Expertise
and insight
for the future

Taryn Baker

The History of Motion Capture Within The Entertainment Industry

Metropolia University of Applied Sciences

Bachelor

3D-animointi ja -visualisointi

Thesis

Date 05.05.2020

Author(s) Title Number of Pages Date	Taryn Baker The History of Motion 29 pages 2020
Degree	Bachelor
Degree Programme	3D-animointi ja -visualisointi
Specialisation option	3D-animointi ja -visualisointi
Instructor(s)	Peke Huuhtanen, Senior Lecturer

This thesis is intended to go over the compact history of motion captures development. Motion capture is a rapidly developing technology in the animation world and is constantly evolving, meaning that there is a lot of history to cover and that it is constantly updating. The goal of this paper is to cover some of the most historical occurrences within the development of motion capture, starting with when movement was first being tracked to produce other moving images, right up to when you can simply track and apply data to a model using a mobile phone.

Some of the difficulties that were predicted to be encountered were that a lot of the technology nowadays is not publically released, as there is a strong market and competition regarding the technology. Development from early on was mostly conducted by universities, and was detailed and widely available. Now that there is a competitive market (mostly due to the entertainment industry) Many discoveries, while announced with great publicity, are often lacking in-depth descriptions of the actual mechanics and developmental processes.

While researching for this paper it was found to be true that much of the data was unavailable, the more modern the technology the harder it was to find data on how things were created. Fortunately, the author of this paper works for a well-regarded entertainment studio that has access to their own large motion capture studio and veteran professionals. It was also a benefit that the name of the company had influence when contacting other professionals about motion capture.

Keywords	motion, tracking, capture, markers, optical, tracker-less, performance capture, motion capture

Contents

1 Introduction

Motion capture (mocap) is a rapidly evolving technology and its history is dense with innovation. This paper is a condensed overview of the major advancements in motion capture. Starting From when people first began to track movement from videos for many entertainment or medical purposes, all the way up to modern technology. Some of the challenges anticipated within this paper is finding the information on a technical level and not just at a commercial sale level. Many new technologies are not as well published in a way that the developmental process and technical side of the technology is easily accessible to the general public. Many new advancements are made by the media/entertainment industry and keeping development under wraps helps maintain an edge in the industry. Something that is easier to access is the information regarding earlier discoveries, as most of the development of motion capture early on was made by the medical industry and was readily published for review by scientists.

This paper starts with an introduction of motion capture to those not yet familiar with the field. It goes over its definition, the basics of how it works and the different categories that it can fall into. Following that is the question on what the difference is between two common terms in industry, motion capture and performance capture. These are two terms that often cause disagreement within the community on if they are simply the same thing under a different name or two entirely different things. This paper hopes to give professional opinions on whether they indeed are or are not the same thing.

After the information on what motion capture is and its terminology, this paper will continue with the first steps and origins of motion capture. From its earliest roots in the 1900's when it was hand drawn with a technique called rotoscope, to a rapid jump in development in the 60's and 80's. Coming through into the 60-80's is where the digital advancements of motion capture really started to take off as the academic and medical community started to show an interest in its development. It's in this time period that the first instances of the different types of motion capture started to surface. From these first developments you will be able to see a rapid development to the technology as it begins to be used further in the entertainment industry.

Other important parts of the history of motion capture do not just include how both the hard- and software has been developed, but how it has become commonplace in the world today in the form of movies and video games. Truly while motion was in many ways developed for academia and the medical field, it is the entertainment industry that has brought this technology into the light and made it known as a more common place

practice. It is in fact now so commonplace that motion capture can be purchased and used by the general public from their mobile phones. That is how rapid motion capture has progressed. From its budding digital beginnings to being readily available in the palm of someone's hands in about 50 years. The history of motion capture is short, but dense. One could write continuously about the history of motion capture and everything that has branched off of it, but hopefully this paper will condense the presumably most notable achievements that it has made.

2 What is motion capture?

2.1 What is motion capture and how does it work?

Motion capture is the term used to describe the digital capture of movement of an object or people. The technology originally came from the studies of movement analysis in living beings, but is now commonly used by VFX studios, sports therapists, neuroscientists, and for validation and control of computer vision and robotics (Vicon, 2019). There are a variety of different motion capture systems available that are able to digitize motion, they fall into three main forms including magnetic, mechanical and optical, with optical further branching with passive, active or markerless systems.

Magnetic systems use sensors located across the body of the motion capture performer that measure the low frequency magnetic fields that are generated by the transmitters. These sensors are cabled to electronic control units that correlate their reported locations within the field (Metamotion, 2019; Lonkar, 2019).

Figure 1. Performers in Xsens magnetic system suits

The mechanical system is often less commonly used, mostly because it doesn't always provide the most natural motion, as well as a slew of other issues. It works by a motion capture performer wearing a human shape set of metal strips, like a very basic skeleton that is hooked on to performers' back. Each joint has sensors which give the position of different body parts and its movement (Sagar Lonkar, 2019).

Figure 2. Performer in Animazoo's mechanical motion capture system

The most common system used is the optical system, which is also the one that most people recognize, as it normally involves people in those distinctive lycra suits with what would appear to be ping pong balls or in some cases LED lights attached to them. This system also works markerless with pre-tracked points being dictated digitally beforehand, i.e. facial points and key body articulations. Multiple cameras (from 16 and well above 100 sometimes) track the reflective markers using infrared light. With all the different camera angles sending data to the software of where any given marker is at any given time, a 3D representation can be created within the software.

Figure 3. Performer in an optical system motion capture suit, also wearing a HMC (Head mounted camera) which tracks facial data optically.

Of course, motion capture isn't just the gathering of raw motion data. There is a long list of things to do afterwards. After the data has been collected, a software will need to be utilized to apply the data to a rig, there are many different softwares available for this including Vicon's Blade, Motion Analysis' Cortex, and Xsens MVN software. There are also a number of third party software solutions like Autodesk MotionBuilder and PeelSolve. Data is often flawed. There isn't any data recording to date that comes out perfect without any editing being required. From feet sliding to lackluster movement or joint popping, it is the work of animators to smooth out the kinks that come with the data and either keep it true to the performance or in some cases add more flare to the movement.

2.2 Motion capture and performance capture

Due to its relatively young and constantly evolving nature, the motion capture industry has created many different terms to refer to processes that are relatively the same or are in fact the same. Even today, among professionals within the industry, terminology is disputed hotly over what is the correct wording to be used in regards to things associated with motion capture. One of those topics that is argued over is the terms motion capture and performance capture. Both terms/practices refer to the capture of motion, be that of the body, face or hands. On one side, it is debated that motion capture is only the capture of body motion and that the capture is considered of low fidelity. In this same camp, it is believed that performance capture is the capture of the face and hands using very high fidelity data, body capture using low data voice recording for accurate lip syncing. It is argued that performance capture captures the truest actions and meaning from the actor being recorded. Performance capture blends motion capture with at a relatively low resolution, typically, the orientation and relative placement of markers placed around limb joints, with more densely placed markers on the face. Facial markers are tracked using a head mounted camera along with any vocal performance provided by the actor. Performance capture allows the facial performance of human actors to drive the facial performance of a digital character. By recording the vocal performance alongside the facial performance, 'lip synch' between the voice and mouth movements of the character can be preserved. (Digital worlds, 2016).

The following quotes have been given by large scale commercial motion capture studios:

"*At Audiomotion Studios, our approach is that Motion Capture is simple body data, including locomotion, walk cycles and basic movements. In some cases, this can be captured adequately with a non-professional actor. Performance Capture involves a deeper level of acting, a greater level of dynamic movement and invariably the use of HMCs (head mounted cameras) where intricate facial expression can make all the difference in the quality that's achieved*." - Philip Morris, 2019, email interview, Audiomotion Studios.

"*In my experience, the term motion capture would refer to body capture only. The kind of thing you'd see in a locomotion or crowd shoot.*

Performance capture would add HMC, audio record and (potentially) facial make-up. The content of the shoot would steer away from physical action and more towards dramatic performance.
Performance capture shoots tend to follow the conventions of television and film. There will be rehearsals, more crew (make-up, script supervisor, AD, etc.) and the amount shot in a day will generally be less than on a standard motion capture shoot." – Stuart Haskayne, 2019, email interview Centroid Motion Capture

"Motion Capture is, in the typical terms the capture of body motions only. In the past this rarely included finger capture either due to technology constraints, rigging contrasts or other issues.
Performance Capture is an evolution of Motion Capture in that it includes Motion Capture but adds facial capture, finger capture and voice capture.
To go further: While much of the translation of the performance in Performance Capture is still done in a post format with heavy reliance on Interpretation and often not 1:1 the industry moves gradually toward Reality Capture.
While the term Reality Capture is still in its infancy it generally refers to capturing with less interpretation of the performance. Head Mounted Cameras that capture geometry on every frame." – Graham Qually, 2019, email interview, Beyond capture

On the other hand, in many articles, the terms motion capture and performance capture are interchangeably used to refer to the same things, sometimes neither term is used and simply the term 'motion tracking' is used. Arguably, any capture of motion, be it of the face, hands, eyes, body or even an inanimate object such as a car, are all motion capture.

"*So our job is about recording editing and playing back movement. The signal we deal with is the "motion signal". So we do Motion Capture.*
Performance capture" name started to appear when some big companies finally understood that motion was very powerful and subtle and was a global thing, involving body, face eyes, fingers, speech etc... and to give a name to that new understanding they called it Performance capture, because there was at last the idea that motion was at the heart of the performance, that movement was so important in a performance that if one

was recording motion it was actually dealing with the whole power of the performance. But if a sound engineer records the singing performance of a great singer, he does not say that he records a performance, he records the sound signal.
So I believe that the name performance capture was useful to make it clearer that we, movement engineers, were recording a signal that was containing the performance. But in the long run, it is a wrong name, that I hope will disappear and will be replaced. We always do Motion Capture or Motion Recording and sometimes we record more than just the body. We include face, eyes fingers, props"- Rémi Brun, 2019 email interview, Mocaplab

"Back in the day when facial mocap was still in it's infancy, the only way to capture facial animation was on a dedicated capture system. The actor would read lines in front of a system consisting of multiple cameras. Body motion was captured in a different studio, usually on a different day. Once cameras became smaller, we started attaching them to the actor's head. That way we could capture facial motion while capturing body motion. Someone coined a term for it, performance capture. The term conveys nicely what it does. It captures the full performance of an actor. Since body motion somewhat affects facial muscle triggering, this is the only way to record authentic facial motion. While to some this is the ultimate goal in transferring actor work to digital medium, it does have some drawbacks. At the time of writing this, camera quality is still lacking on portable systems. Actor head stabilization is still an issue that has not been universally solved. All this means portable systems are still of lower quality, easily causing error on levels that will visibly affect readability of facial features, thus watering down the performance.
Dedicated capture rooms for facial motion still produce the highest quality results. While they do lack the authenticity of simultaneous capture of body and face, they do give you more control over the sub components of the performance and thus the final emotional impact of a scene. At the technical level both ways do exactly the same, the other one just captures the face simultaneously with the body.
It's not uncommon to capture an actor's face while capturing body motion and use that footage as reference for higher quality facial capture sessions in another studio.
In both cases you end up with a performance, the other one is composite of many performances, one is true performance.
Both workflows capture the actor performance, what matters is the end result.
Both are performance capture, both are motion capture." Henri Blåfield, 2019, email interview, Remedy Entertainment

For the basis of this paper. The terms motion capture and performance capture maybe used interchangeably. It is the experience of the author that many people use either to describe the same thing rather often. Some will still debate that they are separate because of the use of audio or the range in fidelity and intention of use, but it is irrefutable that both terms share the same history and technologies to track and digitized motion. Thus this paper will cover the history of both terms as if they are one.

3 The early start of motion capture

3.1 The first motion capture: Rotoscope

The study of motion can be traced back centuries ago, but the art of motion transformed to be digital can be pinpointed to a particular place in time. All things considered, rotoscope was the first instance of taking human motion and then copying that motion to be applied digitally to a character. The studio website for Feleischer Studios states it was in 1915 that Max Fleischer first developed the technique that allowed animators to give a realistic fluid motion to their characters. Max reasoned that if animators were able to use the frames of live action film as a guide, then they could create a series of drawings that, when strung together, would have a more lifelike flow. The question was how to make this logical idea into a functional and practical reality? (Feleischer studios 2016)

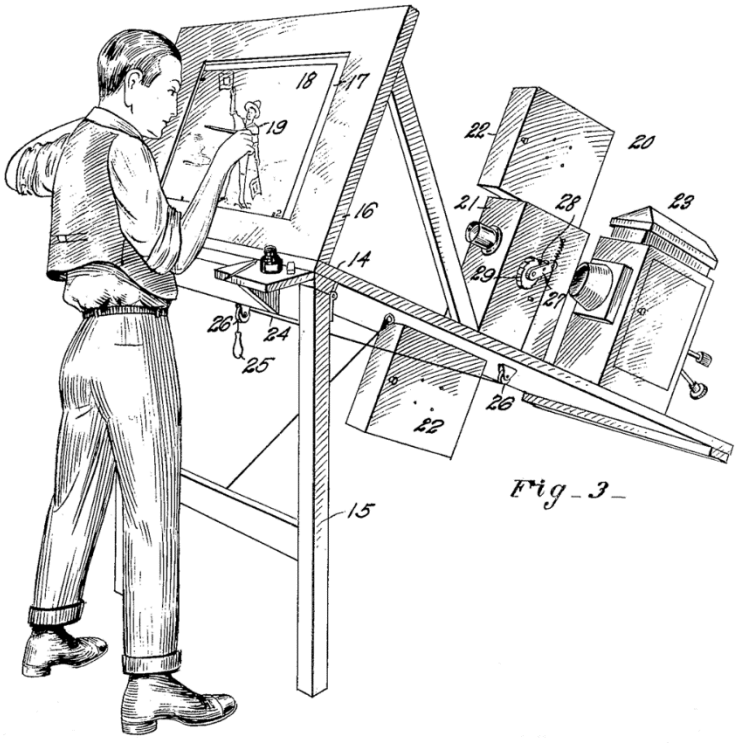


Figure 4. Illustration of a man using a machine to produce rotoscope animation.

What Feleischer ended up creating was the rotoscope. He developed it with the help of his other three brothers and constructed it in his own living room. In their thesis Emmi Rukkila wrote that the first ‘successful’ attempts at animation created with this rotoscope

technique took a year to finish and only lasted for a minute. The brothers improved their process further to produce animations faster, and soon the technique was not only possible but also practical. (Rukkila, 2018)

What they managed to construct after many trials and errors was as described in the book Motion Graphics: Principles and Practices from the Ground Up “projector that produced an image one frame at a time on the pack of a piece of frosted glass. The animator would place his or her paper onto the registration pegs on the glass and trace the projected image, thus producing the rotoscoped animation.” (Crook, Beare, 2016). In 1915, Feleischer created a short series of animations that he named ‘Out of the inkwell’ that focused on the antics of a small clown that would later be known as Koko. Fleischer Studios is responsible for many highly recognizable animations using the rotoscoping method, not limited to, but including Betty Boop and Popeye. Following the great success of rotoscope through the many characters and series that Fleischer produced, the technique was then adopted to be used to make the world’s first full length animated movie: Snow white and the seven dwarfs by Disney.

Figure 5. Example of how real footage is used to portray fluid movement in rotoscope.

Rotoscope is still alive today, although much more technologically advanced using computer software instead of the projector apparatus. During its early beginnings, it suffered many of the same resentment and stigma that motion capture has received from more traditional animators. That it is a form of cheating, as the work does not rely entirely on your own craft.

.

3.2 Mechanical motion capture: Animac

Lee Harrison III is regarded as one of the earliest pioneers in motion capture. During the 1960's Harrison was experimenting with developing a system that consisted of using analog circuits and a cathode ray tube (CRT) to produce animated figures. It consisted of rigging a bodysuit with potentiometers and lincoln logs that relayed a 3D figure in real-time onto his CRT screen.

It was during Harrison's time in college that he was first introduced to an oscilloscope, a type of electronic device that is used to test and measure voltage. It would show the signals from the voltage on a small screen in a 2D image that displayed time and voltage over an X and Y axis. It was after witnessing this and learning how it worked that Harrison was hit with the idea that this could maybe be used in future development with animation. He proceeded to gain employment at Philco Corporation, who worked in the development of batteries, radio and television production. It was in this lab that Harrison was able to form teams who would later help to develop the creation of the 'Bone Generator' in the late 1950's. (Der Apparate Welt, 1992; Funk, 2010)

The Bone Generator was truly revolutionary as it was able to create complex 3D animation without the use of any polygon modelling software or even a CPU, it later evolved into the Animac when the ability to add animation to said bones was added. The Bone Generator would essentially draw out a stick figure digital rig and in synchronization with this was a 'Skin network' that would calculate spinning vectors that would encompass the bones. The skin generator would also be used for scanning and storing different skin features that would include colour, texture and shading, As Walter Funk wrote in his paper regarding the Animac "the function of the Skin Generator is to generate a video signal whose magnitude represents the orthogonal distance or thickness between the bone vector and the surface of the skin. A flying spot scanner is also used to record light intensity signals from specifically prepared photographic transparencies". A Flying spot scanner is a device that uses a scanning source of a spot of light such as a high resolution, high light output, low persistence CRT, to scan an image. The light passes through the image being scanned and then converts it to an electrical signal.(Guimarães, 2019; Eigenwelt Der Apparate Welt, 1992; Funk, 2010)

When combining all of these systems together it became Animac in the early 1960's. It was able to relay a live action image of a dancer in an animation harness that like previously stated was hooked up with potentiometers and Lincoln logs. A mouth was virtually created using algorithms and then controlled by an individual's voice. It would function in a way that it would react to the sound waves and imitate the assumed movements.

Figure 6. Dancer in an animation harness to provide movement for the Animac system as well as the output.

It was used to make numerous experimental films, however it never made it past that, it never branched out into the commercial side of things. Even though Animac never made it out commercially, direct descendants of this innovation did, you can observe them being used in Star Wars, Logan's Run and Electric Company. (Der Apparate Welt, 1992; Walter, 2010)

3.3 Exoskeleton motion capture system

When discussing the history of motion capture, Lee Harrison is often overlooked in favor of Tom Calvert. Both are equally instrumental in the origins of modern motion capture systems and development despite their research and apparatus being similar and removed from each other by at least 20 to 30 years. Following the early invention of ANIMAC, during the 1980's, kinesiology and computer science researcher Professor Tom Calvert from the Simon Fraser University in Canada started to take on the exoskeleton motion capture system for medical usage.

His study was to monitor and document choreographed studies and clinical assessment of movement abnormalities. To track knee flexion for instance from a digital stand point so that data can be accurately compared over time

.

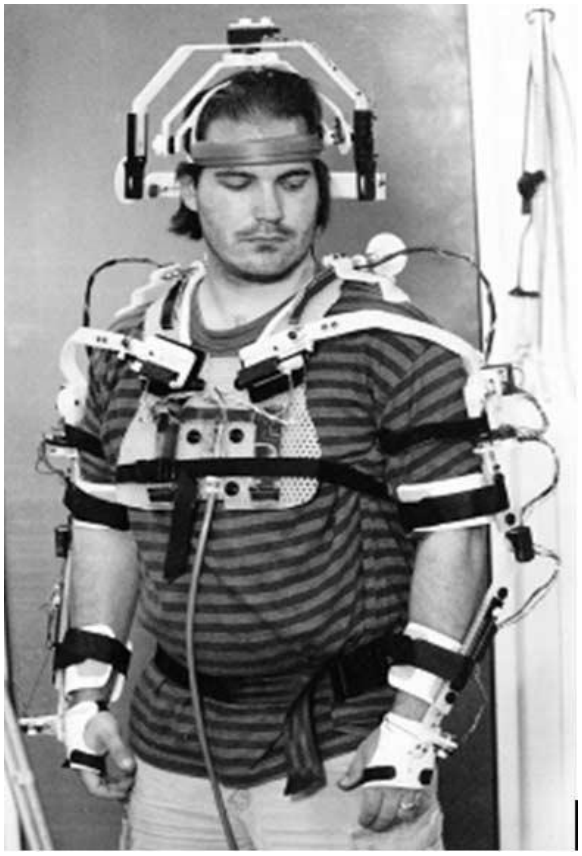


Figure 7. Calverts mechanical exoskeleton being worn by a test subject.

Calverts system was a mechanical exoskeleton that was worn by the test subject. It consisted of the mechanical armature as well as potentiometers along each of the knees so

that they could flex and bend in correspondence to the knee. The data from the analog output was then transformed into a digital readable format and relayed into a computer animation system. Even well into the 21st century we still use the system of potentiometers. Although more refined into a suit rather than a mechanical exoskeleton that encumbers the actor. This is nowadays one of the most popular low budget systems available on the market and is widely used across the board in multiple industries such as entertainment, medical and military. (Sturman, 1999.)

3.4 The Graphical Marionette: Optical motion capture

In the early 1980's, in fact almost coinciding with the mechanical system developed by Simon Fraser University, Massachusetts Institute of Technology and the New York Institute of Technology were both experimenting with the first optical tracking of the human form and movement. It was during the year of 1983 that Delle Rae Maxwell and Carol Ginsberg from Massachusetts Institute of Technology presented the fruit of their research as the 'Graphical Marionette'. The Graphical Marionette used a system called Op-Eye, it relied on infrared LED lights that had been wired into a garment at joints and other bony prominences. It also utilized two cameras that contained a two axis lateral effect diode, as quoted by Maxwell in her paper "these optical sensor cameras detected the position of a spot of light on their surface. Two-dimensional position information is then obtained via four electrode connections at the edges of the detectors".

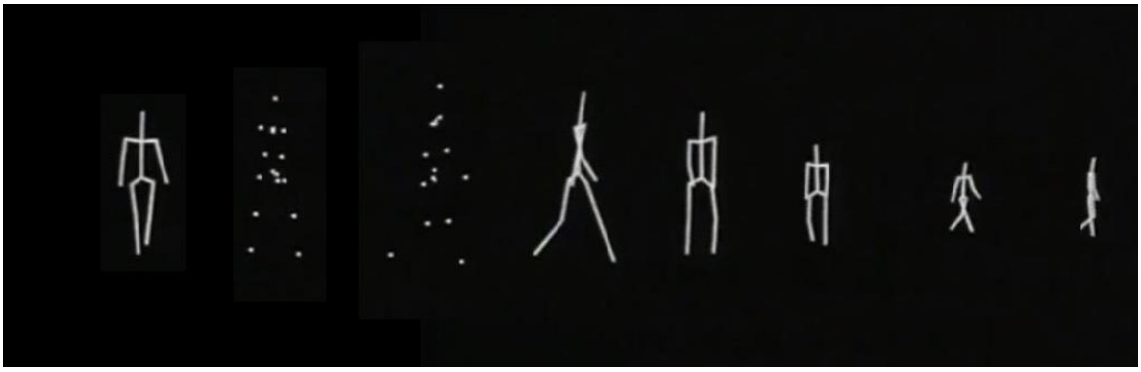


Figure 8. Output of the Graphical Marionette by the Op-Eye system.

The individual who would be wearing the LED rigged garments was known as the acting scripter and they moved about in the designated scripting space. The scripting space was determined by the intersecting fields of view of the two cameras (Architecture Machine group MIT 1988 video. David Sturman, Medialab, 1999). The movements of the LEDS would be captured by the cameras and then digitized in sequence, as Maxwell explained where "the resulting x-y coordinates from each camera and corresponding to a particular LED were used along with focal length information to reconstruct, by triangulation, the position of the LED in object coordinates" within the 3D environment. This coordinate data is then relayed to the animation system that would apply the information

to a basic stick figure representation of a human being with joints in the appropriate locations. While it was considered highly effective at tracking and retaining data for the capture of human locomotion, one of the biggest hindrances that stopped the widespread use of optical set ups in motion capture was the slow rate at which it was at the time able to render the data as well as the enormous expense of the technology. Nowadays this system is highly improved and the methods that had been previously used during the making of the Graphic Marionette are implemented in most optical motion capture systems available in the 21st century. (Sturman, 1999).Demonstration of the Graphical Marionette using the the Op-Eye system.

4 Evolving motion capture

4.1 Puppet motion capture system: First motion capture character

Many other revelations in motion capture came about in the 1980's as well, including Jim Hensons Productions mechanical puppet motion capture system. Tom Sito worded it in his 2013 book 'Moving Innovation: a history of computer animation', that in the year 1989, "Hanson had the idea for a little, floating, seal-type cartoon fellow, sporting a bright red top hat. He did not want the type of animation that was worked on for weeks after recording. He wanted to interact with the character in real time". It was Hensons strong belief that the best performance would be if the other puppeteers and actors could interact with the character in real time, that spontaneous improvisation was how they got the best source of comedy.

By tearing apart a dial box and utilizing the digital pots that live streamed data, they reassembling it into a hand wearable metal and foam puppet rig that they would strap onto the arm of a puppeteer, in this case Steve Whitmire, the puppet master behind Kermit the frog at the time. They named this creation Waldo C. Graphic. (Sito, 2013.)

Figure 9. Jim Henson demonstrating the mechanical puppet used for their 3D puppet Waldo.

When the rest of the cast would interact with ‘Waldo’ multiple screens would need to be present so that the others could see the character and know how they should be interacting with it in the scene. This Waldo character would appear on the screen as a low resolution character, as it too took much time and processing to produce a high resolution character for the sake of recording at the time. It would be at a later stage that the data would be refined and applied as a high resolution image. The process of rendering the high resolution character would take about 120 hour to create a 2 minute render of the charter, it was also during this stage that any additional animation such as secondary actions and overlap could be added as well. (Hensons, 2011; Sito, 2013)

Figure 10. Waldo C. Graphic, as the 3D puppet appeared on the TV show with other puppets.

The character Waldo C. Graphic is an important note in motion captures history, not just because it’s an early predecessor to later mechanical motion capture not adapted for the capture of human body motion, but can widely be considered the first true motion capture character since it was created for widely public accessible TV, and not just as a research experiment. (Sito, 2013)

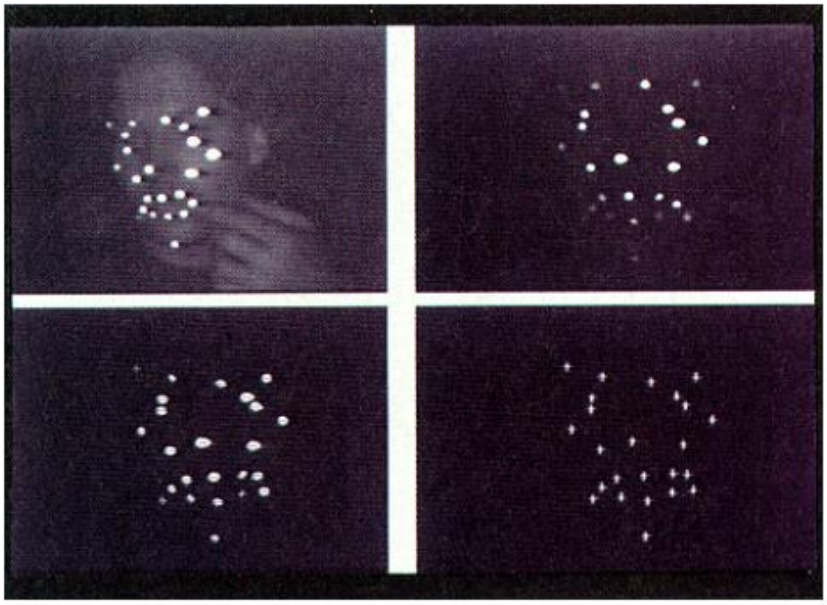
4.2 Facial motion capture

Facial captures development has greatly propelled animation quality in all mediums that utilize any form of animation. It has enabled us to be able to add life to characters faces that were previously missing. Secondary movement, twitches, flaws in human behavior that had not been present before and simply impossible to portray with traditional hand key animation, are all now possible and bring that nuance in life that was previously missing to the world of motion capture. In this field, the technology is ever evolving to the point where it is becoming almost indistinguishable from what is real and what is computer generated.

Like all things it has an origin, this piece of motion capture is not even that old in the greater scheme of things as the earliest records of facial motion capture can be found in a paper written by Lance Williams in 1990. In his paper, Williams brought forth the idea of performance driven facial animation, he stated “As computer graphics techniques raise (sic) to the challenge of rendering lifelike performers, more lifelike performance is required”. And thus a study was conducted by him and his team to see if they could produce a proof of concept. He hypothesized that using what he suggested as an ‘electronic mask’ would allow for the talents of the actors who portrayed characters to be better represented, and that by using a highly resolved human head model with photographic texture as well as a system devised to animate the model by tracking the face of a performer and applying it to the model head. (Williams, 1990)

First Williams had his personal friend Annette White’s head cast in plaster and then had that cast scanned and digitized by the company Cyberware .White later had to stand on a rotating turntable while a panoramic photo was taken of her by Prof. Andrew Davidhazy, so that the images could be converted into a cylindrical projection composite that would be mapped to the scanned head. While they considered using a mechanical system for tracking the facial movements, in the end they used a single camera optical system with passive markers made of reflective tape. The reason behind this was that optical tracking systems Lance said was that it offered the most leverage because it did not restrict the performer, it is simple, widely-available and promised ultimately to permit motion-tracking without special makeup or fidueial marks”. They used a plane of glass that they held between the camera and the actor at a 45* angle and then using light from a slide projector directed light towards the actors face and illuminated it. The reasoning

behind this was that the “light source and camera are coaxial (sharing a common axis). Since the light comes from the camera's point of view, the efficiency of retroreflectors in the camera's field of view is very high. This means that the contrast and brightness of the light could be controlled and adjusted so that the camera could focus its field of view to see the reflective markers and little to nothing else. Tracking the markers required the actor to stay relatively still and change nothing but their facial expression. (Williams, 1990)



.

Figure 11. Reflective tape placed on a models face and reflecting back at the camera.

A human operator would then select and indicate the position of each reflective dot in the first frame captured, after that the marks are automatically tracked in the following frames by following the path of the pixels. Using this method, they were able to map and record some short sequences that showed lifelike twitches and secondary motions which would be unlikely to arise in pure animation. The face animation was very limited to small movements and twitches. It was restricted by the fact that the eyes and mouth of the digital model could not open, so the result was ripples over the eyes and mouth, but this actually proved that mouth and eye movement would be traceable with some future refinement. (Williams, 1990)

The results, while unimpressive by today's standard, were at least a proof of concept during this time when it was thought to be too complex to map and gather data of the human facial performance for animation. It was proof of the possibility and that it was worth future investigation. From these beginnings, facial capture has only further advanced into a realm where it is becoming impossible to tell digital apart from reality. (Williams, 1990)

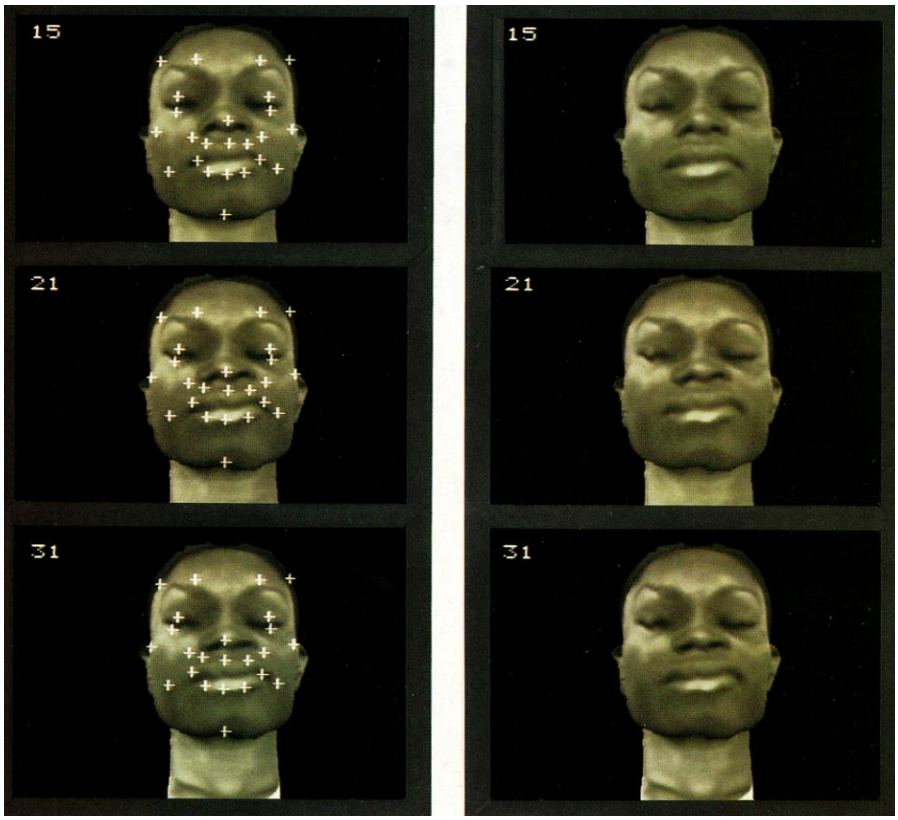
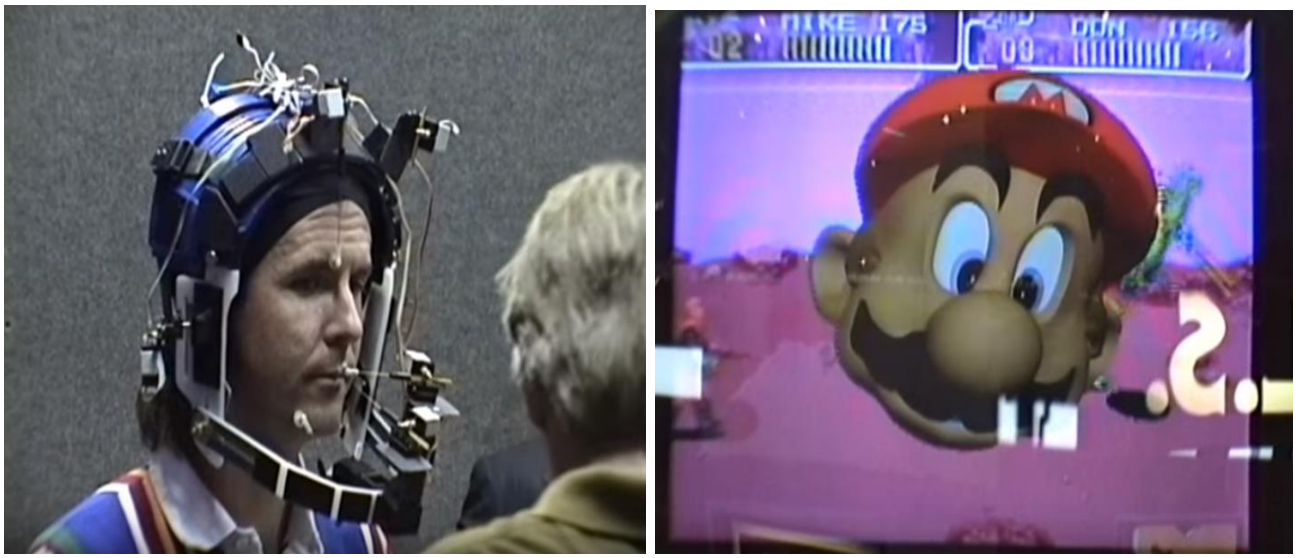


Figure 12. Tracking of the reflective tape on the models head. The markers being tracked to portray some of the facial movements of the model.

4.3 Real time facial motion capture

Many different types of facial animation had been experimented in over the years, some more successful than others. But one that was considered one of the first major successes in mainstream entertainment would have to be the Face Waldo system. It was a real time facial animation motion capture system used to help produce a more lifelike performance in animation, and notably, used to great success by Nintendo of America. (Henson, 2016.)

During the 1990's, a man by the name of Rick Lazzarini who at the time was a renown animatronic and puppetry expert working in the special effects industry with his company 'The Character Shop', had the idea of transferring his knowledge of inventing and using a remote animatronic system that had been called the Facial Waldo (the same system used for helping to develop Jim Hansosns Waldo C.Graphic) into the digital world. The animatronic system Waldo had been used previously so that one person could control all the complicated movements of the puppet while still being out of view. He adapted this into a virtual space by creating floating digital puppets with rudimentary mouth movements using mechanical sensors attached to the chin, lips, cheeks, and eyebrows, and electro-magnetic sensors on the supporting helmet structure. (Sturman,1999.)



.

Figure 13. Actor portraying Mario at the Nintendo booth at the 1992 Winter Consumer Electronics Show using the using the Face Waldo System.

Together with SimGraphics who helped provide the VR software, they developed the technology to at first be used for education purposes within the medical field, after they had been contacted by a Dr. Dave Warner at Loma Linda University Hospital, who was interested in utilizing their performance to comfort ill children who were going to be having procedures. This was probably considered a mild success, however the idea and technology didn't really garner large scale recognition or interest until 1992, when it was utilized at the Winter Consumer Electronics Show (CES) that year at the Nintendo of America booth. It was wildly popular at that year's CES, as it consisted of a screen that would portray Nintendo's company mascot Mario on it and a camera and microphone that was used to view the audience. The actor would be able to see and hear the public and then react to them, causing the face of Mario to react in real time as if it was the character who was reacting to the audience member. (Henson, 2016.)

5 First major uses of motion capture in entertainment

5.1 First uses of motion capture in movies

An important film to note in the incorporation of motion capture for the movies is a little known film known as Sinbad: Beyond the Veils of Mist released in 2000 (figure 13). Made by Pentamedia Graphics and Improvision Corporation in India, with motion capture being conducted in Los Angeles by the House of Moves Motion Capture. While there are some key framed traditional uses of animation, a majority was made using optical systems that tracked using passive reflective markers. The movie had to employ hundreds of animators to do the animation polishing that accompanies motion capture, as very rarely is motion capture entirely perfect in its recording. The film ultimately failed to find an audience and thus was a failure, but that doesn't erase its presence for being one of the first films that did in fact go to the big screen to use motion capture to such a vast degree (Gray,2014.)

Figure 14. Movie still of "Sinbad: Beyond the Veils of Mist" 2000, characters who used motion capture to create their performance.

While Sinbad was definitely one of its first, it wouldn't be the only film around the early 2000's that would try to make its break by using motion capture to conduct all of its animations. Following Sinbad in 2001 was a more widely known film known as Final Fantasy: Spirits Within by Square Pictures .This Sci-Fi was truly a breakthrough film, containing 1327 motion capture acted scenes and with remarkable visuals that are trade mark of the films name sake game counterpart. It took a 960 workstation render farm to render out the film's realistic appearance and a team of over 200 to create it over 4 years. However, like Sinbad, this film also failed to truly capture an audience and ultimately a film that cost $137 million to make only made $85 million in return (Gray, 2014.)

Figure 15. Final Fantasy: Spirits Within 2001 movie still.

Nowadays, one of the most recognized and successful character performances done with motion capture was released only barely 20 years after the first optically based system was developed. This famous character is often credited with truly pushing motion capture to new heights within the industry and setting an example of what could truly be

achieved. During 2002 the release of Peter Jacksons movie Lord of the Rings: The Fellowship of the Ring, the character Gollum played by Andy Serkis was revealed, and motion capture for cinema would never be the same again (Gray, 2014. Romano, 2014)

Figure 16.
Andy Serkis in an optical motion capture suit and the character he lends his acting to “Gollum” from 2002 Lord Of The Rings.

It was during the process of creating Gollum that the technology developed by WETA was created to allow motion capture actors to actually act on set and in real time with their co-actors, although be it in a skin tight lycra suit covered in reflective tracking balls. Before this it had been required that the footage would need to be filmed of the live action actors and then the motion capture would be filmed and added after in a separate studio often far away from the act sets (Gray, 2014.)

It’s crossing the line between clearly 3D characters into the realms where it is starting to pass as believable characters, captivating people in a way that even though they knew this was a computer generated character, they found themselves truly believing in the character and environment. One such powerful example of this is James Camron’s 2009 film Avatar. The thing that has set Avatar apart from its predecessors is not only its hyper realistic capture of movement and facial expressions, but its virtual camera as well. (Gray, 2014.)

Figure 17. Actors from the 2009 movie “Avatar” in full motion capture suit setups along with their onscreen 3D animated counterparts.

Figure 18.

5.2 Motion capture in video games

While over the passing year's motion capture had only been used for non-interactive entertainment or medical purposes, it was during 1994 that the games industry finally got a hold of the product and started to incorporate it into its media. Rise of the Robots was the first video game to incorporate the use of digital motion capture in its product. Rise of the Robots was considered a failure by the market's standards, but it allowed people to get a taste for what could be achieved in the gaming industry. It would set the trail for what would be many more games to follow, that would take a more cinematic approach towards their storytelling and even more believable movements in their characters (Özdem, 2016)

Figure 19. Characters from Rise of the Robots, who use motion capture to provide the animation data for the characters moves.

In the same year, the release of Virtua Fighter 2 was released, also using capture to animate their characters. In an 2013 interview, director Yu Suzuki mentioned how during this time in the world, no one was using motion capture in the video game industry, only really medical and academic fields, so utilizing motion capture in their game was a real first. (Suzuki , 2013; Virtua Fighter motion capture demonstration 2012).

Figure 20. Still from Souls Edge, the characters use motion capture.

Souls Edge was also a notable user of motion capture. Released a year later than Raise of the Robots and Virtua Fighter 2, they were actually the first game that used an optical based motion capture system with a passive marker set. Again, this was another fighting game that wanted to use the smooth fluid motions of motion capture to create a more believable and exciting player experience. (Fischer, 2019)

Figure 21. Characters from Virtua Fighter 2, who use motion capture to provide the animation data for the characters moves.

Possibly one of the revolutionary steps in games was actually the application of facial capture to characters. Nowadays many games utilize facial motion tracking data in their games to help give a more immersive emotional experience. One such game that did this is ‘L.A. Noire’ published by Rockstar Games in 2011.

Figure 22. The 32 camera set up used to capture the face and motion of actors faces for the game “ L.A Noire”.

While facial capture animation had been present in games for a while before the development of this game, what makes L.A. Noire interesting is the way that they chose to do their animation. It was done using ‘Motionscan’. This capture system didn’t collect data in the usual sense of an optical passive system or mechanical one, but instead used a 32 camera set up that used the footage from all of those cameras, situated at various angles, to replicate an actor's face.(Milian, 2011).

Servers automatically map the faces and implement the animation from the actor to the texture and mesh. The goal was to make the animation and faces as realistic as possible. It did though require the actor to stay mostly stationary or else the software loses the full picture, as well as for make up to be applied to the actor depending on the scene since it was mapping images of the actors face rather than tracking the movement data at points. Mark Milian of CNN in 2011 mentions that actors “must plan on several hours to go through hair and makeup… If a character has taken a beating, black eyes are applied with powder and eyeliner; for especially bad smack downs, the actor chomps on a blood capsule… burn victims were in there for 4 hours getting prosthetic stuff.(Milian, 2011)

While interesting, this system has never been used since. It proved to be too expensive, costing 192,000$ for the system as well as creating bulky files, roughly around 1 GB of

data per a second of motion capture recorded. The data also couldn't be edited, which meant that if you wanted to change a line in the script or modify the characters behaviour, you would need to record all new data. (Borkowski, 2017; Milian, 2011.)

Figure 23. The actor, the output and the in game render of the character using the "motion Scan" system in L.A Noire.

6 Modern advancements: Shining a light on motion capture

6.1 The virtual camera

During the production of James Camron's 2009 epic AVATAR, WETA Digital, a frontier company in the art of making digital movie magic, came to the concept of the virtual camera. It was brought about in order to create the same organic and natural feel of traditional movie camera movements for a film that was almost entirely CG. This virtual camera that was made consisted of a 9 inch LCD screen that had a steering wheel attached to it that functioned as a control for moving in the digital environment as well as zooming and panning and passive reflective markers like that used in optical system motion capture. Because of the markers attached to the device, it was possible for the director to be able to walk around the environment and capture organic camera movements, but it was also possible to fully control the camera via the controls in the digital space, not much unlike a video game free camera. (Failes, 2018; Green, 2019)

Figure 24. On screen visualization of an environment with the actors using the virtual camera during the production of the 2009 movie "Avatar".

The benefits from being able to have the virtual camera where that the director and camera operators would be able to view on the LCD screen what was actually happening in the digital world they were trying to create in real time. So instead of seeing the actors in their motion capture suits, they would see the actors as they would appear in the film, as their CG counterparts as well as the environment that the characters would be portrayed

in. Additionally to this, a scene could be captured in the studio and then with the use of the virtual camera in a motion capture studio environment the camera work can be done separately.(Green, 2019)

This does in fact speed up the production process, the actors are no longer needed for that shoot since the performance has already been captured. Now the director and camera operators can modify their camera frames and shoots as they like without having to have the actors redo the performance each time. Also with this advance in technology movie sets do not need to be built so much for scenes only involving the digital characters as it is already constructed digitally. With this kind of advancement Cameron was able to capture what he credited as a “form of pure creation” within the industry (Gray, 2014. Failes, 2018. Green, 2019.)

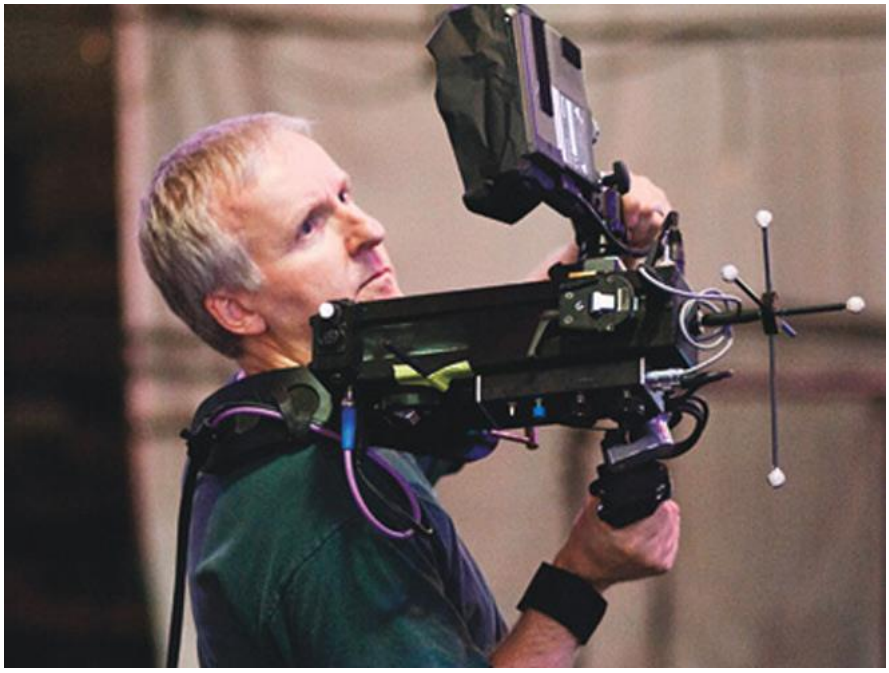


Figure 25. James Cameron using a virtual camera.

Nowadays the virtual camera is a system staple for almost any cinematic experience, be that for movies or video games. The camera can be heavily customized to suit the wielder, be it shoulder mounted like a traditional camera, something that someone can hold off the hip or rather often it is actually nothing more than a screen with diode and joysticks attached to the side that allow rapid changing of digital lenses, focus, pan, zoom and movement in the digital world. (Failes, 2018; Green, 2019)

6.2 Under water motion capture

Underwater motion capture has been about for a while, at least since 2011 (clear obtainable timelines are not currently present). The motion capture development and supplier company Qualisys with collaboration with Chalmers University of Technology in Gothenburg Sweden lay claim to have been the first to have developed a functional underwater motion capture system. Its development was originally primarily for the study of biomechanics of professional swimmers, which is much the same with how many new motion capture systems come to be, as a result of the health and sports industry. Whilst it was created originally for athletic study, it implantation in other fields has been successful, such as the use of gathering data within NASA on how humans deal with microgravity which is conducted underwater, the fishing industry, underwater vessels and of course the entertainment industry which is normally where such systems get their moment in the spotlight, as most developments in entertainment gathers more exposure. (Qualisys 2019; Prof. Westman, 2013)

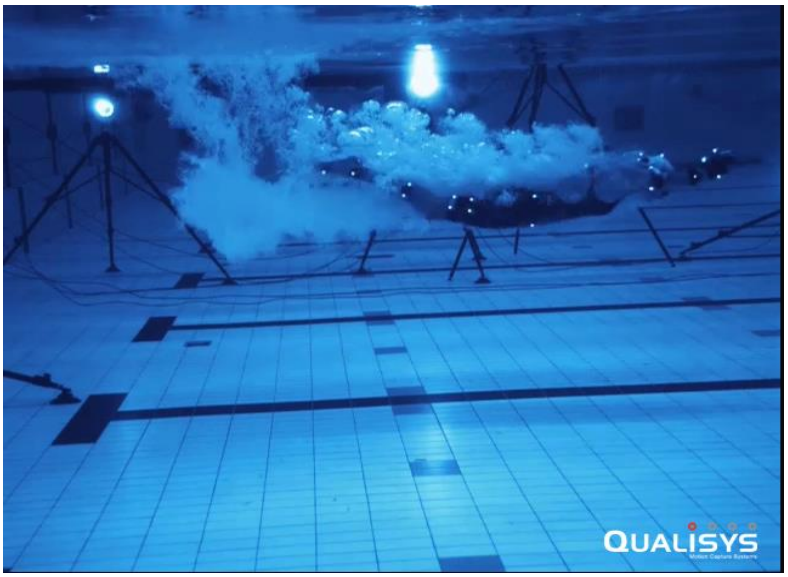


Figure 26. A swimmer using the Qualisys underwater optical system to record their movement underwater.

Typically underwater motion capture systems opt to use an optical passive marker system, marker less systems do exist however the quality of accuracy is not as refined as the passive marker system. How underwater motion capture differs from its dry counterpart is that it needs to navigate the thousands of bubbles created under water that act as reflective spheres in the environment, as well as the reflective surface of the top of

the water as there are often dry mounted cameras placed from above. These two obstacles are the biggest hindrances of underwater capture and how this has been navigated is that the infrared cameras, that are houses in waterproof containments, have strobe cyan lights instead of the typical IR light to help pick up the passive markers with greater accuracy, the cameras also have active filtering software that help to minimize the reflections. (Cameron,2017 ; Bui, 2017; Qualisys, 2019; Prof. Westman, 2013)

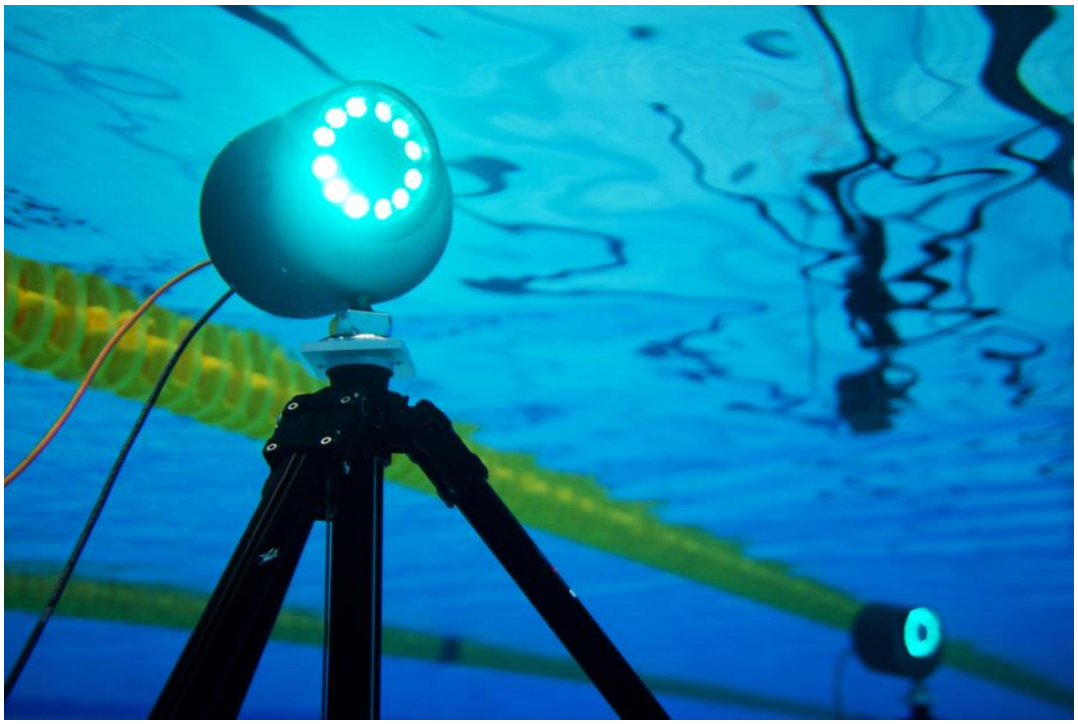


Figure 27. One of the cameras used for the Qualisys underwater motion capture system.

6.3 Mobile phone motion capture

Nowadays, motion capture technology is actually very accessible to the general public and at a very affordable range. It has been compressed to the point that today's most smart phones support some version of motion capture, it can actually be found in some of the most common apps of this time such as snapchat, Instagram and Facebook. During SIGGRAPH 2018, Kite & Lightning revealed that they had implemented the use of the Iphone X to do facial capture, as well as Xsens motion capture suit for body capture for their game Bebylon Battle Royal. At the conference co-founder of Kite & Lightning, Cory Strassburger explained the reasoning behind using the off the shelf technology in their studio to produce their data was because of “challenge is that we’re a very small company. So, I needed a mocap setup that was fast and easy to use, requiring little to no clean up and which would still hit our minimum quality bar”. In 2017 when he started development of this system, he originally started with strapping the Iphone X to a paintball helmet with a gopro mount and an Xsens suit to record the data first, it was then reassembled in Autodesk Maya and exported to Unreal Engine for the game. In 2018 Strassburger was approached by Chris Adamson to see about submitting the setup to Real-Time Live! SIGGRAPH which is a showcase of real time technology for games, movies, simulations, virtual reality and augmented reality (Strassburger,2018).

Figure 28. Cory Strassburger demonstrating their mobile phone motion capture system at SIGGRAPH 2018 for their game “Bebylon Battle Royal”.

The following is a quote from Cory Strassburger from the Siggraph Kite and Lightning 2018 conference video

“Using IKINEMA live action to retarget and retune in the background, as well as take care of mesh clipping and foot sliding, and ARKit installed the Iphone X with 51 blend shapes for the face, they were able to produce results of real time capture of the body and face to a rigged character set up in Unreal Engine. This was not the only thing that they managed to achieve through this new set up. Because all of these components have the ability to stream directly into Unreal and the capture space is only limited by the wifi. It allows for a completely mobile motion capture set up that also permits to have multiple rigs and actors active and streaming into the same scene at the same time. If you so desire, all you need to do in Unreal is hit record on each actor's input sequence recorder and Unreal will capture and retain the data from the performance.” (Strassburger,2018)

Nowadays this technology is also readily available for the everyday consumer and does not always require advanced knowledge or support from a major company/studio to create your own motion capture data. The company RADiCAL also offers a software called MOTION that can be cheaply bought and easily used by anyone who could be possibly interested in motion capture. Although unlike the system that Kite & Lightning developed with Unreal and ARKit, RADiCAL’s MOTION only produces body capture. But it is quick and easy to use, giving you the possibility to capture motion data without a suit as well. (RADiCal 2019)

Figure 29. RADiCal’s demonstration of the mobile phone motion capture system.

7 Reflections

7.1 Authors thoughts on Motion capture

On this personal account of researching for this paper, I speculated that more recent advancements would be harder to find information on, and I was right. With regards for early development, the process and thoughts behind the technology was easy and openly available to me in a well-documented manner, due to the fact that many of the advancements were the results of university studies as well as for medical purposes. I wasn't wrong on assuming that finding the data behind more recent achievements would be harder to gather. Lots of the information was guarded due to the fact that the motion capture industry is nowadays highly commercialized. Maintaining information on the development of technology has competitive financial benefit in today's landscape of selling the technology to the entertainment and medical industry since it helps maintain a competitive edge.

I have been fortunate that I work within the entertainment industry and my place of employment is publicly known to use motion capture extensively. I was able to contact individuals directly for source information that I at first found difficult to locate. It also gave me access to first hand opinions from others that were veteran's regarding the use of motion capture.

Just as mentioned within the introduction of this paper, the history of motion capture has been dense, meaning that the development of this technology in the greater scheme of things is relatively short, but also incredibly rapid. It is my opinion that it will probably slowdown in regards to entirely new developments, but instead focus on improving the current technologies that are currently available. Things like increasing the fidelity of facial capture and making capture suits less and less intrusive or restrictive (perhaps even unnecessary) in the future will all be things that are already in development and will probably be unveiled in the near future.

8 Refrences

Animac test clips 21.7.2013 <https://www.youtube.com/watch?v=TS1WY4H4Dvk> (read 03.02.2019)

Architecture Machine group MIT 1988 <https://vimeo.com/43287234> (read 5.05.2019)

Batchelor, James 2016. MCV <https://www.mcvuk.com/development/what-hollywood-is-learning-from-video-games-mocap> (read 03.10.2018)

Bernal, Yaritza 2017.Development of Underwater Motion Capture System for Space Suit Mobility Assessment, NASA. https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20170009427.pdf (read 10.09.2019)

Cameron, James & Bui, Hoai-Tran 2017. Slashfilm <https://www.slashfilm.com/underwater-motion-capture/> (read 04.05.2020)

Crook, Ian & Beare, Peter 2016. Motion Graphics: Principles and Practices from the Ground Up. London, United Kingdom: Bloomsbury Publishing

Digital worlds,16.09.2016 <https://digitalworlds.wordpress.com/2016/09/07/from-motion-capture-to-performance-capture-sampling-movement-in-the-real-world-into-the-digital-space/> (read 04.05.2020)

Failes, Ian 2018. Cartoon Brew <https://www.cartoonbrew.com/vfx/how-ready-player-one-combined-virtual-production-and-motion-capture-tools-to-create-digital-characters-157821.html (read 02.09.2019)>

Feleischer Studios, 2019. <https://www.fleischerstudios.com/rotoscope.html> (viewed 27.10.2019)

Fischer, Ron 2019_. Industry History. Motion capture society <http://www.motioncapturesociety.com/resources/industry-history> (read 16.07.2018)

Funk, Walter 2010.Animatic: Analog 3D Animation .Veritas et Vius http://www.hologlyphics.com/Animac_Analog_3D_Animation.pdf (read 04.05.2020)

Freeman, Will 2013_.Motion capture: Moving with the times. MCV Develop <https://www.mcvuk.com/development/motion-capture-moving-with-the-times> (read 04.05.2020)

Ganeranx 2017 . La Noire 2011 (facial motion capture) <https://youtu.be/kvCxyDwHeag> (viewed 02.02.2019)

Gray, Ali 11.06.2014. A Brief History of Motion-Capture in the Movies. IGN <http://www.ign.com/articles/2014/07/11/a-brief-history-of-motion-capture-in-the-movies> (read 18.07.2018)

Green,Weston 2019. What It's Actually Like to Use WETA Digital's Virtual Camera, Inverse <https://www.inverse.com/article/52727-what-it-s-actually-like-to-use-weta-digital-s-virtual-camera>
<https://www.youtube.com/watch?v=I-ctZRxLEN4> (read 15.09.2019)

Guimarães Fernando. Puppetry: An Overview <https://books.google.fi/books?id=WK7T50f8i-oC&pg=PA70&lpg=PA70&dq=Lee+Harrison+III:+analogue+circuits+and+cathode+ray+tubes&source=bl&ots=_dxcGdV-VQ&sig=ACfU3U3EmRnpvnb-DQ2rB3-rOcHl5XHreQ&hl=en&sa=X&ved=2ahUKEwi3197ByuDgAhW986YKHbB_AXcQ6AEwAHoECAAQAQ#v=onepage&q=Lee%20Harrison%20III%3A%20analogue%20circuits%20and%20cathode%20ray%20tubes&f=false> (read 04.05.2020)

Hensons, Jim 2011. Jim Hensons Company https://www.youtube.com/watch?v=dP6TUB7KQc4 (viewed 04.05.2020)

Jain,Lakhim C 2008. New Advances in Virtual humans: Artificial Intelligence Environment. New York, United States of America. Springer Publishing

Lonkar, Sagar 2019. Types Of Motion Capture <https://sagarlonkar.wordpress.com/about-2/motion-capture/types-of-motion-capture/> (read 22.07.2018)

Maxwell, Delle Rae 1983. Graphic Marionette. MIT. https://dspace.mit.edu/bitstream/hadle/1721.1/76176/11521682-MIT.pdf?sequence=2 (read 08.07.18)

Meta motion 2020, Animazoo motion capture suits <https://metamotion.com/gypsy/Animazoo.html> (read 29.04.2020)

Milian, Mark 2011. The 'amazing' facial capture technology behind 'L.A. Noire'. CNN <http://edition.cnn.com/2011/TECH/gaming.gadgets/05/17/la.noire/index.html> (read 03.09.2019)

Nogueira, Pedro 2011. Motion Capture Fundamentals. Faculdade de Engenharia da Universidade do Porto
<https://paginas.fe.up.pt/~prodei/dsie12/papers/paper_7.pdf> (read 9.10.2018)

Qualisys 2019
<https://cdn-content.qualisys.com/2018/05/AN_Underwater.pdf>

<https://cdn-content.qualisys.com/2014/02/PI_Underwater-camera.pdf>

<https://www.qualisys.com/hardware/oqus-underwater/>

<https://www.youtube.com/watch?v=jv1jW2BqzI4> (read 9.10.2019)

Qually,Graham, 2019. Beyond Capture <http://www.beyond-capture.com/> (read 27.10.2019)

Radical solutions 2019
<https://getrad.co/> (read 10.092019)

Romano, Nick 2014. History of Motion-Capture Performance on Film .Screen Crush <http://screencrush.com/motion-capture-movies/> (read 04.05.2020)

Rukkila, Emmi 2018. Redesigning Super Mario Creating realistic versions of some characters and animations featured in Super Mario franchise, XAMK South- Eastern Finland University of Applied Sciences. Available to read at: <https://www.theseus.fi/bitstream/handle/10024/146703/Rukkila_Emmi.pdf?sequence=1&isAllowed=y> (read 03.10.2019)

Sito, Tom 2013. Moving Innovation: a history of computer animation. London, United Kingdom:The MIT press

Sturman, David 1999. Motion Capture History. Medialab https://www.siggraph.org/education/materials/HyperGraph/animation/character_animation/motion_capture/history1.htm (read 23.05.2019)

Suzuki, Yu 2013. Virtua Fighter 2 1994 interview with Yu Suzuki (translate from Japanese), 2013 <https://virtuafighter.com/threads/vf20th-anniversary-site-the-interviews.19637/> (Japanese interview) <http://vf20th.sega.jp/interview131211.html> (read 09.10.2019)

Welt, Eigenwelt Der Apparate 1992. Lee Harrison III. Pioneers of Electronic Art, Ars Electronica <http://www.vasulka.org/Kitchen/PDF_Eigenwelt/pdf/092-095.pdf> (read 04.05.2020)
<http://www.vasulka.org/Kitchen/PDF_Eigenwelt/pdf/009-010.pdf> (read 04.05.2020)

Prof.Westman, Gunnar 2013. Qualisys 2019. Chalmers University of Technology <https://www.youtube.com/watch?v=qtXz6qocciM> (viewed 10.09.2019)

Williams, Lance 1990. Performance-driven facial animation. SIGGRAPH '90, Dallas <https://dl.acm.org/citation.cfm?id=97906> (read 28.10.2019)

Vicon, 2019.
<https://www.vicon.com/what-is-motion-capture> (read 08.10.2019)

Virtua Fighter motion capture demonstration 2012.
https://www.youtube.com/watch?v=FWv5olHltqc

Özdem, Anil 2016.Introduction To Motion Capture. Bilkent University. <https://www.academia.edu/28101363/Introduction_to_Motion_Capture> (read 20.07.2018)

8.1 Interviews

Blåfield, Henri 2019. Head of Animation. Remedy Entertainment https://www.remedygames.com/

Morris, Philip 2019. Representative of Audiomotion Studios, Audiomotion Studios http://www.audiomotion.com/

Noone, Tom 2019. Lead Animator. Remedy Entertainment
https://www.remedygames.com/

Haskayne, Stuart 2019. Studio Manager. Centroid Motion Capture https://www.centroid3d.com/

Brun, Rémi 2019. Representative of Mocaplab. Mocaplab
http://www.mocaplab.com/