

I-Seminar 7311a

Animation und Bewegung in Computer Spielen

Eine Seminararbeit an der Berner Fachhochschule Abteilung Technik und Informatik
von Stefan Heinemann* und George Sriranjant[†]
betreut von Professor Rolf Haenni[‡]

8. Januar 2012

*heins4@bfh.ch

[†]srirg1@bfh.ch

[‡]rolf.haenni@bfh.ch

Abstract

Dieses Dokument ist im Zuge des Moduls I-Seminar an der Berner Fachhochschule entstanden. Es soll dem Leser die verschiedenen Animationsarten, sowie Bewegung in Computer Spielen näher bringen. Dafür gibt es verschiedene Arten wie man Bewegungen darstellen kann. Eine häufig verwendete Animationstechnik ist zum Beispiel Motion Capturing auch Mocap genannt. In der Welt der Animationstechnik wird abermals von CGI geredet. CGI steht für Computer generated Imagery und bedeutet eigentlich nur Computer generierte Bildreihenfolgen(Film). In diesem Dokument wird als Schwerpunkt die Animationstechnik Motion Capture betrachtet und genauer erläutert.

Dieses Dokument soll nicht als Tutorial dienen, sondern verschiedene Konzepte und Ideen im Bereich von Animation und Bewegung in Computerspielen aufzeigen, damit der Leser sich ein fundamentales Wissen aneignen kann im Bereich der Bild- und Videobearbeitung.

Inhaltsverzeichnis

1	Einleitung	4
1.1	Geschichte	4
1.2	Bildfrequenz	5
2	Konzepte	6
2.1	Keyframing	6
2.1.1	klassisches Keyframing	6
2.1.2	modernes Keyframing	6
2.2	Motion Capture	6
2.2.1	menschliche Fortbewegung	7
2.2.2	Motion Sequencing	7
3	Mocap-Systeme	8
3.1	magnetische Systeme	8
3.2	mechanische Systeme	9
3.3	optische Systeme	10
4	Schlussbericht	13
4.1	Fazit	13
4.2	Quellenangabe	13

1 Einleitung

In diesem Kapitel möchten wir ein wenig Hintergrundwissen im Bereich der Computergrafik vorstellen. Eine kurze Geschichte über ein trabendes Pferd und eine skurrile Teekanne wird hier beschrieben.

1.1 Geschichte

Eadweard Muybridge

Um die erste Animation in der Geschichte zu finden müssen wir ins Jahr 1872. Eadweard Muybridge (geboren: 9. April 1830 in Kingston upon Thames, England; gestorben: 8. Mai 1904 ebenda) ein britischer Fotograf und Pionier der Fototechnik, bekam als Auftrag die Beinstellungen eines trabenden Pferdes zu bestimmen.

Muybridge meisterte die Aufgabe indem er mit 12, 24 und 36 sukzessive auslösenden Fotoapparaten arbeitete und somit die Serienfotografie geboren war. Natürlich war das vergleichsweise zu heute ein schwieriges unterfangen, das Pferd löste auf einer Pferderennbahn gespannte Zugdrähte aus, welche wiederum die elektrisch betriebenen Hochgeschwindigkeitsblenden der Fotoapparate auslösten. So wurde erstmals der sichtbare Beweis erbracht, dass sich beim galoppierenden Pferd bei gewisser Zeit alle vier Beine in der Luft befinden.

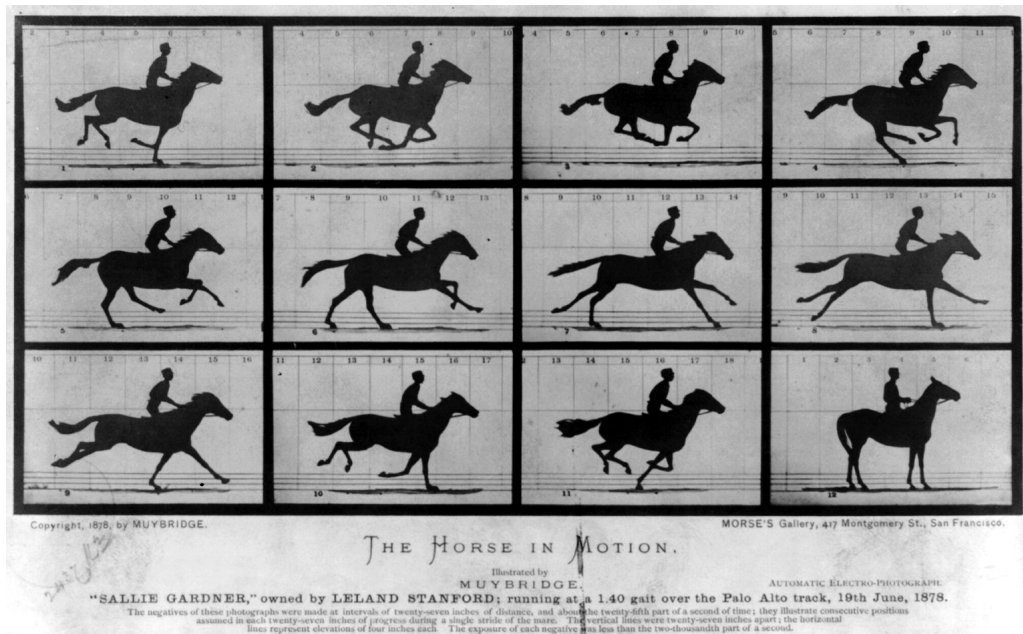


Abbildung 1.1: galoppierendes Pferd

Auf dem 2. und 3. Bild kann man gut erkennen, dass alle Beine gegen die Mitte gebogen und in der Luft sind. Eine weitere Erkenntnis ist auch, dass nie mehr als 2 Beine des Pferdes den Boden berühren.

Utah-Teekanne

Zu jeder Computergrafik gehört selbstverständlich ein 3D-Modell. Eines der ältesten und berühmtesten 3D-Modelle ist die Utah-Teekanne (engl. Utah-Teapot). Hierbei handelt es sich um ein Oberflächenmodell einer Teekanne, bei der der Innenraum nicht modelliert wurde. 1975 entwickelte Martin Newell sie im Rahmen einer Forschungsarbeit an der Universität von Utah.

Als er nach einem Gebrauchsgegenstand für seine Arbeit suchte, fand er die Melitta-Teekanne seiner Frau für geeignet. In Computerzeitschriften wurde über die Jahre immer wieder das Modell übernommen und regelmässig Varianten der Teekanne präsentiert. Mittlerweile wurde es zu einer Art Running Gag in der Computergrafikszene und kommt beispielsweise "versteckt" in den Animationsfilmen Toy Story und Die Monster AG vor. Oder selbst 3D-Animations-Programme wie 3d Studio Max verwenden die Teekanne als Grundkonstruktionsobjekt.

Selbstverständlich nimmt die 3d-Modellierung ein grosser Teil in der Computerspielewelt ein. Jedoch möchten wir uns von diesem Teilgebiet differenzieren und die Animation und Bewegungen in Computerspielen genauer betrachten und erläutern.

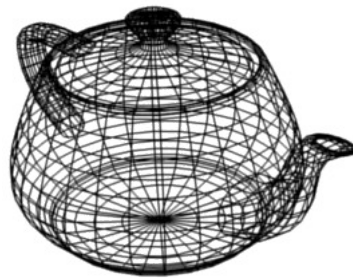


Abbildung 1.2: Utah-Teekanna

1.2 Bildfrequenz

Die Bildfrequenz ist ein Begriff aus der Film- und Videotechnik. Sie bezeichnet die Anzahl der Einzelbilder, die in einem Zeitraum aufgenommen werden. Bildfrequenz wird auch mit fps abgekürzt (für das englische Frames per Second).

Das menschliche Gehirn nimmt ab etwa 14 bis 16 Bildern pro Sekunde aufeinanderfolgende Bilder als bewegte Szene wahr. Zur Stummfilmzeit wurden die Filme mit 16fps erzeugt. Später mit der Einführung des Tonfilms wurde die Bildfrequenz auf 24fps festgelegt, da die Tonqualität bei 16fps nicht ausreichte.

2 Konzepte

Wir beziehen uns hier auf zwei fundamentale Konzepte, wie man eine Animation erstellen kann. Nämlich das Keyframing und Motion Capturing.

2.1 Keyframing

Keyframing ist ein Begriff aus der Animationstechnik und bedeutet auf deutsch Schlüsselbildanimation. Diese Technik stammt aus der Produktion von Zeichentrickfilmen. Es werden grobe Bilder (Keyframes) erzeugt, welche die Animation beschreiben. Zwischen den Keyframes werden detaillierte Bilder gezeichnet (Interframes) bis sich in der Bewegung ein flüssiges Bild ergibt.

2.1.1 klassisches Keyframing

- Regisseur (erstellt Storyboard)
- Schlüsselzeichner (erstellen Keyframes)
- Zwischen(phasen)zeichner (erstellen Interframes)
- Reinzeichner (alle Bilder werden nochmal verbessert und gesäubert)

2.1.2 modernes Keyframing

Die Zwischenzeichner werden nicht mehr benötigt, diese Arbeit wird vom Computer abgenommen. Das zu animierende Objekt steht als 3D-Model in einem Koordinatenkörpers zur Verfügung. Die Interframes werden automatisch mit Rechner erstellt und somit wird die Bildreihenfolge interpoliert. Der Vorteil ist die Senkung des Personalaufwandes. Um die Darstellungen von komplexer Mimik und Gestik wird oft Motion Capturing angewandt, weil es zu Zeit- und Kostenaufwändig wäre dies mit Keyframing zu realisieren.

2.2 Motion Capture

Motion Capture bedeutet übersetzt nichts anderes als Bewegung aufzeichnen. Es wird oft auch MoCap oder MoCa genannt. Motion Capturing wird verwendet um Bewegungen in Echtzeit zu erfassen und auf dem Computer darzustellen. Später kann man Analysen durchführen oder diese Bewegungen auf ein Model projizieren und somit dem Objekt Leben einhauchen.

In den letzten Jahren ist diese Methode zu einer der bekanntesten Animationstechniken geworden. Durch Filme wie Beispielsweise Star Wars, Herr der Ringe und Avatar aber auch durch Computer und Konsolenspiele wie FIFA und Uncharted wurde der Bekanntheitsgrad gestärkt.

2.2.1 menschliche Fortbewegung

Die Idee ist es eine möglichst realistische Bewegung zu erzeugen, jedoch ist dies beispielsweise mit Keyframing nicht so einfach. Der Gang eines Menschen besteht nicht nur aus "ein Fuss vor den anderen setzen", sondern eine individuelle Folge von verschiedenen Bewegungen des ganzen Körpers.

Die künstliche Nachahmung dieser individuellen Bewegungen im Körper gestaltet sich also als schwierig. Mit herkömmlichen Animationstechniken ist dies nahezu unmöglich, jedoch beim MoCap werden reale Bewegungen eins zu eins übernommen und auf das Model beziehungsweise auf den Charakter angewendet. Somit wird die Bewegung nicht künstlich erzeugt sondern eine reale Aufzeichnung des Schauspielers. Dadurch wird es zeitsparender und kosteneffizienter im Vergleich zu anderen Animationstechniken.

2.2.2 Motion Sequencing

Motion Sequencing beschreibt das Aneinanderreihen von Bewegungssequenzen. Erst durch diese Technik entstehen realistische Animationen.

Zum Beispiel: Man hat folgende zwei Bewegungssequenzen, eine schleichende Person und eine rennende Person. Folglich ist das Motion Sequencing das einander hängen der beiden Elemente. Später im Spiel könnte beispielsweise zwischen Schleichen und Sprinten gewechselt werden. Die Schwierigkeit hierbei liegt, das einander knüpfen aller Elemente ohne einen sichtbaren Übergang zu erschaffen. Um einen unsichtbaren Übergang zu erhalten, werden mit Hilfe des Rechners Zwischenbilder generiert, welche einen weichen Übergang der einzelnen Elemente erzeugen (keyframing).

3 Mocap-Systeme

Der Großteil der Systeme kann in 3 Hauptklassen unterteilt werden. Magnetische, mechanische, und optische Systeme. Es gibt es auch Hybridsysteme mit allen Kombinationen der Hauptklassensysteme. Es gibt viele Firmen, die eine Menge von solchen Geräten herstellen. Wir möchten nicht die einzelnen Produkte näher Betrachten und Beschreiben, sondern vielmehr die verschiedenen Systeme erklären.

3.1 magnetische Systeme

Bei den magnetischen Systemen arbeitet man mit Hall-Sensoren. Hall-Sensoren ermitteln ein magnetisches Feld aus, welches von einer Transmittereinheit gesendet wird. Bei diesem Vorgang wird eine Ausgangsspannung bei den Hall-Sensoren erzeugt, welche dem Produkt aus Strom und magnetisches Feld entspricht. Über den ganzen Körper werden Hall-Sensoren verteilt (bevorzugt werden alle Gelenke). Die Sensoren übertragen ihre Daten zu einer Kontrolleinheit, die dann die Positionen im dreidimensionalen Raum berechnet.



Abbildung 3.1: Motion Capture Anzug eines magnetischen Systems

Ein Nachteil ist, dass jeder Schauspieler mit Kabeln und Sensoren umzingelt ist, dadurch können sie sich nicht mehr ganz frei bewegen. Die Verkabelung ist notwendig, weil jeder Sensor eine Spannungsquelle benötigt. Die Akteure tragen eine Batterie, meist als Rucksack oder versteckt im Anzug. All diese unhandlichen Konstruktionen fallen negativ auf die Bewegungsfreiheit der Akteure aus.

Ein weiterer Nachteil ist die Reichweite von magnetischen Systemen. Die Reichweite der

Transmitter ist beschränkt und die Spielfläche muss mit einem Magnetfeld bestrahlt werden. Somit haben die Schauspieler meist eine bis zu Acht mal Acht Meter große Fläche zur Verfügung.

Der Vorteil bei den magnetischen Systemen ist, dass es hier keine Deckungsprobleme gibt wie bei den optischen Systemen. Bei einem optischen System (siehe 3.3 optische Systeme) ist es schwieriger alle Akteure Anhand von Punkten zu unterscheiden und einzelne Personen zu segmentieren. Jedoch können zwei Akteure die unmittelbar nebeneinander agieren, zum Beispiel bei einer Kampfszene bei der die Sensoren vom gegenüberliegenden Akteur gestört werden, das Ausgangssignal ebenso verfälschen. Es können auch andere elektrische Geräte oder Metall in der Umgebung das elektromagnetische Feld stören und somit zu einem schlechten Resultat führen. Der Vorteil ist das die Aufnahmen in Echtzeit stattfinden und verwendet werden können.

3.2 mechanische Systeme

Anders als bei magnetischen Systeme benötigt man hier weder Sensoren noch eine Transmittereinheit. Bei mechanischen Systemen werden die Bewegungen direkt am Körper ausgelesen. Der Schauspieler trägt eine Art Aussenskelett mit jeweils Stangen für Ober- und Unterschenkel und für den Kopf und Rücken. Durch die Bewegungen der Stangen wird die Spannung über einem regelbaren Widerstand (Potentiometer) an den Gelenken verändert und ausgelesen. Diese analogen Spannungsänderungen werden in ein digitales Signal umgewandelt und von der Software interpretiert. Mit Hilfe eines Gyroskop-Sensors (Kreisel) werden die Neigungen gemessen und somit lassen sich alle Körperbewegungen rekonstruieren.

Kalibrierung

Die Kalibrierung ist hier Zeitaufwändig, da sie genau dem Schauspieler angepasst werden muss. Die Abstände zwischen den Gelenken muss genau erfasst werden, damit das Resultat höchstmöglich präzise ermittelt werden kann. Ein weiterer Grund dafür ist, dass der Schauspieler nicht an seiner Beweglichkeit limitiert sein soll und alle Bewegungen ausführen kann. Diese Anpassung des Systems muss also für jeden einzelnen Schauspieler angepasst werden. Im Gegenteil zu den magnetischen und optischen Systemen, wo ganz am Anfang die Kameras justiert werden und gegebenenfalls nachkorrigiert werden müssen, kann hier eine Kalibrierung bis zu einer halben Stunde dauern. Wenn die Einstellungsdaten gespeichert wurden können sie innerhalb von 5 Minuten geladen werden und das System ist einsatzbereit.

Vor- und Nachteile

Hier gibt es keine Probleme mehrere Personen miteinander aufzuzeichnen, da es keine Sensoren gibt die sich beeinflussen könnten wie bei magnetischen Systemen. Aber grundsätzlich ist man eingeschränkt in der Bewegungsfreiheit, wenn man an eine Kampfszene denkt oder Turneinlagen. Mit solchen Aussenskelette ist es nahezu unmöglich solche Szenen nach zu spielen.

Dies ist auch der Grund, warum sich dieses System nicht durchgesetzt hat im Bereich Computerspiele, da man zu wenig Bewegungsfreiheit hat. Aber beispielsweise in der Rehabilitation von Patienten nach einem Unfall kann es durchaus Sinn machen ein solches System zu verwenden. Die Richtigkeit der Bewegungen kann nachvollzogen werden und dem Patienten

helfen seinen Genesungsprozess zu optimieren.

3.3 optische Systeme

Bei einem optischen System trägt der Schauspieler ebenso wie bei den anderen Systemen einen Ganzkörperanzug, der aber hier mit vielen reflektierenden Markern besetzt ist. Mit Hilfe von Kameras werden diese reflektierenden Punkte (Markern) aufgenommen und am Rechner ausgewertet. Es gibt auch Systeme, welche mit Infrarot arbeiten (siehe 3.3 Kinect). Die Marker werden an den wichtigsten Positionen des Körpers angebracht. Dies sind vor allem die Gelenke, weil sie alle Endpunkte der Körperteile darstellen, welche sich tatsächlich auch bewegen können. Hier gilt je mehr Marker vorhanden sind desto genauer kann die Bewegung rekonstruiert werden. Typische Positionen für die Marker sind: Kopf, Nacken, Schulter, Ellenbogen, Handgelenke, Beckenknochen, Knie und Fussgelenke.

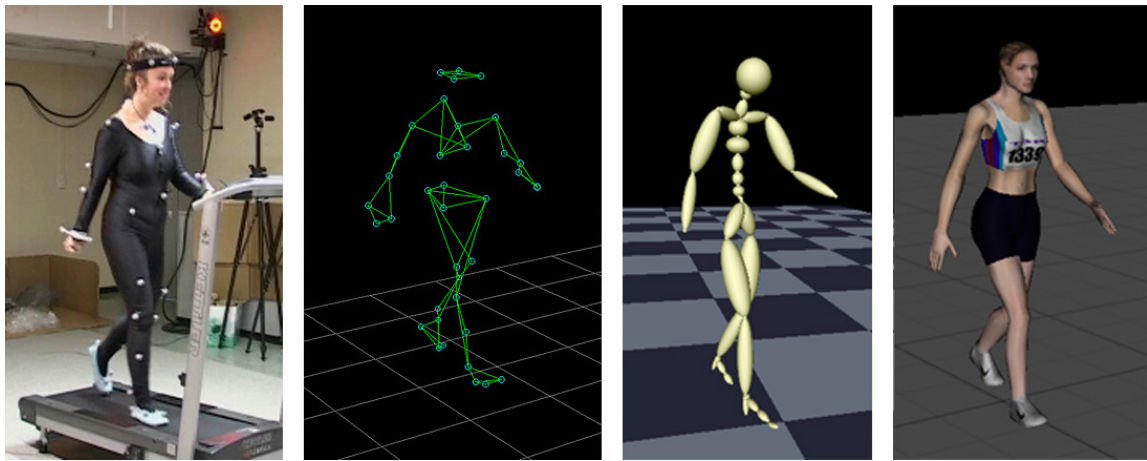


Abbildung 3.2: (vlnr) Schauspielerin mit Markern ausgestattet auf einem Laufband; Kameraaufnahmen in ein Gitternetz-Modell; Übertragung der Messwerte in ein simples 3D-Modell; Animation einer Sprinterin

In dieser Abbildung erkennt man schön alle Schritte, welche notwendig sind um an ein gewünschtes Ziel zu gelangen. Die Statur der Animation muss einigermaßen dem Ursprungsobjekt gleichen. Wenn wir annehmen folgende Frau im Bild wurde als Objekt in einem 3D-System dargestellt und als Model haben wir einen Gartenzwerg. Als Ziel verfolgen wir natürlich eine realistische Bewegungssequenz des Zwerges.

Nun haben wir aber das Problem, falls wir nur die Bewegungssequenzen der Beine in die Animation hinzufügen, der Zwerg zwar geht aber viel zu kleine Schritte macht oder wir verwenden die absolute Distanz, dies führt zu unnatürlichen bis zu unmöglichen Bewegungen des Zwerges.

Bisweilen gibt es keine Zauberformel, welche dieses Problem löst. Meist versuchen die Game-Designern die Animation selbst anzupassen mittels Keyframing und verwenden die Motion-Capture-Aufnahmen als Stütze und Ziel.

Hero Pose

Bevor überhaupt irgendwelche Aufnahmen getätigt werden können muss eine Kalibrierung stattfinden, damit das System die einzelnen Marker unterscheiden kann. Später folgt das System ununterbrochen den Marker und kann diese aufzeichnen. Hierfür steht der Akteur in die Mitte der Bühne und geht in die Heldenposition (die Haltung sieht aus wie ein grosses T). Damit soll gewährleistet sein dass alle Marker sichtbar sind. Als ein modernes Beispiel dient hier Kinect.

Kinect

Kinect ist eine Hardware zur Steuerung der Videospielkonsole Xbox 360. Kinect wurde von Microsoft mit der Zusammenarbeit von der Firma PrimeSense entwickelt. Spieler können statt mit einem Gamepad allein durch Körperbewegungen die Software bedienen.

Kinect ist mit einem Tiefensensor-Kamera, 3D-Microfon und Farbkamera ausgestattet. Die Tiefensensor-Kamera funktioniert folgendermassen:

Die Bühne wird mit Infrarotimpulse ausgeleuchtet und der Infrarot-Sensor in der Kamera misst für jeden Bildpunkt die Zeit, die das Licht bis zum Objekt und wieder zurück braucht. Die benötigte Zeit ist proportional zur Distanz. Das besondere an Kinect ist, dass die Spieler keine Marker benötigen. Durch die Tiefen-Wahrnehmung des Gerätes kann direkt ein 3D-Model der Umgebung im Computer abgebildet werden. Somit wird der Prozess "Übertragung der Marker-Aufnahmen in ein 3D Model" übersprungen, stattdessen wird vom Computer ein eigenes Skelett berechnet, welches in die Szene passt. Kinect kann mit einer Auflösung von 1 cm in die Tiefe und 3mm auf der x und y-Achse genau bestimmen, wo die Objekte sich befinden. In der Tiefe ist es mit 3.5 Meter limitiert.

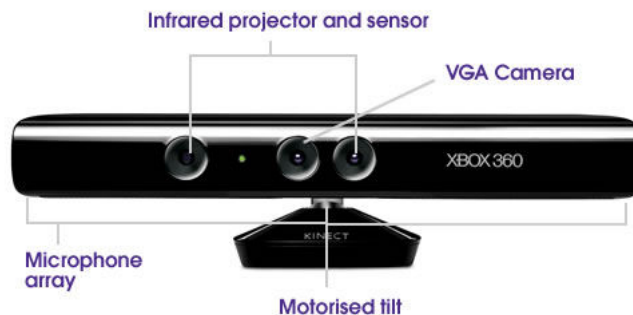


Abbildung 3.3: Kinect-Hardware

Auch ein Motor ist eingebaut, damit kann das Gerät geschwenkt und somit dem Spieler gefolgt werden. Anhand von Kinect können wir erkennen, dass Motion Capturing unser Alltag schon längst bestimmt. Letztens wurde von Microsoft ein Software-Development-Kit vorgestellt um Hobby-Programmierer zu animieren an diesem Gerät zu arbeiten. Weiter wurde ein Open-Source-Treiber unter Windows und Linux freigegeben, welches noch mehr Möglichkeiten hervorruft.

Vor- und Nachteile

Vorteile von optischen Systemen ist die sehr hohe Messgenauigkeit bzw. hohe Abtastrate, grosse Bewegungsfreiheit man hat keine störende Einheiten am Anzug nur kleine Licht-Reflektoren (Markern) und die Möglichkeit mehrere interagierende Akteure gleichzeitig aufzuzeichnen. Das Arbeitsvolumen ist nur durch die Aufnahmefähigkeit der Kameras begrenzt. Ein grosser Nachteil ist die fehlende Identifikation der Marker. Am Anfang bei der "Hero Position" ist es mühelos möglich alle Marker zu unterscheiden doch später bei Bewegungen, bei denen einige Marker abgedeckt sind, wird es schwierig diese zu finden. Hier muss meist von Hand nachgearbeitet werden. Oft werden verschiedenfarbige Marker eingesetzt um die Identifikation zu gewährleisten. Auch ist es so möglich ein vorübergehend abgedeckter Marker später anhand seiner Farbe wieder zu "erkennen" und aufzuzeichnen. Bei völlig verdeckten Markern ist die Nachbearbeitung nicht zu vermeiden.

4 Schlussbericht

4.1 Fazit

Aus den Drei Motion-Capture Systemen hat sich definitiv das optische Motion Capturing System herauskristallisiert. Durch die einfache Integration der Schauspieler und Bewegungsfreiheit, werden Abdeckungssprobleme und somit die Nacharbeitung am Computer gerne in Kauf genommen.

Natürlich ist Motion Capturing nicht für jedes Computerspiel geeignet und muss vorher abgewägt werden, ob es Sinn macht dies zu verwenden. Mocap-Systeme werden schon seit Jahren in Animationsfilmen verwendet. Jedoch erst seit kurzem in Computergames. Der Grund warum bisher in Computerspielen nicht mit Mocap-Systemen gearbeitet wurde liegt definitiv an den Kosten die dabei entstehen (mehr Personalaufwand, Nacharbeitung und Feinschliff am Computer). Doch mittlerweile ist die Computerspielbranche explodiert und ist ein Multi-Millionen-Dollar Unternehmen geworden, vergleichsweise mit Hollywood.

In den letzten Jahren ist in der Computerspielwelt eine neue Genre entstanden. Diese Spiele sind vergleichbar mit einem interaktiven Spielfilm. Der Protagonist trifft Entscheidungen im Spiel, welche wiederum CGI (computer generated imagery) Video Sequenzen auslösen. Somit wird dem Spieler das Gefühl geboten er selbst wirkt in diesem "Spielfilm" mit und taucht in die Geschehnisse des Protagonisten ein. Als Beispiel dient das Computerspiel Heavy Rain.

Aber auch andere Gebiete wie die Orthopädie und Sportmedizin haben von Motion Capture profitiert. Beispielsweise wird der Gang eines Menschen aufgezeichnet und kann analysiert werden. Chirurgen verwenden Motion Capturing als eine Art Gestensteuerung um ihre Software zu steuern, da sie ihre Hände aus hygienischen Gründen nicht anders benutzen dürfen. Aber es gibt auch Profigolfer die ihre Schwungtechnik beim Golfen analysieren und versuchen ihre Bewegungen dabei zu verbessern.

4.2 Quellenangabe

- Bildfrequenz - wikipedia.org
- utah-teekanne - wikipedia.org
- horse in motion.jpg - wikipedia.org
- magneticgirl.jpg - <http://tyrell-innovations-usa.com>
- optmocap.jpg - lukemccann.files.wordpress.com/2010/09/lolornamocap.jpg
- Optical Motion Capture: Theory and Implementation - <http://smile.uta.edu/Guerra-FilhoRITA05optical.pdf>

- motion capturing - <http://animalrace.uni-ulm.de>
- electro mechanic motion capture - <http://www.metamotion.com/motion-capture/electro-mechanical-motion-capture.htm>
- Motion Capture to the People: <http://www.saidi.se/pdf/moe.pdf>
- kinect.png - <http://uatrobotics.blogspot.com/2011/09/led-cube-controller.html>