Bachelor's thesis

Information and Communications Technology

2019

Joonas Törmänen

COMPARISON OF ENTRY LEVEL MOTION CAPTURE SUITS AIMED AT INDIE GAME PRODUCTION

TURKU AMK
TURKU UNIVERSITY OF
APPLIED SCIENCES

BACHELOR’S THESIS | ABSTRACT

TURKU UNIVERSITY OF APPLIED SCIENCES

Information and Communications Technology

2019 | 49 pages

Joonas Törmänen

COMPARISON OF ENTRY LEVEL MOTION CAPTURE SUITS AIMED AT INDIE GAME PRODUCTION

Motion capture is the digital process of translating human movement into other media such as video games, movies, sports and medical applications. Historically it has been expensive, but motion capture equipment has advanced enormously while the cost has gone down, bringing the power of motion capture to the masses.

The purpose of the thesis was to acquire a better understanding of how motion capture can be a powerful tool for character animations in game development. The thesis gives a general overview of what motion capture is, its development through history and what kind of challenges it presents. In addition, currently used technologies are described with their respective pros and cons. The source material for theory was collected from written and electronic sources.

The studied theories were used to compare two different motion capture solutions targeted at the game development indie market; Rokoko Smartsuit Pro and Perception Neuron V1 suit for Turku Game Lab. Three aspects were compared: usability, performance and accuracy. The comparison was achieved by recording identical animations using both suits. The goal was to study these two motion capture systems to find out which one is more suitable for the Turku Game Lab’s production workflow for creating animations that could be implemented and easily used in all of the company’s game projects.

Based on the findings of this study, it could be concluded that although the suits provided very similar and promising results in resulting animations, the other compared aspects provided very different results that must be taken into consideration when choosing either of the compared motion capture solutions.

KEYWORDS:

Game development, Motion capture, Rokoko Smartsuit, Perception Neuron

OPINNÄYTETYÖ (AMK) | TIIVISTELMÄ

TURUN AMMATTIKORKEAKOULU

Tieto – ja viestintätekniikka

2019 | 49 sivua

Joonas Törmänen

INDIE-PELITUOTANTOON SUUNNATTUJEN LIIKKEENKAAPPAUSPUKUJEN VERTAILU

Liikkeenkaappaus on prosessi, jossa kohteen liikkeet muunnetaan digitaaliseen muotoon. Sitä voidaan käyttää mm. videopeleihin, elokuviin, urheiluun ja lääketieteellisiin sovelluksiin. Historiallisesti se on ollut pitkään kallista, mutta koska kaappausteknologiat ovat edistyneet huomattavasti samalla kun kustannukset ovat laskeneet, on yhä useammalla mahdollisuudet tekniikan käyttöön.

Opinnäytetyön tarkoituksena oli saada parempi käsitys siitä, kuinka liikkeenkaappausta voidaan hyödyntää hahmoanimaatioihin osana pelinkehitystä, sekä antaa yleiskuva liikkeenkaappauksesta ja sen kehityksestä sekä millaisia haasteita se tuo mukanaan. Lisäksi opinnäytetyö käsittelee tällä hetkellä käytössä olevia yleisempiä liikekaappaustekniikoita, sekä niiden etuja että haittoja. Teoriaosuuden lähdemateriaaleina käytettiin kirjallisuutta ja sähköisiä lähteitä.

Tutkittuja teorioita hyödynnettiin vertailussa kahden eri indie-pelikehitysmarkkinoille suunnatun liikekaappausjärjestelmän välillä: Rokoko Smartsuit Pro ja Perception Neuron V1. Vertailussa tarkasteltiin kolmea eri näkökohtaa: pukujen helppokäyttöisyyttä, suorituskykyä ja niiden tuottaman animaation laatua. Vertailu suoritettiin tallentamalla eri animaatioita identtisissä olosuhteissa. Työn tavoitteena oli tutkia ja selvittää kumpi puvuista soveltuisi paremmin Turku Game Labin pelikehitysprosessiin ja voitaisiin käyttää yrityksen muissa hahmoanimaatioita vaativissa projekteissa.

Tulosten perusteella voitiin päätellä, että vaikka puvuilla tuotetut animaatiot olivat oikein lupaavia ja samanlaisia, muut vertailut aspektit tuottivat erilaisia tuloksia, jotka on syytä ottaa huomioon tehtäessä valintaa työssä käytettyjen liikkeenkaappausjärjestelmien välillä.

ASIASANAT:

Pelinkehitys, Liikkeenkaappaus, Rokoko Smartsuit, Perception Neuron

CONTENT

PICTURES

TABLES

LIST OF ABBREVIATIONS (OR) SYMBOLS

Abbreviation	Explanation of abbreviation (Source)
2D	Two-dimensional.
3D	Tree-dimensional.
A-pose	A neutral pose where 3D character or actor is standing upright, looking directly ahead with their arms out on the side pointing down in a ‘’A’’ shaped pose.
Blockbuster	A product that is a great commercial success, usually a movie or a (video)game.
BVH	A short form for Biovision Hierarchy, which is a file format used for motion capture data with character animations.
FBX	A short from for Filmbox, which is a file format used to provide interoperability between different digital content creation applications.
IMU	Inertial Measurement Unit is an electronic device that measures linear and angular motion using gyroscopes and accelerometers.
Indie	A short form of independent. Usually refers to a company not belonging or affiliated to a major company or publisher.
Keyframing	Keyframing is the form of animating an object using key frames. Key frame is a mark on a timeline used in animation and filmmaking for holding special information defining where a transition should start or stop.
Marker	An instrument used in optical motion capture system, which allows you to capture the movements of a moving object as 3D motion data.

Potentiometer	A analogue device that can be found and used with electrical and electronic circuits that forms an adjustable voltage divider.
SDK	Software development kit is a set of software development tools used for creating applications.
S-pose	A neutral pose where the 3D character or actor is crouching down, looking directly ahead with their arms out on the front shoulder height pointing away in a ‘’S’’ shaped pose.
T-pose	A neutral pose where the 3D character or actor is standing upright, looking directly ahead with their arms out to the side in a ‘’T’’ shaped pose.
Triple A-game	A classification used for videogames typically having higher development and marketing budgets.
USB	Universal Serial Bus is an industry standard that established specifications for connections and cables between devices.
Virtual reality	A simulated three-dimensional, computer generated environment which can be interacted with by a real person.
Virtual set	A virtual reality environment that allows the real-time combination of people and computer-generated object interactions.
VR	A short form for Virtual Reality.
WiFi	Wireless Internet standard.

1 INTRODUCTION

Due to increasing amount of platforms, and free-to-use game engines, the game industry has become accessible to everyone, resulting in a more populated and competitive field. The number of people employed by the Finnish Game Industry is on the rise and the games are made in larger volumes, as shown in the 2018 report on the Finnish game industry conducted by Neogames (Neogames, 2019). To keep up with the competition, the game developers need to stay efficient while maintaining a high level of professionalism.

As the overall quality in graphics and visuals are also on a constant state of growth, the players and customers expect to see animations with matching quality. If a video game character can not properly mimic realistic body movements or facial expression, much of the immersion is lost (Mattson & Mårtensson, 2014). The traditional digital animation methods are not keeping up with the standards of the industry, providing a challenge for developers to stay relevant but also efficient. This has increased the demand and popularity of motion capture.

Motion capture (sometimes referred to as mocap for short) is the process of recording and digitizing motion as 3D data. It allows to maintain the immersion of realism by copying the movements of the human body, without the need to animate by hand, instead using a combination of cameras, sensors and software to accurately record patterns of body movement. In this thesis, the generic terms “motion capture’’ and ‘’mocap’’ will be used for the sake of simplicity.

Motion capture has traditionally been expensive and hence mostly been used by large companies for triple-A games like ‘’Call of Duty’’ (an on-going first-person shooter game series released by Activision) and blockbuster movies like ‘’Avatar’’ (2009). Having access to such professional systems demands money that the small game studios and indie developers can not afford, which puts them in a problematic situation in the competition to produce the ‘’coolest game’’ possible (Mattson & Mårtensson, 2014; Conditt, 2018).

Due to advances in technology, as well as the growing demand for motion capture, the price is quickly dropping. It is a needed shift in the industry because the amount of animations needed to bring even one character to life is staggering, making the traditional key-framing methods hardly an option anymore. Mocap aims to provide quick and realistic animations, adapting to a variety of budgets in order to animate various digital characters. (Huuhtanen, 2018).

As markerless sensor-based mocap solutions, targeted to the masses, have become available, it is important to research how much value they bring to the actual production pipeline. If budget friendly motion capture can be adopted by small independent game development studios and research facilities, it could give them a competitive edge in the market.

The practical part of the thesis was to become familiar with how mocap can be used to streamline animation production for video games by comparing two sensor-based mocap systems: the Rokoko Smartsuit Pro (Rokoko, 2019a) and Perception Neuron V1 (Perception Neuron, 2019). It examines the preparation, setup, recording and editing of motion data as well as animation cleanup in MotionBuilder using the two suits. The goal was to find out which one of the compared mocap suits should be used for the Turku Game Lab’s future projects as part of their animation workflow. The results from the comparison should help the reader, especially those who are involved in 3D animations and games, to have a better understanding of motion capture and the tested suits before investing into (sensor-based) mocap solutions.

The thesis is not a tutorial covering the entire set up process step-by-step but rather gives an overview of an author’s first-time experience with setting up, testing the suits and figuring out the ups and downs of the both mocap systems. Knowing the fundamentals of setting up and using animations in both MotionBuilder and Unity are expected. For the project, information was gathered from various sources, such as online videos, publications and from the process of testing the mocap suits for the Turku Game Lab.

The thesis begins with a theory section, describing basic concepts of motion capture, its history and the possibilities and benefits. It is followed by the method section, where a detailed description of the thesis case is given, as well as the tools that were used to accomplish the work. This section also discusses the testing strategy this study uses to answer the research question. The third main part of the thesis maps out the work done. Finally, a conclusion is drawn from the key findings of the research and the commissioner’s project identifying some possible future research needs.

2 BASIC THEORY OF MOTION CAPTURE

In his book ‘Understanding Motion Capture for Computer Animation’, Alberto Menanche defines motion capture as ‘’the process of recording a live motion event and translating it into usable mathematical terms by tracking a number of key points in space over time and combining them to obtain a single three-dimensional (3D) representation of the performance. In brief, it is the technology that enables the process of translating a live performance into a digital performance’’ (Menanche, 2010).

Motion capture itself can be divided broadly to areas of body movement, facial capture and hand gestures. Capturing the more subtle movements of face and hands require special capture systems that are more commonly used in professional productions to make the 3D characters more lifelike. A combination of both facial and full-body motion capture is usually referred as ‘’performance capture‘’ (Dent, 2014). The scope of this thesis is on the low-budget production and hence it will focus only on the motion capture of body movements and the generic terms ’’motion capture’’ and ‘’mocap’’ will be used to reference to body movements.

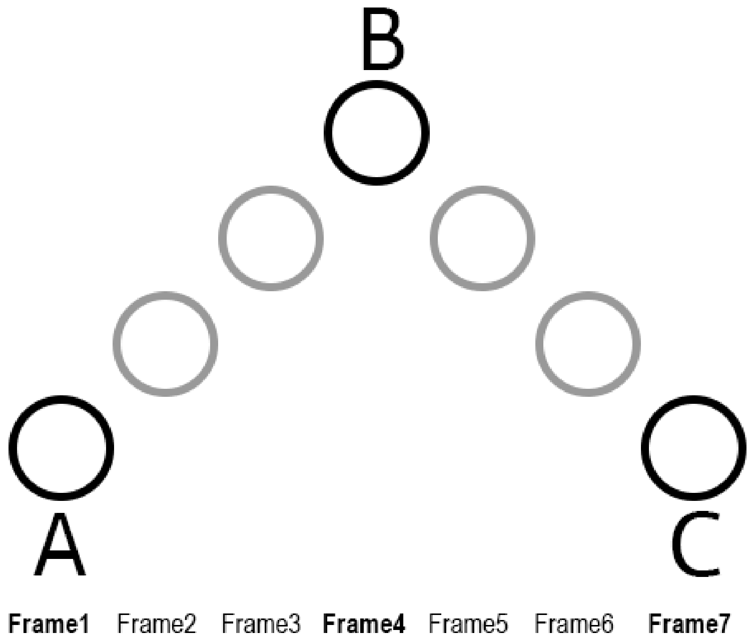
Motion capture systems that use tracking cameras (with or without markers) can be referred to as optical, while systems that measure inertia or mechanical motion are non-optical. Optical systems are the popular form of motion capture, as it generally provides greater results compared to other systems. However, it comes with some important limitations, such as a higher cost due to the need for a specialized studio with fixed equipment, which is why it’s usually related to high-profile movie and game studios. Non-optical motion capture is not yet as accurate as high-end optical systems, but users are typically satisfied with the results, increasing its popularity among those lacking massive budgets (NoFilmSchool, 2019).

Mocap might be a relatively inexpensive tool and is often thought to be a quick and simple solution. However, the actual motion capturing is only part of the whole process. Before even starting the capturing session, you need to plan and decide what and how you are going to capture the motion data. Once the capturing is done, mocap data needs to be processed and cleaned before it can be applied to a virtual 3D character and integrated into your own production pipeline. Greater productivity should not be expected immediately as the key to successful mocap requires establishing a production pipeline that suits your needs and environment (Mocappys, 2019a).

2.1 General pros and cons

Although motion capture might be thought of as a time saver, it is not going to replace all animation. Mocap’s suitability over other animation methods needs to be assessed on a per project basis. It also must be noted that each mocap system has its own strengths as well as weaknesses. There is not one system suited best for all mocap scenarios. The following is a summary of general pros and cons of mocap. More system specific pros and cons will be given in the ‘motion capture technologies’ section.

Creating digital animation is normally done by hand and is known as keyframe animation or ‘’keyframing’’, which refers to the process of filling in the movement of a character or an object between different initial starting and ending poses frame by a frame to create the illusion of the motion. Animator moves the object from point A to point B to point C and computer calculates the movement between these given points and automatically generates the intermediate transition frames (Picture 1) (Computer Hope, 2017). However, some animations were next to impossible to create accurately due their inherent complexity, for example the human walking animation. Walking is very easy for people, but making it look natural and fluid as an animation can be complicated (Rahul, 2018).



Picture 1. Example of a simple keyframe animation.

The biggest advantages motion capture offers, compared to keyframe animation, are production speed and realistic results. Motion capture aims to automate the process of creating the intermediate/key frames, saving hours of painstaking manual animation done by hand. Better motion capture technology means greater production capabilities as less time is spent animating and more time is left for fine-tuning other parts of the production, like game play or aesthetics (Conditt, 2018). With most mocap systems, the visual feedback can be produced in real-time, without needing to first record and then review the data. This allows the motion capture actor, or the director, to see how the digital 3D character moves on a virtual set as the actor moves, increasing the efficiency of mocap (Dent, 2014).

Other great advantage motion capture has over keyframe animations is that it only records the movement, so it can be applied to almost any shaped or sized virtual 3D model (Xsens, 2019a). It is also not limited to only humans, other live subjects like animals or static objects like puppets can also be used for motion capture purposes. A recent example of this is the dog mocap experiment done by Finnish game company Remedy Entertainment (Starcevic, 2018).

Some of the major disadvantages are the sensitivity of the motion capture, required hardware, their price and availability. Motion data is often not ready to be used ‘’out of the box’’, as it might contain unwanted movement that needs to be cleaned up manually in post-production work. Even with the clean-up of data the result may also not be what was expected.

Even though mocap might seem superior, keyframe animation isn’t going anywhere. Both techniques can be blended, allowing great control of both style and quality of the final output. For example, using mocap for body movement and keyframing for other animations, such as facial and finger movements (Kines, 2000).

2.2 A brief history of modern motion capture

The idea of copying human motion for animated characters is nothing new: the different systems are a product of a century of innovation and experimentation (Conditt, 2018). Even before computer technology became available, the invention of zoopraxiscope by Eadweard Muybridge (1830-1904) and its later development by Etienne-Jules-Marey (1830-1904) laid the groundwork for motion studies (Kitagawa & Windsor 2008, p.5).

In 1914, animator Max Fleischer (1883-1972) invented a technique known as rotoscoping, a method of producing animations like ‘’Out of The Inkwell’’ series (1918-1929), by tracing film footage of live actors’ frame by frame. It was first used in a feature film in Disney's (1901-1966) movie ‘’Snow White and the Seven Dwarfs’’ (1937) to copy the real people’s actions as closely as possible, making the animation feel fluid and more alive. These early inventions and pioneers can be thought of as a precursor for the mocap technology we can see in modern movies and games (Conditt, 2018).

Mocap was taken to a digital form as we think of it today in the late 1970’s, although it was mainly being used and developed by the health and military industry. Bio-kinetic researcher Tom Calver was experimenting with early mechanical capture suits using potentiometers for choreographic studies. In the early days, cleaning mocap data often took as long as if the animation was created by an animator by hand. With rapidly developing computers, mocap technology could really start to bloom and make its way into the entertainment industry in 90s (Conditt, 2018).

In the 2000s, motion capture technology has been developing rapidly and the exclusivity of big studios to mocap technology has subsided, opening brand new worlds for motion capture applications. Games like ‘’L.A.Noire’’ (2011), ‘’Hellblade: Senua’s Sacrifice’ ‘(2016) and movies like ‘’Lord of The Rings’’ (2001) and ‘’Dawn of the Planet of the Apes’’ (2014) each played their part by improving the realism further by polishing the use of the motion capture in their own ways (Dent, 2014; Conditt, 2018).

2.3 Motion capture technologies

Nowadays there is large variety of motion tracking technologies available as technology has developed and different disciplines of motion capture have emerged (Failes, 2019). The systems presented in the thesis include optical, inertial, mechanical and magnetic systems. This thesis won’t go deep under the hood in technicalities, such as math and the algorithms behind each system, but rather provides an overview of them with their pros and cons.

2.3.1 Mechanical motion capture systems

Mechanical motion capture systems are typically based on some physical mechanical device that consists of metallic pieces and potentiometers, which measures the angles or lengths between each part of the system using sensors. These systems can be worn on the body, for instance using sensors attached to an exoskeleton (Picture 2). Other types of mechanical motion capture systems include digital armatures and gloves (Kitagawa & Windsor 2008, p.11). As the capture subject moves, the exoskeleton is moving along, and the sensors track the rotations. The actual body posture is determined by using the angle data from these sensors with kinematic algorithms (Xsens, 2019a).

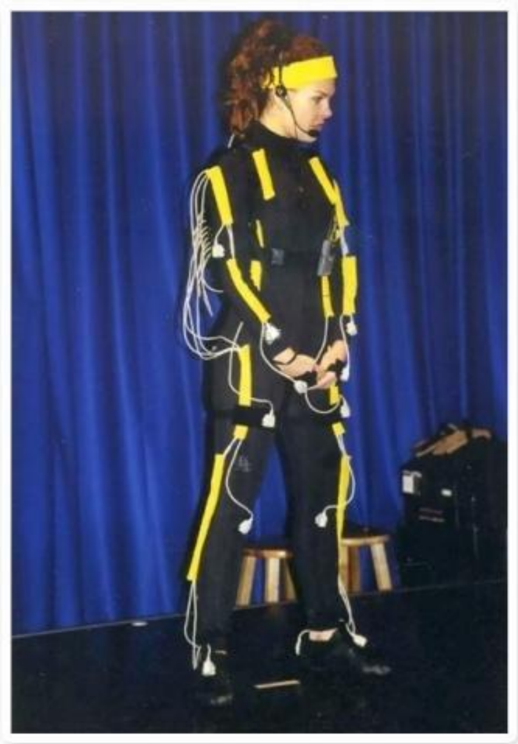
Picture 2. An actor wearing a mechanical mocap suit (Metamotion, 2019c).

This technique offers high precision and is not influenced by external factors such as cameras to track markers or the movement of the subject. They are also free from occlusion or magnetic and electrical interference problems. Other advantages include its relatively inexpensive price, highly portable nature and the possibility of capturing multiple performances simultaneously (Rahul, 2018).

One of the biggest disadvantages of this system is its limited freedom of movement. Even something like jumping can be a difficult animation to capture, since the exoskeletons are usually heavy and use wired connections to connect the sensors with the computers (Rahul, 2018). Attachment and positioning of these body-based linkages provided challenges and the systems must be calibrated often. As the subject moves, the position of the linkages relative to the body might change due to the soft tissue of human body. Also, the positioning and correct alignment of the sensors with body joints is difficult and capture of subject's movement is restricted at their joints. This is usually fixed by adding an offset to each sensor or adding the magnetic sensors to the mechanical motion capture system (Xsens, 2019a).

2.3.2 Magnetic motion capture systems

Magnetic motion capture systems use the magnetic field around its sensors and can measure the direction and strength of the surrounding magnetic field. With this technique, usually 12 to 20 tracking sensors are placed on the subject to measure the low-frequency magnetic fields generated by the transmitter source (Picture 3). More advanced magnetic systems use an active electromagnetic source and a sensor with multiple coils (Kitagawa & Windsor 2008, p.10).



Picture 3. A actor wearing a magnetic mocap system (Jáuregui, 2011).

Magnetic motion systems don’t suffer from the line-of-sight problems (occlusion) like optical motion mocap does, because the human body is transparent for the magnetic fields. This means that the sensors can be hidden under the clothes or inside the mocap suit (Nymoen, 2014). Other advantage of this technique is that the motion data is highly accurate and no further calculations to compute the positions and rotations are needed, so it can be used for real-time applications. (Rahul, 2018).

However, this system also has weaknesses as the system is sensitive to disturbances from magnetic objects, like metal or electrical devices, in the tracking area (Nymoen, 2014). This can cause distortions in the motion capture data, making it noisier and not as good as optical mocap. The magnetic fields used in this system also decrease in power fast as the distance from the generating source increases. (Xsens, 2019a). The performer needs to be connected to a computer via wiring and the need to recharge or replace batteries may also limit the freedom of movement of the subject (Kitagawa & Windsor 2008, p.11).

2.3.3 Inertial motion capture systems

Inertial motion capture systems use Inertial Measurement Unit (IMU) sensors to track and capture motion capture data based on accelerometers, magnetometers and gyroscopes attached to the performer (Picture 4). The sensors record location, acceration and orientation as movement data. Recorded mocap data is transmitted wirelessly to the receiver and from there to the computer using WiFi (Xsens, 2019a).

Picture 4. Performer wearing full-body (left) and strap-based (right) inertial mocap suits (Xsens, 2019c).

Compared to the other motion capture solutions, inertial motion capture systems have a strong advantage. First of all, they are completely self-contained and do not rely on external sources, such as markers or cameras which require line-of-sight. They are also not affected by poor lighting conditions. The sensors are very small and lightweight, meaning they can be used easily in portable devices. Finally, these systems have low latencies so they can be sampled at very high rates, allowing them to be used in real-time in the most popular 3D software programs and applications (Nymoen, 2014).

Inertial systems are not as precise or accurate as other systems, making it impractical to track fine and high-speed movements where accuracy is crucial with all their subtleties. They are also magnetically sensitive to external sources that cause noise and other errors in the recorded data, which means more cleanup will have to be done in post-production. Since every character is usually captured separately, capturing interactions between two or more actors is another limitation of this technology, to name just a few (CGIcoffee, 2017).

2.3.4 Optical motion capture systems.

Optical motion capture systems are mostly based on video cameras and computer vision algorithms that track the motion in 3D space. The cameras used can be either regular video cameras (like a simple web cam), infrared (IR) cameras or depth cameras (like Microsoft Kinect). These optical motion capture systems can be split into a different collection of technologies that range widely in terms of quality and cost. A common distinguish between the optical systems are those that use markers and those that do not, called marker-less (Nymoen, 2014).

Marker-based optical systems use the combination of multiple cameras to track predetermined points (markers), aligned with specific body landmarks where movement occurs in order to determine position (Picture 5). The markers can be either passive (reflective) or active (light emitting) (Nymoen, 2014).

Picture 5. Performer wearing optical motion capture suit (Unreal Engine, 2018).

Marker-less optical systems use images from multiple cameras that track RGB color to evaluate the position of the performer. Rather than using optical infrared cameras to track the markers on the actors' suit, the cameras recognize the color of different body parts of the actor trying to match the silhouette of these colors from all angles using computer algorithms (Swimbanks, 2019).

One of the advantages of optical systems is the freedom of movements of the performers in the capture volume as no equipment or wires are connected to them. This makes larger volume of mocap possible as you can capture multiple actors in larger area at the same time. With marker-based systems anything that can be marked with the markers can be captured, since the system only captures the data coming from the markers (Kitagawa & Windsor 2008, p.10).

However, optical systems notoriously struggle with line-of-sight problems (also known as occlusion), meaning that the camera can't see the marker. This is especially problematic when one performer blocks the markers on another performer's body. (Kitagawa & Windsor 2008, p.9). Other downsides include limited speed at which each marker can be examined, and data forwarded to the computer, affecting the number of positions per second that can be captured. These problems can result in the loss of mocap data, so in worst case the performance might have to be retaken due to useless data (Nymoen, 2014).

2.4 Motion capture applications

Historically, the motion capture technology has been widely used in the medical and entertainment industries but is not limited only to movies and games and offers many other opportunities with applications in various industries. For example, motion capture can be utilized in education, live entertainment, and advertising to name a few. Together these different fields and industries are feeding each other and opening new opportunities by developing motion capture even further.

Movies and TV

Movie industry is still one of the biggest industries using different motion capture techniques. Visual effect studios, like Industrial Light & Magic and Weta Digital among others use motion capture for realtime animations for feature films and television. For some film characters the quality of the animation is especially crucial. When the animation movement is almost, but not quite humanlike, it feels fake which risks alienating the audience. This breaking of immersion can be countered by using motion capture and real human movements (Dent, 2014).

There are many reasons to use mocap for movies and television. It can be used for animating background motions such as scenes with massive crowds or providing animators with a foundation from which to continue the work by hand. Sometimes scenes can be impractical or too dangerous to capture with real actors in its original environment, so motion capture is used to for safety reasons (Metamotion, 2019a; Özdem, 2018).

Health and sports

Motion capture has been used in medical research for centuries to study human trajectories, mobility and body behavior. One of the main areas of motion capture is biomechanical analysis, where mocap is used to analyze how a person performs their movements to identify points in the body where any strain is occurring. This can be used for example to verify the condition of a patient who requires a prosthetic or who is rehabilitating from a stroke (Rahul, 2018).

Like in the health industry, motion capture can be used in sports to analyze the performance of athletes. It is used to analyze the movement to improve performance or prevent injuries of athletes (Rahul, 2018). In the past years the health care systems have been trending towards monitoring and training patients in their own environments, which has promoted for a large leap in development of portable and wearable mocap systems (Xsens, 2019a).

Virtual Reality

A relatively new field is the use of motion capture for virtual reality (VR) applications, where realtime tracking of 3D characters is very essential for realistic user experience (Xsens, 2019a). Adding motion capture will increase realism and immersion by removing the need for physical input devices like typical game pads or joysticks. For example, Xsens provides a fully wireless full-body motion capture system that allows the user to experience the full freedom of movement in VR (Xsens, 2019b).

Game Industry

There are generally two main types of animation used in games for 3D characters: Real-time and cinematics. With real-time the player can choose from pre-created movement options how to control the game characters. Cinematics on the other hand are fully rendered video sequences used mostly for intro and cut-scenes in the games (Metamotion, 2019a).

The main reason to use motion capture systems for video games is that they provide an easily achieved realistic results without spending hours on manual animation. As mocap process has become easier, faster and more efficient, less time it spent animating. This allows the developers to focus on other parts of the production with the extra time they have (Conditt, 2018).

Big studios like, Electronic Arts, Sony and Ubisoft have their dedicated motion capture studios to support their own development processes. These studios have been used for popular game series like FIFA, NBA, Battlefield, Assassin Creed and Beyond Two souls to create realistic animations. Because of the demand for realistic graphics and animations, and motion capture’s contribution to speeding up the production, it has great importance to the game industry (Özdem, 2018).

2.5 Affordability versus accuracy

There are a lot of companies providing different mocap solutions for different price points. ranging from $100 to $100,000 and upwards. Therefore, a cost-benefit analysis is a must before investing in motion capture.

A completed ‘’high-end’’ optical motion capture system, such as OptiTrack, can cost anything between $8 000 to $600 000 and beyond depending on the target application and amount of the hardware (OptiTrak, 2019). Companies that make and provide these types of professional systems can charge thousands of dollars just for one session, not to mention the additional cost of post-processing the motion data. For example, a New York based motion capture studio motionCAPTUREnyc lists $4,000 a day plus $20 a second for solving and re-targeting as a base cost for the motion capture work (motionCAPTUREnyc, 2019). A typical optical system used in the entertainment industry consists of a range starting from 4 and up to 32 cameras connected to the computers, which is hardly an option for those lacking massive budgets or pristine studio setups.

On the low-end of the scale for motion capture, you can do marker-less motion capture with solutions like iPi Motion Capture Software (iPi Soft LLC.) costing $195 to $1 995 (depending of the version). This is only the software, so you need to buy a camera device like Microsoft Kinect, costing $159 to $599 (depending of the version) by Microsoft or Sony Playstation Eye by Sony for just $6 per camera (iPi soft, 2019). These systems use depth sensors instead of conventional video cameras to capture motion data (Mattson & Mårtensson, 2014).

However, the biggest drawback of these low-end systems is the difficulties of getting correct head or hand tracking with fingers (Swimbanks, 2019). Low-end hardware might be also cheaper to start with, but variables like volume and target quality stack up the price quickly. In the price one must also factor in the time-consuming manual tasks like setting up the cameras properly, plenty of post-production work, as well as the possibility of needing to have multiple recording sessions to achieve good results, not to mention the requirement for large enough and well-lit space to do and capture the performance itself (CGIcoffee, 2017). These limitations and drawbacks in both high and low-end optical systems have contributed to the development of sensor-based mocap technologies.

The advantage of sensor-based systems is very portable and easy to set up making them very versatile. They only require a computer, the mocap suit and a power bank to power up the sensors to be able to capture animation. In the smallest set-up you only need a relatively magnetic-free environment and a WiFi connection enabling them to be used almost anywhere, allowing the users to work in a variety of locations and lighting conditions. Recorded mocap data can be streamed in real-time to game engines and animation tools. Some systems can even capture basic finger animations and save the mocap data directly to a memory card, making it useful when in a location with poor WiFi connection (CGIcoffee, 2017).

Disadvantage of these sensor-based systems are the lack absolute positioning, making them less accurate as the camera-based systems that provide more precise positional accuracy with the tracking. Because of this the recorded animations might require polishing and cleanup in post. Essentially, what you gain in saving money you lose in the exact positional accuracy of the mocap data (NoFilmSchool, 2019).

Other major downside is the challenge with multi character interactions, since every character is captured separately, and the interactions are prone to the same limitations with absolute positioning with movements such as clapping hands or folding arms. For example, in a fighting scene the user must re-enact both sides, unless they have an additional suit (CGIcoffee, 2017). Even with the downsides, their accessibly and the possibilities to the masses makes them suitable alternatives between the expensive and cheap mocap solutions for those lacking resources (NoFilmSchool, 2019).

Although there are some websites that give the pros and cons for the different price range suits, there is no real user-claimed evidence about how much better the motion capture is with an expensive, Rokoko Smart Suit Pro (2 500$) vs. a cheaper Perception Neuron V1 suit (1 500$) (NoFilmSchool, 2019; CGIcoffee, 2017). This thesis aims to help indie developers decide whether either of these two sensor-based systems are the right tool for you to get the most bang for the buck when using mocap in your animation workflow.

3 METHOD

3.1 Case

The practical part of the thesis was carried out as a project for Turku Game Lab. The aim was to study how motion capture could be used as a tool for animation workflow for the purposes of improving Turku Game Lab game development pipelines. This was done by producing animations using two sensor-based mocap suits of different price levels and comparing them.

The Game Lab had previously acquired Rokoko's motion capture suit but were inexperienced with motion capture. During the background research, it appeared that Finnish Virtual Reality Association's office in Turku also had a motion capture suit, which provided an opportunity to test two different sensor-based suits.

Turku Game Lab is a joint working environment of University of Turku and Turku University of Applied Sciences founded in 2011. The Game Lab functions as a multidisciplinary entity where students can develop games together and provides services in game education and development specializing in serious games (Turkugamelab, 2019). Finnish Virtual Reality Association (FIVR) is a non-profit organization for both professionals and hobbyist of virtual reality dedicated to advancing the state of virtual reality development in Finland. (Finnish Virtual Reality Association, 2019).

In order to have enough qualitative comparisons between the suits, they would need to be compared on common game animations, that one could expect to see in a video game. Two such animations that had different complexity of movements were selected for the project.

- A common idle action that starts with folding of arms and goes into a seated position with legs crossed.
- A plain walk that transforms into a run and jump.

The points that were selected for comparison based on discussions with the commissioner are:

1) Usability, which is divided into two categories:

   a. Setup: how difficult it is to get the system up and running "straight out of the box".
   b. Application: how easy it is to use the system for doing mocap recordings, after everything is set up.
2) Performance: How frequently mocap data is lost and therefore how much post-production work is required.
3) Accuracy: How precise the tracking is with both fast and slow movements when the user is doing different motions and body positions.

To make the comparison as fair as possible and to rule out most of the variables of the recording sessions, both mocap suits were used in same recording environment with the same circumstances to produce as identical animation sequences as possible. The process was documented by keeping a diary of experiences with both mocap systems as working through the production of the animations.

3.2 Tools

To perform this project, several different softwares and products was used. Although the used software were up-to-date, both of the used systems were older generations and at the time of writing this thesis both manufacturers were already providing improved versions of their mocap suits. Because of this it’s important to keep in mind that some of the findings of this thesis might have already been fixed in the more up-to-date versions of these suits.

There are also some common problems that might occur with sensor-based mocap systems, such as sensitivity to magnetic interference that may affect the performance of the systems. The selected studio environment for recording contained computers and other electrical devices, so some errors in recorded data were expected.

Additionally, Perception Neuron mocap suit has the benefit of finger motion recording and the ability to store motion data onboard using the built-in micro-SD slot, but for this study those was not used as the goal was to produce animations as identical as possible, with the same conditions applied for fair comparison. Even though these are not taken into account with animations, they are considered relevant factors when making recommendations in the final findings of the thesis.

Rokoko Smartsuit Pro is a sensor-based motion capture system to track full-body motion (excluding fingers) that uses on-body IMU sensors placed inside a nylon full-body suit.

Perception Neuron V1 is a sensor-based motion capture system to track full-body motion (including fingers) that uses on-body IMU sensors placed inside anti-slip traps.

Rokoko Studio is a software developed by Rokoko for previewing, recording and forwarding Smartsuit Pro body motion data to 3D softwares and tools. A version 1.12.1 was used in the production.

Axis Neuron is a software developed by Noitom Ltd. for previewing, recording and forwarding Perception Neuron motion data to the most popular 3D software programs. A version 3.8.42.8591 was used in the production.

MotionBuilder is a 3D character animation software with motion capture playback and editing tools developed by Autodesk Inc. MotionBuilder was chosen as post-production tool mainly because of the author has previous experience using it for 3D animation. A version 2018 was used in the production.

Unity is a cross-platform real-time engine developed by Unity Technologies. Unity was chosen as the main platform because it is the main engine used by the customer and therefore we could focus on the differences between the tested mocap solutions. A version 2019.1.7f1 was used in the production.

Mixamo Auto-Rigger is a free web-based tool for quickly creating, rigging and animating 3D characters developed by Mixamo Inc.

4 PRODUCING MOTION CAPTURE ANIMATIONS

4.1 Pre-production and concepting

It’s important to recap that a well-executed mocap can truly help immersion, while poorly executed one can ruin it equally fast. Properly executed and designed motion capture can save you money and time (Kines, 2000). Therefore, one of the key elements of using mocap is the need for careful planning and proper pre-production plan, as Kitawaga and Windsor emphasize in their book ‘‘MoCap for Artists’’. They consider that pre-production gives a “roadmap to how you’re going to organize and accomplish all of your goals” (Kitagawa & Windsor, 2008, p.13).

A plan was created (Table 1) based on the ‘‘Motion Capture Production Pipeline’’ by Kitagawa and Windsor (Kitagawa & Windsor, 2008) that was used to assess the full scale of the project and to plan capture sessions. Capture sessions were split into different days to see how much re-calibration of the suits is needed in between different takes and if it effects the workflow.

Table 1. Project plan to help assess the full scale of the project.

Pre-production	**Production**	**Post-production**
Commissioner brief Research Planning Commissioner feedback Mocap system and software set-up Testing mocap systems	Session 1 Practicing and recording first set of animations: idle, fold arms, sit down Session 2 Practicing and recording second set of animations: walk, run, jump	Animation clean-up with the mocap systems software’s Testing animations in Unity Improvements to animations in MotionBuilder

4.2 Set up

This section includes the unboxing, attaching the sensors and calibration of the suits, as well as setting up the Unity project, Rokoko Studio and Axis Neuron. Both suits had already been used in some extent and removed from their original boxes, so it was not possible to have a true ‘‘out-of-the-box’’ testing experience with the suits. For both suits, the setup was based on the quick-start guides provided by the companies and the goal was to rely on user on the manual or any other source only when needed. Neither of the suits came with an external power source so a standard power bank had to be acquired to use the suits wirelessly.

Setting up the work environment

The testing and recording of the animations were performed in the Finnish Virtual Reality Association (FIVR) Turku office located at the Turku Hive premises. The room was selected because it satisfied the needs of the testing. It had enough space to perform the testing and the recording of animations and had only a few electronic devices that could possibly cause any performance distortions in the mocap data. With sensor-based mocap systems it is recommended to be at least 3 feet away from magnetic objects otherwise the sensors might be influenced and make mocap data imprecise.

Setting up Unity project

In order to test the animations a project following the basic workflow for humanoid avatars (digital representation of a person or beings) in Unity was set up (Unity User Manual, 2019). In theory everything should work if the chosen 3D character uses a humanoid skeletal rig. A default 3D model from Unity’s standard assets package (Unity Technologies, 2019) and from Mixamo’s auto-rigger was used to apply and test the animations produced with the mocap suits. The character from Unity uses A-pose by default whereas the Mixamo character uses T-pose. The different poses were used to see if it affects the re-usability of the animations across different range of characters. Both Rokoko and Perception provided a plugin that allows the user to see the mocap movements in real-time on their own virtual avatars in Unity, but because Rokoko’s real-time streaming requires a paid version of the software, the streaming features were not tested.

4.2.1 Rokoko Smartsuit Pro

Software setup

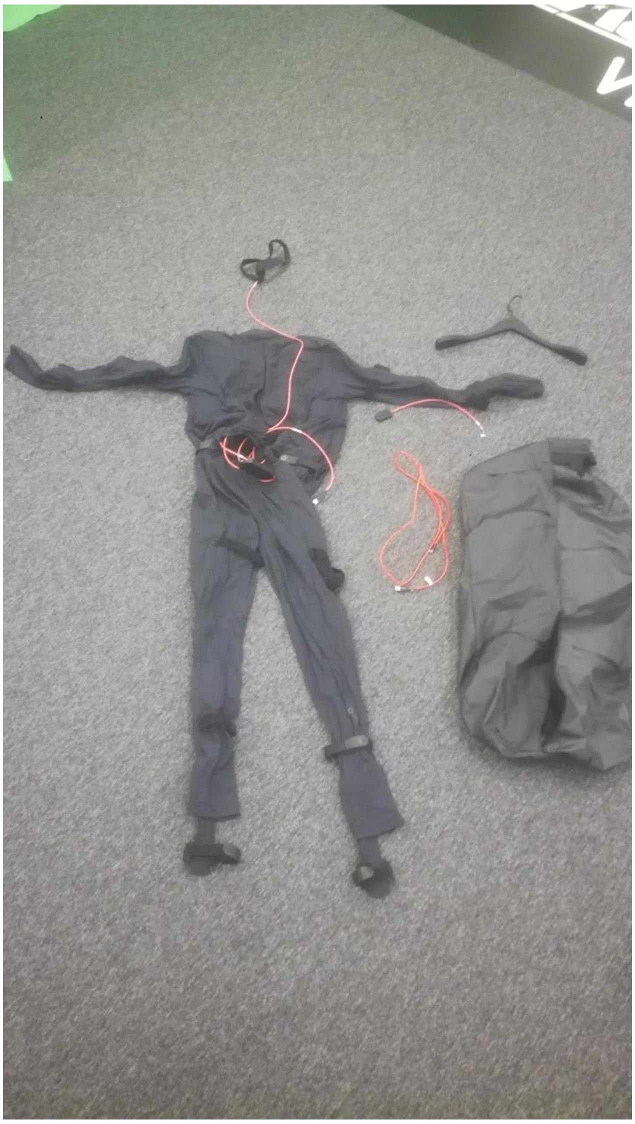
Rokoko Studio requires the user to a sign into a user to get access to Studio, and there was no way to bypass it. The registration took a couple of minutes, after which you had to set up a user and a body profile. Software contained a pre-made demo profile and a sandbox mode for quickly testing out the suit. For the testing a custom user profile with correct body measurements was created in order to increase the accuracy of the displayed data on the virtual avatar and to save the settings for quicker access.

No step-by-step or quick start instruction was given in the software and only a brief intro of the Rokoko Studio features could be found from the settings panel. The getting started guide needed to be read online from Rokoko's website, which described the setup overviewing the most essential functions of Smartsuit Pro and the Rokoko Studio.

Unboxing

The suit is intended to be so intuitive to wear and use that it can be ready for recording in just few minutes to test out the working version of a digital character. All the sensors were already embedded to the full-body suit and only the headband had to be attached by the user (Picture 6). Unboxing the suit took less than a minute.

The Smartsuit Pro system comes with the following items: 19 IMU sensors to track body motions, 19 cables connecting all the sensors to the hub, one USB cable, a hub to network the sensors and computer, a textile full-body suit and a waterproof suit cover for storing the suit.



Picture 6. Rokoko Smartsuit Pro mocap suit unboxed with all the components layed on the floor.

Setting up a WiFi connection

Connecting the hub via USB cable to the computer to configure the WiFi settings prompted a firmware update. The software automatically sets up firewall and other network settings making the process very fast and easy. As for setup, it wasn’t without its difficulties as there were issues with setting up the WiFi connection, even after updating the firmware. Upon troubleshooting with the help of the getting started guide a reason was found: the laptop used for testing was not connected to the same WiFi network than the one to which the suit was connected. Changing both devices to use the exact same network fixed the connecting issue.

Connecting a power bank

Once network settings had been successfully applied, the hub could be unplugged from the laptop and an external power source could be connected to the hub. The indicators on the hub give instant feedback if the WiFi connection has been successful or not. Green light indicates successful connection and red a failed connection. Software configuration was not successful on the first try as no virtual avatar was appearing in the software. Following the troubleshooting tips from the setup guide by updating the firmware and applying WiFi settings again fixed the issue.

Putting on the suit

The suit was easy to put on as it is very flexible. The hub just needed to be connected to an external power source. The sensors were already placed inside the suit in parallel with body segments so all there was to do was to ensure that the straps were tightened to prevent sensors from moving.

Calibrating the suit

Calibration process is required to help orient the sensors to accurately capture motion data. Calibrating the Rokoko Smartsuit Pro was easy as all the user must do is press a button, stand still for 3 seconds while performing a straight pose by having thumbs and nose point forward. After calibrating the suit by following the instruction from the software the suit was finally tracking motion in real-time although there was slight incorrect bending with the left leg and head rotation on the virtual avatar. Adjusting the sensor placement and re-calibrating fixed the incorrect body alignments with virtual avatar. Software has a visualization of the suit sensors indicating magnetic interference. Green means the best conditions for doing mocap, yellow means the suit is sensing magnetic interference and white means a disconnected sensor.

Set up conclusion

After both software and hardware setup the Rokoko’s mocap system was ready for recording, so quick test movements were recorded to test out the system and to verify it was working. At this point it could be confirmed that the suit worked well, and no external sources had to be used to fix issues as all the troubleshooting help could be found from Rokoko’s website. Sometimes when getting too close to a magnetic source it seems that the animation is not stable, so some sliding and drifting happens with the movements of the virtual avatar.

First time, out of the box setup experience took 1 hour 15 minutes by following the guides provided by Rokoko.

4.2.2 Perception Neuron V1

Software setup

The download and installing was a straight forward process. Axis Neuron doesn't require you to make any account and is ready to be used once the installation was completed. Software launched with a quick-start guide so one needed not be looked for online.

Unboxing

The suit is designed to be very adaptable and versatile by permitting multiple working mode combinations from 3 to 32 sensors. For the testing only 18 sensors were used with simple arm setup that did not include finger tracking. Unboxing the parts took less than a minute with additional 30 minutes of connecting the individual suit pieces and then snapping the sensors into the suit one by one (Picture 7).

Picture 7. Perception Neuron mocap suit unboxed with all the components layed on the floor.

Perception Neuron system comes with the following items: 32 IMU sensors to track the body motion, two magnet protection cases for storing the sensors, two USB cables, a hub to network the sensors and a computer, one full-body strap for attaching sensors to the body, two set of wrist/hand/finger straps with additional replacement sockets and sensor calibration tool.

Setting up a WiFi connection

The suit's hub needed to be connected to the computer via an USB to configure the WiFi settings. The hub makes a loud beeping sound indicating it has been connected to a laptop. As for setup, it was straight forward as the suit just needed to be plugged in, WiFi network selected and the settings to the suit applied.

Connecting a power bank

Once network settings had been successfully applied, the hub could be unplugged from the laptop and an external power source could be connected to the hub. There are no different color indicators on the hub for connection status, but after powering the hub it makes a loud beeping sound confirming a connection to a WiFi network.

Putting on the suit

The suit is very mobile and not restrictive. Putting on the suit was easy, the straps to the suit just had to be attached to the suit, the sensors plugged in to the straps and the hub linked to a computer by following the guide provided by Noitom Ltd. The sensors need to be placed in parallel with your body segments as close to the bones as possible to positions where the muscle movement is not going to throw off the motion capture. Attaching and putting together the suit took some time as the poses could not be kept completely steady during calibration.

Calibrating the suit

Calibration process is split into 4 steps: steady pose, A-pose, T-pose and S-pose with 3-5 second countdowns each. Calibrating the Perception Neuron suit was not without its difficulties. Sensors are very sensitive for movement so there were many failed calibrations passes with error messages in the software as the author was not completely steady during calibration.

The software has a visualization of the suit sensors showing transmission quality: green is best quality, yellow means the suit is sensing some magnetic interference, red means low quality and black means disconnected sensor. The on-body sensors also have a blinking light indicating their status. After calibrating the suit for the first time and seeing the virtual avatar it was noticed that two of the sensors were not working as the sensors weren't blinking and they were grayed out in the control panel.

Re-calibration did not fix the issue, so the suit had to be troubleshooted by going through different sensors to see if the problem was with the sensors. Later, it was found out that the issue was not with the sensor, but it was the socket that the sensor goes into. After replacing the socket with a replacement, one provided in the box, all the 18 sensors were working. However, there was still a problem with undesirable movement, such as feet rotation and left arm clipping through the virtual avatar.

Even after multiple re-calibrations and system (both the suit and the software) restarts, the sensor on left arm was still a not working properly so the individual sensors were re-calibrated to see if they had been exposed too much to magnetic fields. After that the system had to be configurated and calibrated again with new sensors, which solved the issue with the incorrect rotation on left arm and the suit was finally tracking motion somewhat accurately.

Set up conclusion

Some quick test recordings were made to test out the system and to verify it was working. There was also lot of time spend with troubleshooting and calibrating the system trying to get the mocap animations look decent. Even with re-calibration of the suit there was still both positional and rotational sliding errors with the virtual avatar. The quick-start guide tells what is needed to be done in order to get the suit working, but a couple of online tutorials that had a slightly different setup had to be watched in order to get the correct sensor placement for calibration process.

Some time had to be spent taking the sensors out for a proper storage to keep them away from magnetic sources in order to keep them safe. Pulling the sensors out of the sockets was sometimes hard as they really lock into the sockets on the straps. This quickly became frustrating as it had to be done each time when the system was set-up for recording.

First time, out of the box setup experience took 2 hours 45 minutes by following the guides and tutorials provided by Noitom Ltd.

4.3 Production

Production includes the recording and acting process of the animation sequences. Recordings were split into rehearsing takes and animations with slow and quick movements. For each take three clips were recorded to make the most use out of the suits and to increase the chance that at least one of the nine animations would be usable. Virtual avatar was reseted to the default position between each take to remove any possible incorrections. At this point it was known how the systems worked, so putting on and calibration of the suits was a much faster process.

4.3.1 Rokoko Smartsuit

Recording can be done either with a mouse click or by pressing a hotkey and the Studio will automatically playback the recording.

Setting up the suit for recording took an average 4 minutes with additional 1-minute calibration on both days. Each recording took an average of 1 minute with additional 30 seconds of resetting errors in between shots. Most of the recordings resulted in usable motion data with few errors here and there that could be further tweaked in post.

Session 1: Recording idle, folding arms and sitting down animations. Slow movements seem reliable and stable with smooth motion data, providing solid animations to work with. The left arm animation needs to be cleaned as it was clipping through the virtual avatar when it was sitting down.

Session 2: Recording walk, run and jump animations. Sometimes when changing direction quickly the mocap data will stutter and some errors could be seen on the virtual avatar, such as the virtual avatar sliding on the virtual space or body parts clipping through each other. Running on straight lines seems to produce better results than running in around the recording environment with less crossing of the legs or other interpenetrations happening.

4.3.2 Perception Neuron

No hotkey for recording could be found and the recording button on the hub was not working, so the button inside the software had to be used to start and stop the recordings. Axis Neuron will not automatically playback the recording, so each recording had to be opened manually from the working directory window to preview the animations.

Setup took an average of 10 minutes on with additional 30 to 45 minutes for calibration on both days. Each recording took an average one minute with additional 30 seconds of resetting errors in between shots. Most of the recordings resulted in terrible mocap data that would have required a lot of cleanup to make the animations useful. It took multiple attempts to get the system ready for recording by adjusting straps and re-calibrations to zero-out rotational errors on sensors.

Session 1: Recording idle, arm fold and sit-down animations. Overall slow movements seem stable with smooth motion data. The sensor on left upper arm was misbehaving by bending incorrectly and the suit needed to be re-calibrated multiple times in order to get acceptable results but still require manual cleanup in post.

Session 2: Recording walk, run and jump animations. The fast movements are decent with some minor sliding and clipping of the body parts in the virtual avatar. Fast movement caused more unnaturally bending of the joints and are less spatially precise than slow movements. In some cases, the sensor on the head went offline after performing jumping, this is probably because of the impact when landing on the floor caused the sensor to move as they are very sensitive.

4.4 Post-production

Post-production includes processing the mocap data, exporting to Unity, application to a 3D character in MotionBuilder for final animation cleanup and then a final export back to Unity. Cleanup with the Rokoko Studio and Neuron Axis software's were made to assess how effective the softwares are on their own at making the animations look production-ready without having to edit them in a third-party application, as this would reduce production time.

Animations contained excess motion including pressing the recording button at the start and at the end of recording. Therefore, those needed to be cut out of the final animations. After choosing the best takes and processing them in the software, the processed mocap animations were exported to Unity to see how animations perform with the 3D models.

4.4.1 Handling the Rokoko system motion capture data

Performance enhancements and other features

To enhance the performance of the Smartsuit Pro suit, the Rokoko Studio has five different filters that can be used with the advanced playback tools to edit the recordings inside the software. These filters reduce the amount of cleanup needed in post-production. Animation cleanup is done with keyframing using a curve editor in the playback window.

The locomotive filters allow moves such as running and jumping. Treadmill will lock the character in a fixed position in 3D space. Foot IK (Inversed Kinematic) will snap the feet to the virtual floor. Toe bending allows the foot to bend and stay above the virtual floor. Drift fix helps with the virtual avatar in the recording drifting around the 3D space by showing start and end position of the animation (Rokoko, 2019b).

Using locomotion without toe bending for running seemed to provide best results for fast movements. First step of was to edit the contact points to fix jumping so that the digital avatar was in the air by taking off the contact points from both feet. Then drift was used to keep the character in place. After editing the recordings and applying the filters, Rokoko Studio did a processing pass by analyzing movement and getting rid of imperfections in the recorded mocap data.

Exporting from Rokoko Studio

Rokoko Studio exports in FBX and BVH formats and supports exporting of recordings to four different skeleton formats: Rokoko Newton, Maya HumanIK, 3ds Max Biped and Mixamo.

Animations were exported as FBX format by using Mixamo skeleton export option to ensure best performance with the test character from Mixamo. Getting the Rokoko mocap animations into Unity was a simple process because Rokoko’s skeletal bone structure and re-targeting is natively supported by Unity, so no additional tweaking had to be done for either of the test characters.

4.4.2 Handling the Perception Neuron system motion capture data

Performance enhancements and other features

Axis Neuron has six different tools for optimizing and improving the quality of the recorded motion data inside the software. The cleanup is done via key framing using sliders and dropdown menus for choosing premade options.

Model alignment allows adjusting the posture angle to accurately capture actor's motion. Body size manager allows the user to apply measurements of body parts to ensure a more accurate experience. Smooth factor influences the accelerated speed and acceleration of gravity. Data visualizer visualizes the real-time changing of displacement and angle data for the body. Parameters offer options for floor contact points and joint characteristics. Contact editing allows the user to adjust contact results by editing single frames or by interval frame values (Noitom, 2019).

To fix some overlapping with the virtual avatars legs a combination of foot lock to snap the feet to the virtual floor and lowering the joint stiffness value was used. This helped with giving some leeway for the knee joints, resulting in a more accurate animation. The virtual avatar was leaning forward during the fast animations so lowering the pitch value with Model Alignment tool straightened it. Running was made look smoother by increasing step constraint value to increase the contact sensitivity between the feet and ground. Finally, sensor measurement errors and body shaking were smoothed out by using the smooth factor. Neuron Axis did a processing pass after applying the changes to the virtual model to get rid of imperfections in the recorded mocap data.

Exporting from Axis Neuron

Mocap animations can be exported in an FBX or BVH format and as a text and calculation files. Animations were exported as FBX and BVH formats, but Axis Neuron had no Mixamo skeleton preset so only the default Neuron Axis skeleton was used for the export.

Getting the Axis Neuron mocap animations into Unity was not as easy as with Rokoko's suit, as the Neuron FBX files are not natively supported by Unity's skeletal bone structure re-targeting. All the FBX export options resulted in similar problems with incorrect movements with the test characters in Unity. Captured mocap animations first need to be imported and re-targeted to a skeleton rig in a standalone third-party application to be compatible with Unity's Mecanim Animation System (Unity User Manual, 2019).

4.4.3 Editing the motion capture data in MotionBuilder

Using mocap data on a virtual model requires rigging, which is the process of creating a digital skeleton where the joint structure matches their real-life counterparts exactly. For instance, if you were to rig a digital 3D character's arm you would want to place the bones for the upper arm, the elbow and for the wrist, which then allows the animator to rotate the digital 3D arm in a realistic way (Özdem, 2018).

Rigging by hand takes a lot of time, requires 3D animation skills and is usually done by a dedicated animator. To speed up the process Mixamo's automated tool was used for creating and rigging the 3D character(s) used for testing the mocap animations. Mixamo character(s) were exported with T-pose using FBX for Unity format (Mixamo, 2019).

Re-targeting refers to the process of fitting the skeleton structure and transferring motion animations from one character to another. The default skeleton structure in either system for exporting has not been specifically designed for MotionBuilder because the bone structures are missing bones, but with little tweaking they can be characterized and used with MotionBuilder's re-targeting system (Mocappys, 2019b).

After re-targeting, the animations were polished by removing and editing the keyframes individually to correct positional and rotational errors with the 3D models. The buffer frames containing the movements of starting and ending recordings were also cut out from the animations.

4.5 Putting everything together in Unity

A test 3D character from Mixamo and Unity's standard assets was used to move around the scene in the example level (Picture 8). After doing re-targeting and quick clean up in MotionBuilder the animations were loaded back into Unity and then a simple controller was programmed to make the characters change animation based on the user input. Animations were looking decent but some additional polishing to blend the animations would be needed.

Picture 8. Test characters doing the idle animation in the Unity test scene: Rokoko (left) and Neuron (Right).

5 FINDINGS

Setup was a straight forward process and doable alone with both suits, but it is recommended to have someone to help with the recording, especially with setting up the Perception Neuron suit for calibration part to ensure proper sensor placement. Because of the longer set-up time with Perception Neuron, a proper pre-planning of the mocap sessions is recommended. With Rokoko’s system the user can jump in the suit, power the hub and be ready for recording in few clicks.

The suits are mobile and comfortable, and the materials of manufactured pieces are good quality. These two suits use a slightly different method for attaching the on-body motion sensors to the performer. Rokoko Smartsuit Pro uses a full-body suit with pockets for the sensors whereas Perception Neuron suit uses a combination of straps and gloves that go around the body parts for the sensors. This is important to note as the full-body suit only fits certain sizes (the one used in the study was fitted for large), so it might not be fitting for all body types, unlike Perception Neuron suit which is more adaptive to the different sizes and body types of the user.

With both suits it is recommending to not be next to magnetic sources, including computer, which is kind of hard when you are trying to operate the suit yourself, so having someone else to operate the computer helps with staying away from magnetic fields. Perception Neuron sensors seem to be more sensitive for magnetic interference, making them very prone to malfunctioning and limiting the kinds of spaces where it can be used.

Setting up WiFi connection and powering the suits' battery was a straight forward process, but there was some trouble getting the suits to connect to the laptop to get it to recognize the mocap systems. This was especially the case when there wasn't any prior experience using the mocap systems. Even though the sensor transmission quality was not the best, there was no input lag between the virtual avatars and the actual movements, but it seemed like the sensors became less accurate as the captures were getting longer.

Neither of the suits performed very well with jumping and seemed struggle with reliably tracking height of the user in 3D space, when the user left the ground. The tracking was not perfect or precise as although it worked well with horizontal movement, moving the core of your body on the vertical axis seemed to cause the suit to realign a bit. However, this could be corrected manually by using the editing tools in both softwares by editing the contact points, which helps the mocap systems to detect when the user is jumping and moving in the air.

Rokoko

Putting on Rokoko's suit is straight forward as sensors and the hub are already embedded inside the suit. Everything is easily accessible when not wearing the suit but accessing the hub while wearing the suit is challenging because it is located on the lower back of the suit. Taking off the suit was as easy as putting it on, it was just matter of minutes of putting it back into the container box. It took an average of 1 to 5 minutes for the initial setup.

Slow movements seem to be stable and accurate with little to zero issues with tracking. With fast movements like running and jumping there starts to be more glitching with the tracking of movement, such as the feet crossing over each other and the virtual avatar stuttering or sliding around the virtual space.

Rokokos skeleton system is not a standard skeleton rig, so the FBX and BVH files were missing some bones. Depending of the software used there might be some workaround done in order to get it working properly with your custom 3D character.

Neuron

The suit is not self-contained, so it cannot be put on straight out of the box and it takes time to setup, which might throw some users off. The user needs to first attach the sensors to the suit individually and then spend more time taking them out for a proper storage to keep them safe from becoming magnetized. The setup process is longer, but very straightforward as all the sockets have an area with an image that tells the correct placement. The initial setup took an average of 5 to 10 minutes with additional 10 to 15 minutes constructing and deconstruing the suit each time when the suit was used.

Slow movements are somewhat stable and accurate with varying results with the tracking. With fast movements the tracking is rougher and more jittery with lots of positional and rotational sliding errors, resulting in arms and legs interpenetrating with the virtual avatar in the recorded data.

With Perception Neuron you want to have a model with the exact same rig as Perception Neuron's default skeletal rig, because re-targeting the mocap data to a different skeleton structure might result in varying results depending of the resources you are willing to spend with proper re-targeting and rigging.

Calibration

Good calibration is the key to good mocap data to avoid any extra cleanup in post-production, so it is very important that the user performs the calibration process properly to tell the mocap system where all the sensors are.

Calibration with Rokoko required only one pose and once calibrated the suit didn’t seem to need much re-calibrating while using it, only after doing the jumping the sensors took some impact causing them to move so they needed to be corrected and re-calibrated. It took an average of 5 to 10 minute to calibrate the suit to make it ready for recording.

Calibration with Neuron required four poses and needed plenty of re-calibrations while using it. Neuron suit had a problem with positional drifting when the sensors cause a body part to move around weirdly causing the virtual avatar sliding around in the virtual space. Adjusting the sensors, re-calibrating and zeroing out the center of the virtual avatar to its origin sometimes helps, but not always which can be frustrating and time consuming when trying to get the suit ready for recording. It took an average of 30 to 45 minutes with multiple re-calibrations to the suit to make it ready for a recording.

Summary table

The summary table (Table 2) lists all the elements tested with the mocap systems. The usability results are rated by using the scale troublesome, complex, somewhat tricky and straightforward. Performance and accuracy are rated on 1 to 10 scale where 10 is the best. When mocap data, nuance or tracking is lost, the score suffers.

Table 2. Performance comparision table of the results comparing the mocap systems.

	Rokoko Smartsuit Pro	**Perception Neuron V1**
Usability: setup	Straightforward	Complex
Usability: application	Straightforward	Somewhat tricky
Performance	8	5
Accuracy	slow animations: 9 fast animations: 8	slow animations: 7 fast animations: 6

The largest difference comes in the usability of the suits: Perception’s suit suffered with usability issues as troubleshooting its problems and inconveniences made the first-time user experience complex. Rokoko’s first-time user experience was a smooth experience exceeding in all the compared areas with only minor usability issues.

Rokoko’s overall performance was also much more consistent with collecting mocap data and the accuracy of modeling and tracking of the user's movement seemed more refined. Although the goal was to test the performance of the mocap systems, it must be mentioned that Rokoko provided better tutorials than Noitom Ltd. on how to clean up and get the mocap data work with most of the popular 3D applications.

6 CONCLUSION

The aim of the thesis was to introduce the reader to the basic concepts of motion capture and the benefits of using it to provide a better understanding how mocap can be used effectively as an animation tool for production. The focus was on comparing two sensor-based solutions aiming to provide professional motion capture to the masses for a budget friendly price.

Based on the findings of this study, it can be concluded that both systems provided very similar results in recorded animations. Overall, mocap animations look decent but the absolute positioning such as crossing legs or arms is not very accurate or precise. Proper polishing from an animator would help immensely to achieve smoother results in both cases. Both mocap systems are good for the reasons they are designed for and work generally good for capturing performance and animations for anything that do not require a high level of precision with tracking or one-to-one re-targeting.

The reasons for buying a particular suit should be based on what you as a game developer are going to do with the motion capture and what are the requirements for your desired workflow. When getting a mocap suit you might want to go for an affordable price, but it is important to know what you are getting for the price. Generally, with a cheaper price you are saving money at the cost of accuracy and performance.

For a $1,500 price Perception Neuron gives a motion capture system capable of putting your body movements onto a 3D character with the added benefit of hand gesture capture and free real-time streaming feature. With Perception Neuron, there is more work involved in order to get started with using mocap and more time was spent troubleshooting its problems than actually recording animations. Additionally, there might be some more work involved in order to get the animations working properly with your custom 3D character depending of the software used, not to mention a great deal of mocap data had to be processed with its attached software and then cleaned up in post.

For a $2,500 Rokoko Smartsuit Pro gives a motion capture system capable of putting your body movements onto a 3D character in just a few minutes. With Rokoko, the users must do very little to get started with using mocap as sensor placements are ready to go and taking the suit on and off is a much quicker process. It took significally less time than

Perception Neuron suit to get the system up and running "straight out of the box" and provided smoother animations with minimal cleanup required in post.

Perception Neuron offers finger motion tracking and possibility to track specific body areas with different setup modes, so if your animations need finger tracking or more flexibility with the tracking setup then Perception Neuron is the right solution for you. At the time of writing this thesis the Rokoko Smartgloves have not been released yet but the Rokoko Smartsuit gives a huge jump in productivity due to its versatility, overall performance and tracking quality, making it more ideal to work with.

The ease of use, the setup time, exporting the animations to Unity, the cost of replacement parts and the amount of calibration to achieve desired results or precise tracking, amount of cleanup that is acceptable are all areas that must be taken into consideration when making the decision if either of these two sensor-based systems are the right solution for your motion capture purposes.

To reduce the number of key frames and to make the animation clean-up easier, the movements should have been recorded in individual takes rather than in one long sequences with a help of someone else pressing the record button. There may have been a way to work around many of the issues with the systems found in during the testing that the author did not know about or were not included in the scope of the thesis.

Further work could include performing an A/B testing to see if the users (testers) can see the difference in the resulting animations. The real-time streaming features in both systems and outdoor performance should also be tested and researched further to see if they can be used to speed up mocap process and possibly increase productivity. Additionally, it would be interesting to see whether the newer versions of both suits produce similar results or still suffer from the same problems than the author had with them.

REFERENCES

CGIcoffee, 2017-last update, Perception Neuron Inertial Motion Capture vs Optical Mocap Systems And The First Production Motion Capture Session Experience. Available: http://cgicoffee.com/blog/2017/04/first-production-motion-capture-session-report [3.5.2019]

Computer Hope, 2017-last update, Key frame. Available: https://www.computerhope.com/jargon/k/key-frame.htm [25.3.2019]

Conditt, J., 2018-last update, 100 years of motion capture technology. Available: https://www.engadget.com/2018/05/25/motion-capture-history-video-vicon-siren/?guccounter=1&guce_referrer_us=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_cs=qN2FkAq6_bLq71-qwaOeZg [12.3.2019]

Dent, S., 2014-last update, What you need to know about 3D motion capture. Available: https://www.engadget.com/2014/07/14/motion-capture-explainer/?guccounter=1&guce_referrer_us=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_cs=7aJ-1sw3j1VULjsfKR4t0w [27.3.2019]

Failes, I., 2019-last update, WHAT MOCAP SUIT SUITS YOU?. Available: http://vfxvoice.com/what-mocap-suit-suits-you/ [22.6.2019]

Finnish Virtual Reality Association, 2017-last update, About. Available: https://fivr.fi/about/ [20.6.2019]

Huuhtanen, P., 2018-last update, Heijastimia syksyn pimeneviin iltoihin: liikekaappaus ja sen mahdollisuudet. Available: https://blogit.metropolia.fi/mediakompleksi/avainsana/motion-capture/ [26.3.2019]

iPi soft, 2019-last update, Express. http://ipisoft.com/software/express-edition/ [30.3.2019]

Jáuregui, D., 2011, 3D motion capture by computer vision and virtual rendering. Available: https://www.researchgate.net/publication/255990108_3D_motion_capture_by_computer_vision_and_virtual_rendering [24.6.2019]

Kines, M, 2000-last update. Planning and Directing Motion Capture for Games. Available: https://www.gamasutra.com/view/feature/131827/planning_and_directing_motion_.php [20.4.2019]

Kitagawa, M. & Windsor, B. (2008) MoCap for Artists. New York: Focal Press.

Mattson, V. & Mårtensson, T., 2014. Visibly of Using Markerless Motion Capture. Available: https://www.diva-portal.org/smash/get/diva2:731552/FULLTEXT01.pdf [26.3.2019]

Menache, A. (2010) Understanding Motion Capture for Computer Animation. 2nd Edition. San Francisco, CA: Morgan Kaufmann Publishers Inc.

Metamotion , 2019b-last update, Motion Capture Systems – How to Choose. Available: https://metamotion.com/motion-capture/motion-capture-systems-how-to-choose.html [2.5.2019]

Metamotion, 2019a-last update, Who Uses Motion Capture?. Available https://metamotion.com/motion-capture/motion-capture-who-1.htm [29.4.2019]

Metamotion, 2019c-last update, Gypsy Motion Capture System Workflow. Available: https://metamotion.com/gypsy/gypsy-motion-capture-system-workflow.htm [20.6.2019]

Mixamo, 2019-last update. Available: https://www.mixamo.com/#/ [1.6.2019]

Mocappys, 2019a-last update, What is motion capture?. Available: http://mocappys.com/what-is-motion-capture/#.XIUfYSIzZhE [27.3.2019]

Mocappys, 2019b-last update, Retargeting Animation In MotionBuilder. Available: http://mocappys.com/retargeting-animation-motionbuilder/#.XPoO3lwzZPY [6.6.2019]

motionCAPTUREnyc, 2019-last update, How much does motion capture cost?. Available: http://www.motioncapturenyc.com/cost [28.3.2019]

Neogames, 2019-last update, Finnish Game Industry 2018 Report. Available: http://www.neogames.fi/wp-content/uploads/2019/04/FGIR-2018-Report.pdf [29.4.2019]

NoFilmSchool, 2019-last update, Rokoko Review: High Quality Mocap for a Fraction of the Cost. Available: https://nofilmschool.com/rokoko-review-high-quality-mocap-fraction-cost [3.5.2019]

Noitom, 2019-last update, Axis Neuron Manual. Available: https://neuronmocap.com/system/files/software/Axis%20Neuron%20User%20Manual_V3.8.1.5.pdf [10.5.2019]

Nymoen, K., 2014, Chapter 3: Motion Capture. Available: https://www.uio.no/studier/emner/hf/imv/MUS2006/v15/litteratur/knkap3-4.pdf [27.3.2019]

OptiTrak, 2019-last update, Build Your Own. Available: https://optitrack.com/systems/ [30.3.2019]

Özdem A., 2016, Introduction To Motion Capture. Available: https://www.academia.edu/28101363/Introduction_to_Motion_Capture [30.3.2019]

Perception Neuron, 2019-last update, Perception Neuron. Available: https://neuronmocap.com/products/perception_neuron [10.5.2019]

Rahul M., 2018, Review on Motion Capture Technology. Available: https://globaljournals.org/GJCST_Volume18/4-Review-on-Motion-Capture-Technology.pdf [28.3.2019]

Rokoko, 2019a-last update, Smarsuit Pro. Available: https://www.rokoko.com/en/products/smartsuit-pro [27.3.2019]

Rokoko, 2019b-last update, Rokoko Guide: Cleaning Up Motion Data in Studio. Available: https://www.rokoko.com/en/explore/blog/rokoko-guides-cleaning-up-motion-data-in-studio [10.5.2019]

Starcevic, V., 2018-last update, The story behind Remedy’s viral dog mocap star. Available: https://www.gamasutra.com/view/news/319105/The_story_behind_Remedys_viral_dog_mocap_star.php [26.4.2019]

Swimbanks, B., 2019-last update, Motion capture on an indie budget. Available: https://www.thehaguetech.com/stories/motion-capture-on-an-indie-budget [1.4.2019]

Turkugamelab, 2019-last update, About Game Lab. Available: http://www.turkugamelab.fi/?page_id=72 [6.5.2019]

Unity Manual, 2019-last update, Animation System Overview. Available: https://docs.unity3d.com/Manual/AnimationOverview.html [5.6.2019]

Unity Technologies, 2019-last update, Standard Assets. Available: https://assetstore.unity.com/packages/essentials/asset-packs/standard-assets-32351 [15.5.2019]

Unreal Engine, 2018-last update, Instructor designs Motion Capture Classroom using Unreal Engine 4. Available: https://www.unrealengine.com/en-US/developer-interviews/instructor-designs-motion-capture-classroom-using-unreal-engine-4 [22.6.2019]

Xsens, 2019a-last update, A history of motion capture. Available: https://www.xsens.com/fascination-motion-capture [12.3.2019].

Xsens, 2019b-last update, Multi-person full-body immersive VR made easy. Available: https://www.xsens.com/press-releases/multi-person-full-body-immersive-vr-made-easy/ [30.3.2019]

Xsens, 2019c-last update, Xsens MVN animate. Available: https://www.xsens.com/products/xsens-mvn-animate/ [23.6.2019]