global burden of disease study 2013”, *The Lancet*, vol. 386, no. 9995, pp. 743–800, 2015.
- [2] R. Oden, “Systematic therapeutic exercises in the management of the paralyses in hemiplegia”, *Journal of the American Medical Association*, vol. 70, no. 12, pp. 828–833, 1918.
- [3] E. Flores, G. Tobon, E. Cavallaro, F. I. Cavallaro, J. C. Perry, and T. Keller, “Improving patient motivation in game development for motor deficit rehabilitation”, in *Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology*, ACM, 2008, pp. 381–384.
- [4] *Oxford dictionary of english*. Oxford: Oxford University Pres, 2015, ISBN: 9780191727665.
- [5] C. Cruz-Neira, D. J. Sandin, T. A. DeFanti, R. V. Kenyon, and J. C. Hart, “The cave: Audio visual experience automatic virtual environment”, *Communications of the ACM*, vol. 35, no. 6, pp. 64–73, 1992.
- [6] M. Slater, “A note on presence terminology”, *Presence connect*, vol. 3, no. 3, pp. 1–5, 2003.
- [7] M. V. Sanchez-Vives and M. Slater, “From presence to consciousness through virtual reality”, *Nature Reviews Neuroscience*, vol. 6, no. 4, pp. 332–339, 2005.
- [8] O. Blanke and T. Metzinger, “Full-body illusions and minimal phenomenal selfhood”, *Trends in cognitive sciences*, vol. 13, no. 1, pp. 7–13, 2009.
- [9] H. Galvan Debarba, “Embodiment sensitivity to movement distortion and perspective taking in virtual reality”, PhD thesis, Ecole Polytechnique Fédérale de Lausanne, 2017.
- [10] F. De Vignemont, “Embodiment, ownership and disownership”, *Consciousness and cognition*, vol. 20, no. 1, pp. 82–93, 2011.
- [11] M. Tsakiris, G. Prabhu, and P. Haggard, “Having a body versus moving your body: How agency structures body-ownership”, *Consciousness and cognition*, vol. 15, no. 2, pp. 423–432, 2006.
- [12] K. Kilteni, R. Grotens, and M. Slater, “The sense of embodiment in virtual reality”, *Presence: Teleoperators and Virtual Environments*, vol. 21, no. 4, pp. 373–387, 2012.
- [13] R. M. Held and N. I. Durlach, “Telepresence”, *Presence: Teleoperators & Virtual Environments*, vol. 1, no. 1, pp. 109–112, 1992.
- [14] M. Slater and M. Usoh, “Representations systems, perceptual position, and presence in immersive virtual environments”, *Presence: Teleoperators & Virtual Environments*, vol. 2, no. 3, pp. 221–233, 1993.
- [15] E. Molla, H. Galvan Debarba, and R. Boulic, “Egocentric mapping of body surface constraints”, 2017.
- [16] D. Djaouti, J. Alvarez, and J.-P. Jessel, “Classifying serious games: The g/p/s model”, *Handbook of research on improving learning and motivation through educational games: Multidisciplinary approaches*, pp. 118–136, 2011.
- [17] S. Chen and D. Michael, “Proof of learning: Assessment in serious games”, *Retrieved October*, vol. 17, p. 2008, 2005.

- [18] R. P. Paul, *Robot manipulators: Mathematics, programming, and control: The computer control of robot manipulators*. Richard Paul, 1981.
- [19] A. Goldenberg, B. Benhabib, and R. Fenton, “A complete generalized solution to the inverse kinematics of robots”, *IEEE Journal on Robotics and Automation*, vol. 1, no. 1, pp. 14–20, 1985.
- [20] D. Manocha and J. F. Canny, “Efficient inverse kinematics for general 6r manipulators”, *IEEE transactions on robotics and automation*, vol. 10, no. 5, pp. 648–657, 1994.
- [21] E. Molla and R. Boulic, “Singularity free parametrization of human limbs”, in *Proceedings of Motion on Games*, ACM, 2013, pp. 187–196.
- [22] N. West, *Motion capture: Magnetic systems*, ser. Next Generation. Imagine Media, 1995.
- [23] R. A. Al-Asqhar, T. Komura, and M. G. Choi, “Relationship descriptors for interactive motion adaptation”, in *Proceedings of the 12th ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, ACM, 2013, pp. 45–53.
- [24] E. Molla, “Precise and responsive performance animation for embodied immersive interactions”, PhD thesis, Ecole Polytechnique Fédérale de Lausanne, 2016.
- [25] A. Aristidou and J. Lasenby, “Fabrik: A fast, iterative solver for the inverse kinematics problem”, *Graphical Models*, vol. 73, no. 5, pp. 243–260, 2011.
- [26] P. T. Inc. (2017). Collaborative data science, [Online]. Available: <https://plot.ly>.
- [27] *Ieee 100 : The authoritative dictionary of ieee standards terms*. New York: Standards Information Network, IEEE Press, 2000, ISBN: 0738126012.
- [28] J. Khoury, *Human perception of guided interaction*, Semester project at IIG., 2015.
- [29] T. Meese, “Using the standard staircase to measure the point of subjective equality: A guide based on computer simulations”, *Perception & Psychophysics*, vol. 57, no. 3, pp. 267–281, 1995.
- [30] E. Burns, S. Razzaque, M. C. Whitton, and F. P. Brooks Jr, *Macbeth: Management of avatar conflict by employment of a technique hybrid*, 04. 2007, vol. 68.
- [31] J. J. LaViola Jr, “A discussion of cybersickness in virtual environments”, *ACM SIGCHI Bulletin*, vol. 32, no. 1, pp. 47–56, 2000.
- [32] H. Galvan Debarba, S. Bovet, R. Salomon, B. Herbelin, and R. Boulic, “Perspective taking and multisensory congruency for embodied interaction in virtual reality”, 2017.