

Institut für Visualisierung und Interaktive Systeme (VIS)
Abteilung Graphisch-Interaktive Systeme (GIS)
Universität Stuttgart
Universitätsstraße 38
D – 70569 Stuttgart

Fachstudie Nr. 130

Avatar Frameworks

Alexander Duschek Philipp Schuster Xi Tu

Studiengang: Softwaretechnik

Prüfer: Prof. Dr. Thomas Ertl

Betreuer: Dipl.-Phys. Michael Raschke

begonnen am: 11.11.2010

beendet am: 03.06.2011

CR-Klassifikation: H.5.2 (User Interfaces)

Inhaltsverzeichnis

1	Einführung	4
2	Erstellung von Avataren	6
2.1	Avatarerstellung	6
2.2	Entwicklungskonzepte für Avataranwendungen	16
2.3	Beispiele für Avatare	20
2.3.1	Avatare in Computer- und Konsolenspielen	20
2.3.2	Avatare im Kino	25
2.3.3	Avatare als Lernhilfe	25
2.3.4	Avatare als Orientierungshilfe	26
2.3.5	Avatare zur Kommunikation	27
2.3.6	Avatare in der Werbung	28
2.3.7	SecondLife und Ableger	28
2.3.8	Avatare im Web allgemein	31
2.4	Zusammenfassung	31
3	Mensch-Avatar-Beziehungen	32
3.1	Kulturelle Unterschiede	32
3.2	Akzeptanzprobleme von Avataren	35
3.3	Zusammenfassung	38
4	Interaktionskonzepte	39
4.1	Interaktionstechnologien bei HCI mit Avataren	39
4.2	Anpassung an das Benutzerverhalten	48
4.3	Spracherkennungsmethoden von Avataren	50
4.4	Mimik und Gestik	53
4.5	Körpersteuerung	54
4.6	Visualisierung von Emotionen	59
4.7	Grenzen der Interaktionsmöglichkeiten mit Avataren	60
4.8	Zusammenfassung	62
5	Evaluation von bestehenden Frameworks zur Avatarerstellung	63
5.1	Untersuchte Kriterien	63
5.2	Vergleichstabelle	65
5.3	Detaillierte Untersuchung der einzelnen Frameworks	65
5.3.1	Voki	65
5.3.2	My Avatar Editor	69
5.3.3	Evolver	72

5.3.4	XBox	77
5.3.5	Yahoo Avatare	81
5.4	Zusammenfassung.....	83
6	Zusammenfassung und Ausblick.....	85
	Literaturverzeichnis	87
	Abbildungsverzeichnis	90

1 Einführung

Als Avatar bezeichnet man eine künstliche Person oder einen grafischen Stellvertreter einer realen Person. Der Begriff trat im Zusammenhang mit Computern 1992 zum ersten Mal auf. Der Begriff *Avatar* stammt aus dem Sanskrit und bedeutet übersetzt *Abstieg*. Im Hinduismus steht es für die Inkarnation des Gottes Vishnus; allgemein für eine herabgestiegene Gottheit. Oftmals wird in der Fachliteratur auch der Begriff *Embodied Conversational Agent* (dt. verkörperter Konversationsagent) verwendet. Dieser beschreibt eine Figur, die mit einem Benutzer oder einer anderen Figur mittels einer implementierten künstlichen Intelligenz so interagieren kann wie es für Menschen normal ist. Es spielt dabei keine Rolle, ob der Avatar physisch existent ist oder lediglich in einem virtuellen Szenario existiert. Die Bandbreite von Avataren kann von einfachen Icons über 3D-Abbilder eines Menschen bis hin zu komplett mechanischen Robotern reichen.

Diese Fachstudie beschäftigt sich mit der Untersuchung existierender und zukünftiger Interaktionskonzepte beim Einsatz von Avataren sowie deren technischer Realisierbarkeit. Das Ziel der Arbeit ist die Klärung folgender Punkte:

- Erstellung von Avataren

Wie werden Avatare erstellt? Welche Konzepte gibt es? Gibt es Beispiele aus verschiedenen Bereichen?

- Mensch-Avatar-Beziehungen

Welche kulturellen Unterschiede sind bei Avataren vorhanden? Warum sind Avatare in der heutigen Zeit trotz ihrer Realisierbarkeit nicht sehr verbreitet?

- Interaktionskonzepte

Welche Interaktionstechnologien gibt es? Welchen Einfluss besitzen Mimik und Gestik auf eine Konversation? Wie erkennt der Avatar das Verhalten des Benutzers? Wie erkennt der Avatar natürliche Sprache und setzt diese für sich um?

- Evaluation

Welche Frameworks zur Erstellung von Avataren gibt es auf dem Markt? Welche Möglichkeiten bieten diese Frameworks? Wie könnten diese Frameworks für eigene Avatarprojekte genutzt werden?

- Zusammenfassung und Ausblick

Was sind die hauptsächlichen Ergebnisse der vorangegangenen Kapitel? Wohin entwickelt sich der Avatar?

Wenn man sich mit Avataren intensiver beschäftigt, stellt man fest, dass diese nicht nur ein Teilgebiet beschränken, sondern einen interdisziplinären Bereich darstellen. Darum

betrachten wir das Thema auch aus verschiedenen Blickwinkeln, um einen besseren Eindruck vom Gesamtverständnis zu erhalten und zu ermitteln.

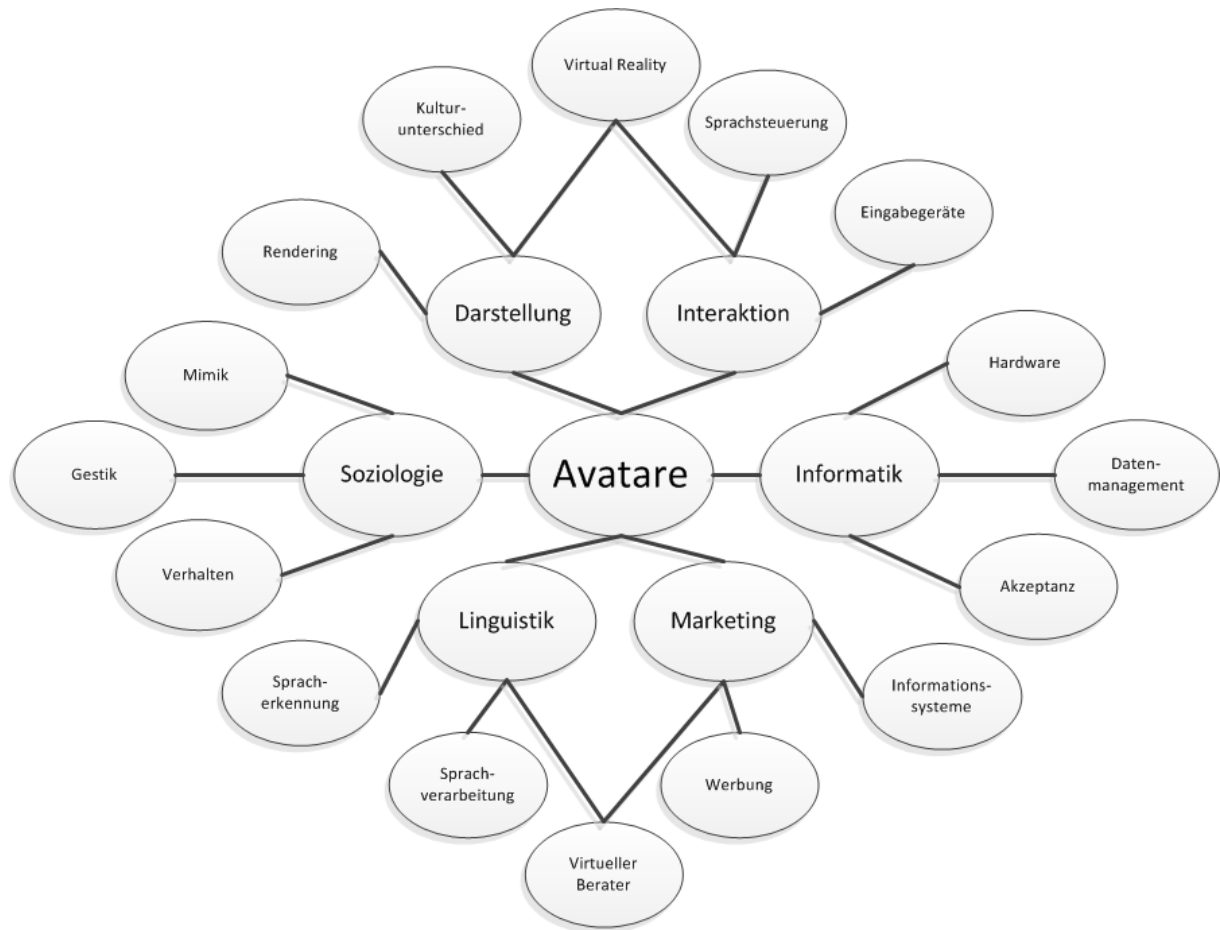


Abbildung 1 – Interdisziplinarität von Avataren. Avatare sind eine interdisziplinäre Kategorie und nicht auf einen speziellen Bereich beschränkt. Bei der Entwicklung von Avataren sind zahlreiche Parteien beteiligt, die alle ihre Vorstellungen und Wünsche in das Endprodukt einbringen wollen.

2 Erstellung von Avataren

Dieses Kapitel umfasst die Abschnitte „Avatarerstellung“, „Entwicklungskonzepte für Avataranwendungen“ und „Beispiele für Avatare“. Im Abschnitt „Avatarerstellung“ wird zunächst erläutert auf welche Arten Avatare erstellt werden können. Dies reicht von einfachen Avatarerstellungsmethoden für den Normalbenutzer über Vorgehensweisen in der Film-, Fernseh- und Computerspielindustrie bis hin zu Verfahren aus der Avatarforschung. Der darauffolgende Abschnitt beschäftigt sich mit Entwicklungskonzepten für Avataranwendungen. Darin wird untersucht, mit welchen Vorgehensweisen eine Anwendung mit Avatar entwickelt werden kann. Der Abschnitt „Beispiele für Avatare“ gibt zahlreiche mit Bildern versehene Beispiele für Avatare in Filmen, Computerspielen, im Internet, in der Werbung, zur Kommunikation und als Orientierungs- und Lernhilfe. Das Kapitel wird durch eine Zusammenfassung abgeschlossen.

2.1 Avatarerstellung

Avatare können auf die unterschiedlichste Weise erstellt werden. Dabei sind die verwendeten Techniken an den jeweiligen späteren Einsatzzweck des Avatars angepasst, sowohl vom Umfang als auch von der aufzuwendenden Zeit und den entstehenden Kosten. Die Vielfalt an Techniken ist sehr groß. Insbesondere im Bereich der Forschung über Avatare existieren unzählige Methoden, die zum einen völlig verschiedene Ansätze verfolgen, zum anderen aber auch Varianten von Ansätzen bilden, Ansätze ergänzen oder unterschiedliche Ansätze miteinander kombinieren. Einen kompletten Überblick über alle Erstellungsverfahren zu geben würde den Rahmen dieser Arbeit sprengen. In diesem Abschnitt werden zuerst weit verbreitete Vorgehensweisen zur Avatarerstellung vorgestellt. Dazu gehören die einfache Bildauswahl, die Erstellung einer Spielfigur in einem Rollenspiel, das Motion Capturing Verfahren, der Einsatz eines 3D Scanners und Animationsprogramme. Daran schließen sich einige ausgewählte Avatarerstellungstechniken aus der Forschung an. Zu ihnen zählen ein Verfahren, das mit Hilfe mehrerer Kameras einen Menschen komplett erfasst und daraus einen Avatar erzeugt, eine Methode die mit einer Tiefenkamera arbeitet und eine Technik die einen Avatar nur anhand einer einzelnen Fotografie erstellt. Weiterhin zählen dazu ein Ansatz der eine Videokamera verwendet, der Einsatz evolutionärer Algorithmen und die Avatarerstellung mittels Microsoft Kinect.

Die einfachste Art der Avatarerstellung ist die Auswahl eines Bilds aus einer Menge von vorgegebenen Bildern, wie in Abbildung 2 am Beispiel eines Internetforums dargestellt. Diese Form der Avatarerstellung findet man beispielsweise auch in Computerspielen. Der Bildinhalt hat dabei mit der realen Person hinsichtlich ihres Aussehens meist nichts gemeinsam. Jedoch besteht des Öfteren auch die Möglichkeit ein eigenes Bild zu benutzen, bei dem dann der Benutzer ein Foto von sich verwenden kann. Eine interessante Möglichkeit dieser Avatar Nutzung bietet Gravatar¹. Gravatar ist eine Abkürzung für Globally Recognized Avatar. Hierbei besteht die Möglichkeit, in vielen verschiedenen Webangeboten, denselben

¹ <http://www.gravatar.com>

Avatar zu verwenden, wodurch der Benutzer eindeutig identifiziert werden kann. Auf der anderen Seite ist dieses Vorgehen unter Datenschutzaspekten höchst zweifelhaft, da dadurch für jedermann ersichtlich ist, welche Webangebote ein bestimmter Benutzer verwendet. Hinzu kommen die eventuell vom jeweiligen Benutzer hinterlassenen Inhalte. Somit wird ein tiefer Einblick in das Leben des jeweiligen Benutzers ermöglicht. Nach Wikipedia [35] ist es dem Betreiber des Gravatardiensts auch möglich ein umfassendes Nutzungsprofil des jeweiligen Benutzers zu erstellen.

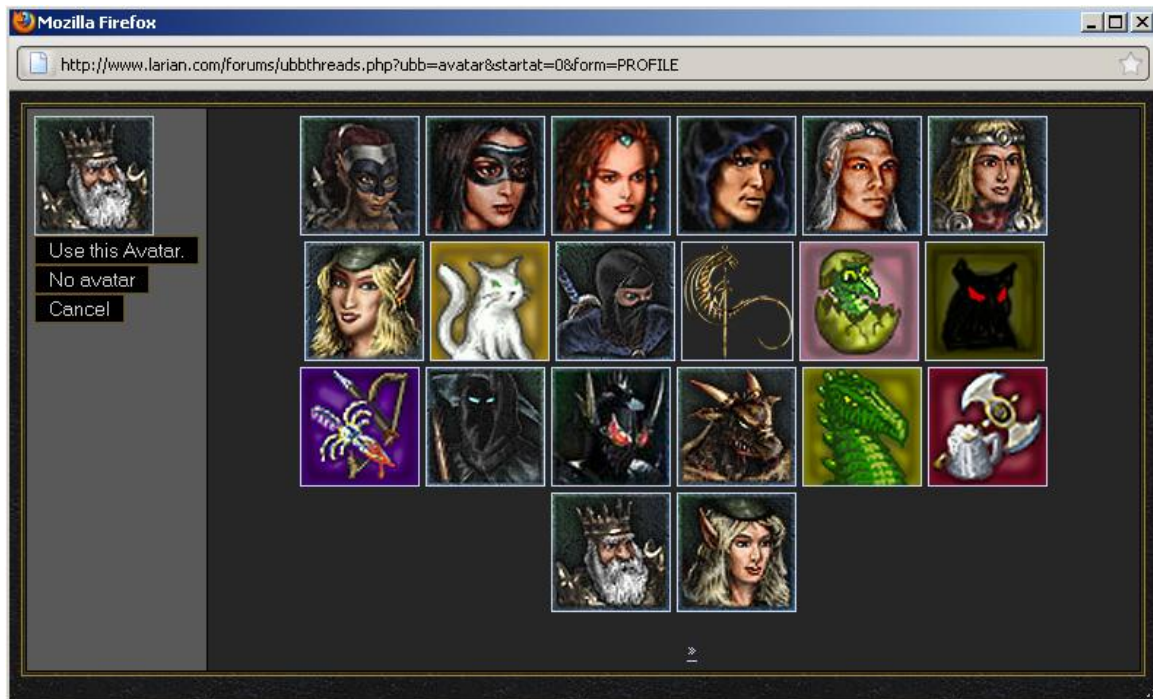


Abbildung 2 – Avatarerstellung durch Bildauswahl in einem Internetforum. Quelle: Forum der Larian Studios Webseite²

Für den Benutzer immer noch recht einfach, aber mit mehr Gestaltungsmöglichkeiten ist die Erstellung eines Avatars bei einem Rollenspiel, wie etwa World of Warcraft³. Die Avatarerstellung bei diesem Spiel ist in Abbildung 3 dargestellt. Der Benutzer kann hierbei aus einer Vielzahl von Wahlmöglichkeiten das individuelle Aussehen seines Avatars bestimmen, beispielsweise die Länge und Form der Haare oder die Hautfarbe. Bei Rollenspielen ändert sich meist das Aussehen des Avatars noch im Verlauf des Spiels, beispielsweise durch erworbene Gegenstände.

² <http://www.larian.com/forums>

³ <http://eu.battle.net/wow/de>



Abbildung 3 – Avatarerstellung bei World of Warcraft. Links unten sind die Auswahlmöglichkeiten für Haar- und Hautfarbe, etc. sichtbar. Quelle: World of Warcraft Webseite⁴

Das Motion Capture Verfahren kommt zum Einsatz, wenn bei animierten Avataren auf möglichst menschlich aussehende Bewegungsabläufe Wert gelegt wird. Nach Noack [24] werden hierbei spezielle Punkte, sogenannte Marker, auf das Original aufgetragen. Die Position dieser Marker wird von speziellen Kameras erfasst. Dadurch wird ein Drahtgittermodell des Originals im Rechner erstellt, welches durch die aufgenommenen Bewegungen animiert wird und mit Texturen überzogen werden kann. Ein Beispiel dafür ist in Abbildung 4 dargestellt. Diese Technik wird häufig zur Animation menschlicher Bewegungen von Avataren in Filmen und Computerspielen verwendet. Das Cyberneum⁵ arbeitet viel mit der Motion Capture Methode. Dort beschäftigt man sich mit der menschlichen Wahrnehmung und dem menschlichen Verhalten im Raum. Dazu wurde eine aufwendige Virtual Reality Umgebung eingerichtet, in der Avatare verwendet werden die mit der Motion Capture Methode animiert werden.

⁴ <http://www.wow-europe.com/shared/wow-com/images/guide/ss04.jpg>

⁵ <http://www.cyberneum.de>



Abbildung 4 – Motion Capture Verfahren. Die zu erfassende Person wird an bestimmten Körperstellen mit Markern versehen, die von Spezialkameras erfasst werden. Quelle: Webseite von Stephan Streuber beim Cyberneum⁶

Auch ein 3D Scanner kann verwendet werden, um einen Avatar zu erstellen. Nach Teutsch [34] kann die Oberfläche eines Objekts mit Hilfe von Laserstrahlen erfasst werden. Dabei entsteht eine dreidimensionale Punktwolke des Objekts aus der dann ein Dreiecksnetz berechnet wird. Dadurch kann die komplette Objektfläche dargestellt werden.

Avatare können auch komplett im Rechner entstehen, etwa durch Benutzung von Computergrafik- und Animationsprogrammen, wie beispielsweise Blender⁷ oder Maya⁸.

Auch die Forschung beschäftigt sich mit Avataren. Es gibt unzählige Arbeiten die sich mit der Erstellung von Avataren befassen. Es folgen nun einige ausgewählte Avatarerstellungsmethoden aus dem Forschungsgebiet der Avatare.

Carranza et al. [7] haben ein Verfahren entwickelt mit dem ausschließlich Avatare von Menschen erzeugt werden können. Dabei wird die abzubildende Person durch viele synchronisierte Kameras von unterschiedlichen Orten aus erfasst, wie in Abbildung 5 dargestellt. Hierbei liegt ein generisches Körpermodell mit einem kinematischen Skelett zugrunde. Aus jedem Kamerawinkel wird durch Hintergrund Subtraktion die Silhouette extrahiert. Sie bildet die Eingabe zu einem Algorithmus, der die Bewegungen berechnet. Dies geschieht durch Optimierung der Überdeckung des Körpermodells mit den Silhouetten. Dieses Verfahren ist ähnlich der Motion Capture Methode, bietet aber den Vorteil, dass nicht in den Aufnahmebereich eingegriffen werden muss, etwa durch Marker.

⁶ <http://www.kyb.tuebingen.mpg.de/nc/employee/details/stst.html>

⁷ <http://www.blender.org>

⁸ <http://autodesk.com/maya>

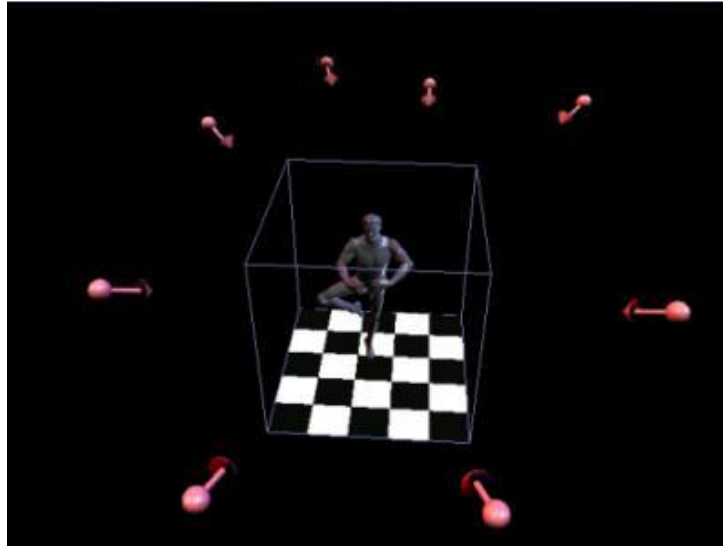


Abbildung 5 – Synchronisierte Erfassung eines Menschen durch sieben Kameras. Die roten Punkte markieren die Position der Kameras, die roten Linien ihre Blickrichtung. Quelle: Webseite zum Verfahren von Carranza et al.⁹

Nach Cho et al. [8] ist es auch mittels einer Tiefenkamera möglich einen Avatar zu erstellen. Eine Tiefenkamera nimmt sowohl visuelle Informationen als auch Tiefeninformationen von bewegten Objekten auf. Dabei geschieht die Aufnahme der Tiefeninformation mit Hilfe von sehr schnell gepulstem Infrarotlicht. Mit diesem Verfahren ist es auch möglich Objekte mit Haaren aufzunehmen, was bei etlichen anderen Methoden ein Problem darstellt. Weitere Vorteile bestehen darin, dass es sich um ein vollautomatisches Verfahren handelt und die Personen, die aufgenommen werden sollen, nicht unangenehm beeinträchtigt werden, da im Gegensatz zum Motion Capture Verfahren keine Marker verwendet werden müssen und das verwendete Infrarotlicht nicht sichtbar ist. Grundsätzlich besteht dieses Verfahren aus zwei Schritten: Alpha Matting und Generierung des Dreiecksnetzes. Die Alpha Matte wird mit den visuellen Informationen und den Tiefeninformationen kombiniert und daraus das Dreiecksnetz berechnet. Der Ablauf ist auch in Abbildung 6 dargestellt. Alpha Matting bezeichnet laut Grady et al. [12] die Trennung von Vorder- und Hintergrund bei Bildinhalten. Für dieses Avatarerstellungsverfahren wurde eine neuartige Alpha Matting Methode entwickelt, die durch Kombination von visueller Information und Tiefeninformation eine sehr präzise Trennung von Vorder- und Hintergrund ermöglicht, wie in Abbildung 7 ersichtlich.

⁹ http://www.mpi-inf.mpg.de/~theobalt/FreeViewpointVideo/free_viewpoint_video.html

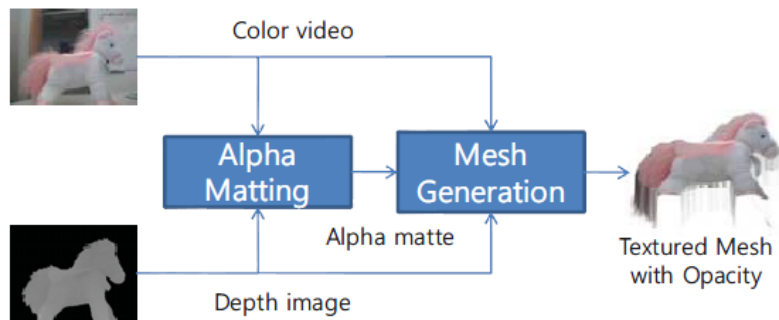


Abbildung 6 – Ablauf des Verfahrens von Cho et al. Zuerst erfolgt die Erstellung der Alpha Matte. Diese wird mit den visuellen Informationen und den Tiefeninformationen kombiniert, um daraus das Dreiecksnetz zu generieren. Quelle: Cho et al. [8]

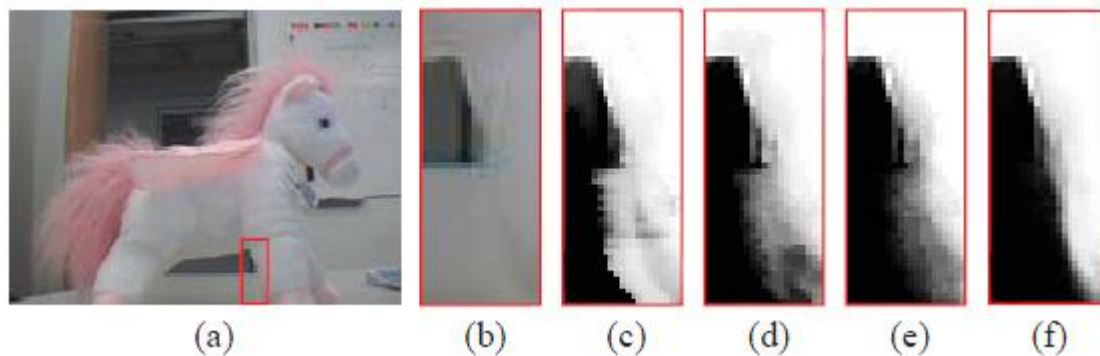


Abbildung 7 – Vergleich verschiedener Mattingmethoden. Bild (a) zeigt die Originalaufnahme, in Bild (b) ist der rot umrandete Ausschnitt aus Bild (a) vergrößert dargestellt. Es ist gut sichtbar, dass die Farben von Vorder- und Hintergrund in diesem Bildausschnitt nahezu gleich sind. In Bild (c) wurde auf den Bildausschnitt Bayesian Matting angewandt, in Bild (d) Robust Matting und in Bild (e) Closed Form Matting. Die fehlerhafte Trennung von Vorder- und Hintergrund ist deutlich sichtbar. Bild (f) zeigt die von den Autoren neu entwickelte Mattingmethode, welche das beste Ergebnis erzeugt. Quelle: Cho et al. [8]

Durch ein von Bitouk und Nayar [6] entwickeltes Verfahren ist es möglich, einen Avatar anhand einer einzelnen Fotografie zu erstellen. In Abbildung 8 ist der Ablauf dieses Vorgehens dargestellt. Diese Methode benutzt ein generisches Gesichtsanimationsmodell, um damit ein prototypisches Gesicht zu animieren. Die Deformation, die bei Sprachbewegungen entsteht, wird dabei durch die Linearkombination einer kleinen Menge von Basisvektorfeldern realisiert. Die Koeffizienten der Linearkombination sind die zeitabhängigen Gesichtsbewegungsparameter. Durch eine Reihe von zu setzenden Korrespondenzpunkten kann das prototypische Gesicht in ein beliebiges Gesicht transformiert werden, also das der Fotografie. Diese Transformation wird dann auch auf die Basisvektorfelder angewandt. Somit kann ein beliebiges Gesicht animiert werden. Durch Motion Capturing werden Gesichtsbewegungsparameter gewonnen die Hidden Markov Modelle trainieren. Dabei handelt es sich nach Dörfler [10] um statistische Modelle, welche verborgene und sichtbare Zustände besitzen. Sie werden häufig in der Sprach- oder Handschrifterkennung eingesetzt. Die nun trainierten Modelle sind dann in der Lage aus Text- oder Spracheingabe die letztendlichen Gesichtsbewegungsparameter zu generieren. Das Verfahren kann auch die Betonung von Wörtern während des Sprechens in die Animation einbeziehen. Zur weiteren Erhöhung des Realismus kann das System auch Augenbewegungen und Blinzeln einbinden.

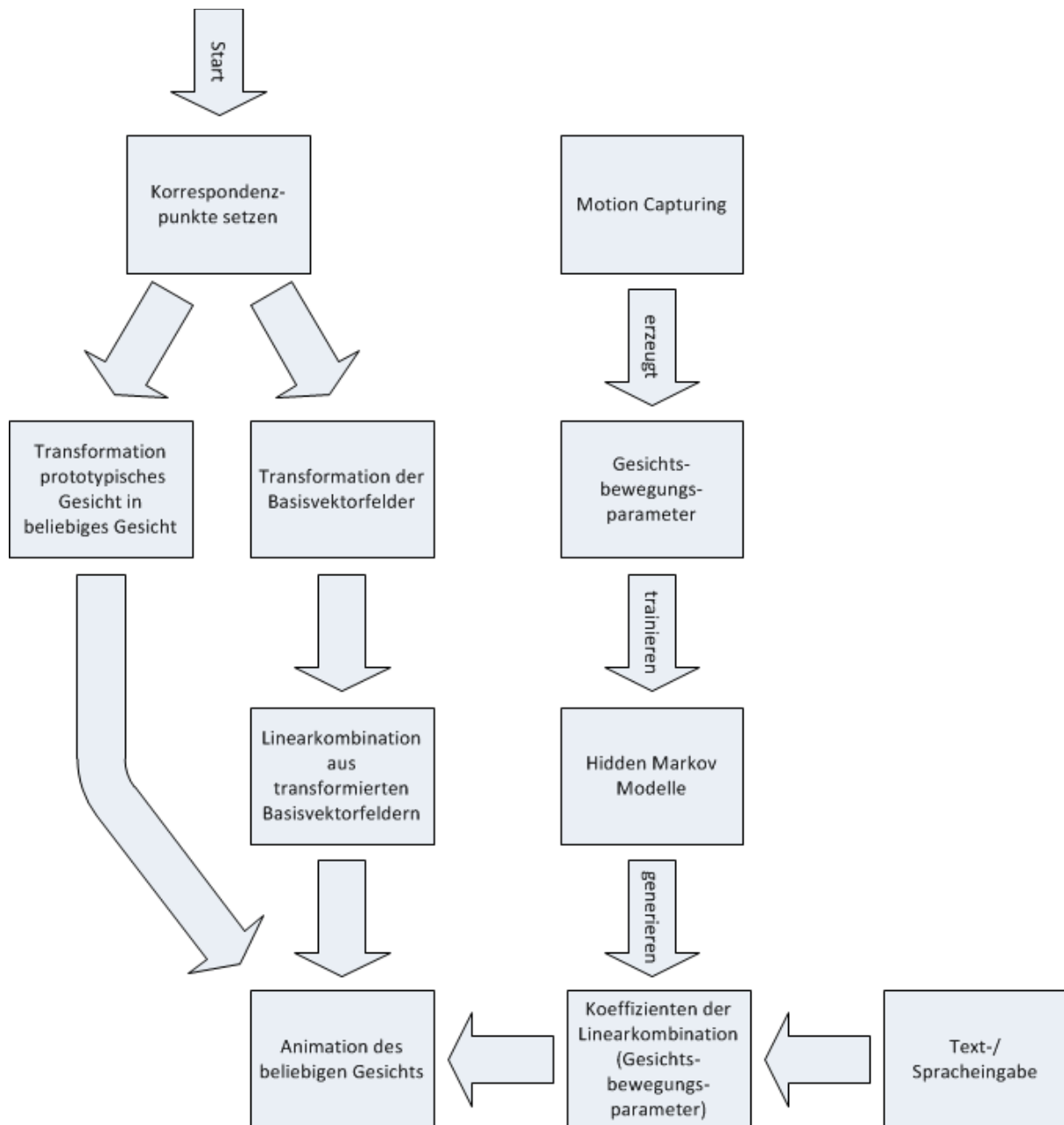


Abbildung 8 – Ablauf des Verfahrens von Bitouk und Nayar [6]. Auf die Transformation des prototypischen Gesichts und der Basisvektorfelder folgt die Animation des beliebigen Gesichts durch die Linearkombination der Basisvektorfelder. Deren Koeffizienten werden von Hidden Markov Modellen generiert, welche durch Motion Capturing Daten trainiert wurden.

Einen ähnlichen Ansatz verfolgen Leung et al. [16]. Hierbei wird mittels einer Videokamera das Gesicht einer Person aufgenommen. Es werden dann manuell in jedem 15. Bild des Videomaterials spezielle Punkte des Gesichts markiert. Laut den Autoren soll dies künftig automatisiert werden. Entsprechende Punkte müssen ebenfalls in einem generischen 3D-Gesichtsmodell gesetzt werden. Dadurch wird der Zusammenhang zwischen den Bilddaten im Video und dem 3D-Gesichtsmodell hergestellt und es wird möglich, die Regionen, die nicht durch einen Punkt markiert sind, mit einer Textur zu überziehen. Bei dieser Textur handelt es sich um eine Cylindrical Texture Map, also einer kontinuierlichen Ansicht eines Objekts über einen großen Winkelbereich hinweg, wie in Abbildung 9 dargestellt. Diese

Textur wird aus verschiedenen Ansichten der Person generiert und auf das 3D-Gesichtsmodell aufgebracht. Der Gesamtprozess ist auch in Abbildung 10 dargestellt. Aufgrund der Tatsache, dass in jedem 15. Bild des Videomaterials von Hand die Punkte im Gesicht gesetzt werden, sind bisher nur statische Avatare möglich. Sobald animierte Avatare angezeigt werden können, soll dieses System auf Rechnern, deren Netzverbindung nur eine geringe Bandbreite hat, trotzdem visuelle Kommunikation ermöglichen. Es soll möglich sein, dass der den Mundbewegungen entsprechend animierte Avatar des jeweiligen Benutzers sichtbar ist, aber kein Echtzeitbild. Ein Nachteil besteht derzeit darin, dass keine beliebige Kopfhaltung möglich ist.



Abbildung 9 – Cylindrical Texture Map. Quelle: Leung et al. [16]

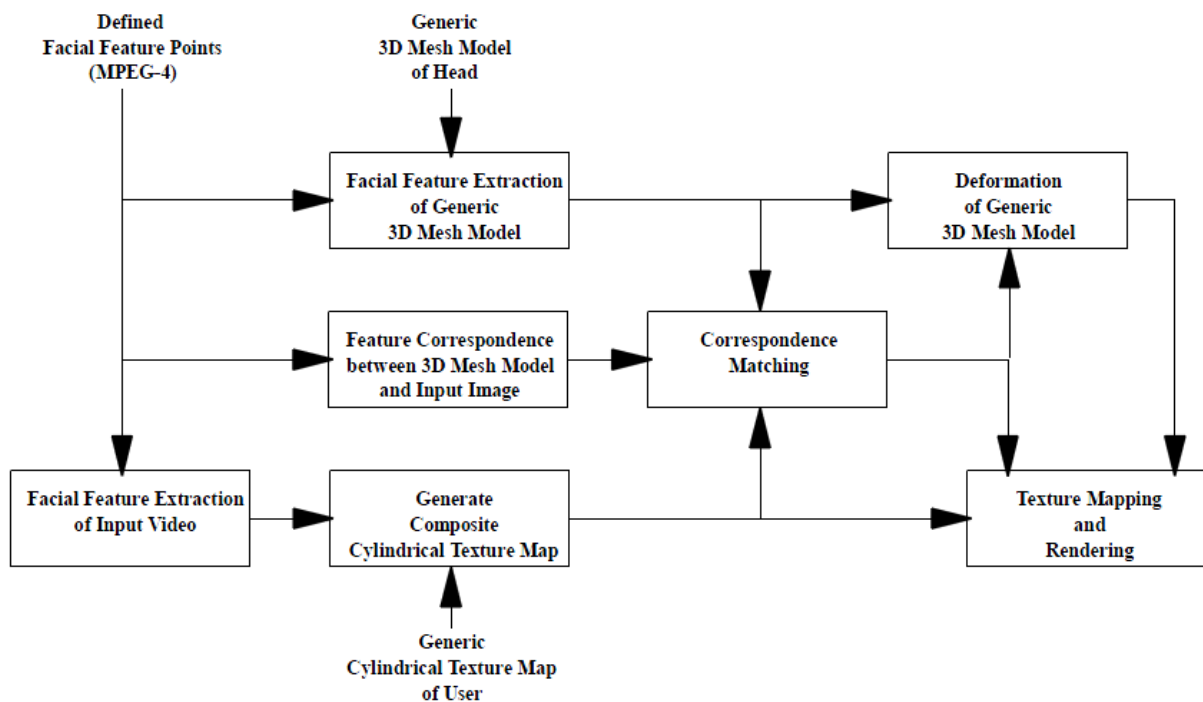


Abbildung 10 – Ablauf des Verfahrens von Leung et al. [16]. Nach der Markierung spezieller Punkte im Gesicht, werden diese ebenfalls im generischen 3D-Gesichtsmodell gesetzt. Dieses wird anschließend mit Hilfe einer Cylindrical Texture Map mit dem Gesicht der zu erfassenden Person überzogen. Quelle: Leung et al. [16]

In Computerspielen oder Filmen besteht häufig das Problem, dass eine große Menge von Avataren benötigt wird. Ein Paper von Albin-Clark und Howard [1] zeigt überblicksartig wie unter Verwendung von evolutionären Algorithmen Avatare entstehen, die sowohl glaubhaft aussehen als auch einzigartig innerhalb der jeweiligen Bevölkerung sind. Das Aussehen der Avatare wird durch ihren Genpool bestimmt, der Elemente wie Armlänge, Körperfett oder Bauchmuskeln beinhaltet. Diesen Elementen wird ein einfacher Wert zugewiesen,

beispielsweise eine Fließkommazahl. Aus diesen Werten wird dann der Avatar erstellt. Dieses Vorgehen ist laut den Autoren etwa bei dem Onlinespiel Second Life¹⁰ der Fall. Die bei evolutionären Algorithmen üblicherweise verwendete Fitness wird dadurch bestimmt, wie stark sich ein Avatar von den anderen in der Bevölkerung unterscheidet. Dazu muss der Fitnessbestimmung ein geeignetes Distanzmaß zugrunde liegen, welches die Unterschiedlichkeit zweier Avatare misst. Fitness bezeichnet nach Beyer et al. [28] bei evolutionären Algorithmen die Bewertung eines Individuums hinsichtlich seiner Reproduktionstauglichkeit.

In einem Video von Zollhöfer et al.¹¹ werden Avatare mittels der Microsoft Kinect Konsole erstellt. Dabei sieht der Benutzer ein aktuelles Graustufenbild seines Gesichts. Zu einem vom Benutzer definierten Zeitpunkt startet er mit dem Drücken einer Taste die Erstellung und kann dann einige Sekunden später seinen fertigen Avatar sehen. Der Ablauf dieser Vorgehensweise ist in Abbildung 11 dargestellt. Es werden folgende Schritte vollzogen: Zuerst werden die Rohdaten aufbereitet und Kennpunkte wie Augen, Nasenspitze und Kinn automatisch markiert. Anschließend wird der Hintergrund abgetrennt um Rechenleistung zu sparen. Anhand der Kennpunkte werden die erfassten Daten nun an einer generischen Gesichtsvorlage ausgerichtet. Diese wird durch Non Rigid Registration so deformiert, dass sie möglichst so aussieht wie die erfassten Daten. Non Rigid Registration bezeichnet nach Sierra [30] die Anordnung von nichtlinearen oder uneinheitlichen Datenmengen. Da die deformierte Fläche nun viele Einbuchtungen und Unebenheiten aufweist wird sie an ein wandelbares Gesichtsmodell angepasst. Dadurch werden die Unebenheiten beseitigt. Zum Abschluss kann noch die RGB Textur vom Kinect Sensor auf die Fläche gelegt werden. Der Aufbau zur Avatarerstellung ist in Abbildung 12 dargestellt.

¹⁰ <http://www.secondlife.com>

¹¹ <http://www9.informatik.uni-erlangen.de/publications/publication/Pub.2011.tech.IMMD.IMMD9.automa>

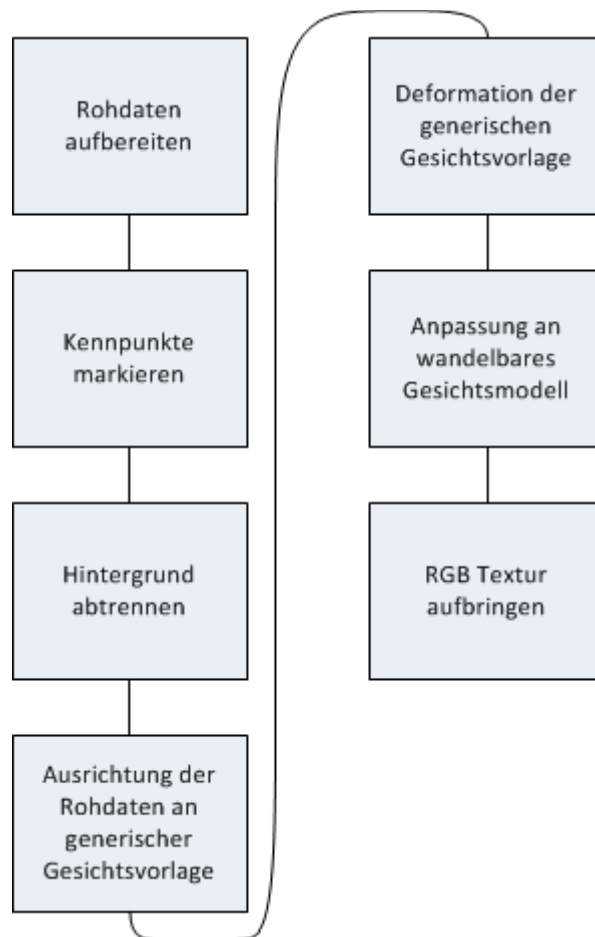


Abbildung 11 – Ablauf des Verfahrens von Zollhöfer et al. Nach der Aufbereitung der Rohdaten erfolgt die Markierung der Kennpunkte und der Hintergrund wird abgetrennt. Im Anschluss werden die Rohdaten an der generischen Gesichtsvorlage ausgerichtet, diese wird deformiert und an das wandelbare Gesichtsmodell angepasst. Abschließend wird die RGB Textur aufgebracht.



Abbildung 12 – Aufbau zur Avatarerstellung per Microsoft Kinect Konsole nach Zollhöfer et al. Auf dem linken Monitor wird das aufbereitete Signal vom Kinect Sensor dargestellt, anhand dessen der Benutzer seinen Kopf positionieren kann. Auf dem rechten Monitor ist bereits das Endergebnis sichtbar. Quelle: Video zum Verfahren von Zollhöfer et al.¹²

Eine weitere Möglichkeit zur Avatarerstellung bieten Avatar Frameworks. Deren Untersuchung bildet einen Schwerpunkt dieser Arbeit und ist in Kapitel 5 zu finden.

Dieser Abschnitt hat verschiedene Ansätze zur Erstellung von Avataren vorgestellt. Den Anfang bildeten Erstellungsmethoden aus dem täglichen Gebrauch, wie die Auswahl eines Bilds in einem Internetforum oder einem Computerspiel. Diese Methoden zeichnen sich durch einfache Benutzbarkeit aus und werden von vielen Benutzern auch häufig eingesetzt. Daraufhin folgte die Vorstellung des Motion Capture Verfahrens, des 3D Scanners und der Animationsprogramme. Im Gegensatz zu den Erstellungsverfahren aus dem täglichen Gebrauch sind diese Vorgehen dem professionellen Bereich der Avatarerstellung zuzuordnen. Sie kommen etwa bei der Produktion von Kinofilmen, Fernsehsendungen oder Computerspielen zum Einsatz. Abschließend wurden sechs Erstellungsmethoden aus der Avatarforschung beschrieben. Bei ihnen bestehen häufig diverse Probleme, etwa dass sie auf experimentellen Aufbauten basieren, nicht von jedermann bedient werden können oder Ergebnisse liefern die noch verbesserungswürdig sind. Ihre Zielsetzung liegt darin neuartige Avatarerstellungsmethoden zu erforschen, zu verbessern und ihre Tauglichkeit für die Praxis zu ermitteln.

2.2 Entwicklungskonzepte für Avatanwendungen

Dieses Kapitel soll beispielhaft zeigen, inwiefern Avatare in den Entwicklungszyklus eines Softwareprojekts integriert werden können. Da es keine Referenzliteratur gibt, wird hierfür das im Software Engineering häufig verwendete Wasserfallmodell (siehe Ludewig und

¹² <http://www9.informatik.uni-erlangen.de/publications/publication/Pub.2011.tech.IMMD.IMMD9.automa>

Lichter [19]) als Orientierung verwendet und anhand dessen Phasen die Eingliederung des Avatars in den Prozess diskutiert.

Das Wasserfallmodell gliedert sich in Phasen, welche nacheinander durchlaufen werden.

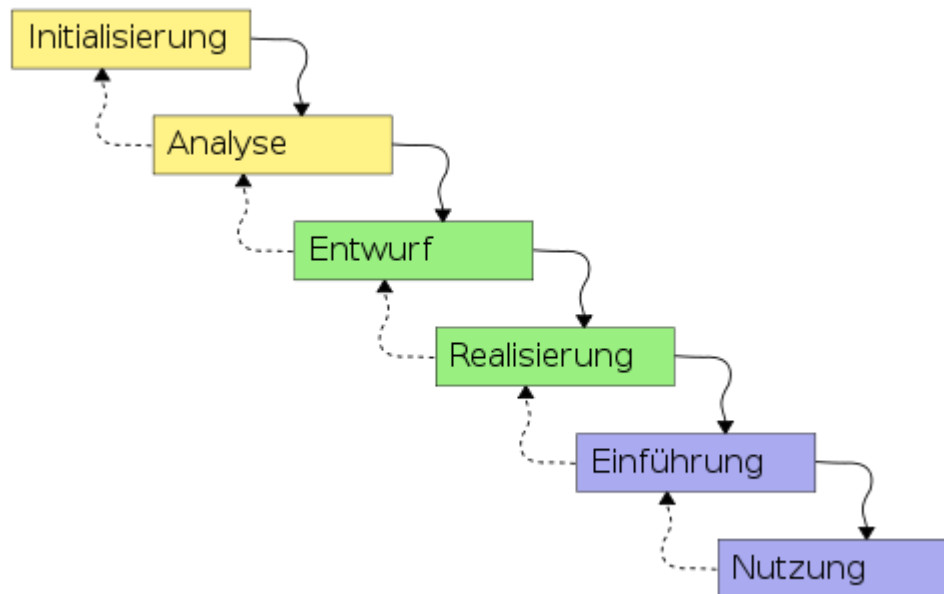


Abbildung 13 – Ablauf eines Softwareentwicklungsprozesses nach dem Wasserfallmodell. Es werden sechs Phasen (Initialisierung, Analyse, Entwurf, Realisierung, Einführung, Nutzung) definiert, die nacheinander abgearbeitet werden bis am Ende das Endprodukt erstellt ist.

Während der Analyse sammeln Auftraggeber (Kunde) und Auftragnehmer (Entwickler) die Anforderungen an das System. Diese werden durch gezielte Befragung in Form von Interviews von Analysespezialisten ermittelt. Bereits zu diesem Zeitpunkt muss der Kunde sich mit dem Entwickler darüber einig sein, ob ein Avatar Teil des Systems werden soll und wenn ja, welche Funktionalität dieser erhalten und abdecken soll.

In der Spezifikationsphase wird die genaue Funktionalität des Systems festgelegt. Hier wird der Avatar unter Anderem in den Punkten Funktionalität, Verhaltensweise, Optik, Sprache, Stimmlage, Gestik, Mimik etc. beschrieben. Spätestens zu diesem Zeitpunkt sollten Usability Tests mit Prototypen mit zufälligen Personen durchgeführt werden, um die Akzeptanz zu überprüfen und Rückmeldungen zu erhalten. Zusätzlich sollten umfangreiche Testfälle erstellt werden, um die Anforderungsabdeckung überprüfen zu können.

Im SESAM Projekt der Universität Stuttgart¹³ wurde eine Architektur entwickelt, die anhand einer eigenen Modellsprache nachträglich Elemente dem System hinzufügen kann ohne dass der Programmcode verändert werden muss. Dabei erstellt der Benutzer mit Werkzeugen neue Modelle unter Einsatz dieser Modellsprache. Der Model Compiler übersetzt dies in eine für das Programm lesbare Form, sodass das Modell einsatzbereit ist. Dieses Architekturprinzip wäre etwa im Zusammenspiel mit der Avatar Markup Language (siehe

¹³ <http://www2.informatik.uni-stuttgart.de/iste/se/research/sesam/index.html>

Kshirsagar et al. [15]) auch in Anwendung auf Avatare denkbar. Die Sprache basiert auf XML (eXtensible Markup Language) und nutzt Parameter aus dem MPEG-4 Standard. Der Avatar wird dabei in einen Kopf- und einen Körperbereich unterteilt. Für jeden Bereich können Abschnitte angelegt werden, die beschreiben, auf welchen Körperteil wann welche Animation ausgeführt werden soll. Dies wäre eine Möglichkeit, einen Avatar in einen Softwareentwurf zu integrieren.

Nach Abschluss der Programmierarbeiten steht der Systemtest an, bei welchem das System als Ganzes getestet wird. Hierbei steht auch der Avatar erneut auf dem Prüfstand, nun aber ebenfalls im großen Kontext. Der Test sollte idealerweise keine Fehler mehr aufdecken.

Die in dieser Ausarbeitung getesteten Frameworks bieten nur geringe Unterstützung für die Integrationen ihrer Avatare in einen generischen Entwicklungsprozess. Die Ausgabe erfolgt meistens in einem Datenformat (etwa ein Bild- oder Videoformat), welches keine Eingriffe zulässt, die wichtig wären (Auslesen der Charaktereigenschaften, Interaktionssteuerung, Verändern der Stimmtonalität, dem Avatar eine andere Haarfarbe geben, etc.). Da sich Avatare somit nicht oder nur sehr schwierig durch Software Dritter in ein Projekt integrieren lassen, beschreibt Lindner [17] einen Ansatz für ein Projekt zur Entwicklung eines Avatars in fünf Phasen (Strategiephase, Konzeptphase, Auswahlphase, Realisierungsphase, Betriebs- und Erfolgskontrollphase)



Abbildung 14 – Phasenmodell zur Entwicklung eines Avatars nach Lindner [17]

In der Strategiephase wird das Projekt analysiert. Ziel ist die Klärung des Projektziels und die Beseitigung von Unklarheiten. Hierbei arbeiten Kunde und Projektleitung eng zusammen, um zu einem gemeinsamen Konsens zu kommen und die anfallenden Fragen möglichst exakt und transparent zu klären. Herrscht Einigkeit über das Projektziel, muss man sich der Frage stellen, wie der Avatar ins Spiel kommt. Dabei gilt vor allem der Grundsatz, dass der Avatar nicht zu breit aufgestellt sein und die gesamte Palette an Funktionen verfügbar haben sollte, da solche Projekte meist scheitern würden. Daher wird geraten, den Avatar einen Teilbereich im Projekt zu besetzen und in diesem sein Potential ausschöpfen zu lassen.

In der Konzeptphase wird aus den Analyseergebnissen das Lastenheft erstellt und daraus die Anforderungen abgeleitet. Man darf hier nicht den Fehler begehen und sich bereits jetzt auf die technischen Umsetzungen festlegen, da man sich damit einschränkt und auf die Gegebenheiten der gewählten Technik festlegen muss. Die Koordination der verschiedenen beteiligten Parteien (Grafiker, Softwareentwickler, Marketing, Management, Analytiker) sowie die Frage, welche Teile des Konzepts an externe Partner vergeben werden sollen, stellen die nächste Hürde dar. Es wird davon abgeraten, die gesamte Konzeptphase auszulagern, da damit eventuell das Ziel des Projekts nicht erfasst wird.

Die Organisation der beteiligten Parteien wird in der Auswahlphase beschlossen. Hierbei hat man die Wahl, das Projekt komplett in Eigenregie durchzuführen oder externe Softwarehersteller, Systemhäuser oder Grafik-Agenturen zu beauftragen. Jeder Partner sollte seine Vorstellungen dem Kunden präsentieren und basierend darauf sollte das oder die Partnerunternehmen ausgesucht werden. Jedes Vorgehen besitzt seine Vorzüge, so sind etwa Grafikagenturen besser für die optische Gestaltung geeignet als Softwarehersteller.

Die Realisierungsphase startet mit der Erstellung eines Pflichtenhefts, aufgrund dessen das System erstellt wird. Mit dem heutigen Stand der Technik ist es verhältnismäßig einfach, Software zu schreiben. Einen Avatar dazu zu entwickeln, der einen Mehrwert für den Benutzer darstellt, ist dagegen weitaus schwieriger. Seine Benutzerakzeptanz hängt vor allem von weichen, subjektiven Anforderungen ab (Stil, Tonalität, Charakter, etc.). Für die Dialogführung müssen ausgebildete Texter herangezogen werden sowie die Konversationsmaxime von Grice beachtet werden. Diese besagen, dass man einen Gesprächsbeitrag so gestalten sollte, dass er dem Zweck der Unterhaltung dient. Erst dann entsteht ein Dialogmodell, das für den professionellen Einsatz tauglich ist.

Eine gute Software sollte lernfähig sein und die Eigenschaft in der Betriebsphase nutzen, um den Entwicklern Rückmeldung zu geben, worauf die Benutzer den Fokus legen (mehr Tiefgang oder mehr Breite im Dialog mit dem Avatar). Ein umfangreicher Praxistest dient der Findung von Schwachstellen im System. Pflege und Weiterentwicklung sind die Daueraufgaben im Entwicklungsprozess.

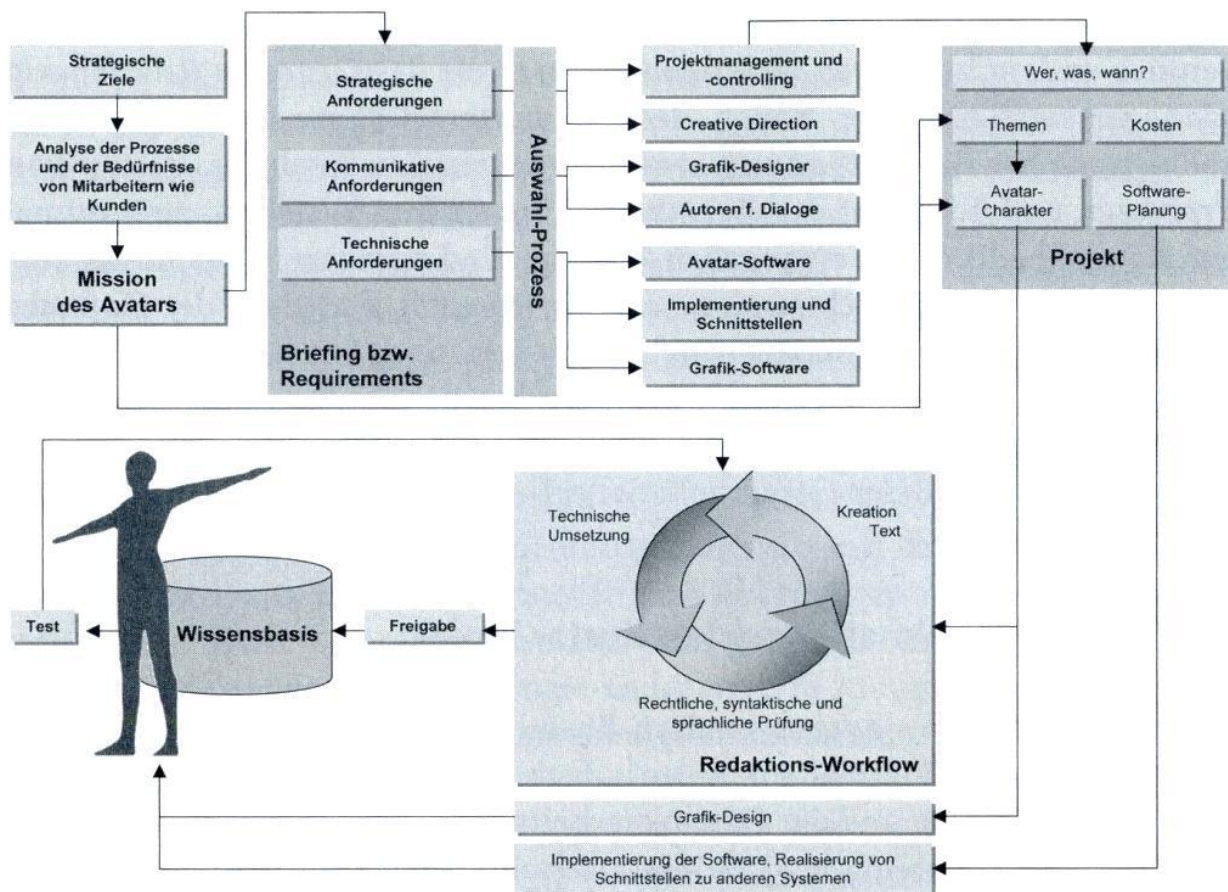


Abbildung 15 – Entwicklungsprozess eines Avatars. Prozess startet bei unsortierten Gedanken und endet nach der Realisierung durch ein Projekt. Quelle: Lindner [17]

Avatare sind ein relativ neuer Begriff im Bereich der Mensch-Computer-Interaktion. Gerade deswegen ist es nicht verwunderlich, dass sich noch keine Standards im Umgang mit ihnen entwickelt haben. Dieser Bereich wird sich allerdings in den kommenden Jahren stark weiterentwickeln und deshalb werden für die Zukunft neue Modelle und Prozesse entstehen, mit denen die Entwickler arbeiten können.

2.3 Beispiele für Avatare

Avatare sind heutzutage häufig in der Praxis anzutreffen. Dieser Abschnitt zeigt an zahlreichen Beispielen die Verwendung von Avataren in der Praxis.

2.3.1 Avatare in Computer- und Konsolenspielen

Im Folgenden werden die Konsolen Microsoft Kinect, Nintendo Wii und Sony Move beschrieben. Daran schließt sich die Beschreibung des Rollenspiels Diablo an.

Microsoft Kinect

„Der Körper ist die Konsole“ ist der Leitspruch der Werbung für Kinect von Microsoft. Bei Kinect kann man ohne Konsole nur durch die Bewegungen des Körpers den Avatar kontrollieren und in der virtuellen Welt kommunizieren.



Abbildung 16 – Microsoft Kinect: „Der Körper ist die Konsole“. Man spielt mit der Kinect Konsole mit dem eigenen Körper und braucht keinen Controller mehr. Quelle: Gamespot Webseite¹⁴

Das Sensorsystem der Kinect Konsole besteht aus einer Kamera, einem Infrarotsender, einem Infrarotempfänger und vier Mikrofonen. In der Mitte befindet sich die normale Kamera mit der die Spieler unterschieden und ihre Gesichtsbewegungen erfasst werden. Die beiden Infrarotgeräte dienen dazu das dreidimensionale Tiefenbild aufzubauen.

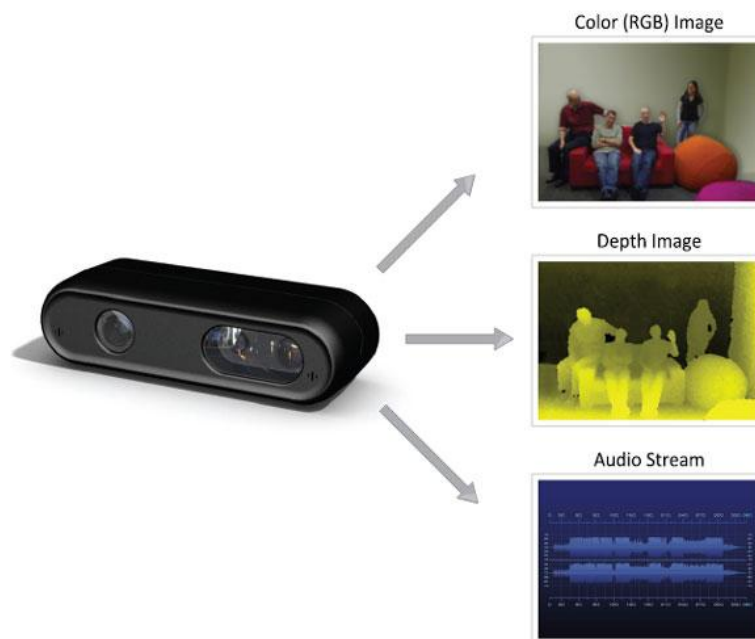


Abbildung 17 – Sensoren der Microsoft Kinect. Die RGB Kamera erzeugt ein normales Farbbild, der 3D Tiefensensor erzeugt das Tiefenbild und verschiedene Mikrofone nehmen die Stimme auf. Quelle: PrimeSense¹⁵

Mit dem 3D Tiefensystem ermittelt die Kinect Konsole die Bewegung der Spieler. Mit den Mikrofonen erfasst Kinect die Stimme der Spieler. Somit entsteht das Farbbild, das dreidimensionale Tiefenbild und die Stimme wird aufgenommen.

¹⁴ http://e3.gamespot.com/special_feature/microsoft/image-feature/index.html?image=14

¹⁵ <http://www.primesense.tw/?p=739>

Nintendo Wii

Am 06.11.2006 zeigte Nintendo sein Spielzeug der neuesten Generation. Die Entwicklung der Wii war eine Revolution im Bereich der Mensch-Rechner-Interaktion.



Abbildung 18 – Nintendo Wii. Wii wurde von Nintendo entwickelt und führte zu einer Revolution der Mensch-Rechner-Interaktion. Quelle: cngba Webseite¹⁶

Die Entstehung der Wii bricht die Grenze zwischen Menschen und Avataren, die in der virtuelle Welt leben und führt Methoden ein, damit man mit Avataren in der virtuellen Welt interagieren kann. Traditionell benutzt man Konsolen um mit Avataren in der virtuellen Welt zu kommunizieren. Die Wii ist das erste Gerät, mit dem man Avatare in der virtuellen Welt durch Bewegungen des Körpers kontrollieren kann. Das heißt, wenn man seinen Körper bewegt, macht der Avatar in der virtuellen Welt die gleiche Bewegung. Die Wii zeigt die Möglichkeit, dass in der Zukunft zwischen virtueller und realer Welt viel mehr passieren kann. Mit Hilfe der Wii Technik kann man in der virtuellen Welt einfach Dinge tun, die in der Realität nur schwer umzusetzen sind.

¹⁶ <http://mall.cngba.com/Upload/wii.jpg>



Abbildung 19 – Nintendo Wii. Das erste Spielzeug der Welt, das mit Avataren in der virtuellen Welt durch Körperbewegungen kommunizieren kann. Quelle: IT Webseite¹⁷

Wiimote heißt die Fernbedienung für die Wii Konsole. Sie sieht aus wie eine normale Fernbedienung für einen Fernseher. Sie kommuniziert per Bluetooth mit der Wii. Auf der Vorderseite der Wiimote gibt es einen Infrarotempfänger, der die Signale des Infrarotsenders der Wii empfängt. Der Sender darf nur vor oder auf dem Bildschirm platziert werden. In der Wiimote gibt es ein NVRAM und ein Kreiselinstrument. Mit Hilfe des Kreiselinstruments kann die Wiimote erkennen, wie die Bewegungen des Spielers jeweils auf der X-, Y- und Z-Achse liegen. Das NVRAM dient dazu Miis zu speichern und zu transportieren.

Sony Move

Um im Wettbewerb des Spielzeugmarkts gegen Nintendo und Microsoft nicht zu verlieren, entwickelte Sony sein eigenes System mit dem Namen Move. Mit diesem System kann man den Avatar in der virtuellen Welt durch den eigenen Körper kontrollieren. Move funktioniert ähnlich wie Wiimote. Der größte Unterschied zwischen Move und Wiimote ist der, dass Move eine Kamera namens Playstation Eye benutzt, die das Licht aus der Lampe des Bewegungskontrollers empfängt um daraus die Bewegungen des Spielers zu berechnen.

¹⁷ www.it.com.cn/f/games/065/10/268985.htm



Abbildung 20 – Playstation Eye Kamera. Sie empfängt das Licht des Bewegungskontrollers, um die Bewegungen des Spielers zu berechnen. Quelle: Techbang Webseite¹⁸

Diablo

Im Bereich der Computerspiele wurden Avatare schon vor Langem eingesetzt, um Figuren zu haben, die in der virtuellen Welt leben und Abenteuer erleben können. Mit der Hilfe von Avataren kann man das Gefühl bekommen, dass man selber in der virtuellen Welt lebt. Man kann dort alles tun, was man in der realen Welt nicht tun kann, wie z.B. Monster jagen, Schätze finden, etc. Das bekannteste Spiel, das die Avatar Technik benutzt, ist Diablo.

Diablo I wurde im Jahr 1996 von Blizzard entwickelt. In diesem Spiel kann man seinen eigenen Avatar in der Welt von Diablo kontrollieren und dem Avatar verschiedene Gegenstände geben, um ihn dadurch zu verbessern. Die Interaktion mit dem Avatar erfolgt durch Tastatur und Maus.



Abbildung 21 – Bildschirmfoto von Diablo I. Es ist möglich den Avatar mit verschiedenen Gegenständen auszustatten, um ihn zu verbessern. Quelle: Diablo1 Webseite¹⁹

¹⁸ <http://www.techbang.com.tw/posts/3421-t-off-defeat-playstation-move-experience>

2.3.2 Avatare im Kino

Die Avatartechnik wurde schon in vielen Filmen eingesetzt, wie z.B. Beowulf oder Avatar. Dabei ist Avatar der bekannteste Film, der die Avatartechnik verwendet hat. Durch Zusammensetzung von 3D Modellierung und dynamischer Erfassung kann bei der Modellierung der Films mehr Genauigkeit geschaffen werden und so der Realismus gesteigert werden.

Im Film Avatar sieht man auf den Gesichtern der Navi, welche mit Hilfe von Avatartechnik modelliert wurden, viele verschiedene Gesichtsausdrücke, wie Freude, Bösartigkeit, Trauer, etc. Alle Gesichtsausdrücke, die Menschen haben können, werden wie echt dargestellt. Ein Navi besteht aus einem 3D Model, bei dem das Gesicht und die Bewegungen durch Menschen modelliert wurden. Zuerst wird das 3D Model erzeugt, dann werden durch eine spezielle Kamera die Bewegungen des Gesichts und des Körpers der Schauspieler aufgenommen. Zuletzt werden die Bewegungen auf das 3D Model übertragen. Im Bereich von Filmen nennt man diese Technik 3D Dynamic Capture.

2.3.3 Avatare als Lernhilfe

Als Lernhilfe haben Avatare natürlich auch Vorteile. Man kann durch sie bessere 3D Charaktere erzeugen, die mehr gewünschte Details präsentieren können. Mit solch einem Charakter wird sich die Motivation beim Lernen erhöhen, weil man einem lustigen Charakter mehr Aufmerksamkeit schenkt. Mit Hilfe von Avataren können die eigenen Antworten mit den schon in der Datenbank gespeicherten Standardlösungen verglichen werden, damit man sofort und genau weiß, ob die Kenntnisse schon erworben wurden oder nicht. In Abbildung 22 ist das Programm Your Shape Fitness Evolved dargestellt. Daraus wird ersichtlich, wie mit Hilfe eines Avatars, eine Kung Fu Haltung erlernt werden kann. Nimmt der Benutzer die Haltung ein, so gibt das Programm eine Bewertung der eingenommenen Haltung aus. Zusätzlich erfolgt durch den Avatar eine Rückmeldung was an der Haltung verbessert werden kann.

¹⁹ www.diablo1.ru/gallery/humor.php



Abbildung 22 - Ubisoft: Your Shape Fitness Evolved. Mit Hilfe dieser Software kann Yoga, Kung Fu, etc. selbst erlernt werden. Quelle: Gamespot²⁰

2.3.4 Avatare als Orientierungshilfe

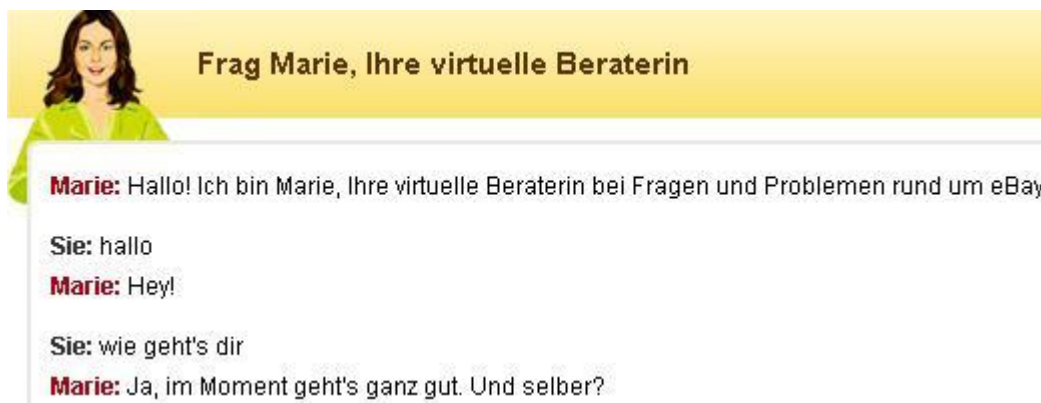


Abbildung 23 – Marie, die virtuelle Beraterin von Ebay. Quelle: Ebay.de Webseite²¹

Ein Avatar zur Orientierung ist Marie, die virtuelle Beraterin von Ebay. Sie wird von Ebay eingesetzt, um mit Hilfe von vordefinierter Logik den Kunden auf der Ebay Webseite eine Orientierungshilfe zu bieten. Es ist möglich mit Marie zu chatten und ihr Fragen zu stellen. Marie ist in der Lage auf sinnvolle Fragen sinnvolle Antworten zu geben. Auf Fragen, die nicht in der Datenbank vordefiniert wurden, erfolgt eine Standardantwort. Beispielsweise wird die Frage „Was ist 3 mal 9?“ von ihr mit „Ich habe Sie nicht verstanden. Können Sie das noch mal anders sagen?“ beantwortet.

²⁰ http://www.gamespot.com/xbox360/sports/yourshapefitness evolved/images.html?sid=6266611#site_nav

²¹ <https://marie.ebay.de/site/marie.aspx>

2.3.5 Avatare zur Kommunikation

Im Jahr 2001 hat die koreanische Webseite SayClub ihr Geschäft mit Avataren begonnen und viele asiatische Firmen folgten ihr nach. Tencent ist eine davon. Tencent hat in seiner eigenen Instant Messenger Plattform QQ in den vergangenen Jahren Avatare erfolgreich benutzt. Man kann in der virtuellen Welt der QQShow, seinen individuellen Charakter erzeugen, damit man im Internet oder in anderen Kommunikationsplattformen wie etwa dem Handy, die angestrebte Persönlichkeit zeigen kann, um sich von anderen zu unterscheiden oder um seine Zufriedenheit zu steigern. In Abbildung 24 ist dargestellt wie Menschen per Avatar in der virtuellen Welt kommunizieren. Sie kaufen mit realem Geld virtuelle Kleidung ein, entweder um ihren eigenen Avatar zu verbessern oder um sie Anderen zu schenken.



Abbildung 24 – QQChat. Hier kann der eigene Stil in der virtuellen Welt gezeigt werden. Quelle: QQ von Tencent²²

²² <http://pc.qq.com/>



Abbildung 25 – QQShow Portal des QQ Messengers. Hier ist die Ausstattung des eigenen Avatars möglich. Quelle: QQ Messenger von Tencent²³

2.3.6 Avatare in der Werbung

Im Jahr 2000 hat die Deutsche Telekom ihren Avatar Robert T-Online erschaffen, um ihre Tochterfirma T-Online zu bewerben. Der Münchner Musical-Schauspieler Matthias Kostya verlieh Robert seine Mimik und seine Emotionen. Mit Hilfe einer speziellen Motion-Control-Kamera wurden insgesamt 90 Einstellungen aufgenommen.



Abbildung 26 – Robert T-Online, der Avatar der deutschen Telekom. Quelle: Netzmafia Webseite²⁴

2.3.7 SecondLife und Ableger

SecondLife ist ein bekanntes Onlinespiel. Das Spiel wurde 1999 von Linden Lab in San Francisco entwickelt und ist seit dem 24.06.2003 online. Die Second-Life-Welt wird von der Client-Software als kontinuierliche 3D-Animation dargestellt. Die Spieler werden in Second-Life als Bewohner bezeichnet, wobei der Spieler die im Spiel vorhandenen Gegenstände

²³ <http://pc.qq.com/>

²⁴ www.netzmafia.de/service/robert-t.html

benutzen kann, um seine individuellen Avatar zu gestalten. Dadurch kann sich der Spieler mit dem Avatar identifizieren. Spieler benutzen in Second Life individuelle Avatar Kleidung, um sich von anderen Spielern zu unterscheiden und mit ihnen zu kommunizieren. In Second Life ist der Avatar wichtig, nur damit kann man problemlos miteinander interagieren.



Abbildung 27 – Avatarerstellung in SecondLife. Mithilfe eines Dialogs ist es in SecondLife möglich, einen personalisierten Avatar zu erstellen. Quelle: JunChen Wu Webseite²⁵

²⁵ www.junchenwu.com



Abbildung 28 – Der Startort von SecondLife. Quelle: JunChen Wu Webseite²⁶

Im Vergleich zu Second Life ist der Schwerpunkt des koreanischen Spiels Audition Dance Battle²⁷ die gesellschaftliche Interaktion. Das Spiel hat ein Hauptthema: Das Tanzen. Der Spieler erzeugt seinen eigenen Avatar, verwendet schönen Schmuck und Kleidung und verwendet spezielle Tanzschritte. Dabei spielt der Avatar eine wichtige Rolle. Mit einem guten Avatar hat man in der virtuellen Welt eine bessere Chance mit anderen Leuten zu kommunizieren.



Abbildung 29 – Szene aus Audition Dance Battle. Man erzeugt in der Welt einen eigenen Avatar, stellt individuelle Tanzschritte zusammen und benutzt verschiedene Kleidung und Schmuck, um mit anderen zu kommunizieren. Quelle: Readnovel Webseite²⁸

²⁶ www.junchenwu.com

²⁷ <http://audition.de.alaplaya.net>

²⁸ http://blog.readnovel.com/article/htm/tid_73661.html

2.3.8 Avatare im Web allgemein

Online Moderator von ModSite²⁹ ist eine aktuell populäre Webanwendung mit Avataren. Webseiten können die Avatare benutzen, um sich auf eine animierte und ansprechende Weise dem Besucher zu präsentieren. Der Moderator kann einerseits den Besucher begrüßen, andererseits den Besuchern das ganze Webportal vorführen. Dadurch wird die Zufriedenheit der Besucher gesteigert. In Abbildung 30 ist ein Moderator von ModSite dargestellt. ModSite verkauft diesen Online Moderator Avatar an andere Webseiten, diese können den Moderator dann verwenden.



Abbildung 30 - Online Moderator von ModSite. Der von ModSite erzeugte Online Moderator ist in der Lage, den Benutzer zielgerichtet zu leiten. Quelle: ModSite³⁰

2.4 Zusammenfassung

Im ersten Abschnitt dieses Kapitels wurden verschiedene Erstellungsverfahren für Avatare vorgestellt. Zuerst wurden Vorgehensweisen für den Normalbenutzer vorgestellt. Daran schlossen sich professionelle Methoden aus der Film-, Fernseh- und Computerspielindustrie an. Den Abschluss bildeten Verfahren aus der Avatarforschung, deren Zielsetzung es ist, neue Avatarerstellungsmethoden zu erforschen, zu verbessern und deren Tauglichkeit für die Praxis zu überprüfen. Der zweite Abschnitt zeigt auf Basis des Wasserfallmodells die Entwicklung einer Avatanwendung und kommt zu dem Schluss, dass sich für diese Art von Anwendung noch keine Standards entwickelt haben, dies jedoch für die kommenden Jahre zu erwarten ist. Der letzte Abschnitt dieses Kapitels zeigte etliche mit Bildern versehene Beispiele für Avatare in den Medien.

²⁹ <http://www.modsite.com>

³⁰ <http://www.modsite.com>

3 Mensch-Avatar-Beziehungen

Dieses Kapitel umfasst die Abschnitte „Kulturelle Unterschiede“ und „Akzeptanzprobleme von Avataren“. Im ersten Abschnitt werden die Unterschiede zwischen Avataren aus verschiedenen Kulturkreisen untersucht. Hauptsächlich dreht es sich dabei um Unterschiede zwischen amerikanischen oder europäischen und asiatischen Avataren. Der zweite Abschnitt beschäftigt sich mit den Akzeptanzproblemen von Avataren und untersucht warum sich in den aktuellen Medien nur selten Avatare finden.

3.1 Kulturelle Unterschiede

Mit der zunehmenden Verbreitung von Avataren in der virtuellen Welt bemerkt man, dass der kulturelle Unterschied im Bereich der Kommunikation Avataren eine große Rolle spielt. Verschiedene Kulturen haben ihre eigene Art Emotionen zu zeigen. Koda et al. [14] haben die Animationen von niederländischen und japanischen Avataren verglichen, die von einem japanischen Künstler entworfen wurden. Es gibt kulturelle Unterschiede bei der Wahrnehmung von Emotionen von animierten Avataren zwischen Japanern und Niederländern. Wenn man die kulturellen Unterschiede bei der Erzeugung von Avataren und Benutzung ihrer Anwendungen kennt, hilft es, Missverständnisse in der Online-Kommunikation zu vermeiden.

Man weiß, dass kulturelle Unterschiede in der realen Welt dazu führen, dass menschliche Emotionen durch unterschiedliche Ausdrücke und Gesten zum Ausdruck gebracht werden. Ein kommunizierender Avatar benutzt die Gesten, die aus der realen Welt abgeleitet sind, um seine Emotionen darzustellen. Deshalb besitzen die Gesten von Avataren auch die kulturellen Unterschiede. Koda et al. [14] haben die Unterschiede zwischen Japan und westlichen Ländern beim Design von menschlichen Ausdrücken gezeigt.

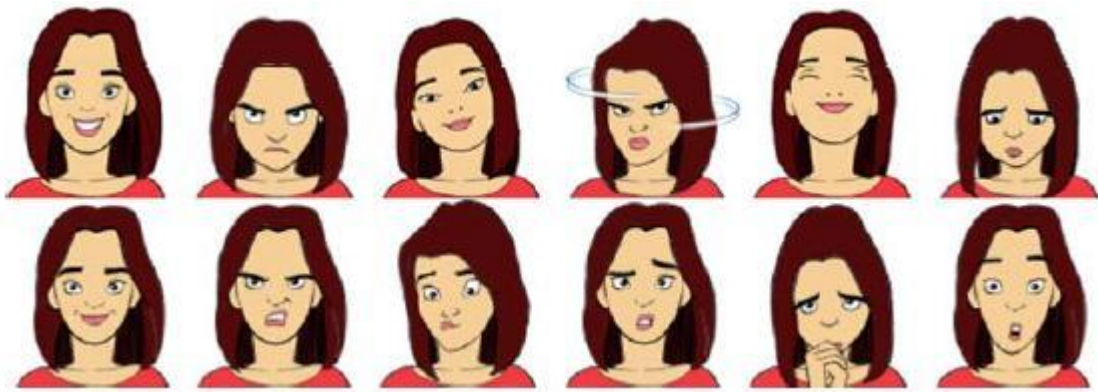


Abbildung 31 – Amerikanische Gefühlskarikaturen. Zwölf Ausdrücke von Amerikanischen Karikaturisten entworfen (Zufriedenheit, Missfallen, Zustimmung, Ablehnung, Stolz, Scham, Dankbarkeit, böse, beeindruckt, verwirrt, reumütig, und verwundet). Quelle: Koda et al. [14]



Abbildung 32 – Japanische Gefühlskarikaturen. Zwölf Ausdrücke von japanischen Karikaturisten entworfen (Zufriedenheit, Missfallen, Zustimmung, Ablehnung, Stolz, Scham, Dankbarkeit, böse, beeindruckt, verwirrt, reumütig und verwundert). Quelle: Koda et al. [14]

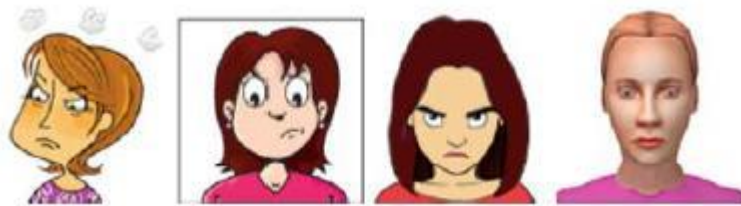


Abbildung 33 – Vergleich der Karikaturen. Ein Vergleich des Gefühlsausdrucks „Missfallen“ interpretiert von Darstellern verschiedener Nationen (von links nach rechts: Frankreich, England, Amerika und Deutschland). Quelle: Koda et al. [14]

Abbildung 31, Abbildung 32 und Abbildung 33 zeigen, dass es größere Unterschiede zwischen Japan und westlichen Ländern gibt. Die Erforschung der Unterschiede zwischen Ost und West verbessert die Kommunikation in der virtuellen Welt. Durch Ausdrücke und Gesten von Avataren in der virtuellen Welt können die Avatare besser verstanden werden.

Avatare erschienen zum ersten Mal vor 20 Jahren in Rollenspielen wie z.B. Diablo (1996), Baldur's Gate (1998), etc. Dabei konnte man in die virtuelle Welt eintreten und den eigenen Avatar mit verschiedenen Gegenständen wie z.B. Rüstung, Waffen, etc. ausstatten, um das gewünschte Aussehen und die angestrebte Verbesserung zu erhalten und um den eigenen Avatar anderen zu zeigen. Bei der Anwendung und Entwicklung von Avataren gibt es zwischen Europäern und Asiaten in dieser Phase keinen Unterschied. Man spielt Avatare in der Welt des Computerspiels, um sie auszustatten und die virtuelle Welt zu erfahren.

Am Anfang der Entstehung des Internets konnte man sich nur über seinen Benutzernamen von den anderen Nutzern unterscheiden. Nachdem immer mehr Onlinespiele entwickelt wurden, schlug man den Weg ein, mit Hilfe von Avataren individuelle Charaktere zu erzeugen, damit man die eigene Persönlichkeit darstellen kann. Man fand heraus, dass viele Onlinespieler sich nicht für das Spiel interessieren, sondern es ihnen wichtiger ist, einen besonderen Avatar im Spiel zu kreieren. Wenn sich dessen Niveau erhöht, verbessert sich auch sein Aussehen, dies kann auch durch erhaltene Gegenstände erfolgen.

Inzwischen haben Avatare sich vom Text zur Grafik, vom Einzelspielerspiel zum Onlinespiel entwickelt. Man hat bemerkt, dass der Avatar alleine existieren kann und nicht unbedingt zusammen mit einem Spiel. Wenn man in einem Onlinespiel einen Avatar entwickelt, muss man viel Energie und Zeit dafür einsetzen. Die Entstehung des individuellen Avatars erfüllt gleichzeitig zwei Anforderungen, die nicht nur leicht verstanden und verwendet werden

können, sondern es auch ermöglichen in der virtuellen Welt durch intuitive Methoden einen eigenen Avatar zu präsentieren. Abbildung 34 zeigt einen Avatar aus QQShow, der mit verschiedener Kleidung, Schuhen, Taschen, etc. ausgestattet werden kann, um einen individuellen Avatar in der virtuellen Welt zu erzeugen.



Abbildung 34 – Avatar aus QQShow. Der Avatar kann mit verschiedener Kleidung, Schuhen, Taschen, etc. ausgestattet werden. Quelle: Tencent Webseite³¹

Aufgrund der verschiedenen sozialen Umgebung und Philosophie entsteht der Unterschied zwischen Europäern und Asiaten bei der Anwendung von Avataren. Im Vergleich zu den Europäern haben Asiaten im realen Leben eine ziemlich enge soziale Umgebung. Es gibt dort im Leben mehr Regeln und Pflichten, mehr Beschränkungen und oft spielt ein Asiate im Sozialsystem eine Rolle, die er sich vielleicht nicht selber wünscht. Also muss er in seinem Leben eine Rolle mit vielen Disziplinen und Spezifikationen spielen. Sein persönlicher Charakter ist also verdeckt, deshalb braucht er einen Kanal, um seinen realen unterdrückten Charakter zu zeigen. Wenn er beispielsweise der älteste Sohn in der Familie ist, muss er von seiner Kindheit an immer hören, dass er das Vorbild für seine Geschwister ist, dass er die ganze Familie unterstützen muss, etc. Dann wird sich normalerweise eine ernsthafte und ruhige, häufig schweigende Persönlichkeit entwickeln. Eigentlich hat nicht jeder älteste Sohn in einer Familie dazu Lust und ist auch bereit eine solche Pflicht zu übernehmen. Die Entstehung von Avataren und des Internets ist die beste Lösungsmöglichkeit. Denn man kann einen von selbst erstellten Avatar erzeugen und ihn in der virtuellen Welt verwenden, um den Anderen zu zeigen, was man will. Dagegen haben Europäer im realen Leben mehr Freiheiten, aber weniger das Bedürfnis in der virtuellen Welt, ihre eigene Persönlichkeit darzustellen, weil es in der realen Welt für sie genügend Spielraum gibt. Deshalb sind Avatare in Asien, besonders im Bereich der Kommunikation, so populär geworden. Sie

³¹ <http://tech.qq.com/zt/2007/QQshow/>

bringen eine gute Möglichkeit hervor, dass man seine gewünschte Persönlichkeit zeigen kann. Heutzutage benutzen die meisten Asiaten Avatare fast nur im Bereich der Online-Kommunikation, sehr selten in anderen Bereichen. In Europa hingegen werden Avatare nicht nur zur Kommunikation, sondern auch oft in Bereichen wie Lernen, Orientierung, Kinofilme, etc. verwendet. Bei solchen Bereichen werden mehr Details benötigt, aufgrund der verschiedenen Anwendungsbereiche. Es bestehen auch verschiedene Bedürfnisse beim Einsatz von Avataren. Deswegen benutzt man in Europa Avatare oft mit 3D Modellierung, 3D Dynamic Capture, etc., während in Asien die 2D Modellierung von Avataren bereits ausreicht, um sich von anderen zu unterscheiden oder seine eigene Persönlichkeit zu zeigen.

Zusammenfassend lässt sich sagen, dass verschiedene Kulturen und soziale Umgebungen unterschiedliche Anforderungen an Avatare stellen und unterschiedliche Anforderungen verlangen wiederum unterschiedliche Entwicklungsrichtungen von Avataren. Nach Meinung der Autoren dieser Fachstudie werden in der Zukunft fortgeschrittene Avatare entwickelt werden, die die Wünsche von Europäern und Asiaten zugleich erfüllen.

3.2 Akzeptanzprobleme von Avataren

In diesem Abschnitt wird untersucht, warum Avatare in der heutigen Medienlandschaft nur recht selten anzutreffen sind. Um die Zeit der letzten Jahrtausendwende waren Avatare häufiger präsent als dies heutzutage der Fall ist. Prominentes Beispiel ist Karl Klammer, die Büroklammer aus Microsoft Office, dargestellt in Abbildung 35. Die Probleme die die Akzeptanzprobleme verursachen, lassen sich verschiedenen Bereichen zuordnen. Es bestehen sowohl technische Probleme als auch psychologische Probleme bei den Nutzern.



Abbildung 35 – Karl Klammer, die Büroklammer aus Microsoft Office. Quelle: Webseite der Werbewoche³²

Laut Lindner [17] sind insbesondere die weichen Faktoren für die Akzeptanz des Avatars bei den Benutzern verantwortlich. Dazu gehören das grafische Design, die Tonalität der Antworten und die Eigenheiten des Avatars. Durch seine Eigenheiten bekommt der Avatar Charakter und gewinnt an Wiedererkennungswert.

Nach Lindner [17] müssen Avatare besondere Bedürfnisse des Benutzers erfüllen, die noch nicht erfüllt werden, beispielsweise eine robustere Spracherkennung. Siehe dazu auch Abschnitt 4.3 über Spracherkennung. Bei der Entwicklung der ersten Avatare fanden ihre

³² http://www.werbewoche.ch/sites/werbewoche.ch/files/imagecache/original/article/images_preview/zgl_1.jpg

Entwickler heraus, dass Glaubhaftigkeit und Lebensnähe nicht zwingend durch die menschliche Modellierung der Avatare erreicht wird. Die Entwickler benutzten Vorgehen von Disney-Trickzeichnern, Karikierung und Übertreibung um den Benutzer dazu zu bringen einem Avatar Lebenswirklichkeit zuzuschreiben.

Dieses Phänomen wurde näher erforscht und trägt den Namen Uncanny Valley Effekt, zu Deutsch Effekt des unheimlichen Tals. Nach Nagayama und Seyama [29] besagt dieser Effekt, dass desto realistischer das Gesicht eines Avatars dargestellt ist, die Benutzer es als umso unangenehmer empfinden. Eine grafische Darstellung des Effekts findet sich in Abbildung 36. Designern von Filmen und Computerspielen wird daher empfohlen den Realitätsgrad einer Figur nicht zu hoch zu setzen, um nicht in den Wirkungsbereich dieses Effekts zu gelangen. Eine Alternative wäre die comicartige Umsetzung von Figuren.

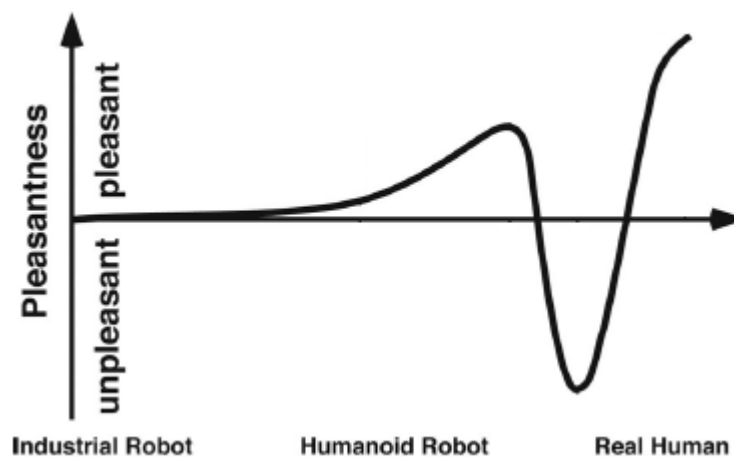


Abbildung 36 – Uncanny Valley Effekt. Umso realistischer ein Avatar umgesetzt ist desto unheimlicher wird er seinen Betrachtern. Quelle: Grafik modifiziert nach Nagayama und Seyama [29]

Ein weiterer Effekt im Zusammenhang mit der Akzeptanz von Avataren ist nach Takahashi et al. [33] das Phänomen der Media Equation. Dies besagt, dass Benutzer häufig mit Computern bzw. Avataren interagieren wie mit einem Menschen. Laut André et al. [2] ist es damit denkbar, dass ein Avatar sogar der Freund eines Benutzers werden kann. Wichtige Voraussetzung dafür wäre eine hohe Sozialkompetenz des Avatars. Der Benutzer erwartet menschliches Kommunikationsverhalten, sowohl verbal als auch nichtverbal. Laut den Autoren ist dies eine wichtige Voraussetzung um die Akzeptanz von Avataren bei den Benutzern zu erhöhen.

Laut Lindner [17] ist ein Technologietrend in der Mensch-Computer-Interaktion die Personalisierung und Individualisierung. Die Kommunikation mit Systemen ist durch einen hohen Grad an benutzerzentrierter Kommunikation geprägt. Anwendungen richten sich am Anwender aus. Somit werden Avatare zur persönlichen Schnittstelle. Der Trend der Personalisierung und Individualisierung drückt sich dabei durch zwei Ausprägungen aus: Zum einen muss die Kommunikation zwischen Mensch und Maschine anderen Ansätzen folgen als

dies heute der Fall ist. Zentrales Element dabei ist die Sprachein- und -ausgabe. Zum anderen muss die Personalisierung weiter vorangetrieben werden. Systeme müssen sich mehr an ihre Benutzer anpassen. Der Avatar lernt etwas über seinen Anwender und kann ihn so besser bedienen. Beide Gebiete, sowohl die Spracherkennung als auch die zielführende Auswertung von Nutzerdaten, sind, nach Meinung der Autoren dieser Fachstudie, hochkomplexe Forschungsgebiete und bleiben sicherlich für Avatare noch für einige Zeit Flaschenhals und Schlüsseltechnologie zugleich.

Ein interessantes Beispiel zu Akzeptanzproblemen in der Kommunikation mit einem Avatar gibt Lindner [17]. Hier geht es um Cora, den Avatar der Deutschen Bank. Deren Kommunikationskomponente ist darauf ausgelegt, komplett ausformulierte Fragen beantworten zu können. Nun tendieren Internetnutzer häufig dazu Fragen in Stichwortform zu stellen. Daraufhin müssen sie von Cora durch umständliche Gegenfragen zum Frageziel geführt werden und ärgern sich dann darüber, obwohl sie die Kommunikationskomponente unterfordert haben.

Nach Meinung der Autoren dieser Fachstudie liegen die Akzeptanzprobleme von Avataren in Anwendungssoftware hauptsächlich darin, dass es bisher nur wenige Einsatzbereiche für Avatare als Softwarebestandteil gibt. Avatare übernehmen in Anwendungen so gut wie immer die Rolle der Hilfefunktion. Dabei werden sie von vielen Nutzern häufig bereits nach kurzer Zeit deaktiviert oder ignoriert, da die vom Avatar angebotene Hilfestellung oftmals nicht zufriedenstellend ist. Das eigentliche Problem besteht in der schlechten Hilfefunktion und nicht im Avatar. Leistungsfähige Hilfefunktionen mit einem ansprechenden Avatar hätten sicherlich ein größeres Erfolgspotenzial, speziell bei älteren Menschen. Weiterhin fehlen derzeit zum einen schlichtweg Ideen für neue sinnvolle Einsatzmöglichkeiten von Avataren in Anwendungssoftware. Auch würde eine robuste Spracherkennung eine große technische Verbesserung beim Einsatz von Avataren darstellen. Die Spracherkennung erfährt derzeit deutliche Verbesserungen in der Erkennungsqualität, siehe Abschnitt 4.3 über Spracherkennung. Wünschenswert wäre, dass Avatare dadurch häufiger zum Einsatz kommen.

Weiterhin sind die Autoren dieser Fachstudie der Meinung, dass Avatare in integrierten Systemen, wie etwa einem Fahrkartenautomat, derzeit weitaus größere Erfolgschancen als in reiner Anwendungssoftware hätten. Ein Fahrkartenautomat mit dem man über einen Avatar interagieren kann, wie zu früheren Zeiten mit einem Bahnangestellten am Schalter, wäre eine enorme Verbesserung gegenüber den heutigen Automaten, insbesondere auch wieder bei älteren Menschen. Hier spielt die Erkennung natürlicher Sprache die entscheidende Rolle, ohne sie macht die Umsetzung des Fahrkartenautomaten mit Avatar keinen Sinn.

Abschließend lässt sich feststellen, dass die Akzeptanzprobleme von Avataren von vielfältiger Natur sind. Sie liegen sowohl im Bereich der Psychologie als auch im technischen Bereich. Wichtig für die Erhöhung der Akzeptanz von Avataren sind ein ansprechendes, nicht zu

realistisches äußeres Erscheinungsbild, sinnvolle Interaktionstechniken, wie robuste Spracheingabe und brauchbare Funktionen des Avatars.

3.3 Zusammenfassung

Dieses Kapitel befasste sich zuerst mit den kulturellen Unterschieden zwischen Avataren aus unterschiedlichen Kulturkreisen. Auffällig war hierbei der Unterschied in den Emotionen zwischen amerikanischen oder europäischen und asiatischen Avataren. Daran schloss sich der Abschnitt über die Akzeptanzprobleme von Avataren an. Dabei wurde festgestellt, dass die Akzeptanzprobleme vielerlei Ursachen haben. Ein nicht zu realistisches Äußeres, robuste Spracherkennung zur Interaktion und hilfreiche Funktionen können die Akzeptanz von Avataren deutlich erhöhen.

4 Interaktionskonzepte

In diesem Kapitel werden zunächst gängige Interaktionstechnologien der Mensch-Computer-Interaktion (HCI) vorgestellt und diskutiert. Anschließend gehen wir auf das Verhalten der Avatare ein. Wie erkennen Avatare Emotionen? Welche Spracherkennungsmethoden gibt es? Wie werden Mimik und Gestik dargestellt? Zusätzlich dazu wird noch die Körperbewegung der Spielekonsolen von Microsoft, Sony und Nintendo vorgestellt.

4.1 Interaktionstechnologien bei HCI mit Avataren

Dieses Kapitel behandelt eine generelle Auswahl an Technologien der Mensch-Computer-Interaktion, welche nicht alle zur Interaktion mit Avataren geeignet sind. Zu Beginn werden traditionelle Methoden wie etwa der Einsatz von Maus und Tastatur beschrieben. Im weiteren Verlauf werden neuartige und teilweise fiktive Konzepte und insbesondere deren Möglichkeiten zur Interaktion mit Avataren präsentiert.

Vor allem im Bereich von Computerspielen ist der Einsatz von Maus und Tastatur noch vorherrschend. Es existieren unzählige Spiele, in denen eine vom Benutzer kreierte Figur durch Klicken (Click and Point) oder durch Tastaturnavigation zu einer Aktion bewegt wird. Ein klassisches Click and Point Adventure Spiel war „The Secret of Monkey Island“, bei der eine Figur durch die Maus gesteuert wurde. Zuerst musste man die Aktion auswählen (etwa „Gehe zu“ oder „Nimm“) und dann auf das Ziel klicken (eine Person oder ein Gegenstand). Dieses Genre war vor allem um 1990 sehr beliebt.



Abbildung 37 – Szene aus dem Click and Point Adventure Spiel „The Secret of Monkey Island“. Der Spieler wählt mit einem Mausklick eine Interaktion aus dem linken Bereich und mit dem zweiten Klick das Ziel (ein Gegenstand, ein Tier, eine Person, etc.). Quelle: Jeuxvideo Webseite³³

³³ <http://image.jeuxvideo.com/images/pc/t/h/the-secret-of-monkey-island-special-edition-pc-037.jpg>

Bei später erschienenen Rollenspielen war eine Kombination aus Maus und Tastatur gefragt. Der Spieler kann meistens zwischen diversen Grundcharakteren wählen und diese nach eigenen Wünschen trainieren, ausstatten oder agieren lassen. Die Tastatur wird vornehmlich zum Steuern der Figur genutzt, die Maus zum Anvisieren von Gegenständen oder feindlichen Wesen.



Abbildung 38 - Maus und Tastatur. Maus und Tastatur haben sich als universelles Eingabegerät im privaten Computerbereich durchgesetzt. Mit der Tastatur lassen sich etwa Figuren bewegen oder Textanfragen und -befehle an Interaktionsobjekte stellen. Die Maus dient zur Fixierung von Zielen, da sie eine größere Genauigkeit als die Tastatur bietet. Quelle: Golem Webseite³⁴

Daneben gibt es noch Lenkräder als Eingabegerät, welche vor allem bei Rennsimulationen ihren Reiz besitzen, da mit ihnen wesentlich genauer gesteuert werden kann als mit der Tastatur. Als universelles Eingabegerät hat sich das Gamepad im Computer- und Konsolenbereich bewährt. Sportspiele, Kampfsimulationen oder auch Rennspiele sind damit deutlich komfortabler spielbar als mit anderen Geräten. Mittels eines Steuerkreuzes wird die Figur oder das Auto des Spielers in verschiedene Richtungen bewegt. Weitere Tasten sorgen für verschiedene Aktionen, je nach Spieltyp (Schlagen oder Blocken bei Kampfsimulationen, Beschleunigen oder Abbremsen bei Rennsimulationen). Ein heute nur noch selten verwendetes Interface ist der Steuerknüppel oder Joystick. Dieser ermöglicht die Richtungsangabe durch Bewegen des Knüppels in die Richtung, die vorgegeben werden soll. Zur Steuerung von Flugzeugen und Hubschraubern in Flugsimulationen werden sie von Profispielern noch eingesetzt. Für die Interaktion mit Avataren eignen sich diese Geräte aber nur bedingt, da sie außer der Steuerung keinerlei Möglichkeiten bieten.

³⁴ http://scr3.golem.de/screenshots/0706/Natural_Ergonomic_Desktop_7000/NED7000_lg_screen.jpg



Abbildung 39 – Gamepad. Gamepads sind bei Konsolen- und Rennspielern bevorzugt anzutreffen, eignen sich jedoch schlecht für Interaktionen, die über die bloße Richtungssteuerung hinausgehen. Quelle: Sevasaonline Webseite³⁵

Die Interaktion durch Gestik hat sich im Spielbereich durchgesetzt. Für die drei meist verbreiteten Spielekonsolen (Nintendo Wii, Microsoft Xbox, Sony PlayStation) existieren Systeme zur Gestenerkennung. Mittels eines Sensors, der in der Hand gehalten wird, oder durch einen Sensor, der in einem Abstand zum Spieler platziert ist, wird die Bewegung registriert und darauf reagiert das System. Beim Bowling etwa muss der Nutzer nun nicht nur im richtigen Moment eine Taste drücken, um die Kugel loszulassen, sondern er muss auch die Schwungbewegung korrekt ausführen, um die Aktion korrekt auszuführen. Somit ist er stärker in das Spiel integriert als wenn er nur Tasten auf einem Gamepad drückt. Lässt man jedoch etwa seinen Hund mit dem Sender spielen, wird das System die falsche Bedienung nicht erkennen. Es kommt immer auch auf die korrekte Bedienung an. Systeme zur Erkennung von ganzen Körperpositionen lassen sich auf vielfältige Weise einsetzen. Mimik und Gestik können wie vom Benutzer vorgegeben auf den Avatar übertragen werden. Möchte man jedoch ein Auto steuern, stößt dieser Ansatz an seine Grenzen, da etwa die Beschleunigung des Wagens nicht intuitiv und mittels Gesten nur unsauber realisierbar ist.

³⁵ <http://www.sevasaonline.com/catalog/images/G-08XU.jpg>



Abbildung 40 – Kinect für Xbox360. Das Gerät erkennt mittels eines Sensors die Gesten eines oder mehrerer Benutzer und überträgt diese auf den Avatar. Handelt es sich dabei um ein menschenähnliches Abbild, gelingt die Interaktion sehr gut. Quelle: Play Webseite³⁶

Für professionelle Grafiker ist ein Grafiktablett das bevorzugte Werkzeug. Dabei wird mit einem Stift mit speziellem Sensor auf eine Platte gezeichnet und die Daten werden an den Computer übertragen. Diese Vorgehensweise erlaubt eine größere Präzision im Bereich der Grafikbearbeitung. Ähnlich dazu funktionieren Touchbildschirme oder Touchpads, welche vor allem in Tablet-PCs, Infoterminals in öffentlichen Einrichtungen und Banken sowie Smartphones zu finden sind. Die Bedienung erfolgt hierbei über Finger und der berührungsempfindliche Bildschirm registriert die Aktionen. Je nach Bildschirm sind dabei auch Aktionen und Gesten mit mehreren Fingern möglich.

Der sinnvollste Ansatz zur Interaktion mit Avataren ist sicherlich durch Einsatz der natürlichen Sprache in gesprochener Form. Abwandlungen davon sind recht populäre Chatbots, welche auf Texteingaben reagieren und entsprechende Aktionen ausführen können. Die Bandbreite dieser Bots reicht vom Erkennen einzelner Befehle bis hin zur vollen Konversation. Einige Unternehmen setzen Avatare als „Portier“ auf ihrem Internetportal ein, um den Benutzer zu leiten und ihm die Möglichkeit zu geben, direkte Fragen zu Produkten, Dienstleistungen oder dem Unternehmen selbst zu stellen. Diese Avatare besitzen die Fähigkeit, dem Benutzer die gewünschten Informationen zur Verfügung zu stellen und auch auf thematisch abweichende Fragestellungen angemessen zu reagieren. In einigen Fällen kommen auch interessante persönliche Details des Firmenvertreters beim Small-Talk ans Tageslicht³⁷.

³⁶ <http://i.cdn.play.tm/s/32349/p/s6.jpg>

³⁷ http://apps.facebook.com/luxoft_avatar



Abbildung 41 – Avatar der Firma Luxoft. Ein Chatbot, der auf einfache und kurze Textanfragen reagieren kann. Er präsentiert seine Firma und versorgt den Benutzer mit Informationen darüber. Smalltalk lässt sich kaum betreiben, es werden ausweichende Standardantworten gegeben. Quelle: Luxoft auf Facebook³⁸

Heute gibt es zahlreiche Software zur Spracherkennung, im Privatkundenbereich aber vornehmlich zum Diktieren von Texten oder Ausführen einfacher Befehle. In den Microsoft Betriebssystemen *Windows Vista* und *7* ist ein Programm integriert, welches gesprochene Texte direkt notiert und auch entsprechende Korrekturmethode für falsch erkannte Worte mitbringt³⁹. Zusätzlich kann das Programm auf Kommando andere Programme starten oder Webseiten öffnen. Es muss etwas Zeit in das Training der Software für den persönlichen Sprachstil investiert werden, der Ressourcenverbrauch ist nicht unerheblich und die Funktionalität recht eingeschränkt.

Entwickler vom Microsoft Research Team präsentierten im März 2011 ein Verfahren, mit dem die Gestenerkennung im Raum ermöglicht wird⁴⁰. Dabei leuchtet ein Projektor auf eine lichtreflektierende Schicht und eine Kamera fängt das zurückgeworfene Licht auf. Wenn sich Objekte (etwa eine Hand) im Weg befinden, erkennt die Kamera dies als Schatten. Wird nun eine spezielle Geste ausgeführt, kann ein Programm entsprechend darauf reagieren. Im Testvideo wird unter anderem gezeigt, wie man einfache Dinge malen kann ohne die Malfläche zu berühren. Diese Technologie befindet sich noch in einem frühen Entwicklungsstadium und praktische Umsetzungen für Endanwender sind noch nicht veröffentlicht. Inwiefern die Steuerung von Systemen oder Avataren damit ermöglicht wird, ist noch nicht absehbar. Denkbar wäre ein System, welches Entwicklern bei der Modellierung

³⁸ http://apps.facebook.com/luxoft_avatar

³⁹ <http://www.microsoft.com/enable/products/windowsvista/speech.aspx>

⁴⁰ <http://www.microsoft.com/showcase/en/us/details/19c7e564-bd1a-4060-ba9d-8133f2ea27c7>

von Entwürfen hilft, indem statt echten Werkstoffen Projektionen benutzt werden. Anhand dieser könnte dann etwa die Haptik veranschaulicht werden.



Abbildung 42 – Szene aus dem Film „Iron Man“. Der Protagonist modelliert seine Erfindungen unter Einsatz von Gesten mit Hilfe von Projektionen anstatt echter Werkstoffe. Die Objekte können frei im Raum positioniert, verändert, neu erstellt oder verworfen werden. Am Ende kann aus dem modellierten Entwurf eine Umsetzung mit den gewählten Materialien durchgeführt werden. Quelle: Ambience of Media Webseite⁴¹

Eine weitere Interaktionstechnik ist in Barrientos und Canny [5] beschrieben. Dabei kann ein Nutzer einen Avatar steuern, indem er mittels eines Stiftes einen Buchstaben auf eine speziell präparierte Oberfläche schreibt. Basierend darauf, welchen Buchstaben und wie der Nutzer schreibt, verhält sich der Avatar. Jedem Buchstaben ist eine Animation zugeordnet, die zuvor in verschiedenen Varianten aufgezeichnet wurde. Je mehr Aufzeichnungen vorhanden sind, desto genauer kann der Avatar die passende Stimmung wiedergeben. Dazu wird das Schriftbild des Buchstabens analysiert. Ist der Buchstabe etwa sehr groß gezeichnet, wird der Avatar eine raumgreifendere Animation ausführen. Die Schreibgeschwindigkeit bestimmt die Ausführungsgeschwindigkeit, der Druck die Betonung und die Neigung des Buchstabens die Richtung, in welche die Animation ausgeführt wird. Viele Leute besitzen die Neigung, während des Telefonierens auf Papier zu schreiben. Diese Daten könnten etwa ein Telefongespräch mittels Avataren visualisieren.

Die ab 1966 produzierte Science-Fiction-Serie „Star Trek – The Original Series“ enthielt kreative Ideen wie einen tragbaren, kabellosen Kommunikator, welcher die ersten Mobiltelefone lange vor deren Markteinführung prägte. Science-Fiction bringt den aktuellen Stand der Wissenschaft und Technik mit fiktionalen Elementen in Verbindung und schafft daraus neue Elemente, die in einigen Fällen zu einem Marktprodukt der Gesellschaft werden können. Es ist nicht auszuschließen, dass weitere fiktive Erfindungen in unserer Zeit umgesetzt werden.

⁴¹ <http://www.ambienceofmedia.com/wp-content/uploads/2010/05/iron-man-2-tony-stark.jpg>



Abbildung 43 – Captain James T. Kirk mit seinem tragbaren, kabellosen Kommunikator. Diese Aufnahme stammt aus der Serie „Star Trek – The Original Series“ (1966 – 1969). Quelle: Dippold Webseite⁴²

Im fiktiven Szenario von *Star Trek* spielt die Spracherkennung eine große Rolle. Jedes Raumschiff ist mit einem Computer mit integrierter Spracherkennung ausgestattet, welche auf Anfragen des Benutzers reagieren kann. Mittels des Befehls „Computer“ wird die Spracherkennung aktiviert und der Computer ist bereit für die eigentliche Aufgabe. Diese kann vielfältig ausfallen, vom Ermitteln des Aufenthaltsortes einer Person bis zur Initialisierung einer Diagnose eines kritischen Systems.

Holodecks⁴³ sind ebenso ein fiktives Interaktionskonzept aus dem Star Trek Universum. Dabei befindet sich eine oder mehrere Personen in einem speziellen Raum, in welchem ein Computer eine Umgebung (etwa eine Landschaft, einen Keller oder den Center Court von Wimbledon) sowie weitere Personen (Hologramme) simulieren kann. Ein Hologramm ist eine durch Kraftfelder räumlich begrenzte Projektion aus Licht und Energie. Die simulierten Personen wirken real und interagieren innerhalb der ihnen vom Computer vorgegebenen Parameter mit dem Benutzer. Es kann praktisch alles vom einfachen Tennismatch mit einer Person bis hin zu historischen Schlachten im Zweiten Weltkrieg nachgestellt werden.

2009 wurde an der Universität Tokio ein Projekt⁴⁴ vorgestellt, bei welchem man 3D Projektionen fühlbar machen konnte. Dieser Ansatz könnte den Interaktionslevel auf ein neues Niveau anheben, da der Benutzer nicht mehr vor einem Bildschirm sitzt oder von technischen Geräten umgeben ist und sich mit der speziellen Bedienung auseinandersetzen muss. Durch Science Fiction können wir uns inspirieren lassen, wie die Mensch-Computer-Interaktion in 10, 50 oder 100 Jahren aussehen kann.

⁴² http://www.dippold.org/startrek/serien/tos/pics/3_1_9.jpg

⁴³ <http://memory-alpha.org/wiki/Holodeck>

⁴⁴ <http://www.heise.de/newsticker/meldung/Fuehlbare-3D-Projektion-750329.html>

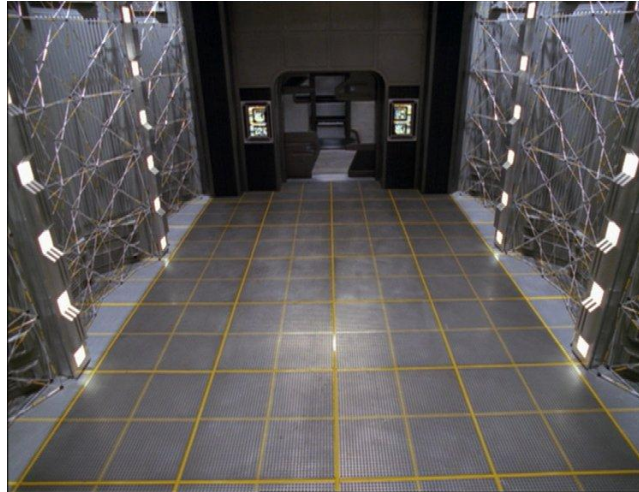


Abbildung 44 – Holodeck. Ein Raum aus dem fiktiven Star Trek Szenario, der mittels Licht und Kraftfeldern eine virtuelle Umgebung mit Personen erzeugen kann. Die Objekte sehen real aus und verhalten sich gemäß ihren Vorgaben. Vom Tennismatch bis zur Nachstellung historischer Ereignisse kann hier alles simuliert und nachgespielt werden. Quelle: Wikia Webseite⁴⁵

Es ist ein Entwicklungstrend erkennbar, der den Menschen von der aktiven Nutzung der Technik abspaltet. Zu Beginn des Computerzeitalters waren Eingabegeräte, die durch Bewegungen oder Drücken vom Menschen benutzt wurden, am weitesten verbreitet. Um die Jahrtausendwende kamen sprachgesteuerte Systeme auf wie zum Beispiel eine automatische Telefonhotline, die den Anrufer nach der Problemkategorie fragt, um diesen mit dem passenden Mitarbeiter verbinden zu können. Der Anrufer musste nur die entsprechenden Befehle zur Interaktion mit dem System nennen. Die nächste Stufe stellten dann körper- oder gestengesteuerte Systeme dar, welche primär im Spielekonsolenmarkt ab 2008 für den Privatanwender verfügbar waren. Das System von Nintendo Wii verwendet dazu einen Sensor, der in der Hand gehalten werden muss, wohingegen der Benutzer bei der Xbox von Microsoft frei in seinen Bewegungen ist. Die Technik ist für den Anwender zwar noch sichtbar (es muss eine Sichtverbindung zwischen Akteur und Empfänger bestehen), aber sonst müssen keine Eingabegeräte bedient werden. Die logische Fortsetzung dieser Entwicklung könnte zur direkten Interaktion in der virtuellen oder realen Welt ähnlich dem Holodeck. Sobald die Simulation läuft, sieht der Benutzer nur noch die Umgebung und keinerlei technische Gegenstände mehr. Die Technik tritt immer mehr in den Hintergrund, wird gleichzeitig aber immer komplexer und erlaubt mehr Möglichkeiten.

⁴⁵ http://images.wikia.com/memoryalpha/en/images/4/43/Voyager_holodeck.jpg

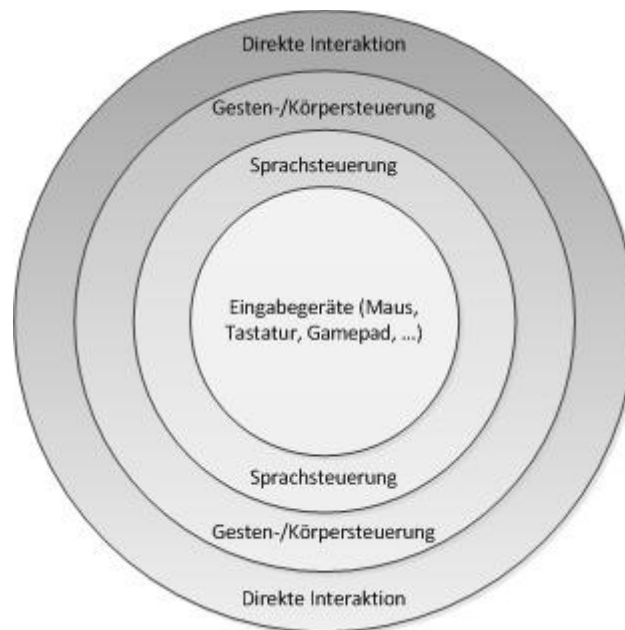


Abbildung 45 – HCI Interaktionslevel. Zu Beginn des Computerzeitalters waren Geräte, die direkt vom Nutzer bewegt oder gedrückt werden mussten, für die Interaktion verantwortlich. Um die Jahrtausendwende wurden sprachgesteuerte Systeme eingesetzt, bei denen die Kommunikation über bestimmte Befehle des Nutzers erfolgte. Anschließend kamen Gesten- und Körpersteuerungen auf, welche die Technik weiter vom Anwender abstrahierten, sodass der Schluss naheliegt, dass in Zukunft die komplexe Technik mehr in den Hintergrund rückt.

Während bei Eingabegeräten die analogen Signale (etwa ein Tastendruck oder eine Bewegung eines Steuerknüppels) auf digitale Aktionen abgebildet werden können (es ist immer definiert, was eine Eingabe als Reaktion hervorruft), sieht es bei der Sprach- und Körpersteuerung etwas komplizierter aus. Um das analoge Signal (etwa ein Wort oder eine Geste) entsprechend umwandeln zu können, muss das System erst mittels Erkennungsalgorithmen und –heuristiken ermitteln, um welchen Interaktionstyp es sich handelt, bevor es einer digitalen Entsprechung zugeordnet werden kann. Kombiniert man mehrere Interaktionsarten, erhält man eine Fülle an möglichen Eingaben, für die eine Reaktion vom System angeboten werden muss. Die Komplexität und die benötigte Rechenleistung steigen bei höheren oder mehreren Interaktionsleveln deutlich an.

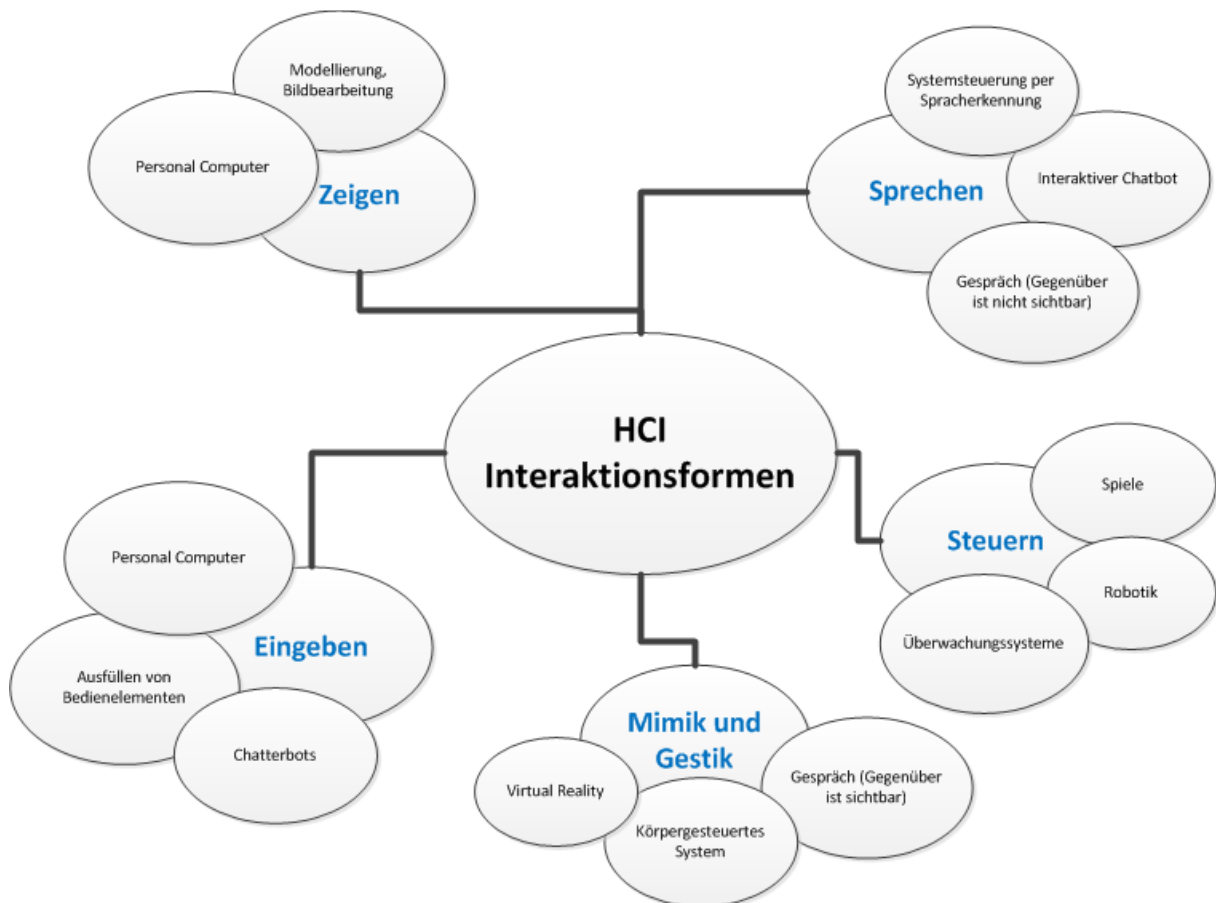


Abbildung 46 – HCI Interaktionsformen. Eine Übersicht über verschiedene Interaktionsformen mit Beispielen für deren Umsetzungen. Manche Zielsysteme basieren auf mehreren Techniken, sodass die Abgrenzung teilweise schwierig ist.

Die Interaktionsformen haben sich seit Beginn Einführung des Computers stark verändert, jedoch werden im Privatbereich noch vielfach Technologien verwendet, die auf den ursprünglichen Eingabegeräten (Maus und Tastatur, Bedienung durch Drücken von Tasten) basieren. Nur in einzelnen Bereichen haben sich andere Formen der Interaktion durchgesetzt.

4.2 Anpassung an das Benutzerverhalten

Dieses Kapitel fasst Methoden und Studien zu den Themen Gesprächsführung, Gesprächsinteraktion, und Untersuchungen über kulturelle Verhaltensunterschiede im Gespräch zusammen und verdeutlicht die Wichtigkeit der Einhaltung gewisser Regeln bei der Mensch-Mensch-Interaktion. Dies ist wichtig für einen Avatar, der das Kommunikationsverhalten eines Menschen angemessen simulieren möchte.

Wenn ein Avatar mit einem menschlichen Benutzer interagiert, sollte das Ziel darin bestehen, dass der Avatar einen Menschen in seinen Eigenarten imitiert, um dem Gegenüber möglichst natürlich zu erscheinen. Ein Gespräch besteht immer aus einem verbalen (Informationsträger) und einem nonverbalen Teil (steuert das Gespräch). Zwei Personen, die sich nur gegenüber sitzen und deren Gespräch sich nur auf der verbalen Ebene abspielt, werden vermutlich keine aufregende und lebendige Diskussion führen. Dafür sorgen erst Mimik, Blickverhalten, Gestiken und Körperhaltung und verleihen der

Konversation die entsprechende Würze. Um einen Avatar als angenehmen Gesprächspartner zu empfinden, muss dieser das komplexe Zusammenspiel von verbaler und nonverbaler Kommunikation beherrschen.

Je nach Einsatzzweck muss der Avatar entsprechende Kriterien berücksichtigen, um möglichst benutzerfreundlich zu sein. Handelt es sich etwa um Hilfesystem mit Sprachausgabe (zum Beispiel eine automatische Hotline), erwartet der Benutzer eine zielgerichtete Führung und eine verständliche und prägnante Hilfe. Ein interaktiver Führer auf einer Rundfahrt muss dagegen auch thematisch abweichenden Fragen auswerten können, Faktenwissen haben und eine angenehme Stimmlage besitzen, um dem Zuhörer ein gewohntes Gefühl zu geben. Besitzt der Avatar ein körperliches Abbild oder ein Gesicht, ist auch die Mimik und Gestik wichtig. Er sollte Emotionen transportieren (vor allem lächeln und interessiert aussehen) und das Gesamtverhalten sollte dem eines Menschen nahekommen.

Ebenso ist es wichtig, auf die Rahmenbedingungen zu achten. Studien haben signifikante Unterschiede im Verhalten und Erkennen von Emotionen zwischen westlichen Ländern (Europa, USA) und Asien (Japan, Korea, China, etc.) gezeigt (siehe dazu Kapitel 3.1). Eine Untersuchung der Gestiken zwischen Probanden aus Japan und Deutschland ergab, dass Deutsche häufiger und flüssig ablaufende Gestiken einsetzen, während Japaner eher weniger, dafür raumgreifendere und unrunde, eckige Gestiken verwenden Rehm und Nakano [23]. Andere Untersuchungen zeigten, dass Gestiken zum Verstärken von Höflichkeitsformulierungen oder Kritik verwendet werden, siehe Rehm und André [26]. In Rehm et al. [27] wurde gezeigt, dass ein Avatar als gleichwertiger Gesprächspartner angesehen wird, wenn sein Verhalten dem eines Menschen im Gespräch entspricht. Der Avatar muss seinen emotionalen Zustand transportieren und dabei von menschlichen Ausdrucksmerkmalen Gebrauch machen (etwa lächeln, zwinkern, kontextbezogenes Blickverhalten, Kopfbewegungen). Ergebnisse in Benutzerstudien belegen, dass Personen sich bei der Interaktion mit Avataren natürlich verhalten, wenn der Avatar sich entsprechend präsentiert und keine Grundregeln der Gesprächsführung missachtet, siehe Rehm und André [25]. Darunter fallen Punkte wie, dass der Sprecher nicht unterbrochen werden soll oder dass sprunghafte Themenwechsel ohne Bezug zueinander vermieden werden sollen.

In Lindner [17] ist ein Verfahren zur Erkennung von natürlicher Sprache beschrieben. Dabei wird ein Ausdruck auf bekannte Muster untersucht, die in einer Datenbank gespeichert sind. Zu jedem Muster gibt es eine entsprechende Antwort, die dem Benutzer präsentiert wird. Um synonyme Ausdrücke erkennen zu können, besteht ein Muster aus komplexen logischen Ausdrücken. Dadurch kann zu den Fragen „Wie hoch ist unsere Fahrtgeschwindigkeit denn?“ und „Wie schnell fahren wir denn?“ dieselbe Antwort gegeben werden. Je nach Größe der Datenbank können unterschiedlich viele Anfragen der Benutzer verarbeitet werden. Die Korrektheit der Antworten hängt stark von der Implementierung ab. Statistische Musteranalysen sind äußerst zuverlässig und schnell, was die Erkennung angeht, da sie auf fest programmierten Algorithmen basieren. Allerdings benötigen diese Methoden eine große Datenbank, in der die Ergebnisse abgespeichert sind. Regelbasierte Systeme erfassen

in der Regel nicht alles, sind aber verlässlich, was die Qualität der Ergebnisse betrifft. Semantische Analysen sind bereichsspezifisch, da sonst die Menge der Synonyme unüberschaubar wird. Allerdings helfen sie, Fehlinterpretationen zu verhindern. Kombiniert man alle aufgeführten Methoden, erhält man die bestmögliche Variante, um eine Eingabe zu analysieren. Ist das System lernfähig, verringert sich auch der Verwaltungsaufwand, da neue Muster automatisch erkannt und integriert werden können. Für unbekannte Muster kann eine Standardantwort definiert werden, etwa mit der Aufforderung, die Anfrage neu zu formulieren.

4.3 Spracherkennungsmethoden von Avataren

Ob Avatare in der Zukunft wieder häufiger in Erscheinung treten, wird stark von der Qualität der Spracherkennung abhängen. Auch Lindner [17] betont die Wichtigkeit der Spracherkennung für den Erfolg von Avataren. Avatare mit Spracherkennung sind dafür prädestiniert an Orten eingesetzt zu werden, an denen die Interaktion auch für ungeübte Nutzer möglichst einfach sein soll. Dieser Abschnitt stellt zuerst ein Wohnungsmanagementsystem mit einem Avatar, der Spracherkennung verwendet, vor und gibt anschließend einen Überblick über die Grundlagen der Spracherkennung.

Moir und Filho [11] haben ein System entwickelt, welches einen sprechenden Avatar und Spracherkennung zum Wohnungsmanagement verwendet. Das Interface dieser Software ist in Abbildung 47 dargestellt. Dem Avatar können verschiedenste Befehle erteilt werden, beispielsweise Musik abzuspielen, auf dem Computer nach Dateien zu suchen oder RSS Feeds vorzulesen. Besonders hervorzuheben ist die Internetrecherchefunktion, welche Fragen wie etwa „Wie viele Planeten hat unser Sonnensystem?“ beantworten kann. Sind Gäste zu Besuch, so kann dem Avatar der Befehl gegeben werden nicht mehr zuzuhören, um nicht versehentlich Teile der Unterhaltung als Befehl zu interpretieren.

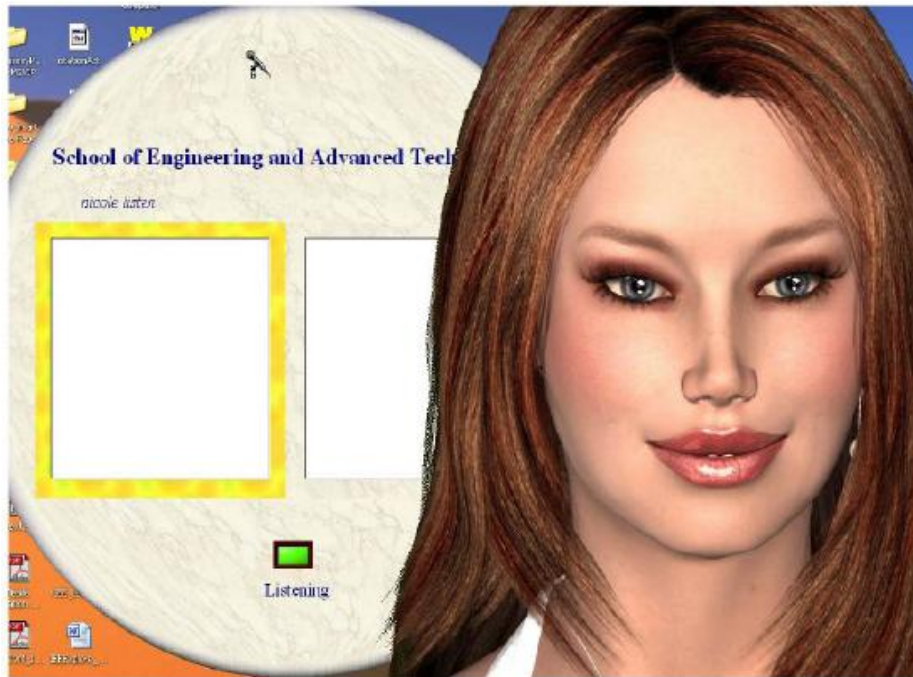


Abbildung 47 – Interface des Systems von Moir und Filho [11] mit dem sprechenden Avatar. Quelle: Moir und Filho [11]

Es folgt nun ein Überblick über die Grundlagen der Spracherkennung. Nach Anusuya und Katti [3] ist Spracherkennung als der Vorgang definiert, bei dem ein Sprachsignal durch einen Algorithmus, welcher als Computerprogramm implementiert ist, in eine Folge von Wörtern überführt wird. Die Hauptkomponenten einer Spracherkennungssoftware sind das akustische Frontend, das akustische Modell, das Sprachmodell und die Suchkomponente. Das akustische Frontend nimmt das Sprachsignal entgegen. Mit dem akustischen Modell wird bestimmt mit welcher Wahrscheinlichkeit ein Teil des Sprachsignals einem bestimmten Laut entspricht. Über das Sprachmodell wird die Wahrscheinlichkeit für Wortfolgen berechnet. Abschließend geben beide Module ihre Ergebnisse an die Suchkomponente weiter, welche die Eingaben verarbeitet und die von ihm erkannte Wortfolge ausgibt. Spracherkennungssysteme können in vier Klassen geordnet werden:

- Systeme zur Erkennung isolierter Wörter: Dies sind Spracherkennungsprogramme, die nur Einzelwörter erkennen können und zwischen jedem gesprochenen Wort eine Pause verlangen.
- Systeme zur Erkennung verbundener Wörter: Diese Spracherkenner sind ähnlich denen, die nur isolierte Wörter erkennen können, jedoch darf die Pause zwischen zwei Wörtern auch nur kurz sein.
- Systeme zur Erkennung kontinuierlicher Sprache: Die Spracherkennungssysteme erlauben es dem Benutzer in nahezu natürlicher Sprache zu sprechen wie etwa bei einem Diktat.
- Systeme zur Erkennung spontaner Sprache: Diese Programme können Sprache erkennen, die viele Elemente enthalten, die etwa bei einem Diktat nicht vorkommen, beispielsweise „hmm“ oder „äh“.

Weiterhin lassen sich Spracherkennungssysteme klassifizieren nach

- Sprachart: isoliert oder kontinuierlich
- Sprecherart: Sprecherabhängig, sprecherunabhängig oder sprecheradaptiv
- Vokabulargröße: Klein, mittel oder groß
- Sprachstil: Diktat oder spontan

Die Erkennungsgenauigkeit einer Spracherkennungssoftware hängt von verschiedenen Faktoren ab. Dazu gehören:

- Die Umgebung: Geräusche, Arbeitsumgebung
- Signalerfassung: Mikrofon, Telefon
- Sprecher: Sprecherabhängigkeit, Sprecherunabhängigkeit, Geschlecht, Alter
- Sprachstil: Lautstärke, Sprechgeschwindigkeit
- Vokabular: Spezifisches oder generisches Vokabular

Grundsätzlich gibt es drei verschiedene Ansätze zur Spracherkennung:

- Akustisch phonetischer Ansatz
- Mustererkennungsansatz
- Künstlicher Intelligenz-Ansatz

Der akustisch phonetische Ansatz basiert auf der ältesten Form der Spracherkennung, die Laute erkennt und ihnen Bezeichner zuweist. Jedoch wird dieser Ansatz selten von kommerziellen Produkten verwendet. Im Gegensatz dazu sind Ansätze, welche Mustererkennung verwenden heutzutage weit verbreitet. Sie benutzen Sprachvorlagen oder stochastische Modelle, wie etwa das Hidden Markov Modell zum Vergleich des Sprachsignals mit zuvor erstellten Trainingsdaten. Der Ansatz der künstlichen Intelligenz kombiniert den akustisch phonetischen Ansatz mit dem Mustererkennungs-Ansatz. Die Performanz von Spracherkennungssystemen wird typischerweise durch ihre Wortfehlerrate und ihre Geschwindigkeit gemessen. Nach Silsbee [31] lässt sich die Fehlerrate bei der Spracherkennung durch Lippenlesen stark senken. Auch Menschen profitieren vom gleichzeitigen Lippenlesen während des Zuhörens. Dieses Phänomen ist allgemein als McGurk-Effekt bekannt, benannt nach Harry McGurk [21]. Es gibt von Massaro et al. [20] einen Avatar namens Baldi, welcher sich dies zu Nutze macht. Bei Baldi handelt es sich um einen dreidimensionalen animierten sprechenden Kopf, der zum Sprachenlernen eingesetzt werden kann.

Nach Spehr [32] wird die Qualität der Spracherkennung derzeit immer besser. Dennoch befindet sich die Spracherkennung noch nicht in einem universellen Benutzungszustand, da die Qualität ihrer Erkennung noch immer vom verwendeten Wörterbuch und dem Umfang vorhergehenden Trainings abhängt. Spracherkennung mit begrenzter Wörterbuchgröße erzielt aber bereits gute Ergebnisse.

Dieser Abschnitt hat einen Spracherkennungs-Avatar und die Grundlagen und Besonderheiten der Spracherkennung vorgestellt. Abschließend wurde festgestellt, dass sich die Qualität der Spracherkennung immer weiter verbessert.

4.4 Mimik und Gestik

Mimik und Gestik spielen eine große Rolle bei der Gesprächsführung. In Barrientos [4] wird die kommunikative Rolle der Gestik genauer betrachtet. Dabei stellte sich in Studien heraus, dass ein Sprecher, der mit Gestiken arbeitet, seltener unterbrochen wird, da der Sprecher nachdenkt und dies durch teilweise unbewusste Gestiken verdeutlicht. Das Verhalten eines Sprechers passt sich immer der aktuellen Situation an. Dies gilt vor allem für unbewusste Handlungen. Eine Frau, die mit ihrem Baby spricht, stellt ihre Sprache sowie ihr Verhalten auf das Baby ein. Wenn mehr Motivation vom Sprecher gefordert ist, verwendet dieser häufiger Gestiken, um den Inhalt überzeugender zu präsentieren. Generell werden Gestiken nur verwendet, wenn dadurch zusätzliche oder verdeutlichende Informationen an den Zuhörer übertragen werden. Ansonsten verzichtet der Sprecher darauf.

Das Gesprächsverhalten unterscheidet sich teilweise deutlich je nach Region. In Deutschland ist es etwa als unhöflich angesehen, den Sprecher zu unterbrechen und ihn seine Ausführungen nicht beenden zu lassen, wohingegen in Spanien vor allem in größeren Gruppen jeder jeden unterbricht, sobald er das für sich Relevante aus den Worten des Sprechers herausgehört hat. Dies kann damit begründet werden, da in der deutschen Sprache das Verb bei Erzählungen meist am Ende steht und der Sprecher lange, verschachtelte Sätze bildet. Somit werden erst am Schluss des Satzes alle Einzelheiten klar. Die romanischen Sprachen erlauben diese Stellung im Satzbau nicht, darum kommt man in Spanien meistens nur dazu, die Hälfte des Satzes zu sagen, bevor man von jemandem unterbrochen wird.

Ein guter Sprecher untermalt seine Aussagen mit Gestiken, sieht sein Gegenüber an und transportiert die Stimmung seiner Erzählung durch die Stimmlage und die Körpersprache. Wer freudestrahlend vom Tod eines Angehörigen berichtet, wird vermutlich nur sehr schwer ernstgenommen werden. Der Zuhörer schaut interessiert aber nicht permanent zum Sprecher, bestätigt diesen durch entsprechende Gesten (Kopfnicken, lachen, Hochziehen einer Augenbraue, trauriger Blick etc...) und bringt vereinzelt kurze verbale Anmerkungen.

Um einen Avatar optisch ansprechend abzubilden und seine Mimik und Gestik natürlich wirken zu lassen gibt es mehrere Ansätze. Einer davon ist in Gröne [13] beschrieben und arbeitet mit einem Polygongitter, welches die Gesichtsoberfläche darstellt. Die Bewegung wird anhand von Interpolation zwischen einem Start- und einem Endobjekt berechnet. Jeder Knoten im Gitter wird stufenweise an seine neue Position geschoben, sodass die Animation flüssig und natürlich erscheint. Das Verfahren lässt sich gewichten, indem man mehrere Zielobjekte definiert und diese mit einem Faktor kombiniert, sodass die Ausprägung eines Zieles stärker zur Geltung kommt als ein anderes. Da das Verfahren auf Berechnungen

beruht, ist es echtzeitfähig, verlangt aber je nach Ausprägung entsprechende Rechenkraft eines Computers.

Ein weiterer Ansatz von Morishima und Yotsukura [22] funktioniert per Sprachsteuerung. Dafür werden für alle Vokale und Selbstlaute Aufzeichnungen einer Person erstellt und durch Parametrisierung digitalisiert. Eine generische Gesichtsmaske wird manuell auf das persönliche Gesicht einer Testperson zugeschnitten, sodass mit Front- und Seitenbilder ein vollwertiges 3D Abbild entsteht. Spricht der Nutzer nun, wird das Gesprochene in Echtzeit analysiert und in die einzelnen Aufzeichnungen umgerechnet. Zusätzlich kann aus dem Gesprochenen auch die Stimmung des Benutzers auf den Avatar übertragen werden. Zur Verfügung stehen bei dem vorgestellten Ansatz aber nur die Emotionen Wut, Trauer, Fröhlichkeit und neutral.

Es gibt einige Untersuchungen und Studien im Bereich der Gesprächsführung und Verhaltensforschung und diese Ergebnisse sollten bei der Erstellung eines Avatars berücksichtigt werden, um den Avatar dem Menschen anzupassen (siehe Abschnitt 3.2 „Akzeptanzprobleme“).

4.5 Körpersteuerung

Dieses Kapitel behandelt die Technik, die die Spielekonsolen XBox, Playstation und Wii verwenden, um den Benutzer das System per Gesten kontrollieren zu lassen.

Microsoft Kinect kann gleichzeitig Farbvideos, dreidimensionale Videos und Geräusche aufnehmen. Es benutzt hauptsächlich einen 3D Tiefensensor, um die Körperbewegungen eines Spielers zu erfassen. Zwei Linsen bauen das 3D Tiefensystem auf. Eine Linse ist der Infrarotsender, die andere ist Infrarotempfänger (siehe Bild)

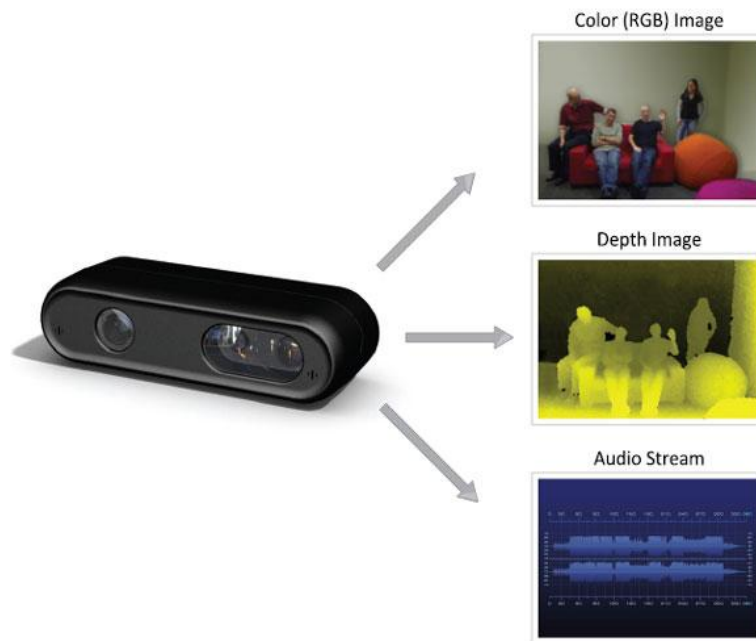


Abbildung 48 - Microsoft Kinect Erfassungsmöglichkeiten. Die Kinect Konsole kann ein Farbbild, ein Tiefenbild und Geräusche erfassen. Quelle: PrimeSense Webseite⁴⁶



Abbildung 49 - Microsoft Kinect Sensoren. Dargestellt sind die verschiedenen Sensoren der Kinect Konsole: Die RGB Kamera, die Tiefensensoren und die Mikrofone. Quelle: Games Blog⁴⁷

Das Kinectsystem benutzt zuerst die Light Coding Technik, um die Körperbewegung des Spielers zu erkennen. Light Coding codiert mithilfe der Infrarotkamera den erfassten Raum um den Sensor. Der Prozessor kann daraus ein räumliches Bild generieren. Wenn die Infrarotstrahlen auf einen Gegenstand treffen, werden die Umrisse erfasst. Da jedes Objekt unterschiedliche Umrisse besitzt und der Sensor der Kinect sehr genau arbeitet, kann jedes sich bewegende Objekt im Raum erfasst und zugeordnet werden.

⁴⁶ <http://www.primesense.tw/>

⁴⁷ <http://www.guardian.co.uk/technology/gamesblog/2010/jun/21/games-microsoft-kinect-pricing>

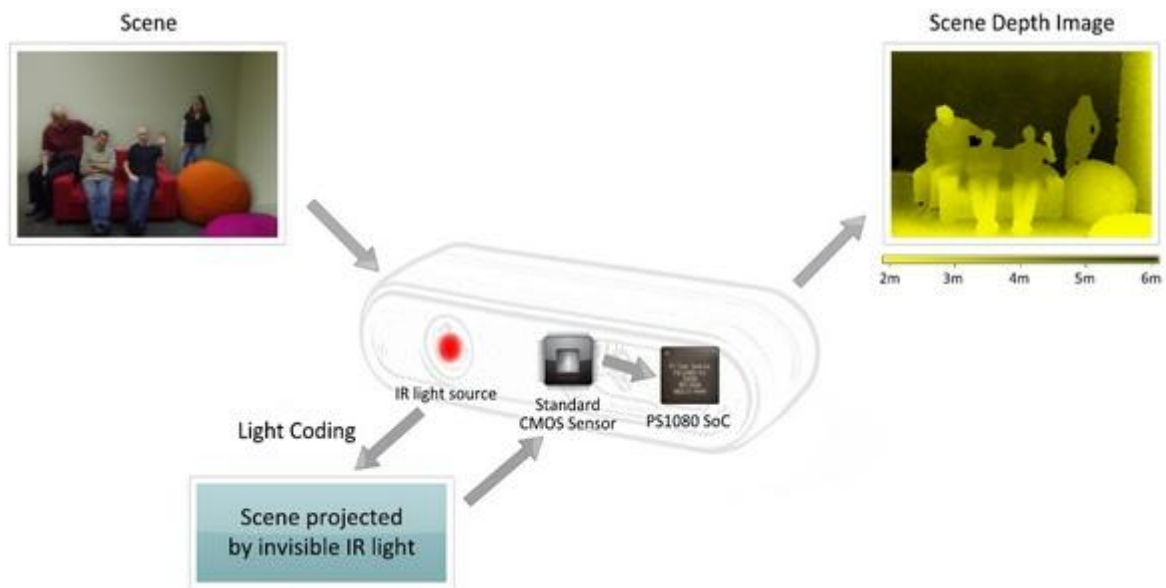


Abbildung 50 – Darstellung der Light Coding Technik. Mittels Infrarotlicht werden die Umrisse der Szene erfasst, ein CMOS Sensor nimmt das Signal auf und ein Prozessor errechnet daraus ein Tiefenbild. Quelle: PrimeSense⁴⁸

Nachdem Kinect eine Grafik erfasst hat, transformiert das System die von Light Coding erzeugte Grafik in ein skelettbasierendes Körperverfolgungssystem. Das System kann parallel 20 Körperteilgruppen von insgesamt zwei Spielern erfassen. Die Bewegungen von Körper, Gliedern und Fingern gehören zu dem erfassten Umfang. Die Haltung jedes erkannten Körperteils wird analysiert und mithilfe der Farbkamera ein vollwertiges Bild erzeugt. Die räumliche Tiefe erzeugt das bereits erwähnte 3D Tiefensystem, sodass am Ende ein vollwertiges Abbild der erkannten Person entsteht.

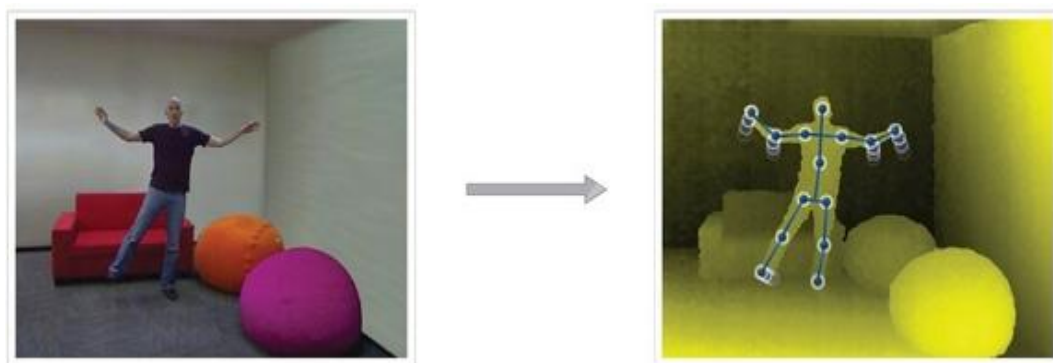


Abbildung 51 – Skelettbasierendes Körperverfolgungssystem. Sichtbar ist die Erfassung des Tiefenbilds, welches ein skelettbasierendes Körperverfolgungssystem für menschliche Körper verwendet. Quelle: PrimeSense⁴⁹

Danach untersucht die XBox Konsole jeden Pixel und entscheidet, welche Pixel zusammengehörige Gruppen von Körperteilen ergeben. Zusammen mit einer Datenbank entscheidet die Konsole, ob ein Pixel einem Körperteil entspricht. Ist dies der Fall, wird die Position im Raum errechnet. Besonders häufig ausgeführte Bewegungen können in der

⁴⁸ <http://www.primesense.tw/>

⁴⁹ <http://www.primesense.tw/>

Datenbank abgespeichert werden, sodass die Berechnungen beim nächsten Mal geringer ausfallen.

Nintendo geht mit der Wii einen ähnlichen Weg. Das Konzept besteht aus drei Teilsystemen, die die Körperbewegungen des Spielers erfassen.

- Kommunikationssystem

Der Handsender (Wii Remote) benutzt das Bluetooth Protokoll, um sich mit der Wii-Konsole (Wiimote) zu verbinden. Die Informationsübertragung findet bidirektional statt. Der Handsender fungiert sowohl als Sender als auch als Empfänger ebenso wie die Konsole. Der Empfänger des Handsenders nimmt die Infrarotsignale der Konsole entgegen.

- Positionierungssystem

Beim Spielen muss die Konsole wissen, wie sich der Spieler bewegt. Das heißt, dass die Position des Handsenders ständig erfasst und ermittelt werden muss. Eine Sensorleiste (Sensorbar), die bei der Konsole platziert wird, ist das wichtigste Gerät, um diese Positionsdaten zu ermitteln.



Abbildung 52 – Sensorleiste der Wii. Mittels Infrarotstrahlen wird ermittelt, wo sich die Handsender befinden. Quelle: Wisegamers Webseite⁵⁰

Seine Hauptfunktion besteht darin, Infrarotstrahlen auszusenden, welche die Handsender empfangen. Diese schicken darauf über die Bluetoothverbindung ihre Positionsdaten an die Konsole zurück.

⁵⁰ http://wisegamers.ch/images/29/aa_sensorbar.jpg

- Beschleunigungssystem

Die Handsender verfügen über einen Drei-Achsen-Beschleunigungssensor, mit dem die Beschleunigung ermittelt wird, damit sie anschließend an die Konsole übertragen werden kann.

Das System Move der Sony Playstation, welches die Gestensteuerung für die Konsole nachrüstet, besteht aus der Bewegungssteuerung (Motion Controller), der Navigationssteuerung (Navigation Controller) und dem Playstation Eye.



Abbildung 53 – Die drei Teilsysteme der Playstation Move. Diese umfassen den Motion Controller, den Navigation Controller und das Playstation Eye. Quelle: Techbang Webseite⁵¹

Die Motion Controller ähnelt dem Handsender der Nintendokonsole. An der Vorderseite ist eine LED Lampe angebracht. Die Lampe kann während des Spiels eine zufällige Farbe annehmen, um sich der Umgebung des Spiels anzupassen. Die Hauptaufgabe der Lampe besteht allerdings darin, dass damit die Bewegungen des Spielers verfolgt werden können, wenn dieser Aktionen ausführt. Das Move-System fängt die Bewegung des Motion Controllers anders als bei Wii nicht durch Infrarot, sondern mit Hilfe der Playstation Eye Kamera. Playstation Eye erfasst das Licht aus der LED Lampe und Move daraus kann die Formen von der Lampe und die Stärke des Licht mit den vordefinierten Daten vergleichen, dann weiß die Konsole, wie groß der Abstand des Senders gegenüber Playstation Eye ist. Move hat eine große Schwachstelle: Bei zu hellem Umgebungslicht werden die Bewegungen nicht sauber erfasst.

⁵¹ <http://www.techbang.com.tw/posts/3421-t-off-defeat-playstation-move-experience>



Abbildung 54 – Playstation Move Erfassung. Die Konsole erfasst die Bewegungen der Spieler mit Hilfe des Playstation Eye. Diese verfolgt die leuchtende Spitze des Motion Controllers. Quelle: Techbang Webseite⁵²

Die interne Struktur des Motion Controllers ist ähnlich der der Wii. Mit dem Drei-Achsen-Beschleunigungssensor hat das System größere Genauigkeit bei der Kontrolle. Der zusätzliche Navigation Controller besitzt Pfeiltasten und Rundblickschalter und benutzt Bluetooth, um mit der Playstation zu kommunizieren.

4.6 Visualisierung von Emotionen

Emotionen sind ein grundsätzlicher Bestandteil menschlicher Kommunikation. Je nach Emotionslage eines Menschen sind unterschiedliche Verhaltensweisen der anderen beteiligten Personen beobachtbar. Dieser Abschnitt untersucht wie Emotionen auf Avatare übertragen und visualisiert werden können. Nach Koda et al. [14] werden Emotionen von Avataren die einem anderen Kulturkreis als dem des Benutzers entstammen häufig fehlgedeutet. Siehe dazu Abschnitt 3.1 „Kulturelle Unterschiede“.

In Guo und Zhang [36] wird ein System vorgestellt mit dem Emotionen laut den Autoren äußerst genau auf Avatare übertragen werden können. Die hohe Genauigkeit wird dadurch erzielt, dass zunächst Benutzerprofile angelegt werden und die Emotionserkennung somit benutzerabhängig durchgeführt wird. Das Gesicht des Benutzers wird dabei per digitaler Videokamera erfasst. Emotionen werden in Gesichtsbewegungsgraphen gespeichert, welche auf bestimmten Punkten im Gesicht und ihren Beziehungen zueinander basieren. Wird eine Emotion erfasst, so wird ihr Gesichtsbewegungsgraph mit dem im jeweiligen Benutzerprofil gespeicherten Gesichtsbewegungsgraph verglichen und bei Gleichheit ausgegeben.

Liu [18] entwickelte ein Modell zur Visualisierung von Emotionen bei Avataren, welches auch die Körperhaltung miteinbezieht. Dieses Modell basiert auf dem OCC-Modell von Ortony, Clore und Collins [9]. Im OCC-Modell werden Emotionen als Reaktion auf Objekte, Aktionen des Avatars und Ereignisse generiert. Das OCC-Modell kennt 22 verschiedene emotionale

⁵² <http://www.techbang.com.tw/posts/3421-t-off-defeat-playstation-move-experience>

Zustände. Die Methode von Liu definiert Emotionen als Vektoren deren zugehöriger Gesichtsausdruck ebenfalls als Vektor definiert ist. Der Gesichtsausdrucksvektor enthält dabei alle Vertices seines Ausdrucks. Wechselt nun der Emotionsvektor, so wechselt auch sein zugehöriger Gesichtsausdrucksvektor.

Der Avatar Eve der YelloStrom GmbH erhielt nach Lindner [17] seine Emotionsvisualisierung durch gegenseitige Beobachtung, Modellierung und Animation. Die Mitarbeiter des Entwicklungsteams beobachteten sich untereinander beim Sprechen und notierten die signifikanten Merkmale der Mimik und Gestik. Dadurch wurden 20 Standard- und noch einige Extrastimmungen definiert. Im Anschluss wurden diese Stimmungen als Bewegungssequenz auf das 3D-Modell von Eve übertragen.

4.7 Grenzen der Interaktionsmöglichkeiten mit Avataren

Dieses Kapitel behandelt die Grenzen der Interaktionsmöglichkeiten mit Avataren. Menschen, die mit Avataren interagieren, brauchen eine entsprechende Umgebung, welche dieses Ziel unterstützt. Die Umgebung besteht aus einer gewissen Hardwareausstattung und Datenlogik. Sie ist der entscheidende Faktor für die Interaktion mit Avataren. Die beschränkten technischen Möglichkeiten stellen darum einen begrenzenden Faktor für die Interaktion dar.

Um eine Mensch-Avatar-Interaktion zu initiieren, benötigt man ein geeignetes Medium zur Übertragung, da der Mensch sich in der realen und der Avatar sich in der virtuellen Welt befindet. Das Medium kann ein Peripheriegerät wie Maus, Tastatur, Spielekonsole etc. aber auch ein Medium aus einem Peripheriegerät wie Infrarot, WLAN, Bluetooth etc. sein. Arbeitet das System innerhalb normaler Parameter und bleibt man innerhalb dieses Bereiches, ist eine Interaktion möglich. Das unten stehende Bild beschreibt die räumliche Abdeckung der Sensorleiste der Wii Konsole von Nintendo. Nur wenn man sich im richtigen Bereich aufhält und das System normal funktioniert, ist eine Interaktion möglich.

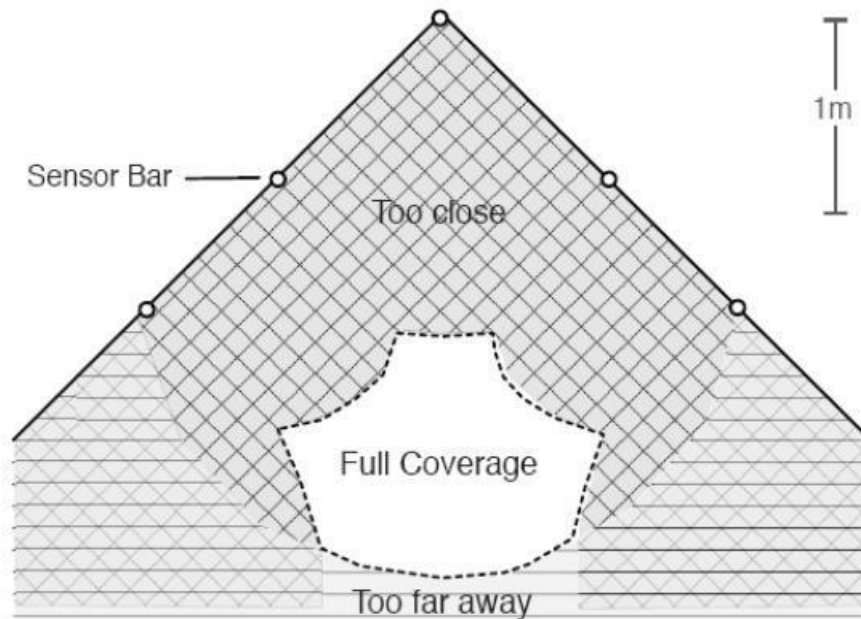


Abbildung 55 – Arbeitsbereich der Wii Sensorbar. In diesem kann man mit den Avataren in der Wii interagieren. Außerhalb dieses Bereichs ist die Grenze der Interaktionsmöglichkeit mit Avataren bei Wii. Quelle: Webseite⁵³

Im ausgezeichneten Bereich kann man mit Avataren in der virtuellen Welt kommunizieren. Man kann auf den Bildern sehen, dass es bestimmte Bereiche gibt, wo der Empfang einen besseren Wert im Wii System erzielt. Das zeigt, dass die Grenze der Hardware bzw. Sensorbar die grundlegende Grenze der Interaktionsmöglichkeit mit Avataren in der virtuellen Welt darstellt.

Die vordefinierte Datenbank und Datenlogik stellen weitere Grenzen der Interaktion mit Avataren dar. Außerdem spielt die hardwareseitige Grenze bezüglich der maximal gleichzeitig verarbeitbaren Spieler bei der Datenbank und Datenlogik hier auch eine große Rolle, weil die Reaktion von den Avataren vor dem Interaktionspartner schon vordefiniert oder vorgeregelt wurde. Man speichert die mögliche Handlung aus Menschen in die Datenbank und definiert dann eine Verknüpfung zwischen den Handlungen vom Menschen und den Reaktionen der Avataren. Deshalb entscheidet die Qualität der Datenbank über die Grenze der Interaktionsmöglichkeit von Avataren. Auf manchen Webseiten etwa werden Avatare als Führer eingesetzt, man kann per Tastatureingabe mit ihnen kommunizieren. Für normale Gespräche wie „hallo“ oder „wie geht's?“ kann ein Avatar so gut wie echte Menschen reagieren, aber wenn unvorhergesehene Fragestellungen auftauchen, kann der Avatar darauf nicht passend reagieren und antwortet meist mit einer vordefinierten Antwort für solche Fälle, da in der Datenbank nicht für jede Anfrage eine passende Antwort vorliegen kann.

⁵³ http://acade.must.edu.tw/upfiles/ADUplod/c23_downmul1331673874.pdf



” Herzlich Willkommen auf dieser Plattform! ... ”

Abbildung 56 – Führer der Yaoti Webseite. Mit der Hilfe dieser Seite kann man einen Avatar erzeugen und in eigene Webseiten integrieren. Der Avatar hat in seiner Datenbank vordefinierte Ausdrücke, Gesten und Sprechsätze. Er kann individuell auf die Anfragen der Nutzer eingehen. Quelle: Yaoti Webseite⁵⁴

Die am weitesten verbreitete Behandlung für nicht definierte Situationen sind Default Antworten. Aber in der Praxis weiß man, dass die Vordefinierung nicht immer ausreicht. Man kann nicht alle Situationen und Szenen, die vielleicht passieren könnten, vorplanen. Darum ist die Fähigkeit, dass die Datenbank selbst lernt und sich entwickelt, ein wichtiges Kriterium, mit der man die Grenze der Interaktionsmöglichkeit mit Bezug auf der Avatardatenbank bewerten kann. Deshalb kann man auch sagen, dass die Grenze der Datenbank und Datenlogik auch die Grenze der Interaktionsmöglichkeit mit Avataren darstellt. Sie definiert mit Hardware zusammen die Interaktionsmöglichkeit mit Avataren.

4.8 Zusammenfassung

Dieses Kapitel sollte zeigen, wie Interaktionstechnologien aussehen und sich über die Jahre entwickelt haben. Es wurde dargelegt, wie Avatare Sprache, Mimik, Gestik und Verhalten erkennen und verarbeiten. Danach wurden die Grundprinzipien der Gestensteuerung an den drei am weitesten verbreiteten Spielekonsolen verdeutlicht und die Grenzen der Interaktionsmöglichkeiten mit Avataren dargestellt.

⁵⁴ <http://www.yaoti.com>

5 Evaluation von bestehenden Frameworks zur Avatarerstellung

Dieses Kapitel beschreibt anhand ausgewählter Beispiele, welche Möglichkeiten aktuelle Frameworks zur Erstellung von Avataren bieten. Zuerst werden die Untersuchungskriterien vorgestellt, auf welche anschließend die Frameworks getestet wurden. Zum Abschluss folgt ein Fazit über alle Testkandidaten.

5.1 Untersuchte Kriterien

Um eine objektive Aussage über die Stärken und Schwächen eines Frameworks machen zu können, wurden die im Folgenden beschriebenen Kriterien herangezogen. Diese orientieren sich speziell an den Charakteristika von Webeditoren und setzen den Fokus gezielt in Bereichen wie Benutzbarkeit und technische Voraussetzungen. Daneben spielen die Gestaltungsmöglichkeiten eine große Rolle, da sie ein zentraler Punkt dieser Arbeit sind. Die einzelnen Kriterien werden im Verlauf des Kapitels erklärt.

Der Punkt *Registrierung* erläutert, ob das Framework ein Nutzerkonto voraussetzt, wenn es sich um ein Online-Framework handelt. Zusätzlich werden problematische Details der Nutzungsbedingungen angesprochen, sofern vorhanden.

Bei den *Voraussetzungen* wird geklärt, welche Vorbedingungen gelten müssen, um das Framework nutzen zu können. Darunter fallen Punkte wie die Beschränkung auf diverse Betriebssysteme, zwingend vorinstallierte Programme (etwa der Adobe Flash Player für ein Webframework) oder ob die Nutzungsbedingung zur Ausführung akzeptiert werden müssen.

Wie das Framework den Nutzer beim Anlegen eines Avatars unterstützt ist im Punkt *Vorlagen* angemerkt. Gibt es Vorlagen für den Bau von Avataren? Können eigene Bilder benutzt werden? Gibt es Vorlagen für Gestaltungsmerkmale (Körperteile, Kleidung, Accessoires, etc.) des Avatars?

Die *emotionale Gestaltung* ist ein wichtiges Kriterium, denn damit wird dem Avatar die Menschlichkeit verliehen. Wie unterstützt das Framework den Benutzer dabei? Die Art der Zustände (Wut, Trauer, Freude, etc.) sowie deren Umsetzung (Zwinkern, Tränen, Handzeichen, etc.) wird hier geschildert.

In dieselbe Richtung gehen die Kategorien *optische Gestaltung* sowie *akustische Gestaltung*. Bei der Optik wird vor allem darauf geachtet, welche Teile des Avatars wie detailliert verändert werden können und ob der Avatar spezielle Stile annehmen kann (etwa Gothic oder Comic). Zur akustischen Gestaltung könnte dem Avatar etwa ein vordefinierter Sprechtext oder eine eigene Sprachaufnahme vorgegeben werden. Die Möglichkeit zur Stimmmodulation des Avatars ist ebenfalls ein untersuchter Punkt.

Der nächste Punkt befasst sich mit der *Darstellung* des Avatars. Sind die Bewegungen bei Animationen flüssig? Stimmt die Lippensynchronität bei der Sprachausgabe? Wie ist allgemein die Qualität der Darstellung?

Ein wichtiger Punkt sind die *Exportmöglichkeiten* des Frameworks. Diese sollten offen für andere Anwendungen sein oder den Avatar in einem Format exportieren, welches für andere Anwendungen lesbar ist. Wie der Avatar weiterverwendet werden kann (etwa Export als Bild oder Video), wird in diesem Abschnitt erläutert. Ebenfalls wichtig ist die *Wiederverwendbarkeit*. Die Erstellung und Gestaltung eines Avatars kann je nach Framework durchaus mehrere Stunden in Anspruch nehmen und diese Arbeit sollte möglichst exportgeeignet gespeichert und verwendet werden können. Ob und wie Avatare gespeichert werden können, ist in diesem Kapitel beschrieben.

Für den Benutzer ist es wichtig, dass das Framework ein hohes Maß an *Benutzbarkeit* bietet, damit er sich voll und ganz auf die Arbeit konzentrieren kann. Eine übersichtliche und intuitive Menüführung und Bedienung ist dafür unerlässlich. Ebenso sind Hilfen und Einführungen wünschenswert, um den Start zu erleichtern. Geringe Ladezeiten sowie ein fehlerfreier Programmablauf sind ebenfalls ein Kriterium.

Es ist angenehm, wenn der Avatar, der gerade erstellt oder bearbeitet wird, mit dem Benutzer bereits interagiert, etwa durch Verfolgen des Mauszeigers mit den Augen oder ein Lächeln, wenn er ein spezielles Kleidungsstück zugewiesen bekommt. Inwiefern die *Benutzerinteraktion* im Framework im Fokus steht, wird hier gezeigt.

Im Abschnitt *Sonstiges* werden weitere Aspekte beschrieben, die das Framework auszeichnen oder erwähnenswert für die Nutzung damit sind. Beispiele wären etwa, dass mit dem Avatar auf irgendeine Weise Geld verdient werden kann oder dass der Avatar als vollgerenderte Figur für den Export zur Verfügung steht.

5.2 Vergleichstabelle

Die nachfolgende Tabelle gibt einen bewerteten Überblick gemäß den Kriterien für die einzelnen Frameworks. Die Bewertung besitzt fünf Stufen von sehr gut (++) bis unbefriedigend (--). Die Gesamtbewertung ergibt sich zu gleichen Teilen aus den Einzelwertungen kaufmännisch gerundet auf die nächste Stufe, wobei die Wertung sehr gut einer Punktzahl von 5 entspricht, während die Wertung unbefriedigend mit einem Punkt verrechnet wird. Kriterien, die das Framework nicht bereitstellt, werden mit n. v. (nicht vorhanden) gekennzeichnet und mit null Punkten verrechnet.

Kriterium	Voki	My Avatar Editor	Evolver	XBox Avatar Editor	Yahoo Avatare
Registrierung?	optional	nein	nein	nein	ja
Vorlagen	++	n.v.	++	++	-
Optische Gestaltung	++	++	++	++	+
Akustische Gestaltung	++	n.v.	+	n.v.	n.v.
Emotionale Gestaltung	n.v.	n.v.	0	0	0
Benutzbarkeit	+	+	+	+	0
Exportmöglichkeiten	++	+	++	--	--
Wiederverwendbarkeit	++	+	++	-	0
Gesamtbewertung	+	-	+	0	-

5.3 Detaillierte Untersuchung der einzelnen Frameworks

Der folgende Abschnitt beinhaltet eine detaillierte Untersuchung aller vorgestellten Frameworks.

5.3.1 Voki



Überblick

Voki⁵⁵ setzt sich laut FAQ seiner Webseite aus „vox“ und „Loki“ zusammen. „Vox“ ist lateinisch für Stimme und „Loki“ ist eine nordische Gottheit. Es handelt sich um ein Onlineframework, welches nicht lokal installiert werden kann. Nach der Erstellung können

⁵⁵ <http://www.voki.com>

die Avatare in verschiedenen Webdiensten und sozialen Netzwerken verwendet werden. Der Avatar befindet sich in einem Rahmen, der Player genannt wird und verschiedene Kontrollmöglichkeiten bereitstellt, wie etwa das Abspielen von voreingestellten Sounds. In Abbildung 57 ist die Bearbeitungsansicht dieses Frameworks dargestellt. Nachfolgend wird Voki detailliert vorgestellt.



Abbildung 57 – Bearbeitungsansicht des Voki-Frameworks. Zu sehen sind der Avatar in seinem Rahmen und verschiedene Kontrollelemente, etwa zur Vorlagen- und Kleidungsauswahl, zur Einstellung der Tonwiedergabe, zur Hintergrund- und Rahmenauswahl. Quelle: Voki Webseite⁵⁶

Registrierung

Es ist auch ohne Registrierung möglich einen Avatar zu erstellen, dieser kann dann jedoch nicht verlinkt, gespeichert oder in soziale Netzwerke eingebunden werden. Nach erfolgreicher Registrierung ist dies aber möglich.

Voraussetzungen

Das Portal kann nur mit dem Flash Player⁵⁷ benutzt werden. Voki darf nicht zu kommerziellen Zwecken verwendet werden, hierzu steht die kostenpflichtige Variante SitePal⁵⁸ zur Verfügung.

⁵⁶ <http://www.voki.com>

⁵⁷ <http://www.adobe.com/de/products/flashplayer>

⁵⁸ <http://www.sitepal.com>

Vorlagen

Es besteht eine große Auswahl an Avatarvorlagen. Der Benutzer kann die Vorlagen aus verschiedenen Kategorien wählen: VIPs, Politiker, Hunde, Katzen, weitere Tiere und noch vieles mehr. Vorlagen können nach Geschlecht getrennt ausgewählt werden. Die Verwendung eigener Bilder ist nicht möglich.

Emotionale Gestaltung

Die Avatare können keine emotionalen Zustände einnehmen. Ihr Gesichtsausdruck entspricht weitestgehend dem der ausgewählten Vorlage und lässt sich bestenfalls durch Größe und Form der ausgewählten Gesichtselemente beeinflussen.

Optische Gestaltung

Die optischen Gestaltungsmöglichkeiten sind äußerst umfangreich. Allerdings kann nur der Oberkörper gestaltet werden, da dieser ab einem bestimmten Punkt abgeschnitten wird. Dieser Punkt ist nicht bei allen Avataren an der gleichen Position. Per Zufallsfunktion kann ein fertiger Avatar erzeugt werden. Viele Detailgestaltungsmöglichkeiten sind von der gewählten Vorlage abhängig und bei anderen Vorlagen nicht auswählbar. Es können die Haarform, die Mundform und bei männlichen Avataren auch der Bart ausgewählt werden. Weiterhin ist die Kopfhöhe und –breite, Größe von Nase und Mund und die Breite des sichtbaren Körperteils einstellbar. Die Farbe von Mund, Augen, Haut und Haaren kann bestimmt werden. Kleidung kann aus verschiedenen Kategorien wie beispielsweise Sport oder Freizeit gewählt werden. Häufig ist eine Kopfbedeckung mit auswählbar. Diverse Brillen und Schmuck können angelegt werden. Für den Hintergrund bestehen zahlreiche Auswahlmöglichkeiten aus verschiedenen Kategorien wie Reisen oder Sport, es können auch eigene Hintergründe eingebunden werden. Einstellbar ist auch die Größe und Position des Avatars vor dem Hintergrund. Die Farbe des Rahmens kann ebenfalls gewählt werden.

Akustische Gestaltung

Auch bei der akustischen Gestaltung gibt es mehrere Auswahlmöglichkeiten. Hervorstechend ist die Möglichkeit anzurufen und seinen eigenen Text aufzusprechen. Dazu wird bei Auswahl der Anrufmöglichkeit eine Telefonnummer und ein Identifikationscode eingeblendet. Desweiteren steht Text-to-Speech, Mikrofonaufnahme und das Hochladen eigener Tondateien zur Verfügung. Die Mikrofonaufnahme benutzt die Mikrofonsteuerung des Flash Players. Tondateien können in den Formaten WAV, MP3, WMA und PCM hochgeladen werden. Text-to-Speech steht in zahlreichen verschiedenen Sprachen zur Verfügung, die von Englisch und Deutsch über Arabisch und Türkisch bis hin zu Russisch und Koreanisch reichen. Zusätzlich können noch verschiedene männliche und weibliche Sprecher ausgewählt werden. Abschließend können alle Tondateien mit diversen Effekten versehen werden, wie etwa Echo oder Chipmunkstimme.

Darstellung

Die Darstellung ist comicartig. Die Text-to-Speech-Ausgabe erfolgt lippensynchron. Ansonsten existieren kaum Animationen.

Exportmöglichkeiten

Der erstellte Avatar kann in drei verschiedenen Größen in zahlreichen verschiedenen Diensten und sozialen Netzwerken wie Facebook⁵⁹, Twitter⁶⁰, MySpace⁶¹, Yahoo⁶², usw. gepostet oder verlinkt werden. Desweiteren steht ein einbettbarer Code zur Verfügung, um den Avatar auf eigenen Seiten einzubinden. Über einen Link ist der Avatar auch auf der Voki Webseite direkt aufrufbar.

Wiederverwendbarkeit

Durch das Anlegen eines Accounts auf der Voki Webseite kann man sich in sein persönliches Profil einloggen und seine erstellten Avatare betrachten, ändern, per Email versenden oder löschen.

Benutzerinteraktion

Der Avatar verfolgt mit seinen Augen die Position des Mauszeigers, solange dieser sich über dem Bereich des Flashplayers befindet. Er dreht auch ein wenig den Kopf in Mauszeigerrichtung. Verlässt man diesen Rahmen richten sich die Augen nach kurzer Zeit wieder geradeaus. Durch Drücken des Abspielenbuttons wird der zuletzt eingestellte Sound abgespielt. Dessen Lautstärke ist regelbar.

Benutzbarkeit

Die Benutzbarkeit ist gut. Alle Einstellungen sind Kategorien zugeordnet und bequem per Mausklick erreichbar. Lediglich in der Verwaltung kommt es derzeit (Stand April/Mai 2011) zu verlängerten Ladezeiten der jeweiligen Benutzerdaten und der erstellten Avatare. Dies wurde durch Test mit zwei verschiedenen Browsern sowie zwei verschiedenen Rechnern mit unterschiedlicher Internetverbindung bestätigt.

Sonstiges

Eine Besonderheit stellt die Möglichkeit für Lehrer dar, Schulstunden im PDF-Format mit Link zum entsprechenden Avatar hochzuladen. Für kommerzielle Zwecke steht das kostenpflichtige SitePal⁶³ zur Verfügung. Es ergänzt Voki um folgende Funktionalitäten: Professionelle Modelle, Verwendung eigener Fotos, bessere Audiofeatures, verschiedene Rahmen, Verwendung in Emails, Client-Server-API, usw. Es gibt verschiedene Pakete mit gestaffelten Merkmalen, deren Preise (Stand April/Mai 2011) von 99,50 Dollar im Jahr bis zu 2499,50 Dollar im Jahr reichen.

Fazit

Das Framework hat viele positive Seiten. Es bietet eine große Auswahl an Vorlagen und optischen Gestaltungsmöglichkeiten. Auch die akustischen Gestaltungsmöglichkeiten sind sehr vielfältig und beinhalten eine umfangreiche Text-to-Speech-Komponente. Nicht

⁵⁹ <http://www.facebook.com>

⁶⁰ <http://twitter.com>

⁶¹ <http://www.myspace.com>

⁶² <http://www.yahoo.com>

⁶³ <http://www.sitepal.com>

vorhanden ist die Möglichkeit zur emotionalen Gestaltung. Hervorzuheben ist die Auswahl an Exportmöglichkeiten, die sehr viele soziale Netzwerke und Internetdienste unterstützt.

5.3.2 My Avatar Editor



Überblick

My Avatar Editor⁶⁴ ist ein Editor für Avatare, der sowohl online verwendet als auch lokal installiert werden kann. Nach der Erstellung können die Avatare in verschiedene Formate exportiert oder auch per Link angezeigt werden. In Abbildung 58 ist die Bearbeitungsansicht des Editors dargestellt. In diesem Abschnitt wird My Avatar Editor detailliert vorgestellt.



Abbildung 58 – Bearbeitungsansicht von My Avatar Editor. Im oberen Bereich sind der Avatar und Elemente zur Eingabe von Name, Geschlecht, Kleiderfarbe, etc. sichtbar. Im unteren Bereich befinden sich die Knöpfe zur Auswahl des aktuell zu gestaltenden Gesichtsbereichs. Quelle: My Avatar Editor Webseite⁶⁵

⁶⁴ <http://www.myavatareditor.com>

⁶⁵ <http://www.myavatareditor.com>

Registrierung

My Avatar Editor kommt gänzlich ohne Registrierung aus. Avatare können aber auch ohne Registrierung weiterbearbeitet werden, siehe Wiederverwendbarkeit.

Voraussetzungen

Die Onlineversion des Editors kann nur mit dem Flash Player 10 oder höher verwendet werden. Für die Desktopvariante wird Adobe AIR 1.5⁶⁶ oder höher benötigt. Beide Editoren sind frei verwendbar.

Vorlagen

Vorlagen zur Avatarerstellung sind nicht vorhanden. Es stehen lediglich ein „Standardjunge“ und ein „Standardmädchen“ zur Verfügung. Weiterhin kann ein per Zufallsfunktion erzeugter Avatar den eigenen Vorstellungen entsprechend angepasst werden. Eigene Bilder können nicht verwendet werden.

Emotionale Gestaltung

Die durch den Editor erstellten Avatare unterstützen keinerlei emotionale Gestaltung. Im besten Fall kann durch Änderung der Gestaltungselemente ein emotionaler Ausdruck erzeugt werden.

Optische Gestaltung

Die erzeugten Avatare ähneln denen der Nintendo Wii Konsole. Die optischen Gestaltungsmöglichkeiten sind äußerst vielfältig. Ausgangspunkt kann ein per Zufallsfunktion erzeugter Avatar sein, der anschließend bearbeitet werden kann. Es können auch weitere zufällige Avatare erstellt werden. Die Zufallsfunktion erzeugt aber zum Teil recht bizarre Ergebnisse. Es besteht auch die Möglichkeit einen „Standardjungen“ oder ein „Standardmädchen“ als Ausgangspunkt zu wählen. Nun kann dem Avatar ein Name gegeben sowie der Name des Erstellers eingegeben werden. Es kann auch erneut das Geschlecht ausgewählt werden. Bei dieser Auswahl ändert sich aber lediglich das Oberteil. Ein männlicher Avatar trägt ein Sweatshirt, welches bei der weiblichen Variante durch eine rockartige Verlängerung ergänzt wird. Der Benutzer kann das jeweilige Geschlecht durch die weitere Gestaltung hervorheben, muss er aber nicht. Es kann die Kleiderfarbe eingestellt und das Geburtsdatum in der Form Tag und Monat eingegeben werden. Die Gesamtgröße und -breite kann eingestellt werden. Gesichtsform und -farbe sowie Gesichtsmerkmale wie etwa Sommersprossen oder rote Wangen sind auswählbar. Zahlreiche männliche und weibliche Haarschnitte stehen zur Verfügung, die unabhängig vom Avatargeschlecht ausgewählt werden können. Sowohl die Haarfarbe als auch die Ausrichtung einer asymmetrischen Frisur können gewählt werden. Für Augenbrauenformen bestehen diverse Auswahlmöglichkeiten und auch deren Farbe kann gewählt werden. Es kann ihr Abstand, ihre Position, ihre Größe und ihr Neigungswinkel bestimmt werden. Dasselben Einstellungsmöglichkeiten wie für die Augenbrauen bestehen auch für die Augen. Es können

⁶⁶ <http://www.adobe.com/de/products/air>

zudem unterschiedliche Brillenformen ausgewählt werden und sowohl die Farbe der Gläser als auch die Position und Größe der Brille bestimmt werden. Dasselbe gilt für den Mund und den Oberlippenbart. Bei der Nase kann die Form, ihre Position und Größe eingestellt werden. Es kann ein Kinnbart ausgewählt und seine Farbe eingestellt werden. Abschließend kann die Position und Größe eines Muttermals gewählt werden.

Akustische Gestaltung

My Avatar Editor bietet keinerlei akustische Gestaltungsmöglichkeiten.

Darstellung

Die Darstellung ist comicartig.

Exportmöglichkeiten

Ein erstellter Avatar kann im MAE Format gespeichert werden. MAE ist ein Akronym für My Avatar Editor. Avatare können auch ins XML Format exportiert werden und im JPEG oder PNG Format abgespeichert werden. Beim Export in ein Bildformat stehen noch Einstellungsmöglichkeiten für die Gesamtgröße, den Zoomfaktor und die Hintergrundfarbe zur Verfügung. Desweiteren kann der Avatar auf dem Bild positioniert werden, nur sein Kopf angezeigt werden und beim PNG-Format der Hintergrund auf transparent gestellt werden. Abschließend besteht die Möglichkeit durch eine manuelle Modifikation der URL eine Ansicht zu erhalten, bei der die Editor Oberfläche ausgeblendet ist.

Wiederverwendbarkeit

Es besteht die Möglichkeit Avatare in eine Datei lokal zu speichern und zu laden. Eine andere Möglichkeit ist die aktuelle URL zu speichern. Somit können abgespeicherte Avatare weiterbearbeitet werden.

Benutzerinteraktion

Benutzerinteraktionen sind keine möglich. Der Editor dient lediglich zur Erstellung eines statischen Bildes.

Benutzbarkeit

Das Framework ist leicht zu bedienen. Durch am unteren Rand angebrachte Reiter können die einzelnen Gestaltungsfelder ausgewählt und die jeweiligen Gestaltungselemente bearbeitet werden.

Sonstiges

My Avatar Editor ist in der Lage Mii Dateien zu laden, wie sie bei Nintendo Wii verwendet werden. Diese Dateien können angezeigt, bearbeitet und exportiert werden. Die Desktopversion des Editors unterscheidet sich von der Onlineversion nur dahingehend, dass keine Standardavatare geladen werden können. Desweiteren stehen My Avatar Editor und seine Quellen komplett frei zur Verfügung. Der Editor kann im Original als Flash Anwendung in die eigene Webseite eingebunden werden. Er kann auch eingebunden werden und über eine Javascript-API oder ActionScript-API angesprochen werden. Letztendlich besteht die

Möglichkeit seine Quellen zu modifizieren und sich so eine eigene Version des Editors zu erstellen.

Fazit

Dieses Framework hat sowohl positive als auch negative Seiten. Es bietet umfangreiche optische Gestaltungsmöglichkeiten. Negativ fallen die nicht vorhandenen Möglichkeiten zur akustischen und emotionalen Gestaltung auf. Weiterhin sind auch keinerlei Vorlagen vorhanden. Positiv hervorzuheben ist der XML Export und die Quelloffenheit, welche es ermöglicht das Framework über verschiedene APIs anzusprechen.

5.3.3 Evolver



Überblick

Auf der Startseite von Evolver⁶⁷ wird man gleich von tanzenden 3D Avataren begrüßt, die sich regelmäßig ändern. Das Framework richtet sich sowohl teilweise an professionelle Modellierer aber auch einfach an Benutzer, die sich einen Ganzkörperavatar zusammenstellen möchten. Wer möchte, kann sofort loslegen oder sich alternativ zuerst in der riesigen Galerie inspirieren.

Registrierung

Es werden verschiedene Möglichkeiten zur Anmeldung gegeben. Man kann sich ein eigenes Evolver Benutzerkonto erstellen, sich mit seinem Facebook Konto anmelden oder falls man über eine OpenID verfügt, kann man sich mit dieser einloggen. Ein Konto ist nicht zwingend erforderlich, jedoch stehen manche Funktionen wie das Speichern von Avataren und einige Exportmöglichkeiten dann nicht zur Verfügung.

Voraussetzungen

Der Editor funktioniert in allen aktuellen Browsern (Microsoft Internet Explorer 9, Mozilla Firefox 4, Opera 11, Google Chrome 10) ohne dass ein Plug-In installiert werden muss. Eine Verbindung zum Internet ist permanent erforderlich, da die Daten nicht lokal bearbeitet werden. Für eine bestmögliche Benutzbarkeit sollte JavaScript aktiviert werden.

Vorlagen

Zu Beginn besteht die Möglichkeit, zwischen vier Optionen zu wählen. Einen vorgefertigten Avatar bearbeiten, einen neuen Avatar erstellen, einen Avatar basierend auf einem Foto erstellen oder einen Avatar eines Benutzers aus der Galerie auswählen. Beim Weiterbearbeiten stehen zahllose Vorlagen für einzelne Körperteile oder Accessoires zur Verfügung.

⁶⁷ <http://www.evolver.com>

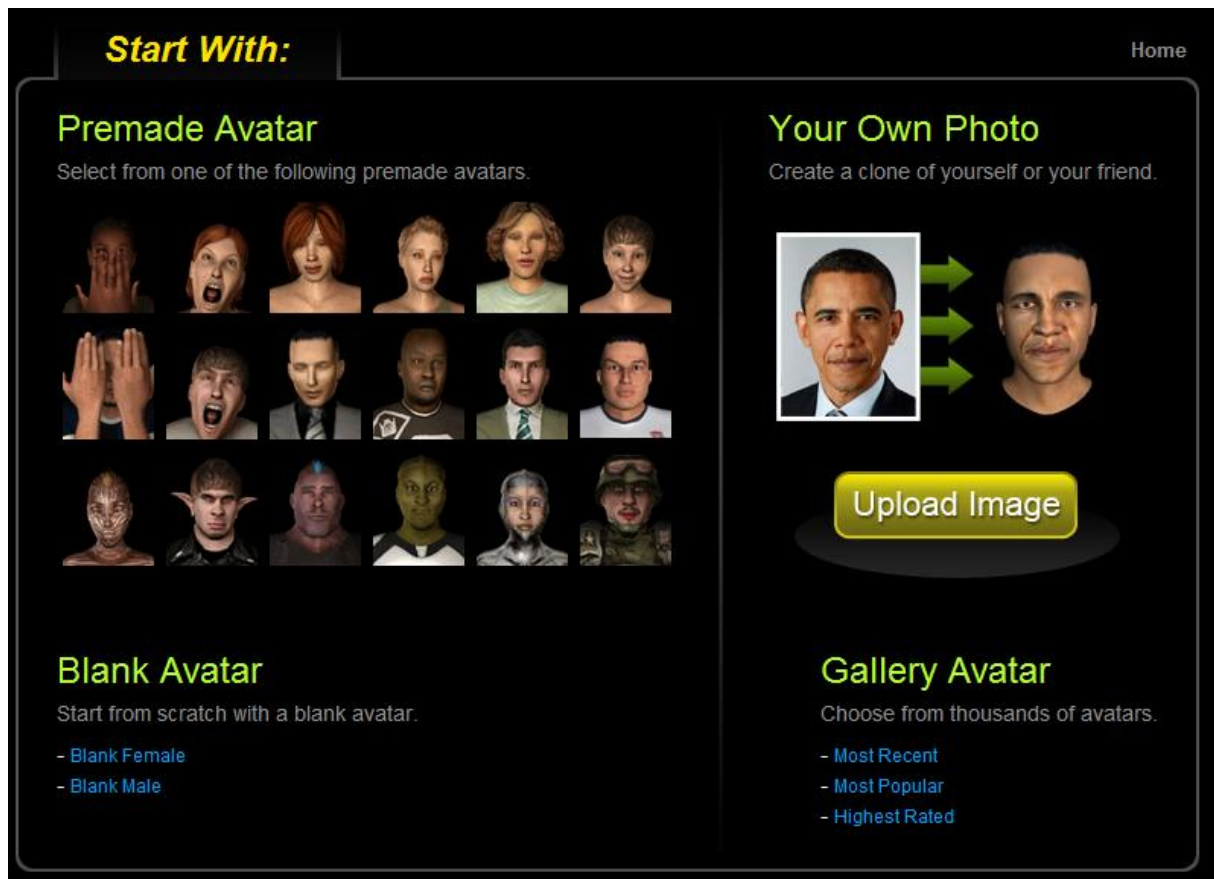


Abbildung 59 – Ausgangspunkt der Avatarerstellung bei Evolver. Man kann vordefinierte Avatare verändern, ein eigenes Foto als Basis verwenden, einen leeren Avatar benutzen oder sich einer Vorlage eines anderen Nutzers aus der Galerie bedienen. Quelle: Evolver Webseite⁶⁸

Optische Gestaltung

Hier spielt das Framework seine Stärken aus. Das Gesicht ist in allen Einzelheiten veränderbar. Augen, Lippen, Nase, Kinn, Haare, Wangen, Ohren etc. Man wählt dabei beispielsweise zwei Augenvorlagen und kann über einen Schieberegler den persönlichen Avatar als Mischung der Vorlagen gestalten. Eine Live-Vorschau informiert den Benutzer fortlaufend, wie das gegenwärtige Endprodukt aussieht. Zusätzlich kann der Avatar in einer 360 Grad Ansicht betrachtet werden. Bei der Gestaltung der Kleidung kann die JPEG-Vorlage in einem externen Grafikprogramm bearbeitet und anschließend wieder dem Avatar zur Verfügung gestellt werden. Die Darstellung des Avatars ist ein detailliertes 3D-Bild, das auch kleinste Änderungen am Schieberegler visualisieren kann. Es ist problemlos möglich, mit einem Foto als Ausgangspunkt eine sehr überzeugende Avatarrepräsentation einer Person zu erstellen, die dem Original sehr ähnlich sehen kann.

⁶⁸ <http://www.evolver.com>

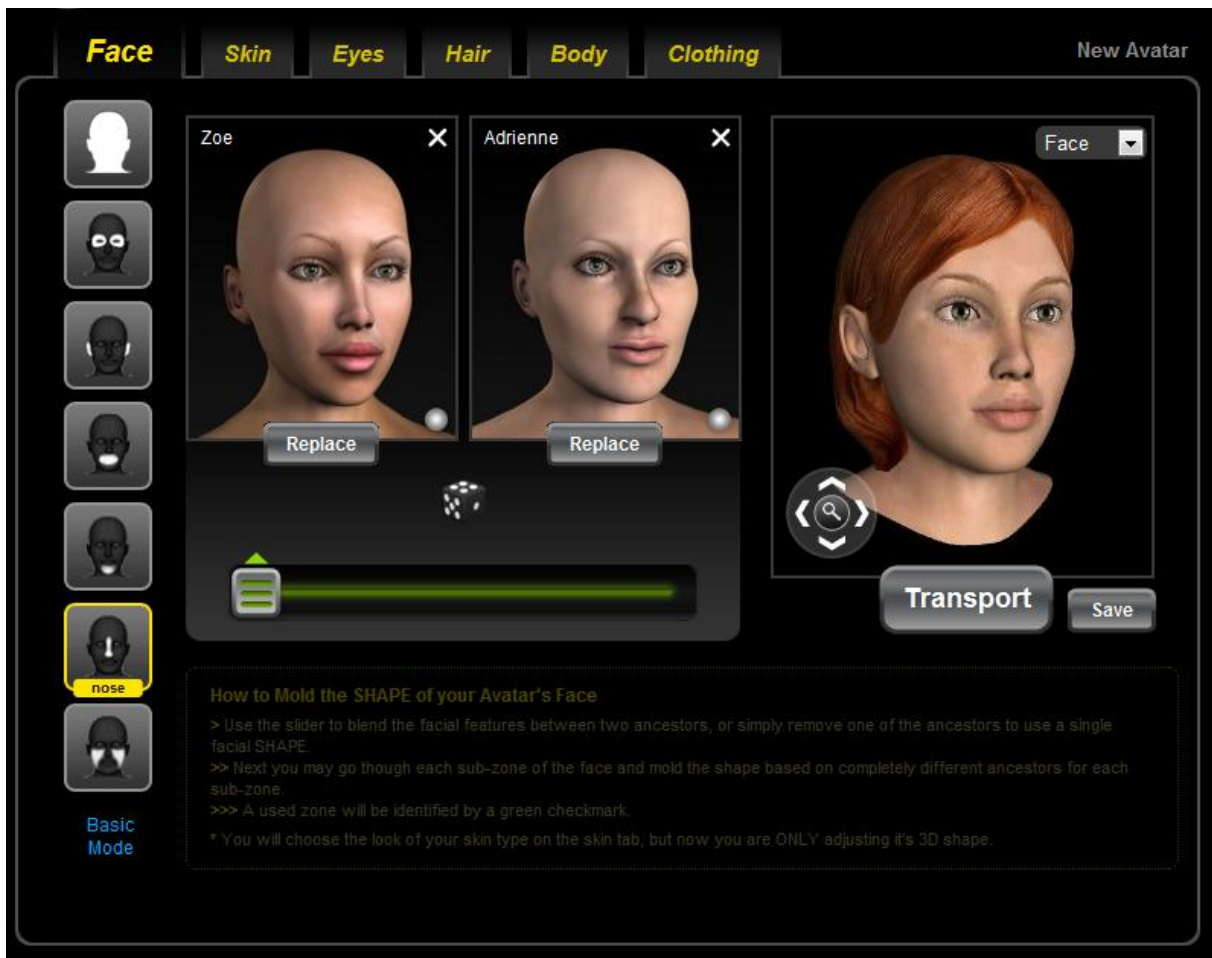


Abbildung 60 – Gestaltungsinterface von Evolver. Basierend auf zwei Vorlagen kann mit einem stufenlosen Schieberegler die gewünschte Mischung für das eigene Gesicht gewählt werden. Quelle: Evolver Webseite⁶⁹

Emotionale und akustische Gestaltung

Die emotionale Anpassung des Avatars ist nur durch vordefinierte Animationen möglich. Der Fokus des Editors liegt eindeutig auf der optischen Gestaltung. Ebenso ist es nicht möglich, dem Avatar einen Sprechtext oder eine selbstdefinierte Animation vorzugeben. Dazu müssen externe Programme herangezogen werden.

Darstellung

Die Animation laufen relativ flüssig und wirken realitätsecht mit Ausnahme der Finger, welche sich in manchen Situationen einfach in einer Standardhaltung befinden. Führt ein Avatar eine vorgegebene Animation aus, so wirkt dieser dabei recht authentisch. Beim Stöbern in der Galerie etwa können die Avatare niesen, traurig sein, schreien, winken, springen oder eine „Daumen hoch“ Animation ausführen.

Diese Animationen stehen in Rohform zur Verfügung, es wird lediglich die aktuelle Optik des Avatars dem Bewegungsmodell hinzugefügt. Trotz der guten Umsetzung ist noch ein deutlicher Unterschied zwischen Avatar und Mensch zu erkennen.

⁶⁹ <http://www.evolver.com>

Exportmöglichkeiten

Nach dem erfolgreichen Erstellen eines Avatars eröffnen sich dem Nutzer unzählige Möglichkeiten, seinen Avatar weiterzuverwenden. Basierend darauf, ob eine normale oder eine Pro-Lizenz vorliegt, stehen Möglichkeiten zur Integration in Netzwerke, Blogs und Videos zur Verfügung.

Die kostenlose Version erlaubt unter anderen folgenden Exportmöglichkeiten:

- Integration in Netzwerke (Facebook, Blogger, MySpace, ...) als Standbild, animiertes Bild oder Flashvideo mit wählbaren Animationen (Winken, tanzen, diverse Fußballeranimationen, typische Posing-Gesten, etc...)
- Integration in FriendsHangout, eine 3D Welt zum Chatten oder Spielen
- Verkaufen der selbstgestalteten Kleidung im Marketplace von Evolver
- Ausstellen des Avatars im Marketplace als Vorlage und Inspiration für andere Nutzer
- Erstellen von kurzen Videoclips, in denen der Avatar vordefinierte Animationen ausführt

Mit einer kostenpflichtigen Pro-Lizenz stehen zahllose weitere Optionen zur Verfügung.

- Integration in diverse Modellierungsprogramme (Maya, 3D Studio Max, etc...) zum Weiterbearbeiten des Avatars
- Professionelles Rendering des Avatars für eine bessere optische Darstellung
- Schnittstelle zu Partnerfirmen (FaceFx, Torque, WildPockets, etc...) und Integration in deren Netzwerke mit speziellen Zielgruppen
- Bestellung einer hochauflösenden Version des Avatars als Poster in verschiedenen Haltungen

Insgesamt stehen aktuell 27 Exportmöglichkeiten mit verschiedenen Einstellungen für viele Zielgruppen zur Verfügung. Alle erstellten Avatare werden im Benutzerkonto abgespeichert und stehen für spätere Verwendung zur Verfügung.

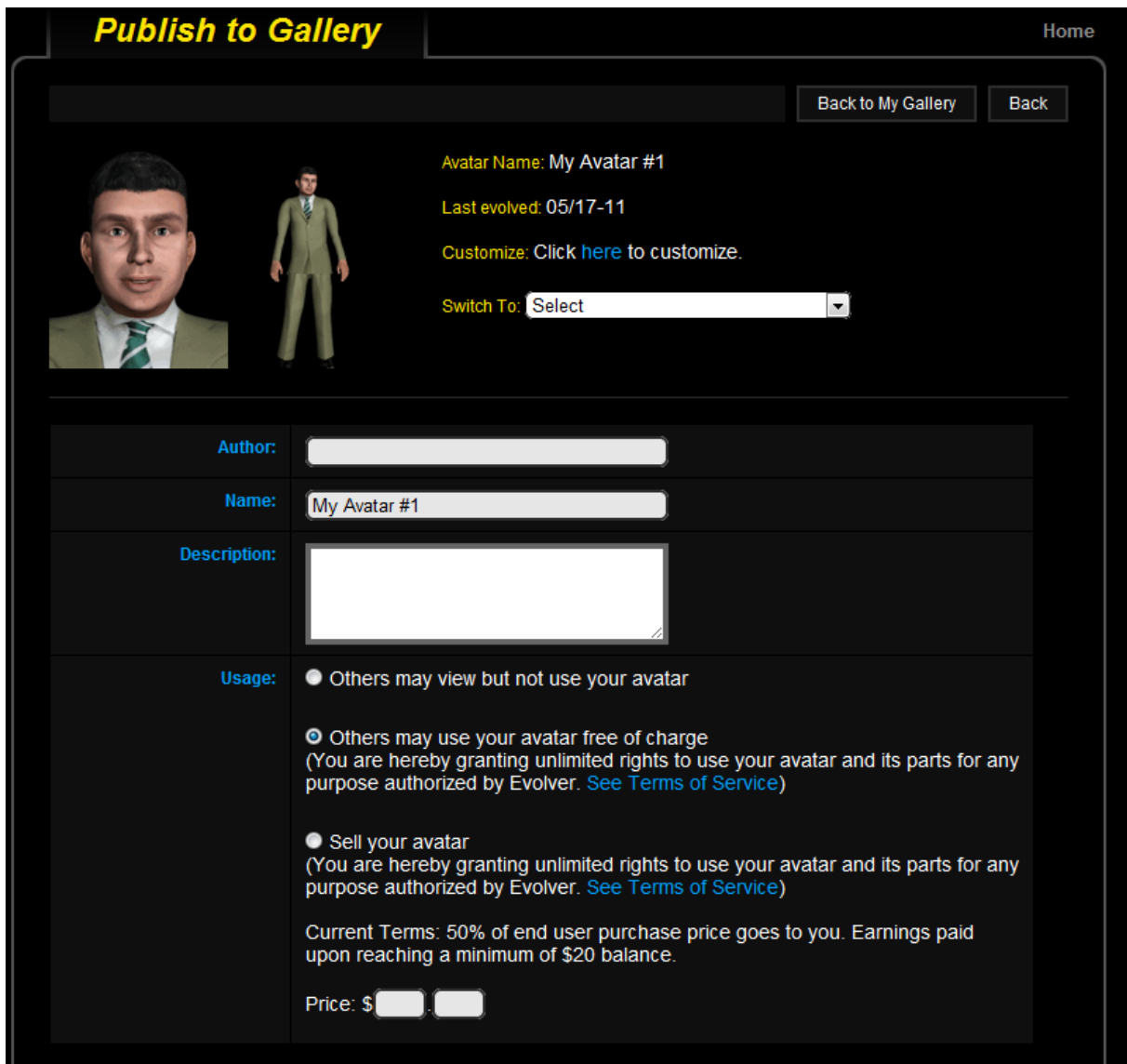


Abbildung 61 – Exportinterface von Evolver. Evolver bietet zahlreiche Exportmöglichkeiten, darunter die Möglichkeit, den Avatar in die hauseigene Galerie einzubinden und ihn den anderen Nutzern kostenlos oder kostenpflichtig zur Verfügung zu stellen. Quelle: Evolver Webseite⁷⁰

Wiederverwendbarkeit

Sofern ein Nutzerkonto angelegt wurde, können alle erstellten Avatare gespeichert und zu einem späteren Zeitpunkt weiterverarbeitet werden. Ein Menüpunkt führt den Benutzer zu einer Verwaltungsmöglichkeit aller von ihm erstellten Avatare.

Benutzerinteraktion

In diesem Punkt enttäuscht das Framework. Der Benutzer kann dem Avatar lediglich vordefinierte Animationen aus einer Liste beim Export vorgeben. Eigene Animationen sind nicht möglich.

⁷⁰ <http://www.evolver.com>

Benutzbarkeit

Aufgrund der Größe und der vielen Möglichkeiten des Editors ist eine Einarbeitungszeit erforderlich. Die Menüführung ist konsistent und bereits kürzeste Verzögerungen im Ladeprozess werden dem Benutzer signalisiert.

Sonstiges

Mit Hilfe des Marketplaces besteht die Möglichkeit, Geld durch Verkaufen des Avatars oder Teilen davon zu verdienen. Wer den Editor in eine eigene Homepage einbinden möchte, kann dies kostenpflichtig tun. Für Studenten steht eine ermäßigte Version mit vollem Funktionsumfang zur Verfügung.

Fazit

Die Vorteile des Frameworks sind eindeutig der sehr gute Editor zur optischen Gestaltung, die unzähligen Vorlagen, die es erlauben, eine Vielzahl an unterschiedlichen Avataren zu kreieren und die Exportmöglichkeiten. Leider fehlen Möglichkeiten zur Erstellung oder Integration von eigenen Animationen sowie zum Hinzufügen von akustischen Effekten.

5.3.4 XBox



Überblick

Das Xbox Live Netzwerk von Microsoft⁷¹ bietet die Möglichkeit, anstelle von einfachen Bildern als Benutzerrepräsentation selbstgestaltete animierte Avatare zu verwenden. Diese Avatare werden einem Benutzerkonto hinzugefügt und ermöglichen so eine Identifikation des Benutzers abseits von Spitznamen oder Nummern.

Registrierung

Das Anlegen eines Benutzerkontos ist nicht erforderlich, jedoch gibt es dann keine Möglichkeit, den erstellten Avatar in irgendeiner Weise abzuspeichern, um ihn später weiterbearbeiten zu können.

Voraussetzungen

Der Editor ist nur online abrufbar und setzt ein installiertes Silverlight Plug-In von Microsoft voraus. Ansonsten eignet sich jeder Browser, für den dieses Plug-In angeboten wird. Eine permanente Internetverbindung ist nötig, da die Daten nicht lokal bearbeitet und gespeichert werden.

Vorlagen

Es existieren viele Vorlagen für alle optischen und emotionalen Merkmale. Bereits zu Beginn wählt man einen Avatar aus und bearbeitet diesen dann entsprechend seinen Vorstellungen.

⁷¹ <http://live.xbox.com/de-DE/AvatarEditor>

Es ist nicht möglich, einen „leeren“ Avatar als Ausgangspunkt zu wählen. Aufgrund der Vorlagenvielfalt kann man jedoch das Endresultat gut an seine Vorstellungen anpassen.

Optische Gestaltung

Die optische Gestaltung unterteilt sich in zwei Abschnitte. Zum einen der Gesichtsbereich (Bereich „Aussehen“), in welchem folgende Elemente wie Gesichtsm Merkmale, Gesichtsbehaarung, Kinn, Mund, Nase, Ohren, Lidschatten, Augen, Augenbrauen, Frisuren oder Hautfarbe angepasst werden können.

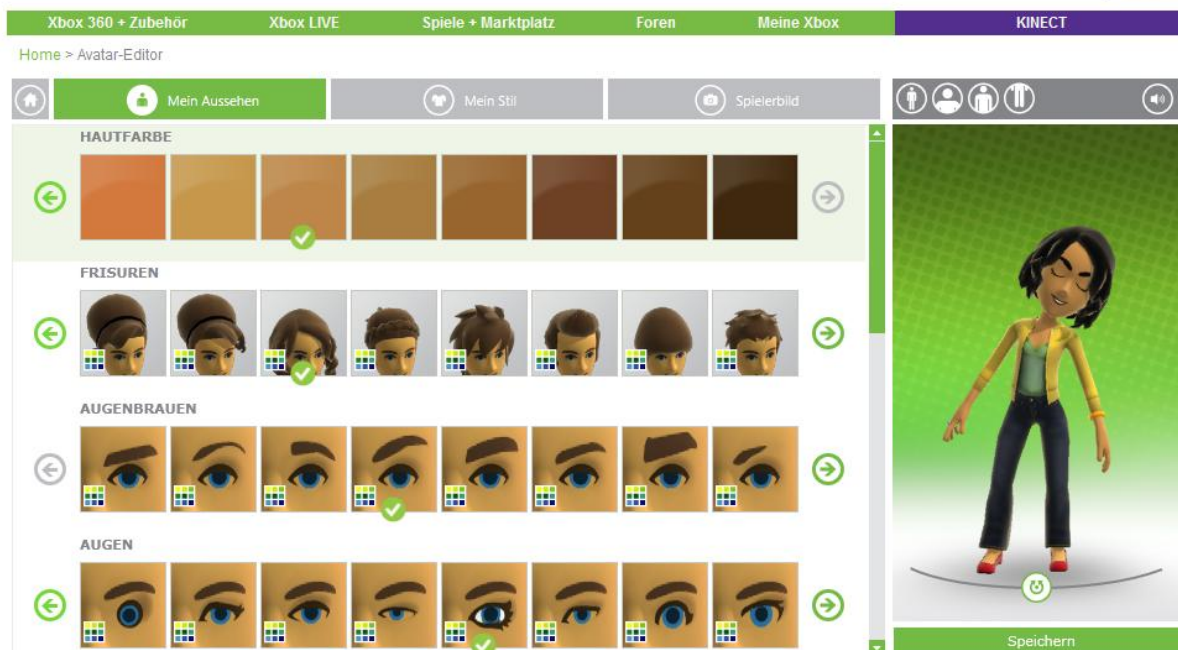


Abbildung 62 – Gestaltungsdialog (Reiter „Aussehen“) des Xbox Avatar Editors. Dieser bietet zahlreiche Vorlagen und viele Gestaltungsmöglichkeiten für den Benutzer. Quelle: Xbox Webseite⁷²

Im Ganzkörperabschnitt (Bereich „Stil“) können dem Avatar Elemente der Kategorien Anzüge, Kopfbedeckungen, Oberteile, Beinkleidung, Schuhe, Tragbares, Brillen, Ohrringe, Armbänder, Ringe oder Handschuhe gegeben werden.

⁷² <http://live.xbox.com/de-DE/AvatarEditor>

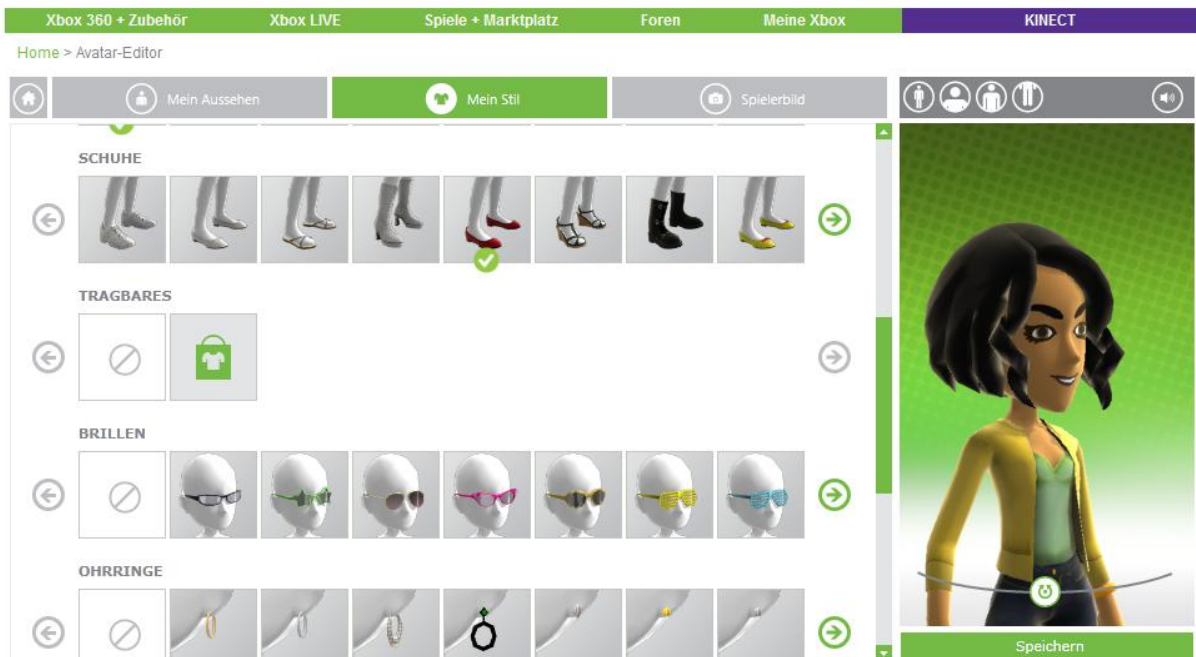


Abbildung 63 – Gestaltungsdialog (Reiter „Stil“) des Xbox Avatar Editors. Dieser lässt den Benutzer das Aussehen des Avatars bestimmen. Quelle: Xbox Webseite⁷³

Die vorgenommenen Änderungen werden sofort übernommen. Der Avatar kann während der Bearbeitung gedreht und seitlich betrachtet werden. Klickt man mit der Maus in das Vorschaufenster, vollführt der Avatar eine zufällige Animation wie hüpfen oder lachen.

Emotionale und akustische Gestaltung

Jeder Avatar kann in drei Animationsposen gespeichert werden. Zur Auswahl stehen zwölf Animationen, die verschiedene Gefühlszustände von Freude bis Enttäuschung transportieren sollen. Aus diesen Animationen kann ein Standbild an einer beliebigen Stelle festgehalten werden. Es können keine eigenen Animationen definiert werden und im Bereich der akustischen Gestaltungen bietet der Editor überhaupt keine Möglichkeit, den Avatar seinen Vorstellungen gemäß anzupassen.

⁷³ <http://live.xbox.com/de-DE/AvatarEditor>

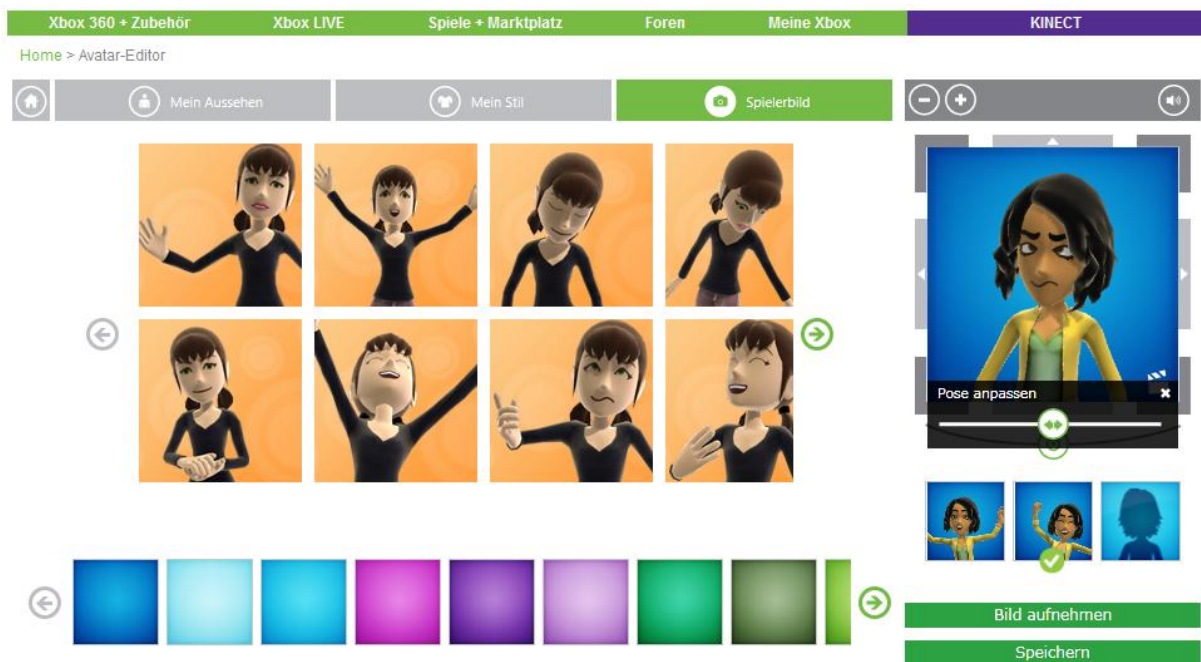


Abbildung 64 – Standbildauswahl. Der Editor bietet die Möglichkeit, den Avatar in drei Posen bestimmter Animationen als Standbild abzuspeichern. Die Animationen sind vordefiniert, der Zeitpunkt der Animation für das Bild kann frei gewählt werden. Quelle: Xbox Webseite⁷⁴

Darstellung

Die Avatare werden in einem comicähnlichen Stil als spielende Kinder dargestellt. Änderungen an sichtbaren Eigenschaften sind trotz der etwas schlechten Darstellungsqualität, die zum Großteil dem Comicstil geschuldet ist, sichtbar. Sehr gut sind die Bewegungsabläufe, wenn eine Animation ausgeführt wird. Der Avatar wirkt dabei authentisch und die Emotionen sind deutlich erkenn- und interpretierbar.

Exportmöglichkeiten

Die erstellten Avatare sind nur im Xbox Live Netzwerk und damit verknüpften Seiten nutzbar. Es besteht von Seiten des Editors keine Möglichkeit, seinen Avatar in anderer Weise zu nutzen.

Wiederverwendbarkeit

Ein erstellter Avatar wird im Profil gespeichert und kann zu einem späteren Zeitpunkt verändert werden. Es kann jedoch immer nur ein Avatar pro Profil gespeichert werden.

Benutzerinteraktion

Der Benutzer kann durch Klicken in das Vorschauenfenster während der Gestaltungsphase den Avatar zu einer zufälligen Animation anregen. Weitere Einflussnahme auf den Vorgang ist nicht möglich.

⁷⁴ <http://live.xbox.com/de-DE/AvatarEditor>

Benutzbarkeit

Der Editor funktioniert so, wie man es von typischen Webanwendungen gewohnt ist. Etwas störend ist die Avataorauswahl zu Beginn, da immer nur 5 Modelle zur Verfügung stehen und das Generieren neuer Modelle bis zu 10 Sekunden dauern kann. Gefallen einem die Vorschläge nicht, kann es längere Zeit dauern, bis man ein passendes Modell gefunden hat. Während der Gestaltungsphase dauert es je nach Internetverbindung einige Sekunden, bis die Vorlagen geladen sind. Danach ist verzögerungsfreies Arbeiten möglich. Hat man bereits einen Avatar seinem Profil hinzugefügt und möchte diesen ersetzen, stürzt der Editor in einigen Browsern reproduzierbar ab.

Fazit

Das Framework hat seine Stärken definitiv darin, dass man bei der Gestaltung des Avatars viele Optionen, Vorlagen und Einzelheiten bedacht hat. Daneben wirken die Animationen äußerst authentisch. Leider sind die Avatare nur für das Xbox Netzwerk als Benutzerrepräsentation zugänglich und eine akustische Untermalung ist nicht möglich.

5.3.5 Yahoo Avatare



Überblick

Bei Yahoo! Avatare handelt sich um ein Onlineframework, wofür zwingend ein Nutzerkonto erforderlich ist. In Abbildung 65 ist die Bearbeitungsansicht dieses Frameworks dargestellt. Nachfolgend wird Yahoo! Avatare detailliert vorgestellt.

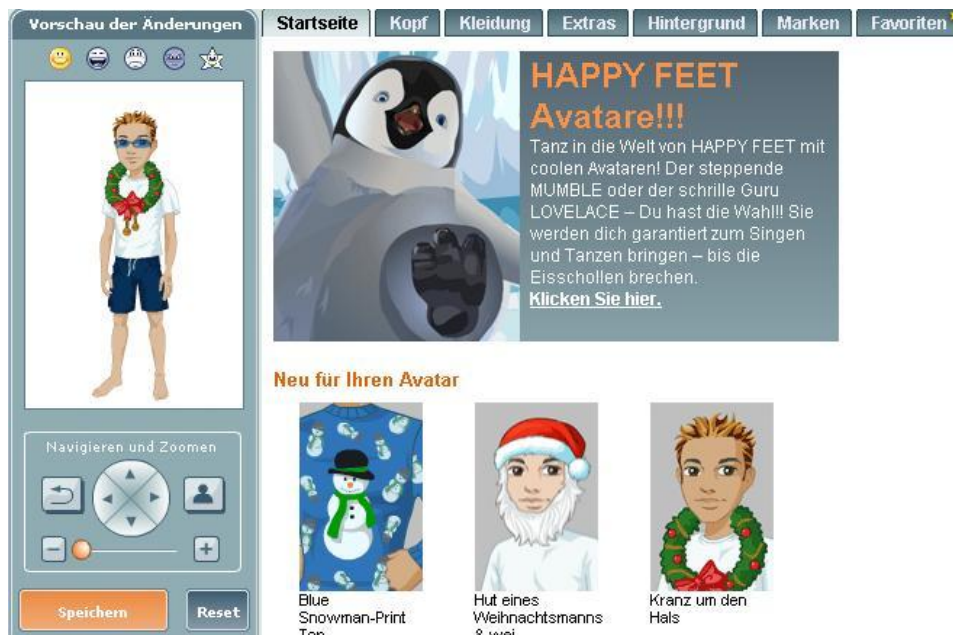


Abbildung 65 – Bearbeitungsansicht von Yahoo! Avatare. Quelle: Yahoo Webseite⁷⁵

⁷⁵ <http://de.avatars.yahoo.com>

Registrierung

Bevor man einen Avatar bei Yahoo! Avatare erstellt, muss man zuerst ein Konto anlegen. Der Avatar wird mit dem Konto und den angebotenen Diensten verknüpft. Deshalb ist eine Registrierung für Yahoo! Avatare notwendig.

Voraussetzungen

Um das Framework nutzen zu können, benötigt man ein Yahoo! Konto, einen aktuellen Browser (Internet Explorer ab Version 5.5 oder neuer) sowie den Adobe Flash Player.

Vorlagen

Es besteht eine geringe Auswahl an Avatarvorlagen. Vorlagen können nach Geschlecht getrennt ausgewählt werden. Die Verwendung eigener Bilder ist nicht möglich.

Optische Gestaltung

Die optischen Gestaltungsmöglichkeiten sind sehr umfangreich. Der Benutzer kann dabei Vorlagen aus verschiedenen Kategorien wählen: Aussehen, Kleidung, Extras, Hintergrund, und Marken. Man kann bei diesen Kategorien verschiedene Elemente auswählen und seinen eigenen Stil kreieren. Es können komplette Outfits gewählt oder eine eigene Kombination aus diversen Oberteilen und Beinbekleidungen zusammengestellt werden. Diverse Brillen und Schmuck können angelegt werden. Für den Hintergrund gibt es zahlreiche Auswahlmöglichkeiten.

Akustische Gestaltung

Es gibt keine Möglichkeit für akustische Gestaltungen.

Exportmöglichkeiten

Man kann seinen Avatar exportieren und ihn auf Webseiten und Blogs verwenden. Wenn man diesen auf der Yahoo! Seite aktualisiert, dann wird seine Avatar auch automatisch auf seine Profilseite aktualisiert. Auf der Seite zum Exportieren des Avatars verwendet man die Pulldown-Menüs, um auszuwählen, wie man seinen Avatar exportieren möchte. Im Menü „Größe“ kann man entscheiden, ob man den ganzen Avatar oder nur das Gesicht exportieren möchte. Für beide Größen kann man das .jpg oder .png Format auswählen. Im Menü „Generieren“ kann man entscheiden, welchen Code man für die Darstellung des Avatars nehmen möchte. Falls man „HTML“ auswählt, erhält man den kompletten Code. Falls man jedoch „URL“ auswählt, bekommt man nur den Bildverweis.

Wiederverwendbarkeit

Durch das Anlegen eines Accounts auf der Yahoo Webseite kann man sich in sein persönliches Profil einloggen und seine erstellten Avatare speichern (zwölf Avatare und 32 beliebige Sachen, wie beispielsweise Kleidung und Sportausrüstung), betrachten, ändern oder löschen.

Benutzerinteraktion

Es gibt keine Möglichkeit im Bereich der Benutzerinteraktion.

Emotionale Gestaltung

Die Avatare können fünf verschiedene emotionale Zustände (glücklich, aufgeregt, traurig, böse und spezial) einnehmen.

Darstellung

Die Darstellung ist comicartig. Es gibt nur bei Ausdrucken einige Änderungen. Außerdem gibt es fast keine Animationen.

Benutzbarkeit

Die Benutzbarkeit ist gut, allen Einstellungen sind Kategorien zugeordnet und sie sind bequem per Mausclick erreichbar.

Sonstiges

Yahoo! Avatar ist kostenlos und benutzt 2D Technik auf Flashbasis.

Fazit

Das Framework eignet sich lediglich dafür, Avatare für das Yahoo Netzwerk zu erstellen. Positiv aufgefallen sind die optischen Gestaltungsmöglichkeiten sowie die große Anzahl an Avataren, welche im Profil gespeichert werden kann. Negativ ins Gewicht fällt der Registrierungszwang und die wie erwähnt mangelnde Verbreitungsmöglichkeit der Avatare.

5.4 Zusammenfassung

Zwei der fünf Testkandidaten (Yahoo! Avatare und Xbox Live Avatare) sind lediglich dafür ausgelegt, Avatare als Profil- und Benutzeridentifikator in ihren Netzwerken zu verwenden. Es gibt keine Möglichkeit, aus diesen Editoren erweiterten Nutzen zu ziehen etwa in der Form, dass der Avatar in ein Softwareprojekt integriert werden könnte. Die Editoren an sich sind gut und man kann sich damit durchaus ein personalisiertes Ebenbild von sich erschaffen, aber da sie von außen nicht ansprechbar sind und keine Schnittstelle nach außen bieten, sind sie für unsere Zwecke nicht geeignet.

Die nächsten beiden Vertreter (Voki und Evolver) bieten umfangreiche Gestaltungsmöglichkeiten für das Erstellen der Avatare in vielerlei Hinsicht. Voki hat den Vorteil bei der möglichen Integration von akustischen Komponenten, während Evolver mit einer größeren Detailliertheit bei der optischen Gestaltung aufwartet. Beide bieten beim Export zahlreiche Schnittstellen zu sozialen Netzwerken, Blogs, Communities und Professionals an, sodass damit eine eventuell weiterführende Arbeit verrichtet werden könnte.

Auch wenn es im Punkt Gestaltung nicht ganz mit den besten mithalten kann, so zeichnet sich die Implementierung My Avatar Editor des quelloffenen Frameworks Avatarcore dadurch aus, dass es von außen über APIs angesprochen werden kann. Dieses Framework würde sich dazu eignen, in einen Entwicklungsprozess eines Projektes integriert zu werden, in welchem ein Avatar erstellt werden soll. Leider sind die bereits erwähnten Gestaltungsmöglichkeiten nicht sehr gut. Es finden sich keine Vorlagen in dem Framework.

Das ideale Framework gibt es nicht, aber eine Kombination aus My Avatar Editor, Evolver und Voki mit erweiterten API Schnittstellen wäre eine gute Basis für weitere Entwicklungen.

6 Zusammenfassung und Ausblick

Das Thema dieser Arbeit sind Avatar Frameworks. Zu Beginn wurde näher erläutert was der Begriff Avatar bedeutet und es wurde festgestellt, dass es sich bei Avataren um ein stark interdisziplinäres Forschungsgebiet handelt. Die Disziplinen Soziologie, Visualisierung, Mensch-Rechner-Interaktion, Informatik und Marketing forschen im Bereich der Avatare.

Daraufhin wurde im zweiten Kapitel vorgestellt wie Avatare erstellt werden können. Den Anfang machten einfache Erstellungsmethoden, wie sie täglich von vielen Benutzern verwendet werden, etwa die Auswahl eines Bilds in einem Internetforum. Daran schlossen sich professionelle Methoden aus der Film- und Fernsehindustrie an. Zahlreiche Erstellungsmethoden aus der Avatarforschung folgten darauf. Deren Zielsetzung liegt darin neue Avatarerstellungsmethoden zu erforschen, zu verbessern und ihre Tauglichkeit für die Praxis zu überprüfen. Auf Basis des Wasserfallmodells wurde anschließend die Entwicklung einer Avataranwendung gezeigt und festgestellt, dass sich für solche Art von Anwendungen noch keine Standards entwickelt haben. Das Kapitel wurde durch zahlreiche bebilderte Beispiele für Avatare in den Medien abgeschlossen.

Das dritte Kapitel untersucht die Beziehung von Menschen und Avataren. Kulturelle Unterschiede und Akzeptanzprobleme von Avataren wurden näher beleuchtet. Der Abschnitt über kulturelle Unterschiede kam zu dem Schluss, dass es sichtbare Unterschiede in der Darstellung von Emotionen und beim Aussehen zwischen westlichen und asiatischen Avataren gibt. Dies ist zurückzuführen darauf, dass Europäer im realen Leben mehr Freiheiten haben als Asiaten und sie daher nicht so ein großes Bedürfnis haben, in der virtuellen Welt ihre Persönlichkeit auszuleben. Zu den Akzeptanzproblemen der Avatare trägt maßgeblich der Uncanny Valley Effekt bei, der besagt, dass Avatare ihren Betrachtern umso unheimlicher werden, desto menschenähnlicher sie gestaltet sind. Abschließend wurde festgestellt, dass für die Erhöhung der Akzeptanz von Avataren daher ein nicht zu realistisches äußeres Erscheinungsbild, eine robuste Spracheingabe und für den Benutzer brauchbare Funktionen vonnöten sind.

Im vierten Kapitel wurden Interaktionskonzepte vorgestellt. Zahlreiche Interaktionstechnologien mit Avataren wurden untersucht, darunter Tastatur, Maus, Gamepad, Gestensteuerung durch Microsoft Kinect, Spracherkennung und auch die Touch-Technologie. Es wurde festgestellt, dass es zwar Neuentwicklungen gibt, die in bestimmten Bereichen auch verwendet werden, jedoch vor allem im Privatbereich nach wie vor Maus und Tastatur zur Interaktion verwendet werden. Daran schloss sich ein Abschnitt über die Anpassung an das Benutzerverhalten, in dem festgestellt wurde, dass Avatare von den Benutzern am besten angenommen werden, wenn diese keine Grundregeln der Gesprächsführung verletzen. Anschließend wurden Spracherkennungsmethoden von Avataren untersucht und dabei festgestellt, dass es nach wie vor Probleme dabei gibt, in jüngster Zeit jedoch große Fortschritte in der Erkennungsqualität erzielt wurden. Im darauffolgenden Abschnitt über Mimik und Gestik wurde deutlich, dass es wichtig ist, dass ein Avatar seine Aussagen mit entsprechender Mimik und Gestik unterstreicht. Im

nachfolgenden Abschnitt über Körpersteuerung wurden die Spielekonsolen Microsoft Xbox, Sony Playstation und Nintendo Wii vorgestellt und ihre Funktionsweise untersucht. Daran schloss sich ein Abschnitt über die Visualisierung von Emotionen an, der verschiedene Methoden dazu vorstellt. Das Kapitel wurde durch eine Untersuchung der Grenzen der Interaktionsmöglichkeiten mit Avataren abgeschlossen. Dabei wurde festgestellt, dass eine solche Grenze im Bereich der Datenbasis von Avataren liegt, beispielsweise die Beantwortung einer nicht zuordenbaren Frage durch eine Standardantwort.

Kapitel fünf schließlich befasst sich mit der Evaluation von bestehenden Frameworks zur Avatarerstellung. Fünf verschiedene Frameworks wurden einer gründlichen Untersuchung unterzogen. Dabei stellte sich heraus, dass zwei Frameworks (Yahoo Avatare und Xbox Live Avatare) sich nur dazu eignen, um Avatare für das jeweilige Netzwerk zu erstellen. Zwei andere Frameworks (Voki und Evolver) bieten umfangreiche Gestaltungsmöglichkeiten und zahlreiche Schnittstellen zu vielen verschiedenen sozialen Netzwerken und anderen Internetdiensten, in die sie eingebunden werden können. Hervorzuheben ist die Ansprechbarkeit von My Avatar Editor über verschiedene APIs. Daher wäre eine Kombination aus My Avatar Editor, Evolver und Voki mit erweiterter API eine gute Grundlage für weitere eigene Entwicklungen.

Abschließend kann festgestellt werden, dass auf dem interdisziplinären Gebiet der Avatare sehr viel Forschung betrieben wird und ihre Einsatzmöglichkeiten bei Weitem noch nicht ausgeschöpft sind. Deutlich sichtbar ist nach wie vor, dass Avatare meistens auf Auskunftsfunktionen beschränkt sind. Dabei nähert sich der im Abschnitt 3.2 beschriebene Fahrkartenautomat durch das Fortschreiten von Forschung und Entwicklung immer weiter seiner Realisierbarkeit. Die Zukunft der Avatare ist nicht eindeutig vorhersagbar. Bis wir jedoch alle unseren eigenen ferngesteuerten Avatar besitzen, wie in James Camerons „Avatar – Aufbruch nach Pandora“ oder Jonathan Mostows „Surrogates – Mein zweites Ich“, wird wohl noch sehr viel Zeit vergehen. Viel spannender ist jedoch die Frage wie wir Menschen dann mit dieser Situation umgehen werden.

Literaturverzeichnis

- [1] Adrian Albin-Clark and Toby Howard. Automatically generating virtual humans using evolutionary algorithms. In Wen Tang and John P. Collomosse, editors, *TPCG*, pages 61–64. Eurographics Association, 2009.
- [2] Elisabeth André, Klaus Dorfmüller-Ulhaas, and Thomas Rist. Embodied conversational characters: Wandering between the digital and the physical world (digitale weltenwanderer: Synthetische charaktere zwischen realität und virtualität). *it - Information Technology*, 46(6):332–340, 2004.
- [3] M. A. Anusuya and S. K. Katti. Speech recognition by machine, A review. *International Journal of Computer Science and Information Security*, 6(3):181–205, January 13 2010.
- [4] Francesca Barrientos. Continuous control of avatar gesture. In *ACM Multimedia Workshops*, pages 5–8, 2000.
- [5] Barrientos, Francesca and Canny, John. Cursive: a novel interaction technique for controlling expressive avatar gesture. In *Proceedings of the ACM Symposium on User Interface Software and Technology*, Papers: Expressive user interfaces, pages 151–152, 2001.
- [6] Dmitri Bitouk and Shree K. Nayar. Creating a speech enabled avatar from a single photograph. In *VR*, pages 107–110. IEEE, 2008.
- [7] J. Carranza, C. Theobalt, M Magnor, and H.-P. Seidel. Free-viewpoint video of human actors. *ACM Trans. on Computer Graphics*, 22(3), July 2003.
- [8] J. H. Cho, H. S. Kim, and K. H. Lee. A 3D avatar modeling of realworld objects using a depth camera. In *3DTV09*, pages 1–4, 2009.
- [9] Bas R. Steunebrink, Mehdi Dastani and John-Jules Ch. Meyer. The OCC Model revisited. http://www.idsia.ch/~steunebrink/Publications/KI09_OCC_revisited.pdf
- [10] Nikolas Dörfler. Hidden Markov Models. http://campar.in.tum.de/twiki/pub/Far/MachineLearningWiSe2003/doerfler_ausarbeitung.pdf
- [11] G. L. Filho and Tom J. Moir. From science fiction to science fact: a smart-house interface using speech technology and a photo-realistic avatar. *IJCAT*, 39(1/2/3):32–39, 2010.
- [12] Leo Grady, Thomas Schiwietz, Shmuel Aharon, and Rüdiger Westermann. Random walks for interactive alpha-matting. In *IN VIIP*, pages 423–429. ACTA Press, 2005.
- [13] Marion Gröne. Gesichtsanimation (facial animation). Seminar Embodied Conversational Agents, 28.06.2001.
- [14] Tomoko Koda, Toru Ishida, Matthias Rehm, and Elisabeth André. Avatar culture: cross-cultural evaluations of avatar facial expressions. *AI Soc*, 24(3):237–250, 2009.

- [15] S. Kshirsagar, A. Guye-Vuillème, K. Kamyab, N. Magnenat-Thalmann, D. Thalmann, and E. Mamdani. Avatar markup language. In S. Müller and W. Stürzlinger, editors, *Eighth Eurographics Workshop on Virtual Environments*, pages 169–177, Barcelona, Spain, 2002. Eurographics Association.
- [16] W. H. Leung, B. L. Tseng, Z. Y. Shae, and F. Hendriks. Realistic video avatar. In *IEEE International Conference on Multimedia and Exposition*, pages 631–634, 2000.
- [17] Christian Lindner. *Avatare - Digitale Sprecher für Business und Marketing*. Springer Verlag, 2003.
- [18] Zhen Liu. Computational emotion model for virtual characters. May 01 2008.
- [19] Jochen Ludewig and Horst Lichter. *Software Engineering*. dpunkt.verlag, 2007.
- [20] Dominic W. Massaro, Ying Liu, Trevor H. Chen, and Charles Perfetti. A multilingual embodied conversational agent for tutoring speech and language learning. In *INTERSPEECH*. ISCA, 2006.
- [21] H. McGurk and J. W. MacDonald. Hearing lips and seeing voices. *Nature*, 264:746–748, 1976.
- [22] Shigeo Morishima and Tatsuo Yotsukura. Face-to-face communicative avatar driven by voice. In *ICIP (3)*, pages 11–15, 1999.
- [23] Yukiko I. Nakano and Matthias Rehm. Multimodal corpus analysis as a method for ensuring cultural usability of embodied conversational agents. In Masaaki Kurosu, editor, *HCI (10)*, volume 5619 of *Lecture Notes in Computer Science*, pages 521–530. Springer, 2009.
- [24] Kai Kajus Noack. Ein Leitfaden zum Motion Capturing. <http://media-it.blogspot.com/2007/07/motion-capturing-guide-ein-leitfaden.html>
- [25] M. Rehm and E. André. From chatterbots to natural interaction - face to face communication with embodied conversational agents. *IEICE Transactions on Information and Systems, Special Issue on Life-Like Agents and Communication*, 88-D(11):2445–2452, 2005.
- [26] M. Rehm and E. André. Informing the design of embodied conversational agents by analyzing multimodal politeness behaviors in human-human communication. In *AISB Symposium for Conversational Informatics*, 2005.
- [27] Matthias Rehm, Elisabeth André, and Michael Wissner. Gamble v2.0: social interactions with multiple users. In *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems, AAMAS '05*, pages 145–146, New York, NY, USA, 2005. ACM.
- [28] Hans-Georg Beyer, Eva Brucherseifer, Wilfried Jakob, Hartmut Pohlheim, Bernhard Sendhoff and Thanh Binh To. Evolutionäre Algorithmen – Begriffe und Definitionen. <http://ls11-www.cs.uni-dortmund.de/people/beyer/EA-terminologie/node9.html>
- [29] Jun'ichiro Seyama and Ruth S. Nagayama. The uncanny valley: Effect of realism on the impression of artificial human faces. *Presence*, 16(4):337–351, 2007.

- [30] Raimundo Sierra. Non Rigid Registration. <http://www.rsierra.com/DA/node10.html>
- [31] Peter Livingston Silsbee. *Computer lipreading for improved accuracy in automatic speech recognition*. PhD thesis, University of Texas at Austin, 1993.
- [32] Michael Spehr. Diktieren ist viel schneller als tippen. <http://www.faz.net/artikel/C31158/dragon-naturally-speaking-11-diktieren-ist-viel-schneller-als-tippen-30001548.html>
- [33] Toru Takahashi, Christoph Bartneck, Yasuhiro Katagiri, and Noriko H. Arai. TelmeA - expressive avatars in asynchronous communications. *Int. J. Hum.-Comput. Stud*, 62(2):193–209, 2005.
- [34] Christian Teutsch. *Model-based Analysis and Evaluation of Point Sets from Optical 3D Laser Scanners*, volume 1. Shaker Verlag, 2007.
- [35] Wikipedia. Avatar. [http://de.wikipedia.org/wiki/Avatar_\(Internet\)](http://de.wikipedia.org/wiki/Avatar_(Internet))
- [36] Guo Yan and Haihong Zhang. Avatar control by facial expressions. In *FUZZ-IEEE*, pages 1351–1354, 2001.

Abbildungsverzeichnis

Abbildung 1 – Interdisziplinarität von Avataren.	5
Abbildung 2 – Avatarerstellung durch Bildauswahl in einem Internetforum.	7
Abbildung 3 – Avatarerstellung bei World of Warcraft.	8
Abbildung 4 – Motion Capture Verfahren.	9
Abbildung 5 – Synchronisierte Erfassung eines Menschen durch sieben Kameras.	10
Abbildung 6 – Ablauf des Verfahrens von Cho et al.	11
Abbildung 7 – Vergleich verschiedener Mattingmethoden.	11
Abbildung 8 – Ablauf des Verfahrens von Bitouk und Nayar [6].	12
Abbildung 9 – Cylindrical Texture Map. Quelle: Leung et al. [16]	13
Abbildung 10 – Ablauf des Verfahrens von Leung et al. [16].	13
Abbildung 11 – Ablauf des Verfahrens von Zollhöfer et al.	15
Abbildung 12 – Aufbau zur Avatarerstellung per Microsoft Kinect Konsole nach Zollhöfer et al.	16
Abbildung 13 – Ablauf eines Softwareentwicklungsprozesses nach dem Wasserfallmodell.	17
Abbildung 14 – Phasenmodell zur Entwicklung eines Avatars nach Lindner [17]	18
Abbildung 15 – Entwicklungsprozess eines Avatars.	20
Abbildung 16 – Microsoft Kinect: „Der Körper ist die Konsole“.	21
Abbildung 17 – Sensoren der Microsoft Kinect.	21
Abbildung 18 – Nintendo Wii.	22
Abbildung 19 – Nintendo Wii.	23
Abbildung 20 – Playstation Eye Kamera.	24
Abbildung 21 – Bildschirmfoto von Diablo I.	24
Abbildung 22 - Ubisoft: Your Shape Fitness Evolved.	26
Abbildung 23 – Marie, die virtuelle Beraterin von Ebay.	26
Abbildung 24 – QQChat.	27
Abbildung 25 – QQShow Portal des QQ Messengers.	28
Abbildung 26 – Robert T-Online, der Avatar der deutschen Telekom.	28
Abbildung 27 – Avatarerstellung in SecondLife.	29
Abbildung 28 – Der Startort von SecondLife.	30
Abbildung 29 – Szene aus Audition Dance Battle.	30
Abbildung 30 - Online Moderator von ModSite.	31
Abbildung 31 – Amerikanische Gefühlskarikaturen.	32
Abbildung 32 – Japanische Gefühlskarikaturen.	33
Abbildung 33 – Vergleich der Karikaturen.	33
Abbildung 34 – Avatar aus QQShow.	34
Abbildung 35 – Karl Klammer, die Büroklammer aus Microsoft Office.	35
Abbildung 36 – Uncanny Valley Effekt.	36
Abbildung 37 – Szene aus dem Click and Point Adventure Spiel „The Secret of Monkey Island“.	39
Abbildung 38 - Maus und Tastatur.	40
Abbildung 39 – Gamepad.	41
Abbildung 40 – Kinect für XBox360.	42
Abbildung 41 – Avatar der Firma Luxoft.	43
Abbildung 42 – Szene aus dem Film „Iron Man“.	44
Abbildung 43 – Captain James T. Kirk mit seinem tragbaren, kabellosen Kommunikator.	45
Abbildung 44 – Holodeck.	46
Abbildung 45 – HCI Interaktionslevel.	47
Abbildung 46 – HCI Interaktionsformen.	48
Abbildung 47 – Interface des Systems von Moir und Filho [11] mit dem sprechenden Avatar.	51
Abbildung 48 - Microsoft Kinect Erfassungsmöglichkeiten.	55
Abbildung 49 - Microsoft Kinect Sensoren.	55

<i>Abbildung 50 – Darstellung der Light Coding Technik.</i>	56
<i>Abbildung 51 – Skelettbasiertes Körperverfolgungssystem.</i>	56
<i>Abbildung 52 – Sensorleiste der Wii.</i>	57
<i>Abbildung 53 – Die drei Teilsysteme der Playstation Move.</i>	58
<i>Abbildung 54 – Playstation Move Erfassung.</i>	59
<i>Abbildung 55 – Arbeitsbereich der Wii Sensorbar.</i>	61
<i>Abbildung 56 – Führer der Yaoti Webseite.</i>	62
<i>Abbildung 57 – Bearbeitungsansicht des Voki-Frameworks.</i>	66
<i>Abbildung 58 – Bearbeitungsansicht von My Avatar Editor.</i>	69
<i>Abbildung 59 – Ausgangspunkt der Avatarerstellung bei Evolver.</i>	73
<i>Abbildung 60 – Gestaltungsinterface von Evolver.</i>	74
<i>Abbildung 61 – Exportinterface von Evolver.</i>	76
<i>Abbildung 62 – Gestaltungsdialog (Reiter „Aussehen“) des Xbox Avatar Editors.</i>	78
<i>Abbildung 63 – Gestaltungsdialog (Reiter „Stil“) des Xbox Avatar Editors.</i>	79
<i>Abbildung 64 – Standbildauswahl.</i>	80
<i>Abbildung 65 – Bearbeitungsansicht von Yahoo! Avatare.</i>	81

Erklärung

Hiermit versichern wir, diese Arbeit selbstständig verfasst und nur die angegebenen Quellen benutzt zu haben.

(Alexander Duschek)

(Philipp Schuster)

(Xi Tu)