

# COMPUTER ANIMATION AND GAMES I – CM50244

## Coursework 2

### Motion Capture:

The early technologies that shaped motion capture as we know it

### Unit Leader:

Dr. Yongliang Yang

### Name:

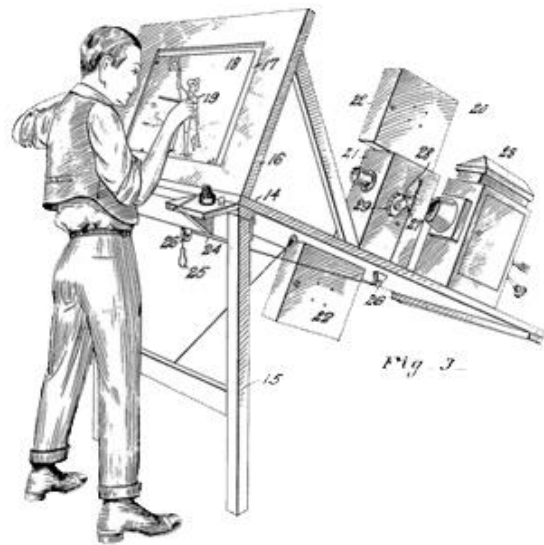
Constantinos Theophilou

### Degree Programme:

MSc Digital Entertainment  
University of Bath

## 1. Introduction

Animation is the concept governing many aspects of today's Digital Entertainment sector, with applications in computer animation, game development and of course visual effects. The notion of animation opened thousands of opportunities for every sector mentioned, since without the process of animating living or even inanimate objects, the contents of the screen would stand still. Motion capture (mocap), even though considerably argued, is responsible for taking animation to the next level, by giving imaginary characters realistic and natural movement. Mocap was introduced well before computers ever existed, with the first steps done completely by hand, calling the procedure "rotoscoping", (Maher, 2017). Max Fleischer,



*Figure 1 – Rotoscoping at its very early stages*

1914, came up with the idea of rotoscoping which is actually tracing over film footage of actors, frame by frame, where each paper represented a new frame, as can be seen in Figure 1. This technique has first surfaced by Disney Studios in their feature animated film "Snow White and the seven dwarfs", in 1937. As one might have understood, rotoscoping is a timely procedure, but with the birth of the computing world, mocap has come a long way since then. This survey will portray the technologies that contributed in shaping the existing motion capture technology with the use of academic literature, which will act as fundamental descriptions of the various mechanisms, as well as commercial products using the said technology, by separating implementations based on their efficiencies. As records of the first technologies are hard to find, certain references will direct the reader to videos of interviews or documentaries about a certain technique.

## 2. Optical Systems

Mocap systems that include any type of visual aid to achieve naturalistic movement are considered optical systems. These could vary from utilising pre-recorded footage, a sequence of images or even real-time streaming mechanisms and live acting. Revisiting rotoscoping, as an optical system, mocap was formally introduced into the digital world by Rebecca Allen's animation, based at the New York Institute of Technology and Computer Graphics Lab at the time. Her work consisted of a 3-dimensional model mimicking the moves of each and every body part of a female dancer to create a sequence of frames that together made Twyla Tharp's "The Catherine Wheel", (Allen, 1983). To produce a semi-realistic mimicking of the dancer's motions, Allen, utilised a half-silvered mirror at an angle of 45 degrees with regards to her CRT monitor. This allowed her to superimpose the actual footage with her computer screen where she could then create the 3D wireframed model of the dancer from scratch, and eventually controlling



*Figure 2 – Rebecca Allen's rotoscoping of "The Catherine Wheel".*

the model's limbs using programming to assign keyboard inputs and joystick rotation to body part joints' translations/rotations. Using the whole sequence of the footage, she then moved on to superimposing frame by frame and with the articulated input mechanism, move the 3D model to fit the pose of the dancer, entailing the use of rotoscoping. The results can be seen in Figure 2. Optical mocap systems can be further categorised to marker-based and marker-less, with the details to be explained in the following subsections.

## 2.1 Marker- based motion capture

Marker-based mocap uses two types of equipment: the camera(s) and the markers. Markers could be any kind of light emitting material or reflective material. This is where a distinction between passive and active markers is made. Since the mechanisms governing these two methods are the same, in some degree, they will appear in alternating order.

During the mid-1980s, mocap was thought to be the centre of attention and therefore researchers were prone to find the best possible approach towards it. Competing university giants MIT Architecture Machine Group and New York Institute of Technology and Computer Graphics, invented the first sightings of optical tracking of the human body using flickering LEDs, sparsely placed on a garment to be worn by the actor, able to emit light, coupled with a series of cameras to record the movement of the emitted light, (Ginsberg and Maxwell, 1984). Ginsberg et al. describe the process of their Op-Eye optical tracking system as follows: "The scripter (actor) moves about in a scripting space wearing a garment with light emitting diodes or LEDs, located at joints and other bony prominences". The scripting space is defined as the space in which intersecting cameras' field of view can record what they call "scripting by enactment". In other words, the method used here is somewhat similar to techniques used in most recent developments of mocap, especially with the use of trackers. The footages recorded by the cameras are compared in a certain degree that can bestow a 3D position of each LED through a time span. Nonetheless, in these early stages of optical motion tracking, there was a limit on how often the trackers' locations could be recorded, thus placing a cap on the frame rate at which the animated character would sequence motion. Additionally, with the limited number of cameras used at the time, some of the trackers were constantly occluded by the body of the actor and with their low-resolution, the trackers could usually be undifferentiable when in close proximity. Their implementation could record positions from about 12 trackers at a time, which is only a fraction of a present development's ability. Once the position of trackers was captured, without any post-processing, the information was directly encapsulated to drive a stick figure, the "Graphical Marionette", Figure 3, into mimicking the same sequence of movements as the "scripter", and stored for future use with a more comprehensive computer character. Mentioned implementations using LEDs or infrared light emitting diodes (IREDs) as markers on the actor, fall into the category of active markers, since they could emit their own light towards the cameras.



*Figure 3 – The Graphical Marionette, active marker motion capture by Ginsberg et al.*

The main distinction between passive and active markers is that passive markers are just simple reflective materials placed on the actor. This entailed the use of another light emitting source behind or around the cameras, so that light could be reflected back to them from the markers, (Frey et al., 1995). Furthermore, passive markers could be any type of material that can be recognised by a computer such as QR codes, even though this is a more recent development, (Sementille et al., 2004).

An example use of such a system for commercial product placement can be witnessed in the TV advertisement called Brilliance, for the National Canned Food Information council, (Allen, 1985). The commercial, aired in 1985, produced by Robert Allen and Associates, played a significant role in the birth of Computer-Generated Imagery (CGI), as well as capturing motion development. They used a technique they called “Brute-Force Animation”, which is directly related to Ginsberg et al. solution. They proceeded with filming a live model, passing the footage into a computer and using software they tracked the position of reference points, painted dots on the model’s limbs, creating a stick figure animation, they denoted as “vector graphics”. Using raster graphics, they rendered their polygonal mesh computer character, the “Sexy Robot”, producing ingenious reflections and together with the positional vectors acquired, the robot seemed life-like on the big screen at that year’s Super Bowl, with natural movements and postures.

Although a motion capture technique used until present time, marker based systems back then lacked reliability. Using a single camera would mean that some of the markers, either active or passive, would be occluded by the actor wearing the garment. Upgrading to two cameras, improves the 3D orientation and position of each marker using triangulation algorithms proposed by the computer vision community, as well as the occlusion of markers problem since the second camera will be setup to capture obstructed body parts. Further increasing the number of cameras may improve results by comparing between the views to accumulate on a marker position, but doing so in a great volume produces an alternative problem which is exponentially increasing the computational complexity. The same stands for the markers. More markers mean more degrees of freedom are captured from the actor and therefore more naturalistic motion recorded, but this may lead to the problem of difficulty distinguishing which corresponds to what body part. Due to the complexity involved, specialists in the field are required to operate optical systems and therefore cost more than other alternatives up to this day.

## 2.2 Marker-less motion capture

Optical motion capture is not limited to marker-based systems. In fact, the most common off-the-shelf motion capture systems today are marker-less. Marker-less systems do not require the user to wear reflective or light emitting trackers on their body, but instead using image processing and depth estimation, such systems are able to record motion using only camera(s). Named systems are not a technology of the past, but nonetheless they constitute a great volume of the mocap success up to this day. It was not until the late-2000s, that a solution was found for such a technology to exist and all papers surrounding this are involving the first fully functional and extremely fast system of its kind,

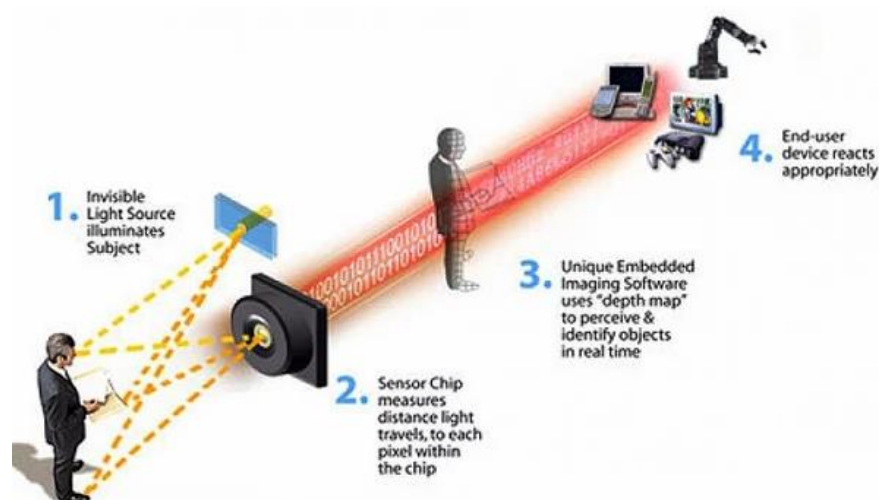


Figure 4 – How the different components of Microsoft’s Kinect establish a depth reading.

Microsoft’s Kinect for the games console Xbox, (MacCormick, 2012) (Cong and Winters, 2018). The Kinect consists of three specialised pieces of equipment; an RGB camera with resolution of 640x480 recording at 30 frames per second, a depth sensor and a multi-array microphone, Figure 4. With the use of

machine learning and artificial intelligence, Microsoft was able to build a huge database of 3D human models, corresponding to the various settings of human body characteristics, such as height, weight and gender, as well as the flow of the clothing that the user wears. It decides on which model to use every time a user wishes to interact with the device using a depth map to identify these characteristics, which are illuminated by an embedded, but yet invisible, light source in real-time. After the CPU-exhaustive building of models and the use of random forest algorithms to achieve the desired data capture, the data collected by the device is cross-compared to the stored 3D models, and it can infer depth from the deformation of the two datasets. This could not be done previously as computer vision was not at its finest during the 1980s or 1990s, but with the aid of Zhang et al. and his principle of 3D reconstruction using structured light, this was made possible in the mid-2000s, Figure 5 (Zhang, Curless and Seitz, 2002).

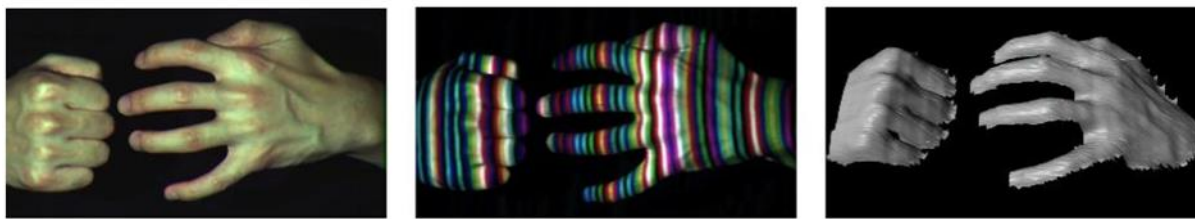


Figure 5 – Zhang et al.'s structured light reconstruction method.

### 3. Non-Optical Systems

Motion capture is not restricted to visual data capturing from images or videos. As a matter of fact, technologies that helped shape the notion of mocap as we know it, consisted of non-optical systems. There are various mechanisms that can be used for non-optical mocap each with a different apparatus approach. Due to the ability to not be limited from visual data acquisition, thus needn't require post-processing and processed in real-time, such mechanisms are still used for haptic feedback in virtual reality systems or for gaming controller input control.

#### 3.1 Mechanical Motion Capture

Mechanical motion capture deals with the recording of limb position using a mechanical suit or by controlling the orientation and shape of a mechanical arm, fixed on an object other than the actual user's body, eventually mimicking the articulation of the joints of a character. This technique focused mainly on real-time applications, applications that could provide translation of joint movement directly to computer generated characters for advertising campaigns during major gaming and animation conventions. Real-time capabilities for mechanical mocap arose from the fact that data acquired did not require further post-processing to be used. Moreover, such technologies, immigrated in the haptic feedback sector for virtual reality immersion, as VR began to dominate the industry further along the ages.

##### 3.1.1 Exoskeletons

Motion capture suits, or in this case exoskeletons, were initially consisting of a series of electrogoniometers; apparatus using resistance to determine joint angles, attached to the bendable joints of a human body, as imagined and implemented by a group of computer scientists at the Simon Fraser University, (Calvert, Chapman and Patla, 1982), Figure 6. Once a joint, say elbow, was bent, the electrogoniometers would bend at the same time, in the same speed, thus recording natural movement passed as input, which could then be used to drive computer generated characters, producing animation sequences. Nonetheless, the exoskeleton was unable to track the real-world coordinates of the limbs, as they could only be used to capture certain movements, e.g. bends. The study focused on

choreographic movements as well as assessing motion abnormalities, and even consisted of integration of the method mentioned (direct input from instrumentation), with notation from movement notation systems (indirect input) such as Kinetography Laban, a type of dance notation. By integrating both mechanisms, Calvert et al. were able to incorporate the output in such a way that could also be used for future reference, producing more life-like animation.

Since then a lot of companies tried to reproduce their own implementation of exoskeleton motion capture techniques, all following the structure of the implementation mentioned above. Mocap focusing companies started calling these types of suits as Waldo devices. In 1988, waldos became the, what thought at the time, next big development in mocap technologies. Specifically, computer animation giants Pacific Data Images, looked to employ a team that could produce an accurate waldo suit for their upcoming animations feature films.

After two tries, Rick Lazzarini's company, The Creature Shop, were able to produce an upper body and head exoskeleton based on optical potentiometers on each joint, (Menache, 2000). Carl Rosendahl, a part of the team working on the waldo, explained the difficulty during production due to the noisy output of the analogue parts of the suit, leading to the replacement of analogue components with digital counterparts. The suit was used in the first ever motion picture that successfully incorporated a digital character called Toys, starring Robin Williams.



*Figure 6 – The first of its kind exoskeleton for motion capture by Calvert et al.*

As years passed, mechanical mocap helmets entered the race for most natural motion in animation and gaming facial expressions. SimGraphics teamed up with Nintendo to produce an interactive Super Mario to showcase at the SCES convention in 1992, (Silicon Graphics and Nintendo, 1992). The major development step taken to produce this new technique was the construction of a helmet, called the "Face Waldo". The helmet consisted of various mechanical sensors attached to the primary facial parts that could reproduce human-like facial expressions, in turn portraying perfectly human emotions. These sensors were placed on the chin, lips, cheeks and eyebrows of the actor, that eventually controlled Mario's corresponding digital facial parts. These types of sensors were used in concert with other electro-magnetic sensors, responsible for tracking the facial parts not covered by the mechanical sensors, to gather the much-anticipated input consisting of real-world coordinates of the facial parts of the actor. The input was directly mapped to Mario's face and together with the "flying mouse", a mouse that when moved in the air could alter the 3D position of Mario's face on the workstation, and the voice of the actor, the famous cartoon and video game character came to life.

### 3.1.2 Digital Puppetry

With increasing interest in real-time motion capturing and animation rendering during the late 1980s, the term digital puppetry became the centre of attention. Digital puppetry is the process of controlling any 2D or 3D computer-generated character's movements and other actions such as speech, in real-time with the use of computers paired with a mechanical arm. During a SIGGRAPH convention in 1988, "Mike, the Talking Head", a digital character, played a leading role in the first ever digital live interactive performance, (Menache, 2000). Silicon Graphics was keen on showcasing their real-time processing and rendering of their so called "4D models", with the fourth dimension being time, where they employed the likes of Wahrman and deGraf to produce an interface in which the input from a puppeteer device, the device that resembles a glove to move the lips and produce the "mouthing"



motion of lips during the exclamation of words, was directly connected to the rendering engine of the polygonal character, Mike, and was able to interpolate frame instances between each motion. The face mesh was produced using a 3D digitizer to scan the face of Mike Gribble, hence the name of the character, and consisted of over 200000 digital data points. With a speech recognition system and the glove mechanism, as well as scans of the real Mike as he mouthed certain phonemes, the lot were able to give digital Mike a personality. The notion of digital puppetry governed the mocap development process with other projects such as Waldo C. Graphic, another digital character controlled and animated in real-time using again the Silicon Graphics 4D workstation. Jim Henson Productions and Pacific Data Images were able to construct a mechanical arm with eight degrees of freedom, to control the position of the character and its movements in the screen, and a two-piece oven-mitt kind of mechanism attached at the very end of the mechanical structure, to control the lips of Waldo C. Graphic, (Jim Henson Productions, 1988), as can be seen in Figure 7.



Figure 7 – Digital puppetry on live television featuring Waldo C. Graphic with the mechanical arm on the left.

### 3.2 Inertial

Inertial mocap systems are considered crucial in the development of gaming controllers, even though their history is not of the same length as the other systems mentioned in this survey. Inertial systems usually consist of inertial sensors and Inertial Measuring Units (IMU), which in turn can track rotational rates and local coordinate positioning using gyroscopes, accelerometers and magnetometers, (Fischer, 2000). As sensors and electrical components such as the gyroscopes or accelerometers, became smaller and cheaper over the last couple of decades, the inertial mocap system was made possible. Just like marker-based mocap, the more IMU sensors on the movable object, the more accurate and natural the data acquired is. A patent for use of inertial sensors for exercising while gaming was claimed in 1999,

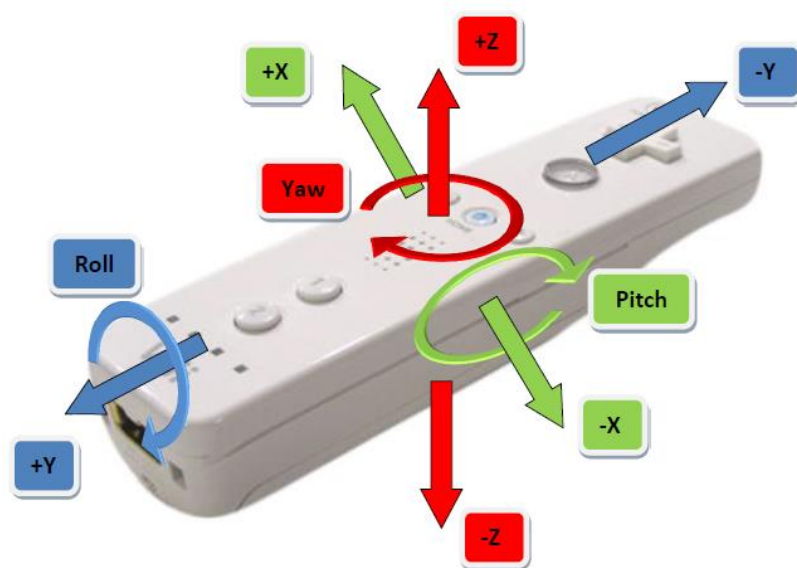


Figure 8 – The Wii-mote and the way orientation is captured using the incorporated IMUs in full 6DOF, (Kirkham, 2010) .

(Walton, 1999). In the patent, Walton suggests that gaming should not just be done while sitting, thus delivered ways of incorporating athleticism with the use of inertial technology. Years later this could be bought in the shape of a game station and more specifically the Nintendo Wii console with the so-called Wii-motes that could provide full 6DOF tracking, Figure 8. The controllers used IMUs to track the orientation of the controllers relative to the IR transmitter which would be

placed on top of the TV unit, used to track the position of the controllers. Big name gaming companies followed the example set by Nintendo and produced such controller or even incorporated accelerometers and gyroscopes into their existing controllers to account for motion tracking. Newer development in this area can be seen in the form of a suit consisting of multiple IMUs to be fully immersed in the VR world with example such that of Xsens MVN, Figure 9 (Roetenberg, Luinge and Slycke, 2009).

This type of system has the disadvantage of suffering from inertial drift, given a magnetic interference from common objects or even the unavailability of gravity, which in turn can ruin the accuracy of results, producing a jittering effect in the data collected. Nonetheless, sensors of this type are quite cheap these days and constitute to many areas of computer science when incorporated with other academic fields.



*Figure 9 – Xsens MVN by Roetenberg, Luinge and Slycke, inertial motion capture suit.*

### 3.3 Magnetic Motion Capture

Researching in the field of mocap technologies, specifically magnetic field mocap, one can find the name Bill Polhemus numerous times. During his studies at Harvard University in the 1960s, he developed an Electromagnetic motion capture technology to work as part of a Head Mounted Display for US fighter pilots, (Fischer, 2000). This was to be used for aiming the weapons to enemy airplanes while not requiring fighter pilots taking their hands off the steering equipment. As every mocap technology mentioned, this also became part of the animation and gaming world. The magnetic mocap works in a unique way to capture motion activity from an actor. It requires specialised apparatus able to transmit a magnetic field, i.e. transmitter, and receivers to detect the generated field as well as other sensors. In summary, it is possible to calculate the distance between receiver and transmitter by the strength of the signal reaching the receiver. Nonetheless, this tells us nothing about the orientation of the limbs of the actor; the only information we get is for the distance between all receivers placed on the actor. This is where sensors are used in concert with receivers. Specialised sensors consist of a detector coil which is able to provide limb orientation relative to the transmitter. Magnetic motion capture has many disadvantages as mentioned above, but when done correctly the results are the most accurate out of all the methods throughout this report other than hybrid methods to be described briefly. A famously renowned magnetic mocap technology is that of Ascension, (Ascension, 1994). In 1990, they first delivered the off-the-shelf transmitter called “The Bird” which could track the position of a moving object in three-dimensional space to a range of up to 24 inches. Later on, they came up with the upgrade “Flock of Birds” which could detect up to 30 tiny receivers in full 6 degrees of freedom (no more are required for perfect motion capture) to a range of up to 28 feet. The flock was actually a series of transmitters that would extend coverage and as they said, it “provides enough coverage for a virtual world to fill a room”, (Ascension Technology Corporation, 1994).

Such technology can significantly suffer from noise by the surrounding environment magnetic materials. This meant that to record motion using magnetic technology, the space had to be limited as the strength of the transmitters weakened the further away receivers were, and could not be set up anywhere as it would suffer from noise from the unknown structure materials that make walls and flooring. Even though the “Flock of Birds” could reduce that error, since it included an antenna that could produce pulsating DC signal reducing magnetic interference, drift of movement data from location receivers was still an issue and only the processing power of a later processor could fix the drift



issue in real-time. Nonetheless, it still constitutes one of the cheaper alternatives of motion capture purchasable from the audience.

## 4. Discussion

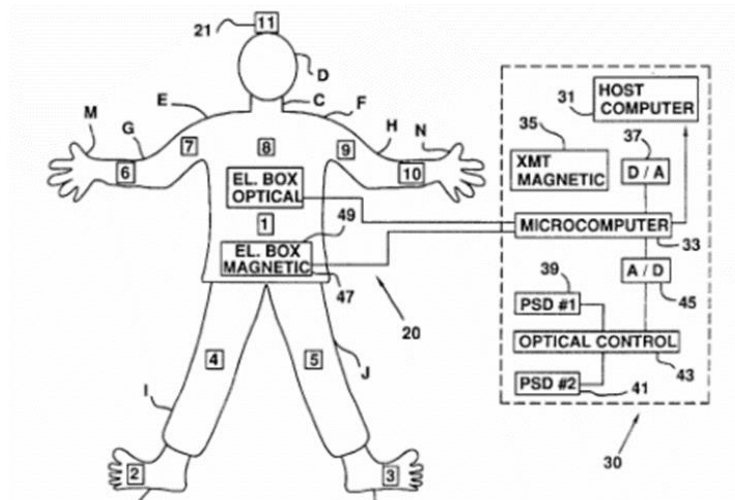


Figure 10 – The idea of hybrid motion capture, using electromagnetic and optical systems with their interface shown on the left, (Hansen, 1998).

smooth and natural. Alternatively, the progress of computer vision algorithms from stereo vision geometry and structure from motion, pushed the motion capture development towards the optical marker-less technology, making it easier for the public to reproduce their own versions, with little or no expense. The mechanisms mentioned above constitute the technology that helped shape the present as we know it for motion capture and newer technologies seem to avoid the limitations of the old technologies based on the fact that computers can now process large data volumes in a few milliseconds. The speed, or frequency, in which a tracking device can gather and process data is directly proportional to the quality of the motion passed to a computer-generated character.

Moreover, using these types of motion capture in concert can increase the efficiency of motion tracking significantly. For example, magnetic and mechanical systems have the advantage of tracking the entirety of a human body in real time but are incapable of tracking multiple objects simultaneously, whereas optical systems are capable of tracking multiple objects but in the past, were incapable of providing real-time motion tracking. Combining these two technologies into a hybrid one can lead to each system cancelling the disadvantages of the other, eventually producing robust enough systems that are fully functional. Hansen claimed, in his patent, that the hybrid system he derived, using magnetic and optical systems in concert, could increase the accuracy and dynamic performance of the overall system when compared to cases using a single technology, (Hansen, 1998).

## 5. Conclusion

Technological development has never reached its peak, always improving on an existing system or forming a new one altogether. The same stands for motion capture technologies. In the present, as one is able to buy off-the-shelf products that can perform the capabilities of any system mentioned multiplied exponentially and in a fraction of the time required to process back then. However, as this was more of a history of motion capture survey rather than explaining the current state-of-the-art, one can understand the significance of the pioneering technology mentioned above and how it aided in reaching the current state of development in the field of motion capture. Every technology has its merits and liabilities but given their unique nature, the results were crucial to building the world as we know it today.

Considering that every motion capture mechanism studied above has its own flaws and strengths, one could think that they can be further improved by combining their abilities into a single mechanism. Firstly, to account for the data gaps during tracking, most of these mechanisms use a post-processing process involving the use of algorithms, such as inverse kinematics, that can potentially fill in these gaps by interpolating between the end effector's positions that are missing, to keep the motion supplied

## 6. References

- Allen, R. (1983). The Bionic Dancer. *Journal of Physical Education, Recreation & Dance*, 54(9), pp.38-39.
- Allen, R. (1985). The Making of Brilliance (1985). [video] Available at: [https://www.youtube.com/watch?v=HZY5\\_ZzRdbk](https://www.youtube.com/watch?v=HZY5_ZzRdbk).
- Ascension Technology Corporation (1994). Flock of Birds info. [online] Biomch-l.isbweb.org. Available at: <https://biomch-l.isbweb.org/threads/2392-Flock-of-Birds-info>.
- Ascension, A. (1994). The Flock of Birds Installation and Operation Guide. Ascension Technology Corporation.. Ascension.
- Calvert, Chapman and Patla (1982). Aspects of the Kinematic Simulation of Human Movement. *IEEE Computer Graphics and Applications*, [online] 2(9), pp.41-50. Available at: <http://ieeexplore.ieee.org/document/1674484/>.
- Cong, R. and Winters, R. (2018). How It Works: Xbox Kinect. [online] Jameco.com. Available at: <https://www.jameco.com/jameco/workshop/howitworks/xboxkinect.html>.
- Fischer, R. (2000). Chapter 2: Motion Capture Process and Systems. "Motion Capture and Motion Editing." AK Peters, [online] pp.1-58. Available at: <https://pdfs.semanticscholar.org/e399/84b1e08f5a98e03e83f2e4d6bac3e997e0d8.pdf>.
- Frey, W., Zyda, M., McGhee, R. and Cockayne, B. (1995). Off-the-Shelf, Real-Time, Human Body Motion Capture for Synthetic Environment. Postgraduate. Monterey, California.
- Ginsberg, C. and Maxwell, D. (1984). "Graphical marionette" (abstract only). *ACM SIGGRAPH Computer Graphics*, [online] 18(1), pp.26-27. Available at: <https://dl-acm-org.ezproxy1.bath.ac.uk/citation.cfm?doid=988525.988549>.
- Gleicher, M. (1999). Animation from observation. *ACM SIGGRAPH Computer Graphics*, [online] 33(4), pp.51-54. Available at: <https://dl.acm.org/citation.cfm?id=345409&CFID=820320792&CFTOKEN=27448133>.
- Hansen, P. (1998). Hybrid motion tracker. US5831260 A.
- Jim Henson Productions (1988). Waldo - The Jim Henson Hour - The Jim Henson Company. [video] Available at: [https://www.youtube.com/watch?time\\_continue=114&v=dP6TUB7KQc4](https://www.youtube.com/watch?time_continue=114&v=dP6TUB7KQc4).
- Kirkham, J. (2010). INERTIAL SENSORS FOR VISUALIZATION CONTROL. Master of Science. University of Manchester.
- MacCormick, J. (2012). How does the Kinect work?.
- Maher, M. (2017). Rotoscoping: From Early Animation to Blockbuster VFX. [online] RocketStock. Available at: <https://www.rocketstock.com/blog/rotoscoping-from-early-animation-to-blockbuster-vfx/>.
- Menache, A. (2000). Understanding motion capture for computer animation and video games. San Diego: Morgan Kaufmann, pp.11-12.
- Roetenberg, D., Luinge, H. and Slycke, P. (2009). Xsens MVN: Full 6DOF Human Motion Tracking Using Miniature Inertial Sensors. *Proceedings*, [online] pp.1-9. Available at:

<https://www.semanticscholar.org/paper/Xsens-MVN-Full-6DOF-Human-Motion-Tracking-Using-Mi-Roetenberg-Luinge/cc2ba84a4d6e06fd85ad434f5b1a8545c1cc993c>.

Sementille, A., Lourenço, L., Brega, J. and Rodello, I. (2004). A motion capture system using passive markers. Proceedings of the 2004 ACM SIGGRAPH international conference on Virtual Reality continuum and its applications in industry - VRCAI '04, [online] pp.1-8. Available at: <https://dl.acm.org/citation.cfm?id=1044684&CFID=820320792&CFTOKEN=27448133>.

Silicon Graphics and Nintendo (1992). Real Time Mario at the 1992 SCES. [video] Available at: <https://www.youtube.com/watch?v=R1t6iNG28zl>.

Walton, C. (1999). Exercising system with electronic inertial game playing. US5989157 A.

Zhang, L., Curless, B. and Seitz, S. (2002). Rapid shape acquisition using color structured light and multi-pass dynamic programming. Proceedings. First International Symposium on 3D Data Processing Visualization and Transmission. [online] Available at: <http://ieeexplore.ieee.org/abstract/document/1024035/>.