TECHNISCHE UNIVERSITÄT WIEN
VIENNA UNIVERSITY OF TECHNOLOGY

MASTERARBEIT

Exploring Interaction in Collaborative Virtual 3D Environments by Performing Online Theater Plays

Ausgeführt am

Institut für Softwaretechnik und interaktive Systeme
der Technischen Universität Wien

unter Anleitung von

Ao.Univ.Prof. Mag. Dr. Dieter Merkl

durch

Markus Edlinger
Forstnerweg 11K
2630 Ternitz

Wien, am 11.12.2007

Zusammenfassung

Virtuelle Welten nehmen derzeit eine steigende Bedeutung als Interaktionsmedium im Marketingbereich und zwischen Privatanwendern ein. Obwohl bereits sehr viele Experimente und Projekte in diesem Bereich durchgeführt wurden, sind immer noch Potentiale vorhanden, die noch nicht genutzt werden.

Die vorliegende Masterarbeit befasst sich mit einem dieser Potentiale und erkundet die Möglichkeiten und Grenzen in Bezug auf Interaktionsformen innerhalb virtueller Welten. Als exemplarische Plattform für die Untersuchung wird Second Life herangezogen, weil dies eine der derzeit größten virtuellen Onlinewelten ist.

Die Grundlage der Arbeit sind praktische Aufführungen eines kurzen Theaterstücks in Second Life, wobei reale Benutzer mittels ihrer Avatare als Schauspieler fungieren. Dabei wird untersucht, welche Techniken und Verfahren eingesetzt werden können, um die Aktionen nach den Anforderungen eines Theaterstücks zu gestalten.

Als theoretischen Hintergrund liefert die Arbeit Überblicke über Avatare als Akteure für virtuelle Interaktion, allgemeine soziale und technologische Interaktionsformen in virtuellen Welten sowie Computeranimationstechniken, mit denen grundlegende Elemente für soziale Interaktionen geschaffen werden können.

Abstract

The importance of virtual worlds as a medium for interaction in different fields, such as marketing or private online networks is currently growing. Although much work has been conducted in this area, there are still some potential uses left which have not yet been examined.

This thesis deals with one aspect of virtual 3D worlds and points out the possibilities and limitations concerning virtual interaction. As an exemplary platform, Second Life, one of the currently most used online environments, is explored.

The basis of this work are performances of a short theater play in Second Life, whereat real users act via their avatars. By these means, the possible techniques for designing the avatar's interactions, which are needed for the performance, are examined.

The theoretical background is composed of chapters about avatars as actors of virtual interaction, general social and technological interaction forms in virtual worlds and computer animation techniques, which are the means for enabling interaction.

Acknowledgements

I would like to thank several people who supported me during this thesis and contributed to the project.

First of all, I owe special thanks to my supervisor, Dieter Merkl who promoted the thesis with competent feedback and indispensable ideas for the practical part. His encouragement was very helpful.

Furthermore, I want to thank the members of the iSpaces research group of EC3 for the inspiring meetings, their positive feedback and their creative suggestions concerning the theater project.

Finally, I want to express my particular gratefulness to Christoph Gluszko, Christian Schromm and Christian Traunwieser. Without their participation in the theater performances and the high amount of their precious time they spent rehearsing the play, the project would not have been possible.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

So there's an interaction there between social, structural and cultural. And if you want to fully capture the environment, you have to show how these things interact.

William J. Wilson

In many areas of everyday life interaction plays a crucial role. Be it in form of communication between people or in form of interactions with surrounding items, without these concepts, our society would not work in the current way. As the initial quotation by William J. Wilson, a professor at Harvard shows, a complete examination of the environment has to include interactions. This points out the importance of these relations and explains the basis for this thesis.

1.1 Motivation

Interaction in general is defined as "*the combined or reciprocal action of two or more things that have an effect on each other and work together*" (Encarta, 2007). Thereby, the "things" which interact with each other can be of different types. If they are people, this form is called social interaction, if they are objects, one aspect of interaction takes place in form of physical forces.

These forms of interaction have been examined by a variety of researchers in the field of physics or social sciences (e.g. Argyle (1969) or Hogan (1985)), but a majority of this work is related to interactions in the real world. However, in the last years, accompanied with the technological advance of computer technology, another type of environments for interaction has been growing in importance: virtual worlds. These environments, as they are improving concerning their level of realism, include more and more features of real life. Interactions between users represented by virtual characters are just as possible as manipulating virtual objects. But interactions in virtual worlds do not only comprise those which are possible in reality. As virtual environments can be designed in a way that is not restricted to limitations of reality, such as physical laws, a multitude of additional interactions can be implemented. As an example, one could design a virtual world without gravity, so that users can move through it as if they were astronauts in outer space.

As can be seen, interactions in virtual environments offer a variety of interesting questions for research. The idea of this thesis is to explore which interactions are implemented in current virtual worlds and which limitations they have. Thereby, the focus lies on interactions between users in form of their virtual representation and interactions between users and the virtual environment as a whole.

This examination is conducted by means of virtual theater performances. The reason for this method is that a main aspect of theater in general is based on interactions between actors or between an actor and the setting. Of course, in real theater, the interactions between actors and the audience also play an important role. This, however, is beyond the scope of this thesis. Thus, a virtual theater performance can show in how far implemented interaction techniques work. As a platform for the examinations, Second Life, a currently very popular and powerful virtual world was chosen. In this environment, real users are performing via their virtual representations and in this way, short existing theater plays are put on stage. By these means, the actions of the actors can be evaluated, compared with the original play and conclusions can be drawn as far as virtual interactions in general are concerned.

The results of this exploration can be relevant for different scopes. Designers of new virtual worlds might consider the problems of interactions in currently existing platforms and introduce new concepts which facilitate social and technological interaction in these environments. Furthermore, enterprises might use some of the presented techniques for applications in virtual worlds. Especially in the area of collaborative work, the strengths and weaknesses of interactions are important. One example, which is topic of many research projects, is education in virtual environments (see e.g. Dwyer et al. (1995)). In this field, the interplay between the users and the learning facilities is vital.

1.2 Thesis overview

At the beginning of this thesis, interaction and its theoretical background will be described. Afterwards, the organization and performance of the virtual theater plays will be illustrated and finally, the findings of the examination will be given. In detail, the chapters of this thesis cover the following aspects:

Chapter 2 will give a definition and an overview of virtual worlds. In addition to this, it will contain classifications of different terms in the context of virtual environments. Thereafter, existing virtual worlds are compared and important features of them are further explained. As a particular example, some selected aspects of Second Life will be described in detail.

A description of avatars as representation of users in virtual environments can be found in chapter 3. After a short review of history, a classification of virtual characters will be given. Then, some features of avatars which are important for virtual social interaction, such as body language, will be described and their presence in Second Life examined.

Thereafter, chapter 4 will explain interaction techniques in detail. Different approaches for the classification of virtual interaction will be compared and the ideas of movement and object manipulation will be explained. As in the previous chapters, the implementation of these techniques in Second Life will be discussed.

Another crucial part of virtual interactions will be described in chapter 5, where an overview of computer animation will be given. Not only a list of animation principles, but also a detailed explanation of animation methods, such as keyframe animation and motion capturing will be included in this chapter.

Finally, chapter 6 and will comprise a detailed description and analysis of the virtual theater performances. After a survey of similar projects, the whole process from the selection of the plays to the final live performance will be illustrated. After that, this chapter will be concluded with a description of the results of the analysis. Thereby, a list of issues regarding virtual interaction in general will be given.

To summarize this thesis, chapter 7 will conclude the most relevant points and try to give a short outlook on future developments in the field of virtual theater performances and interaction.

Chapter 2

Virtual Environments

Virtual environments or virtual worlds play an important role for social relations via computer networks. Their uses include entertainment, teaching, marketing and many others which are currently explored by researchers. To get an impression about this topic, this chapter gives an overview of different virtual environments and their uses.

2.1 Definition

Bartle (2003) defines virtual worlds as follows: *"A virtual world is an interactive simulated environment accessed by multiple users through an online interface"*. This definition includes some of the main characteristics of a virtual world:

- **Simulation:** The environment is an artificial world which only exists on computers and imitates some aspects of reality.
- **Interactivity:** Users can interact either with each other or with the environment itself.
- **Interface:** To be able to interact, an interface between the users and the virtual world is necessary. Its type depends on the type of the environment, but in most cases, standard devices such as monitor, keyboard and computer mouse are used.

Another definition was developed by Bloomfield (2007), who defines a virtual world as *"... a computer-mediated environment that simulates real-world physics with sufficient fidelity, and in which one or more human participants can control* *one or more actors"*. This statement also contains the main features as explained above, and adds the concept of actors which are controlled by users. These actors are the representation of the users in the virtual environment and also called avatars[1].

Finally, a different approach by Book (2006) lists the key features of virtual worlds as follows:

1. **Shared Space:** Different users can act concurrently in the same world.
2. **Graphical User Interface:** Virtual worlds have a graphical interface which can be two - or three dimensional.
3. **Immediacy:** The user's actions take place immediately, i.e. in real time.
4. **Interactivity:** Users can influence the world by manipulating objects, and/or they can communicate with each other.
5. **Persistence:** The effects of the user's actions are saved persistently. Thus, when the user logs in again, her or his previous manipulations still exist[2].
6. **Socialization/Community:** Virtual worlds typically allow the creation of social networks via group concepts.

These three definitions point out the main aspects of virtual environments. The focus of this thesis is the interaction feature, which will be examined in detail in the following chapters.

2.2 Categories and terms

Virtual environments in general can be categorized according to their features. Hence, a variety of synonyms and different names exist which shall be explained in this section[3].

Early versions of virtual environments are text based multiplayer computer games, also called *multi-user dungeons (MUDs)*. These worlds include most of the features mentioned above, e.g. interactivity. However, as they are completely text based, a graphical representation is missing.

[1]A detailed description of avatars can be found in chapter 3.

[2]Unless, of course, these manipulations are further changed by other users.

[3]These categories are usually not defined clearly and as a result, some of the terms are used as synonyms.

A modern version of MUDs are *Massively Multiplayer Online Role-Playing Games (MMORPGs)*. They are based on 3D environments, where the users have specific roles. Furthermore, each user has to fulfill tasks in order to reach certain goals. Most of these games are based on a hunting and collecting concept. This means that the users have to collect items which can be obtained by e.g. fighting against enemies.

The expression *Collaborative Virtual Environments (CVEs)* is used for virtual environments with many users who can interact. Compared to MMORPGs, these worlds usually offer more freedom, as there are no roles and no goals to fulfill. CVEs are mainly designed for social interaction and collaboration.

Immersive Virtual Environments (IVEs) are virtual environments which offer the user the sense of being in it. To increase the feeling of presence, these systems use special devices to provide the user with interaction possibilities similar to reality. For instance, users of these environments wear a head mounted display or are situated in a special room where their movements are tracked. However, due to the fact that this special equipment is rather expensive, the use of IVEs is currently restricted to research institutions or enterprises.

2.3 A survey of collaborative virtual environments

Virtual environments recently have gained popularity. Due to the increasing speed and sinking costs of Internet access, more users are participating in virtual worlds, which leads to a growing number of different CVEs. This section gives an overview of popular virtual environments and compares them according to several criteria.

Figure 2.1 shows a comparison of some virtual worlds. As can be seen, each of the listed environments allows its users to customize the avatar. This feature is inevitable to offer a high level of interactivity and will also be explained in the next chapter. As far as the usage fees are concerned, most CVEs offer a basic account which is free of charge but usually has restrictions. A premium account costs about $10 per month[4], whereby additional pricing models are possible (e.g. a quarterly paid fee in Second Life).

[4]As of November 2007

	Active Worlds	Habbo Hotel	Second Life	There	The Sims online	Whyville
Creator	Activeworlds	Sulake Labs	Linden Labs	Makena Technologies	Electronic Arts	Numedeon
Launch	1995	2001	2003	2003	2002	1999
Cost per month	Free/ $6.95	Free	Free/$9.95	Free/$9.95	$9.99	Free
Idea	Building, Exploration	2D hotel for teenagers, chatting, furnishing	Building, exploration, earn money	Socializing, exploring, building	Entertaining, building	2D education for children
Economy	Buy in shops	Buy furniture	Buy and sell items, services and land	Buy and sell objects	Earn virtual money, buy items	no
Object design	yes	no	yes	yes	yes	yes
Scripting	yes	no	yes	yes	no	no
Own Land	yes	no	yes	yes	yes	no
Edit Avatars	yes	yes	yes	yes	yes	yes

Figure 2.1: Comparison of popular virtual worlds

Furthermore, most virtual worlds offer their users to design the environment by their own. This means that users own and design a part of the world and are able to create individual content. Usually, a general policy defines the rights of content generation in order to maintain a "clean" world.

As far as economy is concerned, Second Life and There offer the most opportunities of the presented CVEs. The currencies are named Linden Dollars and Therebucks and are both convertible to real currency. Moreover, in these worlds users can buy and sell generated content, such as objects or clothing.

To complete the overview of virtual worlds, the rest of this section will point out main ideas which are present in most or all popular CVEs as listed above. As an example, Figure 2.2 shows the main motives for participation in Second Life, based on a study conducted by de Nood and Attema (2006).

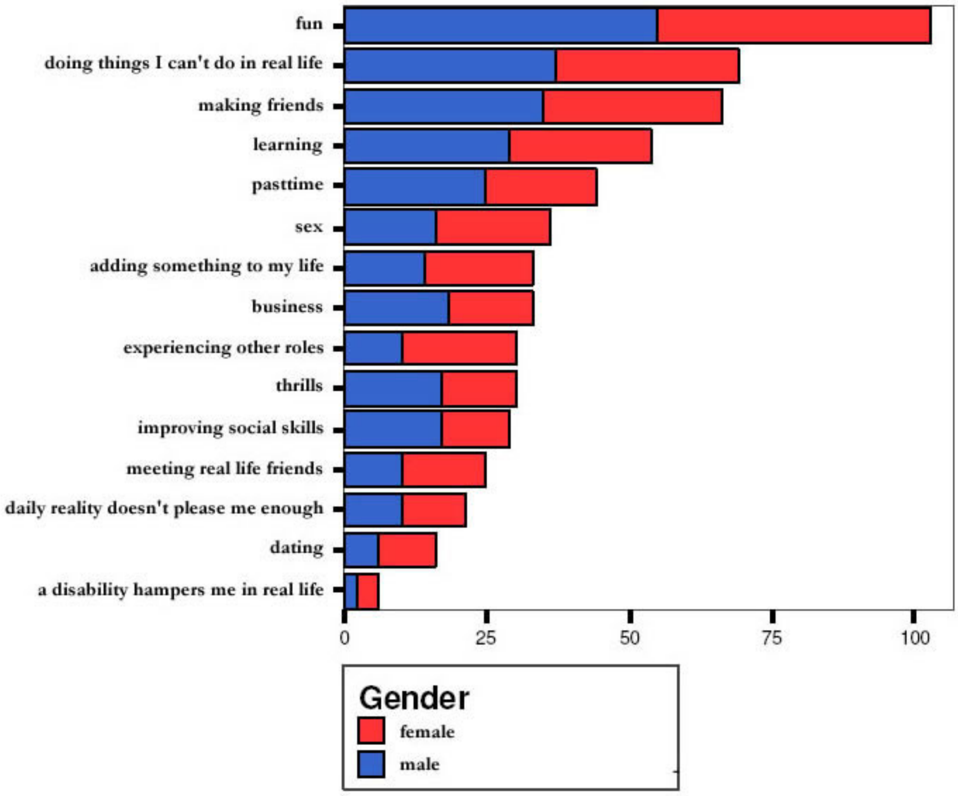


Figure 2.2: Motives for participating in Second Life

As can be seen, the main reasons for participation in virtual worlds are fun, experimentation, socialization, learning and business.

2.3.1 Socialization

As mentioned above, the social aspect is very important in virtual environments. Users communicate with each other, create communities and make friends. Therefore, CVEs provide features which facilitate the socialization process.

Usually, most virtual worlds contain a text based communication interface. One form of representation is a chat-like, line-based area, where the users can type messages which are displayed to all users who are present in a specific area. *Habbo Hotel* for instance uses its hotel room concept to define chat rooms for communication. Another approach is used in *Second Life*, where the chat messages have a range and thus can only be read by users who are nearby.

There displays the chat messages differently. In this environment, speech bubbles are shown over the avatar's head (see Figure 2.3). If the messages are quite long, they are split into different parts to avoid huge bubbles (Brown and Bell, 2004).

Figure 2.3: Speech bubbles in There

Some virtual worlds, such as *Second Life* or many modern MMORPGs additionally have a speech interface. Hence, the users are able to communicate with their voice which is easier, but reveals more of the user's identity.

Another community feature which can be found in virtually all CVEs is group building. For example, Second Life includes a group management interface, where users can create or join groups and define roles. Each user can be member in various groups. The advantages of these groups are that communication is facilitated, objects can be shared easily and news can be spread via a publish-subscribe approach.

2.3.2 Content design

Creating own content is another main feature of virtual worlds. Most of the described CVEs provide functionality for building objects. Restricted forms of content design can be found in *Habbo Hotel*, where the users can buy predefined furniture items which can be placed in a hotel room. Although this allows an individual design of locations, the same objects are used in the whole world, which doesn't lead to much individualism.

Other virtual environments, such as Second Life or There include powerful object manipulation interfaces. As a result, complex individual objects can be created and thus, the whole world contains a variety of different styles. A disadvantage of this freedom is that the environment can have a messy look, because every user designs its objects in different sizes, colors and styles.

In addition to objects, some CVEs include music and videos which can be defined by the users.

2.3.3 Economy

An important aspect of virtual worlds is their economic potential. These days, more and more enterprises are interested in participating in virtual worlds to advertise or even to sell their products.

Besides standard business models, virtual worlds can offer additional ways to make money. For instance, in MMORPGs it is a common custom to pay other users to play some hours and solving quests in order to level up[5] an avatar (Castronova, 2004). Currently, accounts with high leveled avatars for the game World of Warcraft are traded at Ebay. The prices of the most valuable avatars reach up to more than 900 Euros[6]. Furthermore, valuable items for this type of games are sold. Thereby, a combination of real world channels and virtual channels are used.

An example of a CVE with a very powerful economy is Second Life. There, the users can buy and sell objects, services or land on a free market. The currency for the transactions is the Linden Dollar, which can be converted into real currency on special currency exchanges. Thus, money earned in the virtual environment is real. Actually, more than 50% of the users quote that they earn money in Second Life (de Nood and Attema, 2006).

The main business models for earning money in Second Life are the following (Goodliffe, 2006):

[5]Levelling up an avatar means to earn points which give the character additional powers and thus allow it to fight against stronger enemies.

[6]retrieved on http://www.ebay.de, 04-Nov-2007

- **Object design and sales**: The typical income source of Second Life residents are sales of objects. Thereby, the majority of sold items are articles of clothing or other attachments for the avatar.
- **Jobs**: Services are another important sector of Second Life's economy. Enterprises hire users e.g. to advise customers in their shops. Other forms are escort services for users or adult entertainment.
- **Camping**: As the listing of places in the search engine depends on the number of users who visit a place, many land owners pay users to visit their land. This model is called camping and uses objects on which the users sit or stand and receive a small amount of Linden Dollars for a specific amount of time.
- **Weekly stipend**: Those users, who have a premium membership and hence pay a monthly fee, automatically receive a weekly amount of virtual money from Linden Lab.

Additionally, Second Life is used as an advertising channel for the presentation of products or the elaboration of a corporate identity. Book (2005) describes how virtual brands are created and maintained. The **PREEN** clothing store and *Abbot's Aerodrome* are examples of successful private brands in Second Life.

Similar economic concepts, though currently not as popular as Second Life, can be found in There.

2.4 Second Life as an example of virtual worlds

In the first part of this chapter, some characteristics of Second Life, such as its economy were explained. The goal of the following part is to give an overview of other functions of this virtual world, as it will be the exemplary platform for the analyses in the rest of the thesis. Furthermore, different features will be examined in detail in other chapters.

2.4.1 The world

Second Life was started online in 2003 by Linden Labs. The idea of its creators was to "*allow residents to control nearly every aspect of their world. From the shape of* *their avatars to the design of their homes, from how they spend their time to what types of affinity groups they form; Second Life's design was focused on fostering creativity and self-expression in order to create a vibrant and dynamic world full of interesting content.*" (Ondrejka, 2004). Thus, the whole world is created by its users.

As Second Life is a 3D world, an important resource is virtual land. This is the basis for the creation of content, because objects have to be placed on a piece of land. The management of this resource is controlled by Linden Labs, whose initial idea was to create an earth-like map, which represents the organization of the land. Thereby, the land is simulated on a grid of computers, and each of them is connected to 4 other ones in order to allow a continuous travel (Ondrejka, 2004). Each land region represents 65,536 m^2 (Linden Labs, 2007). Later, additional independent islands have been added. The actual land consists of major land parts and a variety of small islands. A travel between these parts is possible via teleporting.

Another relevant aspect of Second Life's world is that it uses a complex physics engine to generate a realistic feeling. Objects can be declared either as fixed or physical, where the latter option has the effect that the object reacts to forces such as gravity or collisions. Furthermore, the avatars are also animated by the engine. Additional features, such as a changing altitude of the sun increase the realism.

2.4.2 Policy

The creators of Second Life give the users as much freedom as possible. As a result, the set of rules and restrictions are limited to the ones needed in order to guarantee an efficient society.

Important restrictions are related to digital rights management. This policy is crucial for the in-world economy, because property rights are necessary to avoid theft and duplication. Hence, objects in Second Life by default belong to their creators. Other users initially have no rights to copy, delete or manipulate them. If needed, these protections can be removed so that other users can manipulate objects too, which is necessary for collaboration. Similarly, land owners can protect their land by banning residents, restricting the creation of objects on it or limiting the access to specific users.

In addition to the protections managed by the users, Linden Lab itself intervenes, if one of their community standards is violated. These rules aim to protect the user's privacy and life in the virtual world. Abuse can be reported by every resident which leads to an investigation by Linden Labs[7].

2.4.3 Scripting

Consequently, another important feature of Second Life, its scripting language, will be described. The language LSL is the basis for user-created applications and thus plays an important role for the creation of content. For instance, in the virtual theater play scripts were used to create a virtual interface for the control of actions.

LSL is a C-like, procedural scripting language which allows adding functionality to objects. It is built on a finite-state machine concept and therefore uses states and events as a programming model. Complex applications, which use several objects, make use of a messaging approach. This means that objects can listen and spread messages on several channels in order to communicate with each other.

The built-in library includes more than 300 functions for a variety of purposes. They include e.g. functions for object manipulation, sales or building vehicles. For complex systems, a communication with external software is possible via XML-RPC or HTTP requests. With these means, powerful systems can be created, such as Svarga, a simulated ecosystem in Second Life (see Calongne and Hiles, 2007).

To avoid that LSL scripts overload the server performance, some functions, as the sending of emails have a delay. Moreover, the memory of each script is limited to 16 KB.

[7]For further information see http://secondlife.com/corporate/cs.php, where the community standards of Second Life can be found.

2.4.4 Growth

Since the start of Second Life it has been growing steadily. Due to the fact that a basic membership is free, many users register to explore the possibilities, which are a common topic in all types of media.

Table 2.1 shows an overview of key figures for growth between 2006 and 2007 (Linden Labs, 2007). It is difficult to measure the number of active users correctly, because many accounts are idle. As an alternative to the population figures, the usage hours published by Second Life can be a more reliable indicator of growth.

Year	**Month**	**Land Size (km^2)**	**Population**	**Hours used**
2006	October	218,25	778.465	6.079.914
2006	November	254,73	1.095.959	6.988.233
2006	December	293,65	1.431.144	7.337.424
2007	January	360,79	1.990.849	10.817.668
2007	February	411,98	2.632.041	12.006.156
2007	March	504,28	3.256.371	15.346.784
2007	April	579,42	3.779.566	18.135.548
2007	May	651,39	4.370.810	20.767.557
2007	June	712,60	5.224.582	21.815.700
2007	July	779,95	5.706.958	23.640.980
2007	August	839,72	6.164.951	23.455.451
2007	September	871,32	6.736.832	24.138.413

Table 2.1: Growth of Second Life

The figures show that population and land size has been growing rapidly. However, the growth in the number of usage hours has been slowing down in the last months. An interpretation of this slow-down can be that on the one hand the novelty of this CVE is declining, because most of the opportunities are already exhausted. On the other hand, changes in policy, such as the introduction of VAT on land sales for European users, also might have reduced the activities of Second Life residents.

Nevertheless, virtual worlds such as Second Life are still settings for many activities, either in the business or private sector. Due to the fact that Second Life is currently the most popular CVE, it will be examined in detail in the following chapters, as far as its avatars and their interactions are concerned.

2.5 Summary of virtual environments

This chapter gave an overview of virtual worlds. It was shown that there are several key features, especially interactivity, which characterize different types of virtual environments. Particularly, collaborative virtual environments, which in general offer a high degree of freedom, are the focus of the thesis.

In comparison to similar virtual worlds, Second Life was shown to be one of the most extensive ones, as it includes a variety of features for its users, e.g. a scarcely limited economy and a powerful scripting language. This, among other aspects, is likely to be the reason for Second Life's popularity and growth which was described in the last part of this chapter.

Chapter 3

Avatars

An avatar is the representation of a user in virtual environments, such as in the Internet or computer games. The areas of use of avatars in the Internet comprise web forums, chat rooms and other types of communities and their idea is to represent the user's personality and to distinguish this user from others. In computer games and other virtual 3D environments avatars have the additional and important function to represent the user physically. This allows an interaction between the user in form of its avatar and the virtual environment.

To avoid confusion with software agents, which are a similar concept, one should keep in mind that avatars are a virtual representation of a human being. This means that an avatar is always controlled by a human. An (embodied) agent, in contrast, is the virtual representation of an algorithm. Very often the representation of the two concepts is similar.

3.1 History of Avatars

The term "avatar" originates from the Sanskrit word "avatara" and has the meaning of "descent" (Bailenson and Blascovich, 2004). The origin of avatars lies in the Hindu religion where an avatar is an incarnation (bodily representation) of a god on earth, which assumes some material form such as a man or an animal. Very important in the mythology of Hinduism are the avatars or incarnations of the god Vishnu. The ten most important ones are called "Desavatara" and can be seen as the main examples of avatars in history (Egen, 2005). One well known incarnation is Krishna, which is a frequent motif in arts.

The type of avatar that is referred in the initial definition, although it has a different function, comprises one main aspect of the Hindu avatars: Modern avatars are "incarnations" of users in virtual environments.

The first time that the expression "avatar" was used in digital environments was already in the late 20th century (Bailenson and Blascovich, 2004). Players of the first MUDs could use a simple visual representation of themselves. However, the publicity of these worlds was restricted to a quite small amount of players. The promotion of the meaning of the term avatar to a bigger audience was influenced by the novel *Snow Crash* (Stephenson, 1992). In this book the author describes avatars as a simulation of humans in the Metaverse, a virtual reality form of the Internet.

From that time on, avatars have been used in almost every virtual environment. Most computer games include a character that represents the player. Depending on the type of the game, simple images, two-dimensional or three-dimensional presentations are used. Additionally, as the Internet has been transforming into a network for social contacts, avatars have been introduced to allow a presentation of the user's person in the Web. Nowadays, most common types of Web-applications that enable communication, e.g. online forums, instant messengers or chat rooms, contain a graphical representation of the user.

3.2 Classification and types

As explained above, avatars are predominant in many different application areas. The restrictions and requirements of each of these areas call for variable types of representation. To get a better overview of these types, a classification by two aspects can be taken into account: dimension and realism.

3.2.1 Dimension

A main feature of representation is the dimension. Two-dimensional avatars are illustrated in form of images. The most common form is a quadratic picture of a person, an animal or a creature and is mainly used in internet forums. An enhancement of this type is an animated picture that consists of a sequence of single images[1].

More sophisticated avatars have a three-dimensional representation. This type is used in virtual three-dimensional environments, such as computer games or non-gaming universes. These avatars have the additional function to display the user's local position in the virtual environment. Furthermore, a 3D avatar typically includes a set of predefined animations that are started when an interaction with the environment, e.g. movement or communication, takes place.

3.2.2 Realism

The second classification aspect that is considered in this thesis is the level of realism. Two main dimensions of avatar realism are explained in literature (Bailenson and Blascovich, 2004). *Behavioral realism* defines the level of similarity to human behavior. This includes e.g. movements. The other dimension is *photographic realism*, i.e. the level of graphic similarity between the representation and reality. Consequently, the focus of this classification lies on photographic realism, because it is measurable in both 2D and 3D environments. Behavioral realism will be discussed in later sections, where gestures and mimics are explained.

Gerhard et al. (2001) define three different avatar types according to realism:

- Shape style / simple style avatars
- Cartoon style avatars
- Humanoid avatars

The *shape style* avatar consists of basic icons that do not depict living beings and is the simplest form. The icon can show a photo, geometric shapes, text messages and other motifs. Due to its lack to conform to the surrounding, this type is usually not used in 3D environments.

A more sophisticated type is the *cartoon style*. It is a simplified depiction of a living being and can include a presentation of emotions through facial expressions.

[1] An example format is the *Graphics Interchange Format* (GIF).

Finally, the most realistic avatar type is the *humanoid style*. The extent of realism of this type reaches to photorealistic avatars that are a realistic representation of a human being. Modern rendering techniques offer the possibility to create 3D bodies which are hardly distinguishable from photos or videos. However, performance limits require a simpler form of representation in common virtual environments.

3.2.3 Examples

Figure 3.1 shows an overview of examples of avatars. Each example is categorized according to the above defined dimensions.

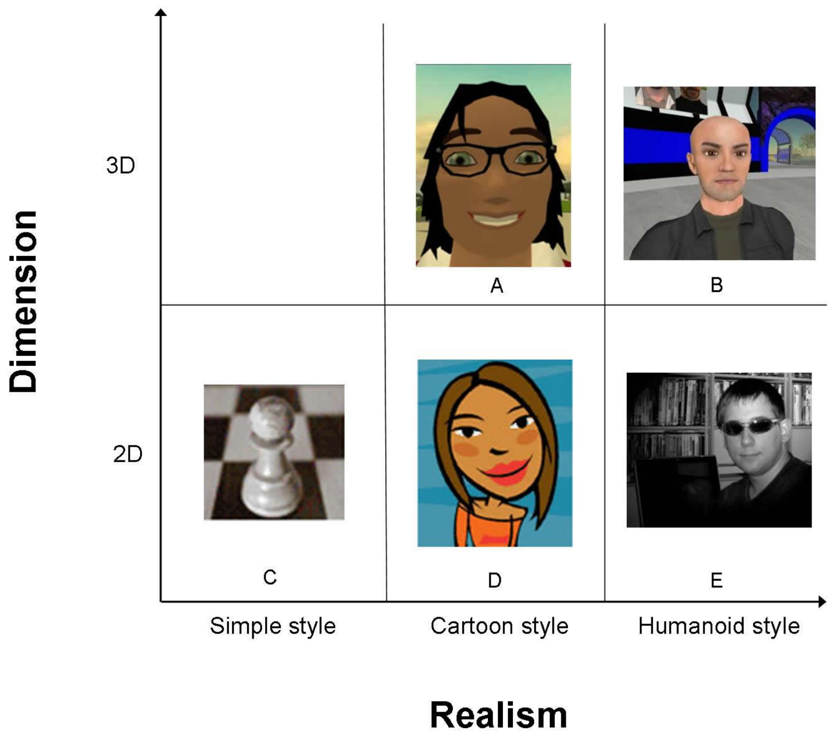


Figure 3.1: Classification of different avatars

The different examples are the following:

- **Avatar A**: Example of a 3D cartoon style avatar in the CVE *There*[2].
- **Avatar B**: A realistic 3D character in *Second Life*[3].

[2]http://www.there.com
[3]http://www.secondlife.com

- **Avatar C**: One simple avatar option in the IM software *Pandion*[4].
- **Avatar D**: A cartoon style 2D avatar from *ICQ*[5].
- **Avatar E**: A 2D realistic avatar in the online community *Studivz*[6].

These examples only show a small selection of existing avatar types and do not include fantastic creatures or animals, which are also very frequent.

3.3 Features of avatars for virtual social interaction

As explained above, avatars can have different forms. However, to enable a sophisticated mode of interaction, the three-dimensional humanoid style is the only possible option. Hence, the following chapters of this thesis solely take this representation form into account.

When one talks about social interaction, different channels of communication have to be considered. In real life, human communication not only assumes the form of speech. According to Argyle (1988), more than 65% of communicated information is transmitted via non-verbal signals. Therefore, to enable a successful virtual communication, avatars have to include features for the transmission of non-verbal signals (Guye-Vuillème et al., 1999). In this section three of the most important means of non-verbal communication in CVEs are described. Gestures, mimics and the eye gaze are considered to be the best means for transmitting emotions and other types of information (for a more detailed description see Ringard (2006)). Finally, the feature of avatar creation is demonstrated, because it is fundamental to present the user's personality and therefore another form of communication.

Technical aspects of the demonstrated features are covered in other chapters of this thesis.

[4]http://www.pandion.be
[5]http://www.icq.com
[6]http://www.studivz.net

3.3.1 Gestures

One of the most extensive channels of non-verbal communication are gestures, because they include movements of different parts of the body, in particular of the arms. Additionally, gestures comprise movements of the whole body, such as walking or jumping. However, most commonly the term gestures is used for movements of the arms that support speaking.

In literature (Beattie, 2003), different types of gestures are distinguished:

- Emblems
- Illustrators
- Regulators
- Affect displays

Emblems or symbolic gestures are consciously used to replace words. Hence, they are relatively independent of oral communication and can be used to simplify communication in general. One example of this type are greeting gestures. As many symbolic gestures can have different meanings in different cultures, one has to be careful with equipping avatars with these symbols.

Another common kind of gestures are iconic gestures or *illustrators*. They have a strong connection with speech and enrich the verbal information. This means that they usually cannot be used separately. Normally, these gestures are used unconsciously and have a big influence on the quality of communication. The speaker's intentions can be made clear through these gestures, when the arms describe visually what is explained orally.

Regulators are gestures which are used to control the turns of a conversation. They are important for the temporal flow of a discussion. One example is putting the hands onto the table to indicate that there is nothing more to say.

Finally, there exist gestures which display *affects*. These are the most unconscious ones and consist of postures that show the emotional state of the speaker.

As far as virtual environments are concerned, the most relevant gestures are the ones of the first three types. Emblems are usually implemented as a set of predefined animations which can be activated by the user. Thereby, the initialization is carried out by pressing a button or using a textual command.

As illustrating gestures strongly depend on the semantic content of communication, an implementation in CVEs is difficult. Hence, most state-of-the-art environments only include standard animations which are activated automatically during speech. Nevertheless, this simple form of non verbal communication improves communication significantly. Moreover, the stopping of the automatic gestures is a form of regulator and illustrates other users that their turn of speaking has started.

In current CVEs emotions are typically represented by iconic gestures, i.e. consciously. Due to the lack of computational identification mechanisms of the user's emotions, the last type of gestures is currently not included in virtual 3D worlds.

3.3.2 Mimics

Similarly to gestures, mimics support a person's communication. More precisely, the human face is considered to be the main channel for communication of emotions (Knapp, 1978). However, due to the fact that facial expressions are controlled by an immense set of different muscles, the simulation of mimics is very complex. Therefore, current CVEs contain this feature in a very limited way. Nevertheless, many research projects have been conducted which investigate means to create avatars that are able to express emotions via mimics.

The effect of emotionally expressive avatars on the richness of user experience has been examined by Fabri and Moore (2005). They developed an instant messenger which includes the presentation of computer generated facial expressions. Four different user experience aspects were tested:

- Involvement
- Enjoyment of the experience
- Sense of presence during communication
- Sense of co-presence (the sense of being together in a virtual environment).

The results of the study showed significant higher scores for involvement and copresence when emotionally expressive avatars were used. However, the researchers stated that expression of emotions may not be appropriate for all types of interaction contexts.

A fundamental work in this area that is the basis of many research projects was conducted by Ekman et al. (1972) who identified a set of basic emotions which can be represented by facial expressions. The restriction to six emotions (Happiness, Anger, Fear, Sadness, Surprise and Disgust) allows a simplification of the complexity of mimics. These basic emotions are independent of the social and cultural context which means that they can be recognized all over the world.

Fabri et al. (2002) adopted the ideas of Ekman and Friesen and examined if the findings about real faces can be implemented on avatar faces. They developed a set of distinct facial expressions on a computer generated face and conducted an experimental study to find out if these expressions were recognized by tested persons. Thereby, they focused on efficiency by using a quite simplified representation. The following figure shows the set of best, i.e. most distinctive representations of the different basic emotions.

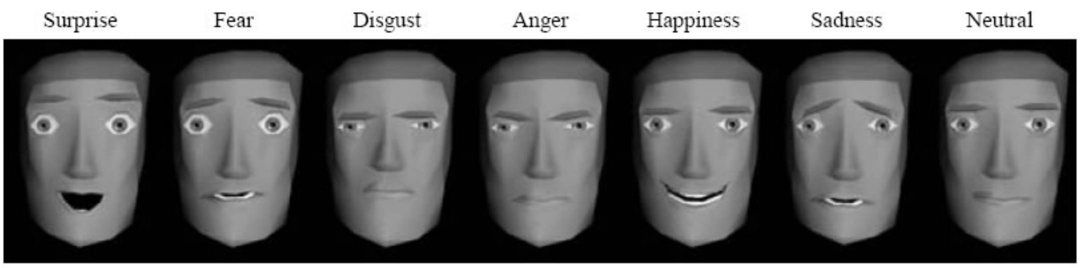


Figure 3.2: Basic facial expressions

The result of the study showed that the recognition rates of most of the computer generated expressions were relatively high (over 60%) and only slightly lower than recognition rates of photographs of facial expressions. As a consequence, reducing the complexity of mimics to a simplified set of basic expressions seems to be a possible way to make avatars emotionally expressive and hence improve virtual interaction.

Another way to deal with the complexity of human emotions and facial expressions is proposed by Imbert et al. (2001). They suggest delegating emotional expressions to an agent so that the user is free to interact without having to activate the expressions deliberately. This approach requires a highly evolved AI system and a more or less complex emotional model.

3.3.3 Eye gaze

Another relevant channel of non-verbal communication is the eye gaze, provided that a face-to-face interaction takes place. It consists of a combination of head and eye orientation. According to Argyle and Cook (1976), it has different functions. Examples of these are:

- Regulation of the communication flow
- Provision of feedback
- Communication of emotions

As regards virtual interaction, eye gaze is an additional aspect to improve social communication.

Garau et al. (2001) analyzed the importance of eye movements in humanoid avatars for virtual interaction. They compared different conversations between two persons and measured the effect of eye gaze on the level of user experience. Four conditions were tested:

1. **Audio only**: The two participants solely communicated via an audio connection and could not see each other
2. **Random gaze**: The opponent was represented by an avatar whose eyes moved randomly.
3. **Inferred gaze**: The avatar's head and eye movement were controlled by the flow of the conversation
4. **Video**: The participants could see each other on a monitor.

The results of the study showed a strong outperformance of the video communication, as can be expected. Furthermore, the inferred-gaze avatar achieved significantly better responses than both the random-gaze avatar and audio only communication. Thus, it can be inferred that an avatar whose eye gaze is related to the conversation allows a markedly better interaction experience. Nevertheless, an implementation seems difficult for similar reasons as previously explained.

3.3.4 Avatar construction

In addition to the above described features for non-verbal communication, an avatar has to provide the possibility of customization as it is a channel for communication of a user's identity. Avatars are a means to tell other users something about one's personality (Schroeder, 2002).

Many users spend very much time on creating a virtual identity. Thus, their visual representation plays a central role in social interaction. On the one hand, it serves to identify and remember users and on the other hand it is considered to increase self focused attention (Vasalou et al., 2007). This means that a user is more aware of his or her presence in the environment.

For these reasons, virtual environments allow their users to choose a visual representation. In some CVEs the user can choose from a set of predefined avatars. This simplified form of avatar construction is problematic for the creation of a virtual identity, because often different users have the same appearance. Differentiation in these platforms is conducted through names. To avoid this problem, more sophisticated CVEs allow their users to define their representation individually.

As far as the choice of avatar appearance is concerned, Vasalou et al. (2007) examined in how far the virtual representation resembles the real appearance of users. They asked subjects to evaluate the similarity of Yahoo! 360°[7] avatars to photos of the corresponding users and found out that more than half of the avatars were similar to the user's appearance. On the other hand, many users want to experiment with their visual representation and choose animals or fantastic creatures for their virtual self.

3.4 Avatars in Second Life

After having explained different theoretical aspects of avatars, the last part of this chapter describes the use of avatars in Second Life as a practical example.

One basic concept of Second Life is that users can experiment in a high degree. One example is that they can generate different visual appearances and save

[7] Yahoo! 360° is an online communication portal. See http://360.yahoo.com/

them in their inventory. When needed, they can be simply activated with just one click. This means that a user can change its avatar's appearance within a second to be e.g. a monster, a little girl or any other form. The only restriction is the avatar's last name which has to be chosen once and cannot be changed. This allows Linden Lab to avoid an abuse of names. Furthermore, Second Life enables the users to customize a multitude of different features of their visual appearance in the virtual environment, whereby the base avatar type is a three-dimensional humanoid or cartoon-like avatar (see section 3.2.2).

Table 3.1 shows an overview of the customizable elements[8]. As the list indicates, the shape of each avatar can be designed very flexibly. In addition to the listed features, users can customize their avatar's look by creating a custom skin or hair texture. To have even more freedom, many users extend some body parts, such as the hair with objects so that these parts are not restricted to the standard settings.

Category	**Subcategory**	**Features**
Shape	Body	Height, Thickness, Body fat
	Head	Size, Stretch, Shape, Egg Head, Length, Face shear, Forehead angle, Brow size, Upper/Lower Cheeks, Cheek bones
	Eyes	Eye size, Eye opening, Eye spacing, Outer/Inner Eye corner, Eye depth, Upper eyelid fold, Eye bags, Puffy eyelids, Eyelash length, Eye pop
	Ears	Ear size, Ear angle, Attached earlobes, Ear tips
	Nose	Nose size, Nose width, Nostril width, Nostril division, Nose thickness, Upper bridge, Lower bridge, Bridge width, Nose tip angle, Nose tip shape, Crooked nose
	Mouth	Lip width, Lip fullness, Lip thickness, Lip ratio, Mouth position, Mouth corner, Lip cleft, Lip cleft depth, Shift mouth

continued on next page

[8]Client version 1.18.4(3)

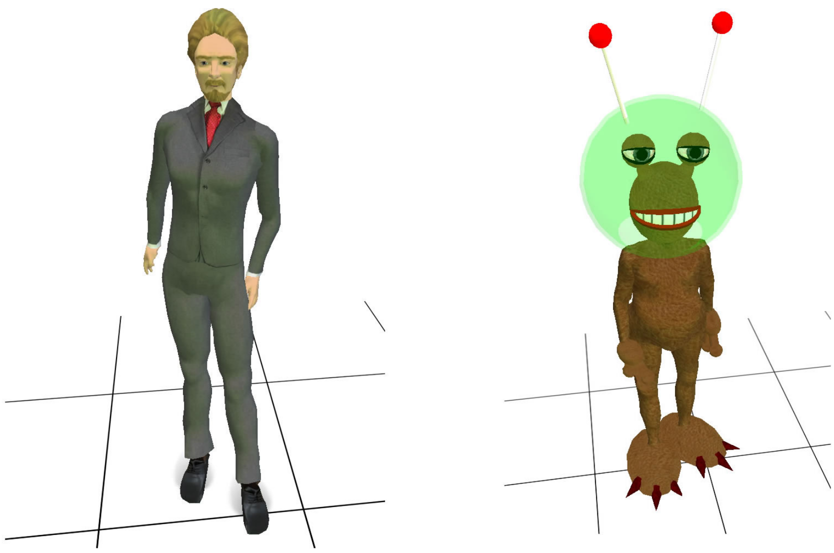
continued from previous page		
Category	**Subcategory**	**Features**
	Chin	Chin angle, Jaw shape, Chin depth, Jaw angle, Jaw jut, Jowls, Chin cleft, Upper chin cleft, Chin-Neck
	Torso	Torso muscles, Neck thickness, Neck length, Shoulders, Breast size, Breast buoyancy, Breast cleavage, Arm length, Hand size, Torso length, Love handles, Belly size
	Legs	Leg muscles, Leg length, Hip width, Hip length, Butt size, Saddle bags, Knee angle, Foot size
Skin	Skin Color	Pigment, Ruddiness, Rainbow color
	Face Detail	Facial definition, Freckles, Wrinkles, Rosy complexion, Lip pinkness
	Makeup	Lipstick, Lipstick color, Lipgloss, Blush, Blush color, Blush opacity, Inner/outer shadow, Inner/outer shadow color, Inner/outer shadow opacity, Eyeliner, Eyeliner color, Nail polish, Nail polish color
	Body Detail	Body definition, Body freckles
Hair	Color	White hair, Rainbow color, Blonde hair, Red hair
	Style	Hair volume, Hair front, Hair sides, Hair back, Big hair front, Big hair top, Big hair back, Front fringe, Side fringe, Back fringe, Full hair sides, Hair sweep, Shear front, Shear back, Taper front, Taper back, Rumpled hair, Pigtails, Ponytail, Spiked hair, Hair tilt, Middle part, Right part, Left part, Part bangs
		continued on next page

continued from previous page

Category	**Subcategory**	**Features**
	Eyebrows	Eyebrow size, Eyebrow density, Eyebrow height, Eyebrow arc, Eyebrow points
Eyes		Eye color, Eye lightness

Table 3.1: Customizable avatar elements in Second Life

As far as the kinds of appearance are concerned, Jones (2006) distinguishes between two different groups of avatars in Second life: normative and fantastic. Figure 3.3 shows example avatars of the different groups.



(a) Normative avatar (b) Fantastic avatar

Figure 3.3: Different avatars in Second Life

Normative avatars look like real human beings and are quite often designed according to an ideal physical appearance. Very often the body shape is designed to give emphasis on the gender, e.g. a muscular body for a male avatar (McKeon and Wyche, 2005). On the other hand, many avatars in Second Life are a reflection of the real life appearance of the according user. Additionally, it seems that a majority of normative avatars have a typical Caucasian and young look, as it is considered a stereotype for a typical human.

The fact that humanoid avatars are designed with a strong expression of their gender can be explained with the idea that gender is relevant for virtual interaction. O'Brien (1999) states that "*Gender is one of the first means by which persons introduce and represent themselves to others in electronic communications*". Accordingly, in the field of virtual worlds the gender of avatars has a strong influence in the first contact.

Furthermore, according to a survey carried out by de Nood and Attema (2006), the majority of users keep their real gender for their avatars. Figure 3.4 shows the results of the survey[9].

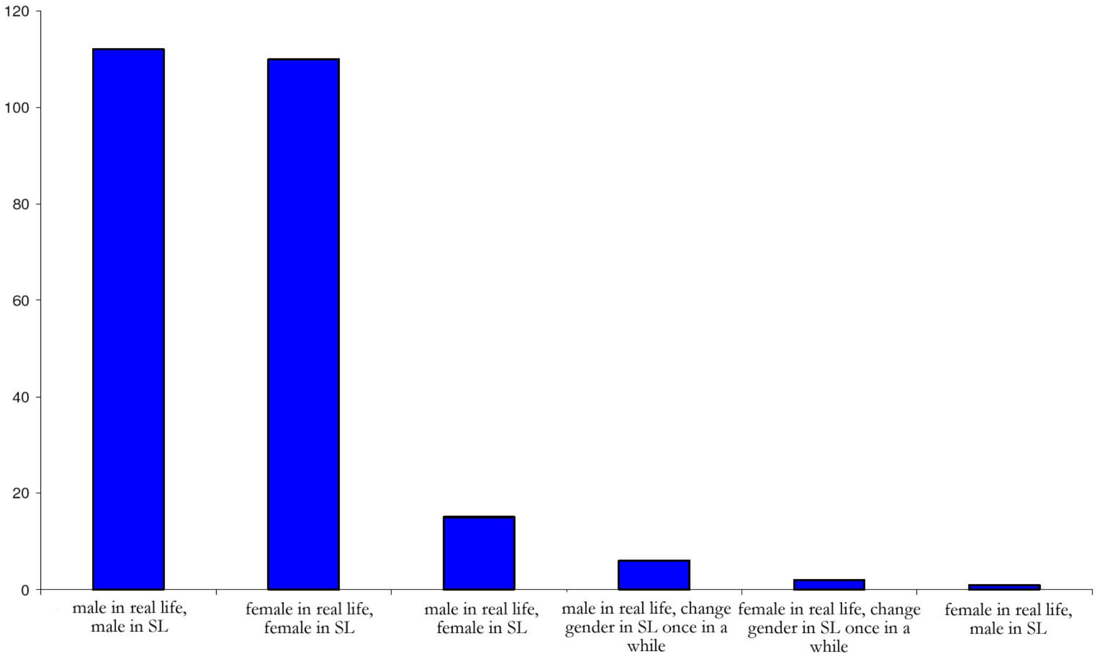


Figure 3.4: Gender in Second Life

As can be seen, the vast majority of the asked users state that they keep their gender in Second Life. Only about 10% change their gender or swap it regularly.

The other group of avatars, fantastic ones, shows a high grade of diversity. Creatures such as fairies or vampires can be encountered easily in the whole environment. Nevertheless, due to the physical avatar model, these creatures have always human characteristics.

[9]The number of female and male respondents in the survey was roughly equal. This is not a representation of the real distribution of genders in Second Life, beacause more than 60% of the users are male (Linden Labs, 2007).

Regarding non-verbal communication, Second Life avatars are provided with basic motions which are executed automatically. Eye gaze is implemented in the way that the head and eyes of an avatar follow the owner's mouse position, which is also used to look around the environment. Hence, the avatar's gaze is usually directed according to the visual display of the screen.

Furthermore, there are gestures that are included to give the avatar a more lively appearance. As an example there are standard animations which are executed after a specific time of idleness, such as movements of the arms.

Regulator gestures (see section 3.3.1) are implemented in a quite unique way. While a user types something into the chat line, the avatar imitates the movements of typing in the virtual world. This is a means to indicate that one is currently communicating. After the avatar has stopped performing the typing gesture, other users can realize that their turn of conversation has arrived. Figure 3.5 shows this gesture.

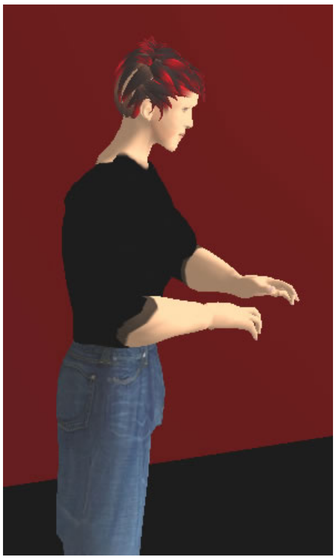


Figure 3.5: Typing gesture

In addition to the automatically executed gestures, a user can also activate custom gestures individually. Thus, symbolic gestures, such as shaking the head or waving can be used for communication.

As far as mimics are concerned, Second Life restricts facial expressions to a set of predefined ones which have to be activated like other custom gestures via a text command[10].

To sum up, Second Life allows a relative broad range of features for virtual social interaction, even though detailed expressions of emotions are not realized.

3.5 Summary of avatars

This chapter showed that avatars, the representations of users in virtual environments, are a rather old concept which originates from the Hindu religion. As far as social aspects are concerned, avatars are the basis of different forms of interaction and therefore have to provide a variety of features.

Not only is the visual appearance, which was classified by means of dimension and realism, a crucial factor, but the body language of avatars plays also an important role. In detail, gestures, mimics and eye gaze were analyzed in this chapter and possible implementations in virtual environments were discussed.

Finally, the theoretical concepts from the first part of the chapter were examined on the basis of Second Life. It was shown that a majority of the features are implemented and thereby, especially the construction of avatars was emphasized.

[10] Other forms of gesture activation are discussed in chapter 6.

Chapter 4

Interaction Techniques

The goal of the following chapter is to give an overview about interaction in virtual 3D environments as far as relations between avatar and the environment are concerned. The focus lies on motion and object manipulation, which are the most important interaction forms of this type. As the previous chapter showed more or less social aspects of virtual interaction, this chapter concentrates on technological aspects. Different techniques are presented and classified according to various systems. Finally, to conclude this chapter and to provide the basis for the explanation of the virtual theater play, a survey of available interaction techniques in Second Life, as one example of modern virtual 3D environments will be given.

4.1 Overview of virtual interaction

To point out different aspects of interaction, various classifications were developed. Robinett and Holloway (1992) classify virtual interaction according to whether another user or an object is involved. Actions that can be conducted include translation, rotation and scaling (see table 4.1).

	User	**Object**
Translate	fly through the world	grab and move object
Rotate	tilt the world	grab and turn object
Scale	expand or shrink the world	scale object

Table 4.1: Classification of interaction by Robinett and Holloway (1992)

This classification distinguishes between actions that manipulate objects and those that manipulate the user's relation to the environment. For instance, when tilting the world, the angle between the environment and the avatar is changed.

Another classification of interaction was developed by Slater and Usoh (1994). They distinguish between magical and mundane interactions (see table 4.2).

Interaction	Manipulation Examples	Navigation Examples
Mundane	object selection and placement, transformations, deformations	walking, driving or flying a vehicle
Magical	scaling the environment, psycho-kinesis	flying under own volition, teleportation

Table 4.2: Classification of interaction by Slater and Usoh (1994)

Mundane interactions are analogous to actions which are possible in the real world and comprise manipulation of objects or the world and movements. Magical actions are not possible in the real world but can be implemented in virtual reality.

Although these two classifications show important aspects of interaction forms, in the following section another one is used to mark off the most important types of interactions in virtual worlds. Thereby, interactions between a user and its environment in CVEs are classified into two groups: movement and object manipulation (see e.g. Mine (1995) and Belleman (2003)).

4.1.1 Movement

Moving through a virtual world allows the user to explore the three-dimensional structure of the environment. As in real world, the feeling of movement is generated by a change of the perspective on objects surrounding a person. This means for example that objects appear bigger when one moves towards them and vice versa or that different parts of an object are visible depending on the relative position to oneself.

Movement via input devices

In virtual environments, movement is usually controlled by some input devices[1]. The most common forms are *physical controls*, such as buttons, computer mouse, joystick etc., because they are standard devices for human computer interaction. However, a drawback of these devices is that movement cannot be mapped naturally. For instance, determining the speed of walking by pressing a special button is very distinct from the real movement.

Another option is to provide *virtual controls*, e.g. virtual steering wheels. This allows more flexibility, but lacks haptic feedback.

Direct user interaction devices allow the most intuitive movement in a virtual world. Although their use is not very common because of rather expensive components, they are included in this overview to demonstrate their capabilities for alternative interaction concepts. Technical devices for this type of interaction include systems for gesture tracking, gaze detection or visual recognition (see Figure 4.1).

Figure 4.1: Examples of devices for direct user interaction

As regards direction of movement, direct user interaction devices enable efficient forms of control by determining the position of the user's hand or the direction of eye gaze.

Hand directed techniques determine the direction of movement by tracking the position and orientation of a hand independently of the gaze. This means that an avatar can move into one direction while looking into another one. In contrast to this, *eye gaze directed techniques* depend on the current orientation of the user's

[1] This statement refers to the direct control of movement with input devices. Of course, other forms of control are also indirectly controlled with those devices.

eyes and head. Consequently, the direction of movement is always bound to the direction the avatar is facing.

Alternative movement forms

In addition to the way of moving around via directing the avatar directly through the environment, there are also alternative movement forms.

A passive form of movement can be achieved by moving the avatar with other objects. These can be vehicles that are controlled automatically or other objects, such as elevators. Additionally, a virtual world can contain attracting objects, which pull avatars to themselves or repelling objects, which push avatars away. None of these movement forms requires any interaction of the user.

As virtual worlds are not restricted to the rules of reality, additional means of movement can be implemented. One of these is dynamic scaling. Thereby, the user reduces the scaling of the virtual world so that a wider area is visible. Usually, this is enabled by a sort of a map view. After that, the avatar is moved to the target position and the view is scaled back to the original size. This enables a quick and simple movement through the virtual world.

Similar to dynamic scaling techniques, goal driven movement allows covering long distances very quickly. It uses a set of destinations from which the user can select. By clicking on the desired target location, the avatar is teleported instantly.

Speed control

Another aspect of movement that has not yet been explained in this section is speed. In virtual worlds several forms of specification of speed are possible:

- Constant speed
- Constant acceleration
- Controlled speed.

A simple form of specification of motion speed is to limit it to a constant level. This leads to an easy to understand control mechanism, but has the disadvantage that either for long distances the speed is too low and travelling is annoying, or for very small distances it is too fast so that performing tasks is difficult.

To avoid the problem of overcoming a long distance while keeping the basic motion speed low, acceleration can be an auxiliary instrument. With this concept, the speed changes according to the duration of motion. At the beginning, the avatar moves relatively slowly to be able to control its position precisely. If the motion continues for a longer time, the speed gets automatically faster. This is useful, if the virtual world is large and users want to visit different locations that are far away.

Finally, motion speed can also be controlled via physical or virtual controls. Simple forms include the keyboard where the user presses buttons to select between different speeds. More sophisticated forms use input devices such as sliders or accelerator pedals to allow a more efficient control.

4.1.2 Object manipulation

In contrast to movement, which describes an interaction between the user and the virtual world as a whole, there are also considerable interactions between avatars and specific objects in the virtual environment. These require a selection of a single object which then usually can be moved or rotated (Mine, 1995).

Object manipulation in virtual environments is problematic, because it normally has to be implemented differently than real manipulation. This means that virtual worlds require additional concepts, because input devices restrict the possibilities for interaction.

Typical problems of input devices regarding interaction are the lack of haptic feedback and limited input information (see Mine, 1997). More precisely, users need feedback of the physical features of objects that are manipulated. In other words, when interacting with an object, it is necessary to get a feeling of the form and size of it in order to be able to perform actions successfully. Standard input devices, such as computer mice or alphanumeric keyboards have very limited capabilities to transmit this feedback.

As a consequence, to enable an effective manipulation of virtual objects, more powerful input devices have to be used, as explained in section 4.1.1. These facilitate interaction by a variety of features (Mine, 1997).

Direct manipulation of objects can be realized by tracking the position of a user's hand. Consequently, the user can use its real hands to interact with the object, which is more efficient than pressing buttons. Moreover, gestures can be used to perform actions in the virtual world. These can be executed by moving the hand in a special way, which enables a very efficient kind of calling commands.

To sum up, usual CVEs such as Second Life or There are restricted in the methods of object manipulation, because they use standard input devices. Nevertheless, interaction is possible and most object manipulation tasks in these environments can be conducted without big problems.

The remaining part of this section will give an overview of both types of manipulation: interaction with standard devices and techniques with direct user input devices.

Selection

Selection of a virtual object is necessary to define on which object further actions are executed. Therefore, different objects in the virtual world have to be identifiable. Furthermore, a user selection technique has to include some form of highlighting of the selected object. Typical forms of indicating which object is selected are to give the object another color or to display a box around it.

Another issue that is considered in modern CVEs is the selection of multiple objects. Usually, the same action is performed on related objects, which requires a method of selecting these objects simultaneously. A common implementation is to keep a special key pressed while selecting single objects.

The simplest form of selecting objects is to choose an entry from a list that displays all available objects. The advantage of this method is that the object does not have to be visible. However, the ID of the object has to be displayed which is not very easy to recognize. Moreover, most interactions with virtual objects require them to be visible for the user.

Most virtual 3D environments use the computer mouse as interaction device. As a consequence, object selection is performed by clicking on the respective object.

If more powerful input devices are available, object selection can be performed by the user's hand or eye gaze. *Gaze directed selection* requires an accurate tracking of the user's eyes. Then it is possible to identify the object which the user currently looks at. *Hand directed selection* is performed by tracking the position of the user's hand. For this interaction technique different complex selection concepts have been developed (for an example see Poupyrev et al. (1996)).

Translation and Rotation

Translation and rotation of objects are other important forms of virtual interaction. Examples include virtual ball sports or the creation of new buildings. As shown in the previous sections, interaction can be controlled via physical and virtual controls.

While interactions via virtual controls offer much flexibility on object manipulation, these controls limit the sense of presence in the virtual world, because manipulation in real world is conducted differently. Normally, people move and rotate objects with their hands or use tools. Therefore, researchers propose the use of physical controls that allow a six degree of freedom input[2]. Nevertheless, these controls generate additional issues that have to be considered.

Another type of object manipulation is the tracking of the user's real hand movements. This is a more intuitive form of interaction, because gestural actions can be performed to manipulate the object more efficiently. A restriction of this approach is the limited arm reach.

To solve this problem, one possible method is automated scaling of the world (see Mine, 1997). Thereby, the user selects objects that shall be modified and grabs them. While grabbing them, they are automatically scaled to a size that allows manipulation and furthermore moved towards the avatar's arms. As a result, the modifications can be conducted and by releasing the respective objects, they are scaled back to their original size and moved to their original location.

[2]Six degree of freedom devices allow the control of up/down, left/right and forward/backward movements.

To position an object when using arm tracking, an amplification factor can be used that allows a larger range. Similar to the motion techniques explained in section 4.1.1, the speed of translation can be adapted to the user's needs.

4.2 Interaction in Second Life

Second Life, as most other virtual worlds, uses standard user input devices, such as an alphanumeric keyboard and a mouse for interaction. Therefore, it does not provide haptic feedback as previously described. Nevertheless, it offers a wide range of possible interactions that can be performed with the given devices.

4.2.1 Movement

Motion is implemented by both physical and virtual controls. Thereby, avatars in Second Life not only can move on the ground, but also fly through the air. Figure 4.2 shows the virtual controls for movement, which have a correspondence as physical controls, i.e. keys on the keyboard. Furthermore, to provide the possibility to look in another direction than the direction of moving, the PC mouse is used for looking around independent of the avatar's motion.

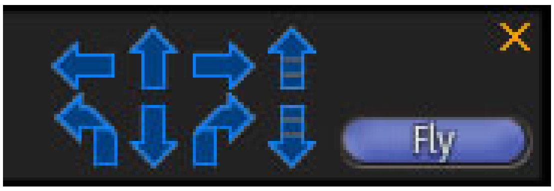


Figure 4.2: Movement controls in Second Life

The controls for upward/downward movement are used to define the height while flying. While walking, they are used to execute jumps or to stoop.

As regards the speed of motion, Second Life by default uses a constant speed for walking. To move faster, avatars can use the flight mode. Additionally, there exist user generated scripts that allow an additional acceleration of motion.

Alternative movement forms as explained in section 4.1.1 are also implemented in Second Life. Typical examples of passive movement are elevators that transport the avatar between different floors of a building, or vehicles with predefined routes that are used to provide tours through an area. In addition to automated vehicles, avatars can steer custom vehicles, such as cars or planes to move around the world (see Figure 4.3). Thereby, the same controls as for avatar movement are used.

Figure 4.3: Example vehicle in Second Life

Another form of motion which is implemented is goal driven movement. So called "teleports" allow a fast transportation to every point that is accessible in the virtual universe. The target locations can either be selected from a list of searchable places or from a map that shows an overview of the whole virtual world. To define the exact target position, a combination of region names and 3D coordinates is used.

4.2.2 Object manipulation

Various interaction techniques for object manipulation are included in Second Life. As users are able to create and define objects with a quite high degree of freedom, interactions are not only restricted to the ones explained above. Hence, a complete description of the implemented features for object manipulation would be beyond the scope of this chapter. Instead, the following section will explain the most important ones.

Selection

In Second Life, each object can be selected by any user. To protect objects from unwanted manipulation, further interactions with selected objects are controlled via a rights management system.

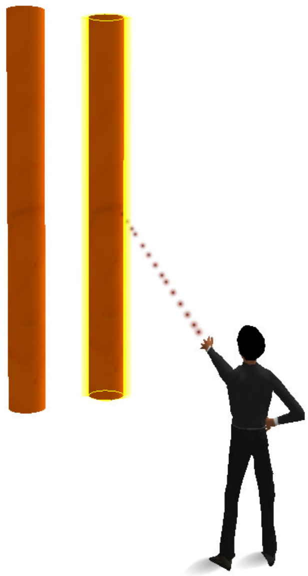


Figure 4.4: Object selection in Second Life

Figure 4.4 shows how selected objects are highlighted. A colored box designates the selected objects and additionally, particles connect the avatar's hand with the respective object. This allows other users to identify who is currently performing actions on an object.

For the selection of multiple objects, several options are available. As in many other programs, standard keys, i.e. Shift and Control, can be held pressed while selecting different objects concurrently. This permits an accurate choice of the items.

If the objects are located close to each other, the computer mouse can be used to define a squared selection area. Each object within this area is selected.

Additionally, objects can be grouped together. This usually makes sense for related objects. The effect is that when the user clicks on one of the grouped objects, the whole group is selected.

Translation

Moving or translation is the most used form of interaction with objects. In Second Life, it is needed primarily to build but also to play games or perform other tasks.

Hence, different concepts of moving objects are realized to satisfy these needs respecting the restrictions of standard input devices.

The most accurate specification of an object's position can be achieved by specifying the coordinates. This permits an exact adjustment of related objects and it is necessary for building. However, this method is not very intuitive and there is no analogy to real world.

A more intuitive kind of translating an object in Second Life is implemented by virtual controls (see Figure 4.5). Thereby, anchors are displayed around the selected object which can be moved along a specific axis. Thus, the object follows the anchor and can be placed more or less exactly. A disadvantage of this approach is that the items only can move along one axis at a time.

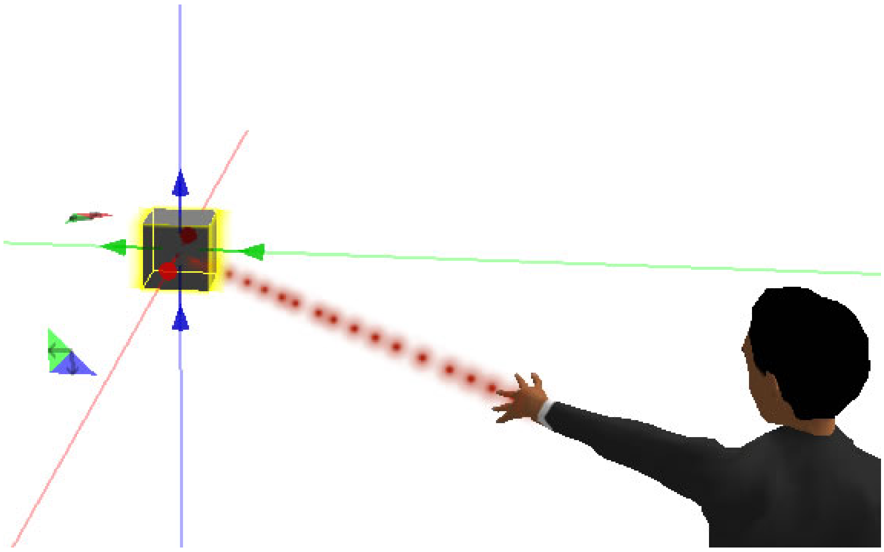


Figure 4.5: Object movement in Second Life

The third method is the most intuitive one. The avatar can grab an object and move it (remotely) to a target position. To perform this form of translation, a combination between the computer mouse and special keys on the keyboard has to be used, because computer mice alone only track two-dimensional motion. Nevertheless, placing objects accurately is quite difficult with this method.

Rotation

Similar to translation, rotation of objects can be performed by specifying the accurate number of degrees or by using virtual controls.

The virtual controls for rotation are similar to the ones for translation. For each axis, a colored circle is displayed around the object. If the user grabs one of the circles and moves the mouse, the object rotates around the according axis.

Other forms of interaction

In addition to the above explained interactions with objects, Second Life includes other forms of manipulating or using objects.

To create a variety of complex buildings, a user can combine several basic geometrical objects, so called primitives, to form a complex object. Each of these primitives has specific properties that define its appearance. By changing the values of these properties, a user can modify the shape or texture of objects.

Furthermore, another manipulation task is implemented: Scaling objects. Similar to translation or rotation, virtual controls facilitate the interaction with objects. To define the size of an object, either the proportions are maintained or just one dimension is changed.

Other interactions between avatars and objects are sitting or attaching. Each object, which provides an adequate surface, is a potential seating for avatars. This special form of interaction binds the avatar to the object. As a result, the avatar undergoes the same manipulations (i.e. rotation, translation) as the object while sitting on it. This method is also used in vehicles which were mentioned above.

In contrast to this, binding an object to an avatar makes it dependent of the actions of the avatar. This means that the object rotates and moves the same ways the avatar does. This method is mainly used to create clothing, e.g. hats.

4.3 Summary of interaction techniques

This chapter focused on different interaction techniques. After a comparison of different classification approaches, a distinction between movement and object manipulation was made. Both categories were presented in detail and the problems and restrictions current virtual environments are facing were identified.

Afterwards, the chapter included an analysis of the implemented interaction techniques in Second Life. It was shown that different movement forms, such as teleports or passive movement, and a variety of object manipulation techniques, e.g. translation and rotation, are available.

To summarize, it can be noted that there are a variety of different interaction techniques in virtual worlds and even with standard input devices, many actions are feasible.

Chapter 5

Animation

Animation plays an important role in virtual interaction, because it is needed to enable social interaction (see chapter 3). Furthermore, other interactions (see chapter 4) are augmented with animations to raise the level of realism by giving the avatar a more lively appearance. Thus, the following chapter concentrates on computer animation methods and gives an overview of existing techniques to animate avatars.

Though computer animation is a very wide field and includes animation of every object in a computer generated scene, the context of this chapter is the animation of humanoid characters, because this is relevant for interaction with avatars. The presented techniques are a relatively small part of the variety of existing methods, but they shall serve to get an insight into this aspect of virtual interaction. Moreover, some of the described methods were used for the creation of the virtual theater play as described in chapter 6.

5.1 Animation principles

As an introduction to this chapter, some basic principles for animation will be explained which were developed in the times of 2D hand drawn animations for Disney films, but are still relevant guidelines for modern 3D animations.

Lasseter (1987) identified 11 principles of traditional animation and applied them to computer animation:

1. **Squash and Stretch** - defining the rigidity and mass of an object by distorting its shape during an action
2. **Timing and Motion** - spacing actions to define the weight and size of objects and the personality of characters
3. **Anticipation** - the preparation for an action
4. **Staging** - presenting an idea so that it is unmistakably clear
5. **Follow Through** and **Overlapping Action** - the termination of an action and establishing its relationship to the next action
6. **Straight Ahead Action** and **Pose-to-Pose Action** - the two contrasting approaches to the creation of movement
7. **Slow In and Out** - the spacing of the in-between frames to achieve subtlety of timing and movement
8. **Arcs** - the visual path of action for natural movement
9. **Exaggeration** - accentuating the essence of an idea via the design and the action
10. **Secondary Action** - the action of an object resulting from another action
11. **Appeal** - creating a design or an action that the audience enjoys watching

Squashing and stretching an object is used for different reasons. One of them is to indicate the rigidity of it compared to other ones. Thus, when a character moves, some body parts are deformed because of gravity. For instance, when an avatar jumps, the extension of its legs is changed.

Timing is an important aspect of animation, especially regarding character animation. Speed of movement can imply different meanings of an action. Regarding avatars, different timings can have different emotional meanings. Slower actions stand for a longer reflection, whereas quicker actions are used for affective behavior. Another function of timing is to indicate the weight of an object. Heavier characters typically move slower than smaller ones.

Preparation of an action, also referred to as *anticipation*, is another aspect that has to be considered. To make actions look less abrupt, a specific time before the proper action a visual indication has to be given. For example if an avatar has to kick something away, the foot has to be pulled back before the kick to anticipate the action. This leads to a better perception of the kick by the viewers. Especially for very quick actions it is important to anticipate them, because otherwise they could be overlooked.

The leading of the viewer's eye so that actions are not overlooked is also called *staging*. Different methods have to be used to make sure that the animation is clearly understandable. As an example, if the focus of the viewers should be on an avatar, it can move faster while the environment stands still.

Like the anticipation of an action, the termination or *follow through* of it has to be kept in mind. This means that after an action has finished, additional animations are needed. For instance, after the avatar has performed the kick mentioned above, the foot has to continue moving some time. Furthermore, leading parts and following parts for an action can be identified. For example when moving a hand, the fingers follow the wrist.

The difference between *straight ahead action* and *pose-to-pose action* refers to the technique of creating an animation. The former means that an animation is created straight from the beginning to its end, whereas the latter means that selected frames are generated first and afterwards the in-betweens. A more detailed description of these techniques can be found in following sections of this chapter.

Arcs and *Slow in and out* are animation principles that have to be considered to get smooth animations. If objects move along arcs, the animation is perceived better than if the object moves along a straight line. Additionally, the motion should be slowed down next to extreme positions of an action.

Another important principle is *exaggeration*. Especially in the context of virtual interaction, animations should be exaggerated to highlight their essence. Thus, emotions can be communicated more effectively. Lasseter (1987) puts it in the following words: "*If a character is sad, make him sadder; if he is bright, make him shine; worried, make him fret; wild, make him frantic*".

The term *secondary action* refers to actions that are subordinated to other main actions. This means that some actions are effects of other ones and less important. For instance when an avatar shows a gesture that expresses its emotions, the facial expressions are secondary. Hence, animating the face does not have to be very complex in this situation, because the viewer's attention lies on the gesture.

Finally, another important animation principle is *appeal*. Animations have to be created so that the viewers like them. Neither a too simple nor a too complex style appeals to the audience and therefore, the reason behind the animation may not be conveyed. An example rule for appealing animations is to avoid symmetry. Each half of a character's body should be animated differently in order to please. Otherwise, the animation looks artificial.

5.2 Introduction to the computer animation process

The above described principles derive from traditional animation, which includes many features that are still valid for modern computer animation. Traditional animation is created by painting single images, also called frames, which show different steps of a motion. If the frames are displayed successively, an illusion of motion is created (Enderton, 2003). Usually, the minimum number of frames per second to avoid that the frames are perceived as single images is 14. The standard for modern animated films is 30 frames/second (Magnenat-Thalmann and Thalmann, 1993). Computer animation builds on the same concept as traditional animation, but some or all frames are generated automatically.

Hodgins et al. (1999) define the general steps of computer animation as follows:

1. Model generation
2. Motion generation
3. Rendering

In the following sections, the modeling and motion generation steps are explained in detail. As the rendering of an animation is mainly needed for computer animated films and the focus of this thesis lies on virtual environments, this final step will not be explained in detail.

5.2.1 Model generation

The basis of the animation of a computer generated character is its model. Thereby, different types of models can be generated to define a virtual body.

Typical types are deformable objects, particles and articulated bodies (Badler et al., 1993).

Deformable objects can be used to model hair or clothing and describe physical features of an object. Thereby, the shape of the surface is important. *Particle systems*, on the other hand are a collection of points. Each of these points can be moved to generate an animation.

The most common modeling technique for virtual characters is *articulated modeling*. This model is a representation of the relevant parts of a human skeleton and includes definitions for the length and rotation possibilities of each body part. Furthermore, the single parts are connected via joints which define points on which the different parts can move and rotate. Figure 5.1 shows a typical graphical representation of an articulated model.

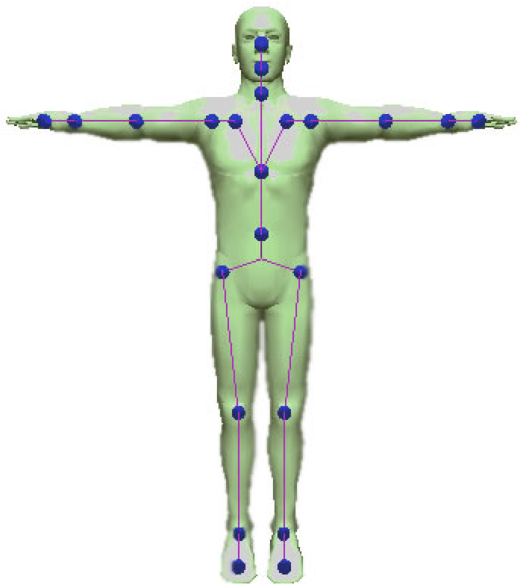


Figure 5.1: Skeleton of an articulated model

The relevant parts of the body are highlighted. Usually, each of the marked points indicates a position for which rotation or movement parameters are stored. Moreover, an articulated model includes a hierarchy of joints. This means that parts in a lower level of the model depend on body parts in a higher level. For instance, when an elbow is stretched, the lower arm and the hand rotate and move accordingly.

Figure 5.2 shows a typical hierarchical order of the different body parts. Changing one element influences elements in a lower section of the tree but none in higher sections or different branches.

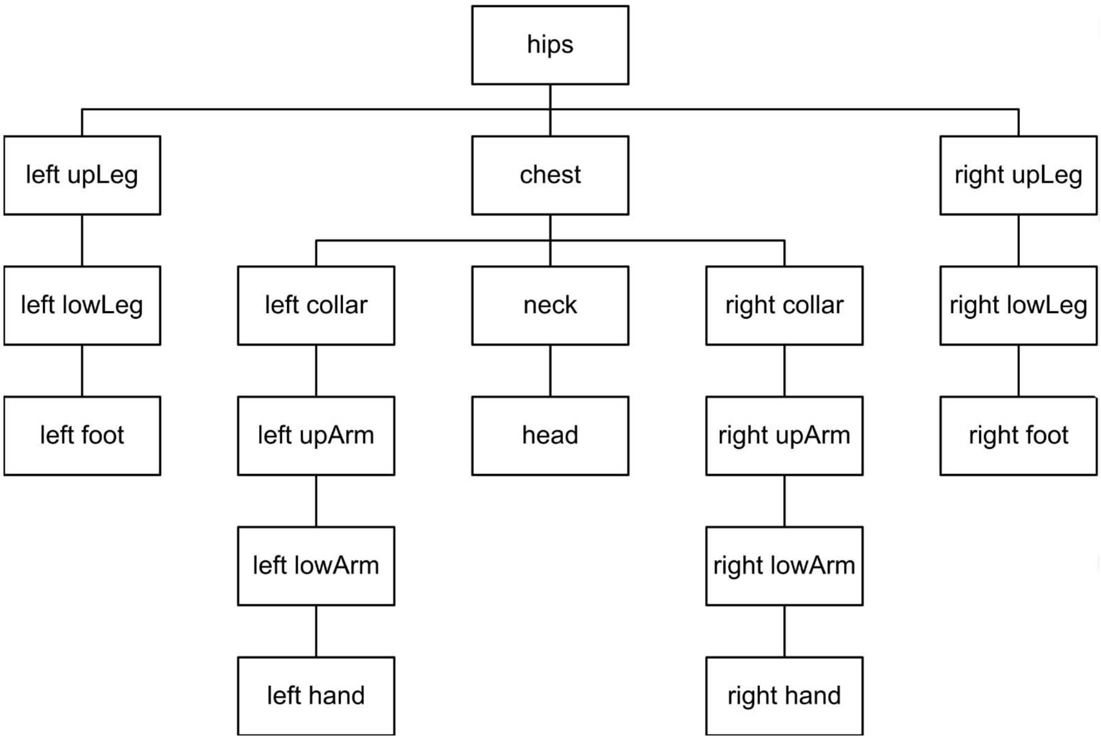


Figure 5.2: Hierarchy of joints in an articulated model

As can be seen, the hips of the figure are the root of the model. This implies that moving or rotating the whole body can only be achieved by changing the respective parameters at the hip point.

To be able to work with an articulated model in computer animation tools, a computer understandable language for representation has to be chosen. One of the most used formats for modern animation tools is *Biovision Hierachical Data* (BVH) (Meredith and Maddock, 2004).

As the name implies, this format represents a hierarchical model. In general, BVH is designed to represent a human skeleton and therefore contains elements named like human body parts (which are also depicted in Figure 5.1 and 5.2). Furthermore, a BVH file consists of one part in which the initial parameters of the different elements are specified (HIERARCHY) and another part, where the motion for each frame can be found (MOTION).

The following code shows a part of a typical BVH file:

```
HIERARCHY
ROOT Hips
{
        OFFSET  0.00    0.00    0.00
        CHANNELS 6 Xposition Yposition Zposition Zrotation Xrotation
            Yrotation
        JOINT LeftHip
        {
                OFFSET   3.29    0.00    0.00
                CHANNELS 3 Zrotation Xrotation Yrotation
                JOINT LeftKnee
                {
                        OFFSET   0.00   −16.57   0.00
                        CHANNELS 3 Zrotation Xrotation Yrotation
                        JOINT LeftAnkle
                        {
                                OFFSET   0.00   −16.55   0.00
                                CHANNELS 3 Zrotation Xrotation
                                   Yrotation
                                End Site
                                {
                                        OFFSET   0.00   −3.30
                                           0.00
                                }
                        }
                }
        }
}
MOTION
Frames:     30
Frame Time: 0.033333
10.87 36.65 13.54 8.70 1.67 91.18 10.89 9.41 −1.80 −2.30 7.50 12.08
   0.00 8.63 13.67
```

The first part usually contains a definition of one single skeleton hierarchy, although BVH files can also contain different skeletons. The ROOT keyword designates the main part of the skeleton. This element is usually the hip as it is the center of the body. Furthermore, the elements of the hierarchy which represent the points of the model are connected with the keyword JOINT.

Each joint includes definitions of the offset, i.e. the difference between the coordinates of the point and those of its parent. Moreover, an additional line defines the channels, i.e. the parameters of each joint. All parts except the root have parameters for rotation along each axis. The root additionally includes parameters for the position of the whole body.

The motion part of a BVH file contains the number of frames, the frame rate and the cannel data. The latter is composed of parameters for each channel as defined in the first part. These data are specified for each frame.

5.2.2 Motion generation

After a model for the object to animate is defined, the next step can be started. Motion generation refers to the process of generating an animation by changing different parameters of the model for every frame.

The definition of the parameters of the model can be a quite complex task. In each frame, the position and rotation of each part have to be set so that the motion looks as wanted considering the basic animation principles explained above. To simplify this task there exist two categories of techniques: *model based* and *motion capture based* methods (Enderton, 2003).

A further classification of model based methods distinguishes between kinematics based, procedural and behavioral methods (Magnenat-Thalmann and Thalmann, 1990).

In *kinematics* or geometric approaches, the definition of motion concentrates on specifying the geometrical parameters of body parts, such as angles or coordinates. Typically, for this approach the method of keyframing is used. This technique will be described in detail in a later section.

Procedural methods generate animations automatically based on algorithms. In contrast to kinematics based approaches, the whole motion is generated by the computer. Thus, physical features of objects can be implemented with this method. The task of the animator is to define the rules of physics which are applied to the animated object in order to simulate behavior in the real world. Hence, additional features, such as mass or elasticity have to be included in the model.

Typical examples of procedural animations in the field of characters are collisions. This means that the computer character is animated in special ways if it collides with other objects. For instance, if the avatar jumps onto the ground, its legs are bended according to the height of the jump and the weight of the character.

Another model based animation concept is *behavioral animation*. Like procedural methods, animations of this kind are fully computer generated and follow steps of algorithms. However, behavioral animation is settled on a higher level and uses artificial intelligence to calculate the necessary animation steps. The generation of the motions is initialized by specifying tasks. Examples could be walking from one point to another or grasping of an object. The algorithms generate the motion sequence according to the task and also take the state of the environment into account. This means e.g. that obstacles are considered.

In contrast to model based techniques, which make use of computational help to generate motions, motion capturing does not necessarily include a model. This technique simply tracks motion data of a real human performer and applies them to a computer character. A detailed description of this approach will be given in section 5.4.

5.3 Keyframe animation

The technique of keyframes originates from traditional animation in Disney films, where the productivity rate of frame generation had to be high (Enderton, 2003). In this era, some skilled animators designed the most important frames of the animation, which are crucial for its composition. The frames in between these key frames were drawn by less skilled animators. Thus, the time and money necessary for the creation of animated films could be reduced significantly.

Nowadays, keyframing is supported by the computer. The animator designs the keyframes and the computer generates the in-betweens automatically. To calculate the in-betweens, the parameters of the animated figure, i.e. rotation, shape, size, etc. are taken into account. This means that the computer calculates the values of each parameter as a value between the start and the end value according to an interpolation[1] algorithm. Basically, two methods of interpolation are possible: linear or spline interpolation.

5.3.1 Linear interpolation

With linear interpolation, the in-betweens are generated by simply dividing the difference between the parameter values of two keyframes into fractions. At each in between frame, one fraction is added to the initial value of the parameter. This generates a linear function:

$$p(f) = p_0 + (f - f_0)\frac{p_1 - p_0}{f_1 - f_0}$$

where p is the value of the parameter and f the number of the frame. The index 0 indicates the previous keyframe and the index 1 indicates the next keyframe.

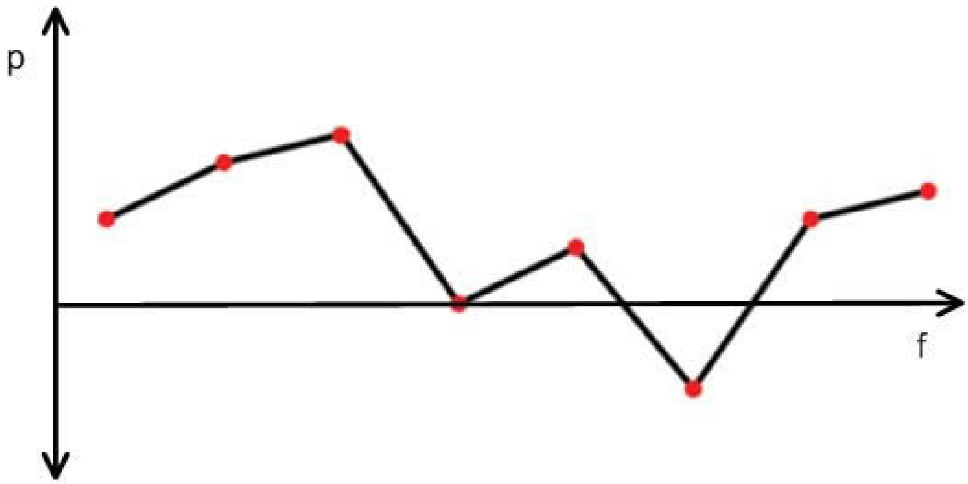


Figure 5.3: Linear interpolation

Figure 5.3 shows one example function of a parameter with linear interpolation. The marked points of the function indicate the keyframes. As can be seen, the curve is not very smooth and therefore, the movements can appear abrupt. This is a violation of the principle of arcs explained in section 5.1.

[1]Interpolation is a mathematical method to calculate a function which goes through a set of discrete data points.

5.3.2 Spline interpolation

To get a smoother animation, the parameter values of the frames between the keyframes have to follow a smoother function. In mathematics, different interpolation methods have been developed to generate a better approximation to complex functions, e.g. quadratic or cubic spline interpolation (Bartels et al., 1998).

An example of a wide used cubic interpolation method is the *Hermite spline*. This approach uses four basis functions that are used to generate a polynomial function between two keyframes.

The basis functions are defined as:

$$
\begin{aligned}
&h_1(t) = 2t^3 - 3t^2 + 1 \\
&h_2(t) = t^3 - 2t^2 + t \\
&h_3(t) = -2t^3 + 3t^2 \\
&h_4(t) = t^3 - t^2 \\
&t \in [0; 1]
\end{aligned}
$$

Thereby, t indicates the frame number relative to its neighbor keyframes. The predecessor keyframe has the number 0 and the successor the number 1. Each frame in between has therefore a value between 0 and 1.

To calculate the parameter values for each frame, the values of the two neighbor keyframes are needed. Additionally, tangent vectors are needed to define the border gradient for the functions. An easy and effective method to calculate the vectors was developed by Catmull and Rom (see DeRose and Barsky, 1988). Thereby, the vector on a specific keyframe is the average of the parameter values of the previous and next keyframe. The mathematical notation of this method is:

$$m_i = \frac{1}{2}(p_{i+1} - p_{i-1})$$

M represents the vector on the keyframe i. P is the value of the parameter at the respective keyframe. The effect of the vectors is that the segments between the keyframes are connected smoothly.

Finally, all of the components are combined to generate a polynomial spline function:

$$p(t) = h_1(t)p_0 + h_2(t)m_0 + h_3(t)p_1 + h_4(t)m_1$$

where the index 0 indicates the previous keyframe and the index 1 indicates the next keyframe.

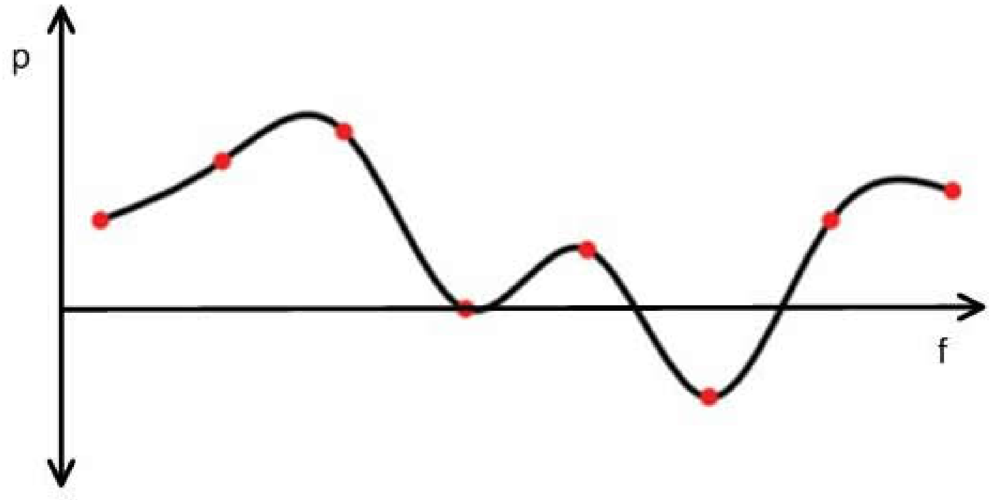


Figure 5.4: Hermite spline interpolation

Figure 5.4 shows the same keyframes as in the previous section. In contrast to linear interpolation, spline interpolation generates a far smoother function. Therefore, modern keyframe based animations usually use spline interpolation for the creation of the in between frames.

Kochanek and Bartels (1984) extend the Hermite spline interpolation for keyframe animation. They include additional parameters to influence the spline, because in many cases the curve has to be more exaggerated or straighter than the normal curve in order to generate a more realistic look of the animation. For this fine tuning, their approach allows the specification of tension, continuity and bias. These or similar parameters are also included in modern character animation software.

5.3.3 Additional issues for skeleton models

As mentioned earlier, articulated hierarchical models are common for animation of virtual characters. However, this representation brings up additional problems for keyframe animation. Due to the setup of a skeletal hierarchy, the positioning of extremities is a difficult task. When an animator e.g. wants to move the foot of the character forward, first the hips have to be moved, then the upper leg, the lower leg until finally the foot can be positioned as wanted. This approach is called forward kinematics and can be quite time consuming.

A better positioning of the body parts can be achieved with inverse kinematics (Zhao and Badler, 1994). This approach was originally developed for robots and nowadays aids the animation of human models. Thereby, the animator specifies the positions of the extremities and an algorithm computes the joint angles for the other body parts automatically so that the constraints of the model are preserved. Thus, the process starts in the lowest levels of the hierarchy and the higher levels are computed recursively.

5.4 Motion capturing

In contrast to keyframe animation, motion capturing follows a straight ahead approach. This means that the frames are created straight from the beginning to the end of the animation and no special keyframes are designated. The main idea of motion capturing is the recording of real human motion which is then transferred to a computer character.

First approaches in this area were developed in the time of traditional animation. At that time, animators filmed the motion of a real performer and projected it onto a glass plate. So they could trace the motions of the performer for the drawn characters. This method is also referred to as *rotoscopy* (Magnenat-Thalmann and Thalmann, 1990).

Modern motion capturing enables a direct transfer of the human motion to the computer. The advantage of this approach is that frames do not have to be created manually and thus, very complex animations can be produced, which would be expensive with other methods, such as keyframe animation.

The general process of motion capturing includes not only the recording of the motions, but also a detailed editing of the resulting data. Gleicher (2000) identifies the steps of the motion capturing process as follows:

1. Plan the motion capture shoot and subsequent production.
2. Capture the motion.
3. Clean the data.
4. Edit the motions.
5. Map the motions to the animated characters.

This shows that an intensive planning of the needed motions should be done before the capturing. Hence, the human performer has to know well how the actions should look like. Moreover, after the capturing it is necessary to clean the data. Errors in the capturing process have to be identified and corrected. If some motions do not look as wanted, they have to be edited manually. Finally, the cleaned and edited data can be transferred to the computer character model.

The capturing step in the process can be accomplished with several techniques. Each of them has advantages and disadvantages, but all lead to appropriate results. The three possible capturing techniques are *mechanical, magnetic* and *optical* systems. Each of them will be described in detail in the following sections.

5.4.1 Mechanical systems

The oldest form of tracking of human movements is mechanical tracking. Thereby, a mechanical skeleton is attached to the performer. During the movements, the construction measures the angles of rotation or the position of the body parts and stores the data for further use. A disadvantage of this method is that the mechanical sensors have influences on the mobility of the performer and thus cause problems for some movements. However, the technique is relatively cheap compared to other ones.

Although this method is usually not used any more for the capturing of whole body movements, there are still applications of this concept for the animation of special parts. One example is *CyberGlove* (Kessler et al., 1995), which uses mechanical tracking to capture the motion of a human hand.

5.4.2 Magnetic systems

Magnetic systems work with electromagnetic sensors which are attached to the most important joints of the human performer. Additionally, there is a transmitter that emits signals which are received by the sensors. Thus, the spatial position of each sensor can be calculated. Due to the fact that each sensor is independent from the other ones, the data received has to be interpreted separately after the tracking step and related to joints of a character model.

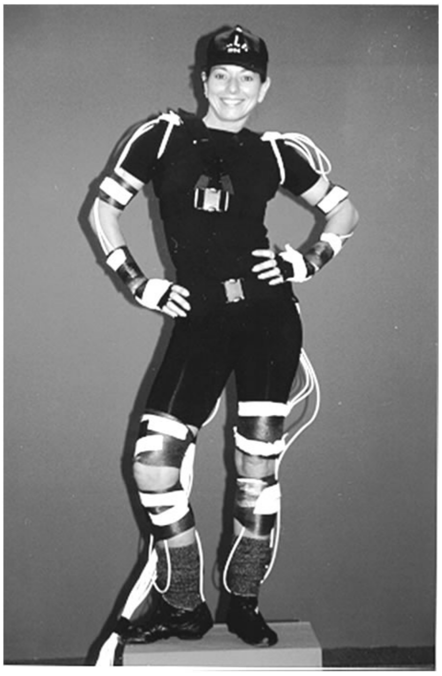


Figure 5.5: Magnetic motion capture system

Figure 5.5 shows a configuration of magnetic sensors on a performer. 18 sensors are placed on the relevant joints to allow a clear interpretation of the motion.

5.4.3 Optical systems

Optical motion capturing systems include approaches with and without markers. The former methods are similar to magnetic systems, because they use reflecting markers that are attached to joints of the performer. The motions are recorded by a set of high speed cameras which are positioned in different angles to be able to record all sides of the body. Afterwards, software processes the recorded sequence by identifying the markers and relating them to a character model.

Herda et al. (2000) point out the importance of the character model, because the tracking of the markers is often difficult as some are occluded by other body parts. Hence, a detailed skeleton model can help to calculate the possible positions of each marker automatically. The calibration of the model according to the performer's anatomy can be achieved by a set of special movements which involve all major body parts.

Another optical motion capturing approach is markerless tracking. This facilitates the capturing, because no markers and no special cameras are needed. However, the absence of markers makes it difficult for the software to identify the performer's body parts. To solve this problem, a variety of approaches have been developed. (For further information see Kehl (2005)).

5.5 Summary of computer animation

An important basis for virtual interaction are computer animation techniques, which were described in this chapter.

At the beginning, a list of principles of animation, which originates from traditional hand drawn animation, was given. This enumeration included a multitude of useful guidelines, which can be applied to modern computer animation. Especially the guiding of the viewers eyes and the exaggeration of movements are concepts to keep in mind.

Thereafter, the general process of computer animation was explained. Articulated hierarchical models were presented in detail, as they are the standard for many character animations and also used in the practical part.

As far as the generation of motions is concerned, this chapter compared the two most common techniques: Keyframe animation and motion capturing. Thereby, different interpolation algorithms were explained, which are the basis for keyframe techniques. It was shown that spline interpolation is the standard for the animation of characters, because it generates the smoothest motions.

As regards motion capturing, in the last part of this chapter, different methods were compared and their advantages and disadvantages were identified.

Chapter 6

Virtual theater

The practical part of this thesis consists of the organization and performance of virtual theater plays. The goal of this project is to analyze virtual interactions between the actors and investigate in how far these interactions are perceived by the audience. Theater plays are a suitable basis for the analysis of interaction, because a crucial part of their idea is to transmit interactions between actors in form of body language and dialogues. A theater play without interplay of the character's actions is not likely to please the audience. Furthermore, the limitations of modern virtual worlds are examined and it will be shown which restrictions for interaction were faced during the implementation of the plays in Second Life.

As the thesis mainly focuses on technological aspects, the analysis does not include questions about theatrical issues. The aim is to examine the potential, but due to the constraints of resources, a high-quality drama performance was never strived for.

For these reasons, the following chapter describes the whole process of creating and performing two selected short theater plays. Each step from the selection of the plays to the final performances will be illustrated to give an impression, how such a project can be conducted. After this, the findings related to virtual interaction are given.

6.1 Similar projects

The idea of performing virtual theater plays is quite new and related to the development of virtual worlds in the last years. However, similar ideas were discussed earlier and generally considered the enhancement of real theater plays with virtual elements. To have an impression of these ideas, some of the projects will be presented.

6.1.1 Augmented Performance

Several projects have been carried out which combine real performances with virtual reality elements. A crucial characteristic of these performances is that each element is generated in real time. This means that there are no prerecorded virtual sequences. All virtual elements are controlled and influenced live by the actors (Evert, 2002).

One of these theater projects was conducted by Sparacino et al. (1999). Their *Improvisational TheaterSpace* included computer agents which played a role in an emergent theater play. [1] As regards the virtual part, media actors were used. These actors were software agents, which were represented as text.

Figure 6.1: Performance with a text based virtual actor

The idea of the theater play was to create a dynamic interaction between a human actor and a text based virtual actor who expressed e.g. the human actor's

[1] The term emergent means that the script of the play is not fixed, but can be varied in a certain degree in order to allow the actors to influence the story.

thoughts. This agent was controlled by the actor's voice or gestures. The text messages were projected onto a screen on stage, which allowed both the actor and the audience to read them. Thus, the text was designed as another actor on the stage. Figure 6.1 shows impressions of the performance.

This project showed an extension of a typical theater play by virtual agents, whereby the main idea was to create the script dynamically. This allowed a flexible design of the virtual elements and reduced the probability of errors.

Another project which consisted of both a human and a virtual actor was conducted by Pinhanez and Bobick (1998). They developed a special theater play named *"It/I"* where "It" is played by an autonomous agent and "I" is played by a human actor. In contrast to the previously presented project, *"It/I"* is based on a given story which is defined with a special scripting language. Hence, the human performer has to follow the script in order to get the planned sequences. Nevertheless, the actions of the computer character depend on the actions of the actor. As the play is designed as a pantomime, the gestures of the actors are tracked and used to control the agent. Figure 6.2 shows how the human actor interacts with "It".

Figure 6.2: Performance of the play It/I

The computer character "It" can communicate in the form of images which are projected onto a screen on the stage, music and stage lights which allows the representation of interaction between the human and the computer generated actor.

But the concept of the play has disadvantages. During the performances in November of 1997, some of the gestures of the human performer were not recognized correctly. Due to the fact that the script relied on a correct sequence of gestures, they had to be told manually to the agent in order to make sure that the play could be finished correctly. This shows that a reliable tracking of the gestures is crucial for this type of performance.

6.1.2 Virtual puppet plays

Another type of virtual theater performances is similar to puppet plays. In this kind of performances, the whole play usually takes place in a virtual environment. In contrast to augmented performances where the virtual part is played by computer agents, these projects only include virtual characters, i.e. avatars which are controlled directly by users. Thus, this type is compared to puppet theater, where human actors control puppets on a stage. In the virtual form, humans control avatars on a virtual stage.

In the play "*Incarnation of a Devine Being*" (Jalkanen, 2001) actors at different locations were participating via tracking input devices in an ancient Greek play. Thus, they interacted and communicated with other actors in real time on the same virtual stage. The representation of the other actors assumed the form of 3D avatars which performed the same movements as their human players. As a result, an interaction between the actors took place as if they were in the same room.

A similar project was carried out by Matsuba and Roehl (1999) who set up a virtual theater play with the virtual reality modeling language (VRML). As a template they chose Shakespeare's Midsummer Night's Dream and designed a virtual 30 minute version of this play. The project with the name "*VRML Dream*" is considered the first live streaming entertainment project of this type with duration of more than 2 minutes (Matsuba, 2000).

Due to the fact that the performance should be accessible to a wide range of users, the data size had to be kept as low as possible in order to be accessible with a 28.8K modem. Hence, the character models generated with VRML were designed with a low number of polygons. As far as the motions of the characters are concerned, motion capturing systems were used which streamed the motion data directly to the internet. Additional to the performers for the motions, other human actors were responsible for the dialogues. For this purpose, they used third party software to stream the audio data over the internet.

Figure 6.3: Performance of VRML Dream

Figure 6.3 shows a screen capture form a performance of VRML Dream. Each of the virtual actors was played by one or more humans in real time.

6.1.3 Theater performances in Second Life

Second Life as a virtual 3D online world offers all important features for theater performances. Thus, several projects were conducted where types of virtual or augmented performances in Second Life were created. In each of these performances, humans controlled avatars as virtual actors in real time.

One example of the performances in Second Life is a virtual concert of the band U2, which was carried out several times by some fans since 2005[2]. They created avatars which resembled the real members of the band, a virtual stage and animations for singing and stage performance of the avatars. With these means, the team gave several virtual concerts, where songs of U2 were streamed into Second Life and the actors performed via their avatars. Figure 6.4 shows an impression of the performance.

[2]For detailed information see http://www.u2insl.com/

Figure 6.4: Virtual concert of U2

A similar project is *Second Life Ballet*[3], where virtual ballet plays are performed. The focus of this performance lies on the animations as they are crucial for dancing.

In addition to these projects, there are also performances in Second Life which gear more towards a drama-like play, such as the play "*From The Shadows*" by the marketing agency *Millions of Us*[4]. The story is about a crime which was presented on a typical theater stage built in Second Life. The voices of the actors were streamed via an audio server, because at the time of the performance, voice chat was not yet included in Second Life.

But not only puppet plays have been staged in Second Life. Augmented performances were also tried out, as a recent example of the theater house "*Schaubühne*" in Berlin shows[5]. They designed a performance of Alice in Wonderland with avatars in Second Life and a human actress on a real stage. The play was designed in such a way that the performer acted as if she was inside the virtual world and could interact with the avatars. In this performance, the virtual actors had no voices, but communicated via gestures and text chat. The views of the avatars were projected onto several screens on the stage so that the audience could follow the play from different perspectives.

[3] http://slballet.org/

[4] http://www.millionsofus.com/

[5] For detailed information about the play see http://slwunderland.wordpress.com/

6.2 The virtual performance

The project conducted in the course of this thesis combines several aspects of the presented similar works and deals with a performance of the type of a virtual puppet play. The reason behind this decision is that interactions in virtual spaces should be examined with those resources which are available to a broad range of users. As a result, the input devices for the actors were limited to a keyboard, a computer mouse and a headset. Furthermore, each of the actors was played by laypersons that did not have any special experiences with virtual worlds and theater performances. These preconditions enable an analysis which is valid for average users and led to the decision for a puppet play like performance, because this type does not require much talent for acting and is possible with standard input devices.

For a detailed understanding of the project, a description of all steps and decisions is helpful. Hence, figure 6.5 shows a process diagram of the theater project.

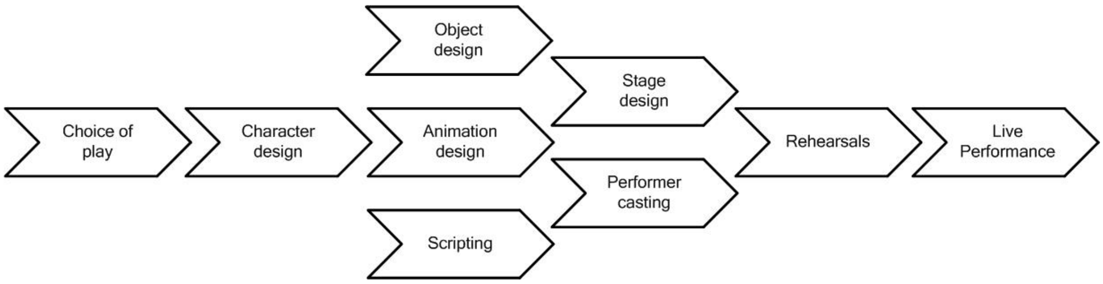


Figure 6.5: Process diagram of the theater project

Each step will be explained in detail in the following sections.

6.2.1 Choice of the play

A crucial decision at the beginning of the project was the choice of the play. Due to the fact that no scriptwriter was available for the project and none of the project members had experience in writing theater scripts, the decision for the selection of an existing play was made. Thus, several criteria had to be considered in order to make sure that the project was feasible within the given time frame and resources.

As since the beginning of the project it has been clear that the creation of the animations would take the most time, the play should not contain too complex motions. On the other hand, the play had to contain gestures and motions, because simple dialogues would be boring for the spectators and the examination of interaction required body language. As a result, the final play had to be performable with a small set of predefined animations, which represent the body language and interactions with objects.

Another important factor was the duration of the play. Due to the limitations regarding preparation time and resources, the selected plays had to be quite short. A duration between 5 and 10 minutes seemed appropriate and sufficient to reach the research goals. Moreover, the appeal to the audience, which was also an important factor, could decline if the play was too long. Likewise, the selected plays had to be entertaining.

Finally, an additional goal of the project, which influenced the decision for the play, was the potential for audience interaction. The play had to contain extra roles which could be played by users from the audience after a short briefing.

Taking into account all of these requirements, a first preselection of possible plays was conducted. A set of plays, which fulfilled most of the preconditions, was "*How to irritate people*" by John Cleese[6]. This set of short sketches was published in 1968 for television and was performed by actors who later formed the British comedy group Monty Python.

Out of the 14 plays in this set, two of them were selected for the further work, because they fulfilled all of the requirements. In the original work they are called "*Opportunistically Irritating*" and "*Gruellingly Irritating*". For the sake of easier and better memorable reference, these sketches will henceforth be referred to as "*Job interview*" and "*Indian restaurant*", respectively.

To give an impression about the contents of the two plays, a short synopsis of both will be given.

[6] see http://www.imdb.com/title/tt0063100/ (retrieved 14-Nov-2007)

Job interview (Opportunistically irritating)

This sketch is about a job interview where the interviewer exasperates the applicant by asking silly questions and behaving strangely. The interviewer first confuses the applicant with questions about greeting and takes notes about his reactions as if he evaluated him. Then, the actions of the manager get weirder and weirder while he continuously takes notes. Finally, some people enter the office who rate the reactions of the applicant with point cards, as if they were skating judges.

Indian restaurant (Gruellingly irritating)

The second sketch, which will only be partly considered in the following sections, takes place in an Indian restaurant. A couple enters the restaurant and the manager treats them with exaggerated care. He punishes himself every time he makes a mistake and in the end, he even jumps into an oven.

6.2.2 Character design

The character design was the next step after the decision for the plays, because some of the animations and requisites depended on the shape and size of the avatars.

One goal of the design of the characters was to make them look quite similar to the original actors of the play. Although in Second Life it is possible to create avatars with complete different styles, such as fantasy creatures, the intent of this project was to use humanoid avatars, because they seemed to be the best choice for the examination of interaction (see also chapter 3). Nevertheless, no restriction for the look of the audience roles was set, because these users should be able to perform spontaneously without having to change their appearance.

Figure 6.6(a) shows the avatar for the role of the interviewer and figure 6.6(b) the one for the applicant.

Every part of the avatars was designed manually, which means that no predefined parts were used. A special part of character design in Second Life is the clothing,

(a) Interviewer (b) Applicant

Figure 6.6: Avatars for the job interview play

which can be produced by creating a texture which is then mapped to clothing objects, that are treated like avatar body parts, i.e. they can be customized with a set of parameters.

As can be seen, the character of the applicant has a thinner and bonier shape compared to the interviewer. This difference should allow on the one hand a better differentiation between the two characters, and on the other hand emphasizes the personalities of the two roles.

6.2.3 Animation design

As explained in section 3.3.1, animations are crucial for the representation of body language. Thus, a high level of attention was given to the creation of the animations for the play. The choice of the plays already took the amount and kind of necessary animations into account and led to the decision that a small set of movements which could be used several times during the play had to be created.

A very important question was how the animations should be executed. The initial approach was to develop a type of virtual control in form of a HUD[7], which resembled the graphical user interface of Second Life. This display contained buttons for each animation and a click on each of it started the respective one. Figure 6.7 shows the control of the applicant.

Figure 6.7: Virtual control for animations of the applicant

Another initial question was which animations would be needed. A first idea was to use already available animations, which are either included as standard in Second Life or were created by other users and are freely available. However, after a detailed analysis of these animations, it was concluded that they were unusable for the selected plays. Most of the animations were dances or other movements, which did not fit into the concept of the plays. Nevertheless, a few of the standard animations in Second Life could be used, such as the laughing and the bow animation.

Animation	**Actor**
Write at notebook	Interviewer
Invite Applicant to take place	Interviewer
Shake head	Interviewer
Laugh	Interviewer
Ring bell	Interviewer
Make face	Interviewer / Applicant
Fix the suit standing	Applicant
Straighten the pants sitting	Applicant
Stare rapidly at Interviewer	Applicant
Bend over desk	Applicant
Dispair	Applicant
Wonder	Applicant
Hold cards up	Evaluators
Bow	All

Table 6.1: Animations of the Job interview play

[7]The Heads-up display (HUD) is a custom designed user interface, which can be attached to avatars.

Due to these findings, it was made the decision that the majority of the animations would be created manually. For this purpose, each of the two plays was analyzed first in order to get a list of required animations. Table 6.1 shows an overview of the animations which were decided to be necessary for the job interview.

Each of the actors had a set of 7 - 8 animations. This number was fixed due to the fact that a higher number seemed to be difficult to manage for the performers.

To give the performance a certain degree of freedom, the "make face" animation was implemented in two different ways. The actors should be able to select spontaneously between these alternatives, which allowed them to improvise a bit.

Another question was which technique should be used for the creation of the animations. As it was clear that the animation generation would be a quite time consuming task, the use of motion capturing was considered. This technique would simplify this step significantly, as described in section 5.4. However, after a survey of the available techniques, it was seen that none of the products was affordable with the available resources. Even though there is cheap software for optical tracking, the technical equipment necessary, i.e. high-speed infrared cameras and reflectors, is quite expensive. On the other hand, programs, which allow motion capturing with just two standard video cameras, are costly themselves.

For these reasons, another approach, which is also described in chapter 5 was used. A keyframe based animation tool named QAvimator[8], which was developed especially for the generation of animations for Second Life avatars, was the basis for all animations created for the play. Figure 6.8 shows a screenshot of the program.

The software is a simple tool to generate BVH files. Each part of the skeleton can be rotated along each axis, which is indicated on the right part of the screen. The left part contains on the one hand a graphical view of the avatar body and on the other hand a timeline which shows all body parts and their respective keyframes. The lines between the keyframes represent the interpolated states. Additionally, the program allows defining the number of frames, i.e. the length of the animation and its speed, i.e. frames per second. Other functionalities, such as inverse kinematics are not included.

[8]http:// www.qavimator.org/, retrieved Nov-15-2007

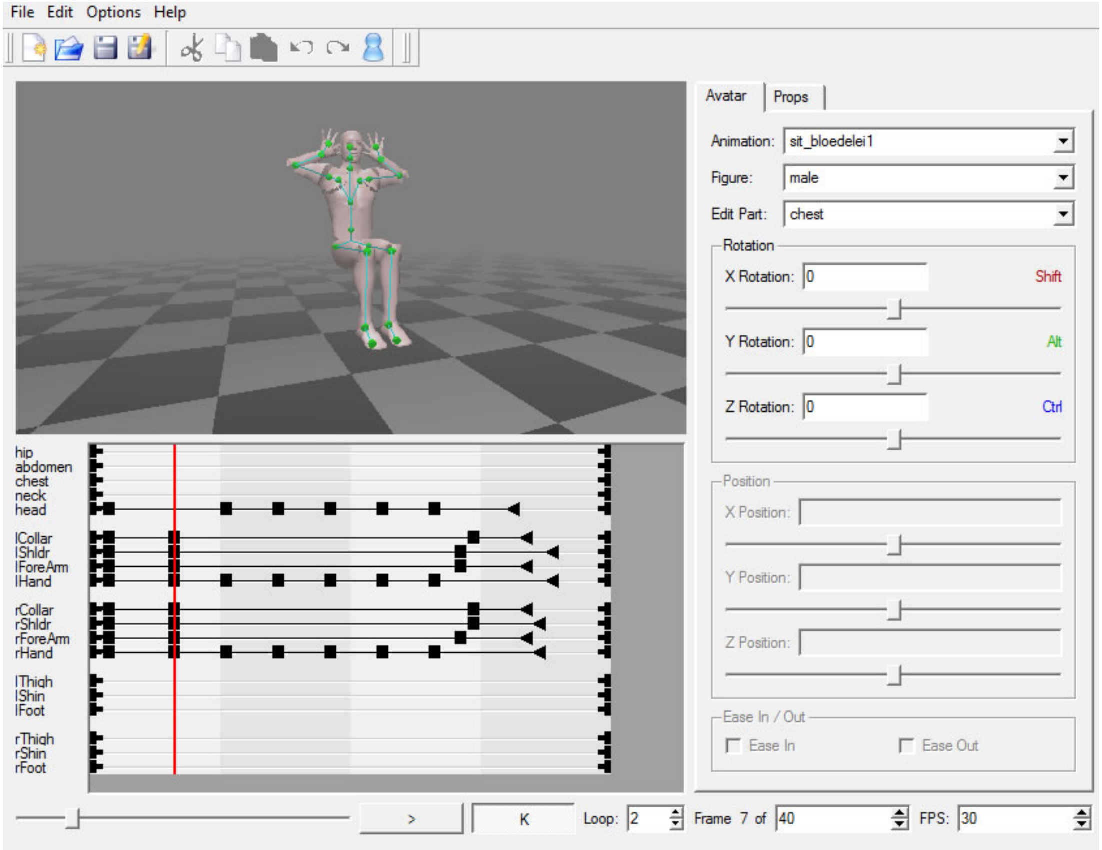


Figure 6.8: Screenshot of QAvimator

After the animations were created with QAvimator, they had to be uploaded to Second Life in order to be used for the performances. Each upload in Second Life costs 10 Linden Dollars which made a detailed testing before the upload necessary. After a BVH file is uploaded, some important parameters have to be specified. Figure 6.9 shows the screen for the animation upload.

As already explained in chapter 3, the mimics and hand gestures of an avatar have to be chosen from a set. Unfortunately, the facial expressions did not work correctly after the upload. For this reason, the created animations did not contain any facial expressions and emotions had to be expressed via exaggerated body movements. Another important parameter is the priority level. This number defines which animations override others if they are played concurrently. For all created animations, the highest priority was used to make sure that the built in idle animations were overridden.

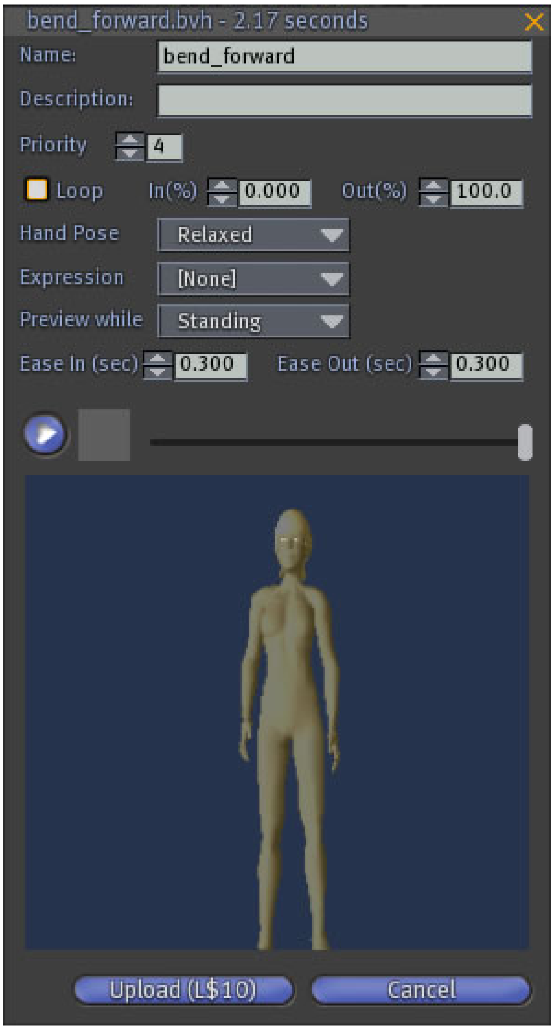


Figure 6.9: Upload screen for animations in Second Life

6.2.4 Scripting

To enable a simple activation of the animations, several controls had to be created with the built-in scripting language LSL. Each actor should have a control bar for all animations which were needed. This HUD simply handled touch-events and executed the respective animation.

In addition to the virtual controls for the actors, the job interview play called for the creation of grading cards for the evaluators. These cards should contain more functionality than just playing an animation. Due to the requirement that the evaluators should be able to show cards without needing to be teached how to use the controls, the cards had to do most of the actions automatically. This means that the performer should just have to execute a command and to specify a 2-digit number for the evaluation. The script had to display the cards, attach them to the avatar, play a hold-up animation and display the correct grades.

The design of the application for the grading cards is limited by the capabilities of LSL. One root object handles the commands executed by the user and shows (in LSL terms "rezzes") the two cards. Furthermore, it starts the hold animation and propagates the grades to the cards. The following code snippet shows the source code of the root object.

Listing 6.1: Script of root object

```
default
{
    state_entry()
    {
        llListen( 55, "", NULL_KEY, "" );
    }

    listen( integer channel, string name, key id, string message )
    {
        if(message=="stop"){
            llStopAnimation("stand_taferl3");
        }
        else{
            llRezObject("Ka_right", llGetPos() + <0, 0, 7>, ZERO_VECTOR, ZERO_ROTATION, 123);
            llRezObject("Ka_left", llGetPos() + <0, 0, 7>, ZERO_VECTOR, ZERO_ROTATION, 123);
            llSleep(0.2);
            llShout(80, message);

            llRequestPermissions(llGetOwnerKey(llGetKey()), PERMISSION_TRIGGER_ANIMATION);
        }
    }

    run_time_permissions(integer perm) {
        if (perm & PERMISSION_TRIGGER_ANIMATION)  {
            llStartAnimation("stand_taferl3");
        }
    }
}
```

As can be seen, the cards are created at height 7 (see lines 14 and 15). The reason for this is that Second Life's rights system requires the confirmation of permissions to attach objects to an avatar. As a result, the cards are shown floating in the air before the user confirms the attachment. To hide the floating cards, they are positioned 7 meters above the avatar so that the audience cannot see them.

Each card attaches itself automatically to the avatar and listens on a special channel for the command and grades. Each digit is predefined on a specific texture which is selected accordingly. The following snippet shows the code for the cards[9].

Listing 6.2: Script of the left card

```
1  default
2  {
3      state_entry()
4      {
5          llListen( 80, "", NULL_KEY, "" );
6      }
7  
8      listen(integer sender_num, string num, key id, string
           message)
9      {
10         if (llGetSubString(llToLower(message),1,1) == "0")
11         {
12             llSetTexture("54b6f6f8-abe9-e0fc-1fe8-285547654efc"
                   ,2);
13         }
14 
15         llSleep(15);
16         llRequestPermissions(llGetOwnerKey(llGetKey()),32);
17     }
18 
19     run_time_permissions(integer perm) {
20         if (perm & PERMISSION_ATTACH)  {
21             llDetachFromAvatar();
22         }
23     }
24 }
```

After a time of 15 seconds, which is the duration of the animation, the cards are detached automatically from the avatar (see lines 15 and 21). This allows a

[9] To shorten the snippet, only the texture of the digit 0 is printed.

repeated execution of the command, where in the meantime the cards are not displayed.

6.2.5 Stage and requisite design

Before the first rehearsals could be conducted, a last preparation step was missing: the creation of the stage and the objects. Although in Second Life almost every location for the performance is thinkable, this project was performed in a theater setting. This means that a small theater house was designed with an audience room that contained benches to sit down and watch the performance. The reason for this is that this design should create a similar atmosphere as in real theaters, i.e. the audience is separated from the actors who are on the stage. The advantage of this concept is that the audience is less likely to disturb the performance and the view angle can be defined. A similar approach is also used in other projects as explained in section 6.1.

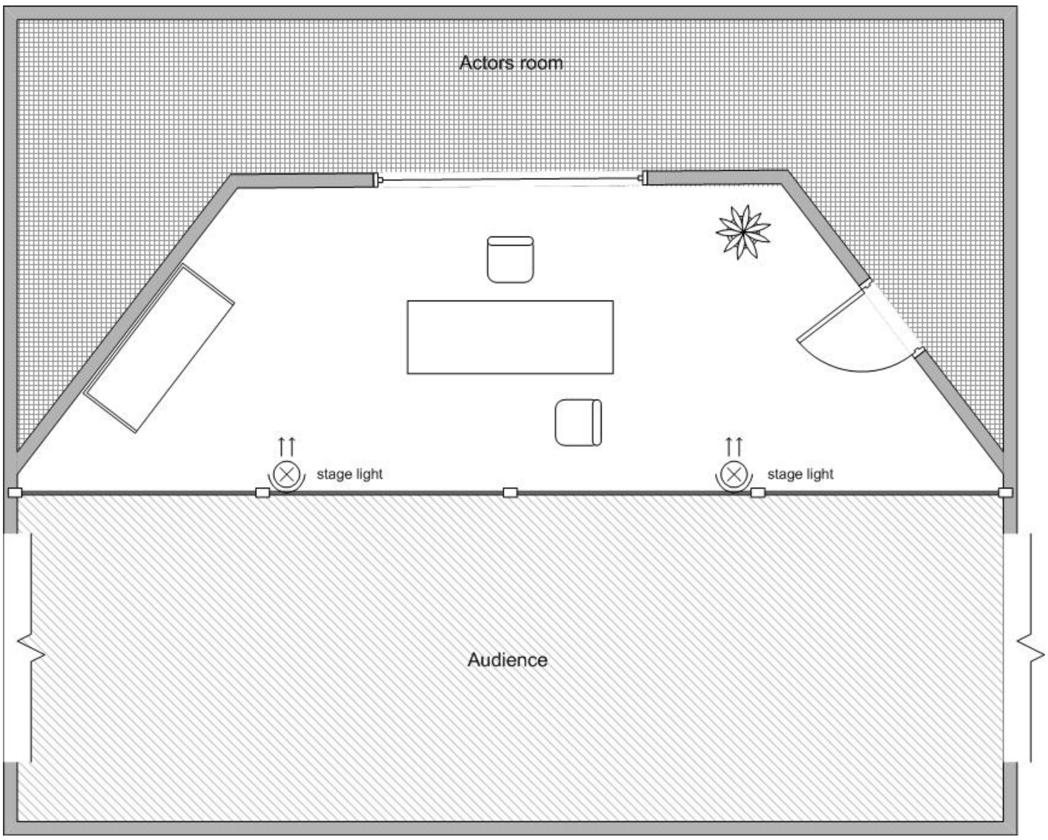


Figure 6.10: Sketch of the stage

The design of the virtual stage was planned in a way that the audience could see each aspect from the seats. As a result, e.g. the side walls were installed diagonally. Figure 6.10 shows a sketch of the stage design. Moreover, the virtual stage contained all props which were necessary for the play, e.g. a desk with two chairs and a small bell on it.

Figure 6.11: The stage in Second Life

In addition to the props, the stage contained a set of lights which were necessary to cope with the automatic daylight changes in Second Life. Furthermore, a backstage area was defined as a room for the actors waiting for their entry and for the direction of the play. To be able to watch the performance from this area, the walls were transparent on the back side. Figure 6.11 shows the final design of the stage in Second Life for the job interview play.

6.2.6 Rehearsals and live performance

As in every theater play, in this project rehearsals were necessary to practice the timing and interactions of the actors. Thereby, especially at the beginning, some problems appeared which had not been considered before.

One problematic factor was the technical infrastructure. The client of Second Life requires a more or less modern PC with sufficient memory and performance. Hence, each of the performers had to have access to a computer which fulfilled the requirements. Additionally, a headset or at least a microphone in combination with speakers was required for the audio streaming. Some performers had problems due to limitations of the performance, a driver problem of the headset or a low network bandwidth. This delayed the first rehearsals markedly.

Another issue that came up during the initial trials was the absence of a detailed sequence of instructions for the play. Although each actor got a script of the play including the dialogues, it seemed to be difficult to identify the correct sequence of animations. To solve this issue, a storyboard of the play was created which included a detailed description of the sequences, actions and animations for each character. In addition to this, a graphic illustration showed the content of each sequence. The whole storyboard of the job interview can be found in appendix A.

As far as the live performances are concerned, several users came to the virtual location and took part. Before the beginning, some of them were asked to take part in the play by performing the role of the evaluators. Thereby, the instructions were rather short, i.e. the supporting actors only got the cards and instructions how and when to use them.

The resulting performance was received quite well by the audience. Those who had seen the original play before cited that the virtual version was very similar to it. Nevertheless, some restrictions were noticed, which will be described in the following chapter.

Finally, to get an impression about the performance, figures 6.12 and 6.13 show some photos of the job interview play.

Figure 6.12: During the job interview

Figure 6.13: The members of the project as audience of the performance

6.3 Interaction issues

During the rehearsals and the live performance of the virtual theater play, potentials and issues regarding interaction were analyzed and evaluated. The observation showed that virtual theater performances are viable in current CVEs, but underlie several restrictions which can impair the appeal to the viewers compared to performances in reality. Thus it can be implied that applications in virtual environments in general have to cope with limitations and it has to be kept in mind that currently, interactions in virtual worlds cannot take place in the same extent as in reality.

The interaction techniques used in the virtual theater performance were evaluated in two different ways. On the one hand, the rehearsals and performances were recorded and later analyzed and on the other hand, the actors gave feedback about their options for acting. This resulted in a list of problems, which not only is relevant for the conducted project, but also might be applied to other application areas and to other virtual worlds:

- Missing mimics/lip synchronization
- Control of gestures with standard input devices
- Limited self-perception
- Delays

6.3.1 Missing Mimics

One striking issue regarding the interaction between the actors was the missing of facial expressions and lip synchronization for speech. Especially when an actor said something, it seemed that an important aspect of the communication was missing. Currently, Second Life in general does not include standard animations for the lips, which are played when a user uses the voice chat. Instead, an icon which is displayed above the head of the avatars indicates if the respective user is currently speaking.

Figure 6.14 shows an impression of the performance where the interviewer talks to the applicant. The icon above the avatar shows, that this character is currently

Figure 6.14: Missing lip synchronization makes the avatars look like puppets

speaking, but the face does not show any evidence of communication. As a result, the virtual characters appear similar to puppets that are not able to express mimics.

The absence of mimics also impaired the humor of some scenes. For instance, it was very difficult to express surprise or some other reactions of the applicant without mimics. Hence, those parts which were built on these expressions appeared somewhat mechanical.

As a result it was found out that facial expressions are crucial and their absence hinders virtual interaction significantly.

6.3.2 Gesture control

Another problem that occurred during the project is related to the control of gestures. As already explained, it was planned that the created animations could be executed with a small control bar. However, this concept appeared to be problematic, because it required mouse clicks on the buttons, but the mouse is also used for the control of the eye gaze of an avatar. As a result, every time a performer clicked on a button, the respective avatar's head moved undesirably.

For this reason, an alternative form of control of the animations was used, which is, however, less intuitive. Second Life supports the execution of animations and sounds with special commands via the text-chat input form. This allowed the definition of a set of custom commands which could be used to start the animations. With this method, the actor could use the mouse to control the eye gaze and at the same time execute animations with text commands.

This issue shows that for virtual interaction it is an important question how to implement controls in a user friendly way. As nowadays, typical users only have access to a keyboard and a computer mouse as input devices, the control of gestures, motion, mimics and eye gaze has to be designed considering these restrictions. Thereby, a difficult task is to manage the concurrency of each of these channels.

6.3.3 Limited self-perception

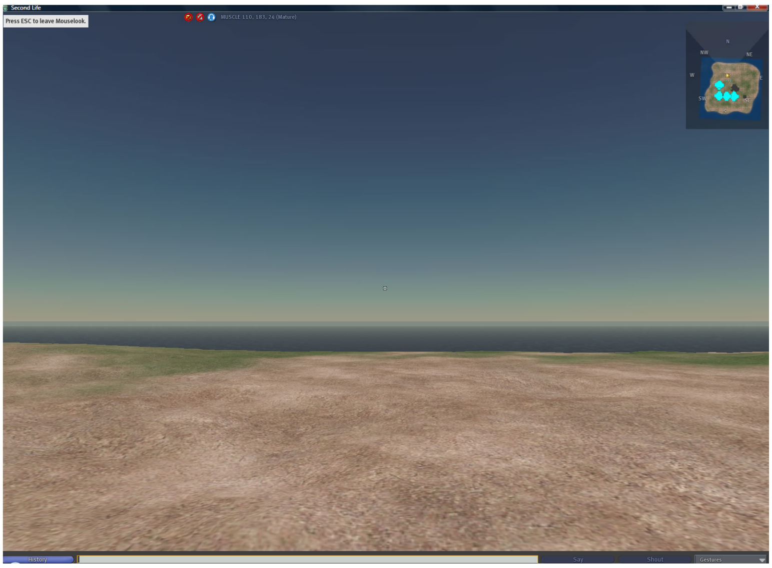


Figure 6.15: Missing avatar body parts in mouselook mode

The next issue which came up during the performance is that the actors sometimes were not aware of their state in the virtual environment. Especially due to the fact that they were not able to perceive their virtual representation, interaction seemed to be difficult. As an example, the performers could not see their body parts when they were in the "mouselook-mode"[10], which was necessary to control the gaze. Figure 6.15 shows the view of an avatar in Second Life, when the mouselook-mode is activated.

As can be seen, the body parts of the avatar are not displayed. For this reason, the actors could not see the execution of an animation of their avatars and thus had to memorize the timing.

Hence, it was found out that the basis of virtual interaction is the self-perception in the environment. If one is not able to put oneself into the virtual body, a successful interaction is difficult.

6.3.4 Delays

Finally, some members encountered problems with the performance of their hardware and network connection. This influenced interaction, because limitations of network bandwidth sometimes delayed the audio stream. As a result, the dialogues were slower and breaks between different parts occurred.

Another effect, which influenced the performance markedly, was the delay in the movement due to a slow network connection. This led to a kind of uncoordinated movement of some avatars. For instance, when the applicant entered the office, performance issues made the steering so difficult that sometimes, the character walked over the office chair.

These issues showed that performance is another important factor at virtual interaction. Thus, techniques that shall be usable for a wide range of users have to be implemented in a way that performance limitations can be handled. Otherwise, these techniques might not work correctly, which can lead to misconceptions.

[10] The mouselook mode is a first person perspective in Second Life. The standard mode is a third person perspective.

6.4 Summary of the practical part

This chapter included a detailed description of the development and performances of the virtual theater play which was conducted in the course of this thesis.

After an overview of similar projects that included the idea of using virtual reality for theater performances, each step of the project was described in detail. It was explained which factors influenced the decision for the play, how the characters and requisites were designed and how the necessary animations were created.

In the end of this chapter, the results of the analysis of virtual interaction during the theater performances were presented. Different issues were pointed out, which hamper interaction in Second Life and might be the basis for future improvements.

Chapter 7

Conclusion

The last chapter of this thesis will recapitulate the most important points of this thesis and try to give an outlook on possible future developments. Thereby, the findings of the practical part will be taken into account.

7.1 Summary

This thesis described the basic concepts of interaction in collaborative virtual environments. It was shown that virtual worlds in general are growing in importance and thus, interaction techniques in these environments are an important research field. Two main areas, social interaction and obejct interaction, were explained in detail. As regards the first one it was explained that avatars as virtual representations of users have to be capable to show body language, either automatically or manually and that their visual apperance plays an important role.

Thereafter, the thesis gave an overview of the different types of interaction and how they are implemented in virtual worlds. It was shown that for these techniques, on the one hand the used input devices have a crucial impact on the types of possible interaction and on the other hand, computer animations are necessary to represent e.g. body language. For their creation there exist several approaches, where two of the most important ones are keyframe based methods and motion capturing.

In the last part of this thesis, the concepts which had been explained initially were analyzed by means of a virtual theater performance. All relevant steps of the process were described to give an impression how a theater performance in virtual worlds can be developed. Thereby, several restrictions were identified, which might hinder virtual interaction and call for improvements in future virtual worlds.

7.2 Future developments

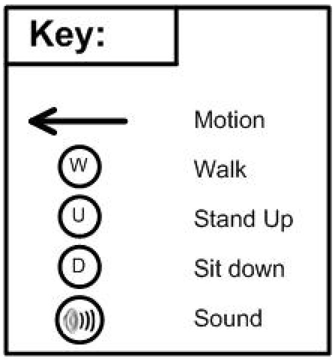
A prediction regarding future developments in the field of virtual worlds is difficult, because technological evolution takes place quite rapidly. Nevertheless, the trend seems to go towards a higher relevance of social aspects, as the presented research projects, recent developments, as the implementation of voice chat in Second Life, and the general movement of Web 2.0 indicate.

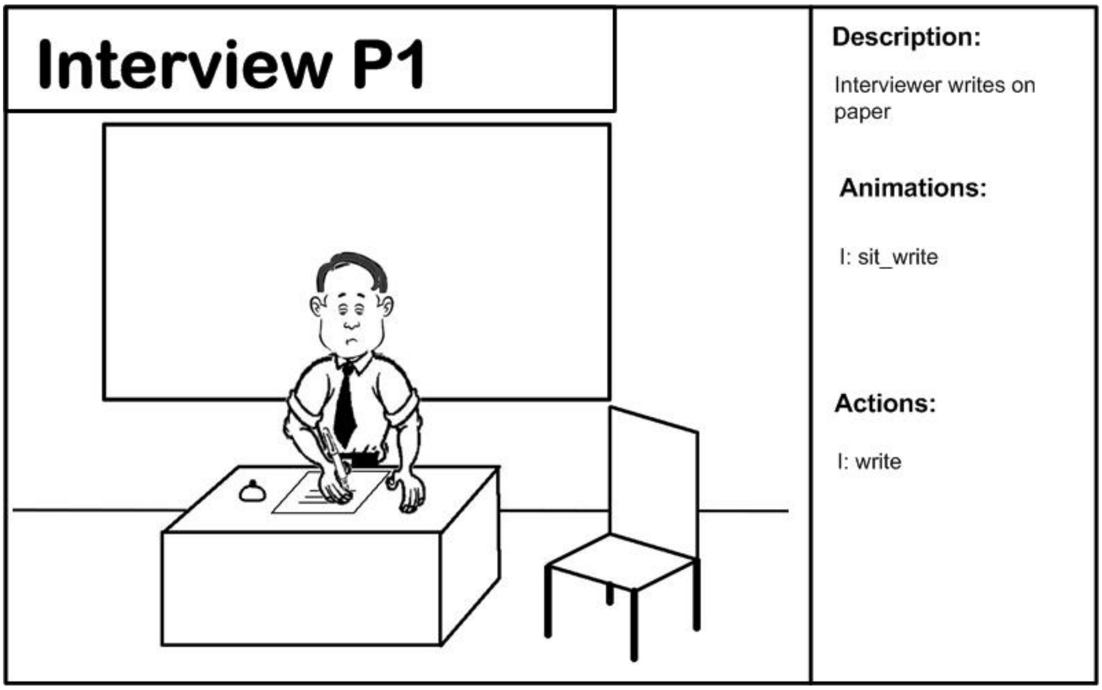
Furthermore, technological advancements in the field of interaction devices and artificial intelligence might support users to interact in virtual environments either by allowing a more extensive control of their visual representation or by controlling specific channels automatically (e.g. by analyzing the user's speech).

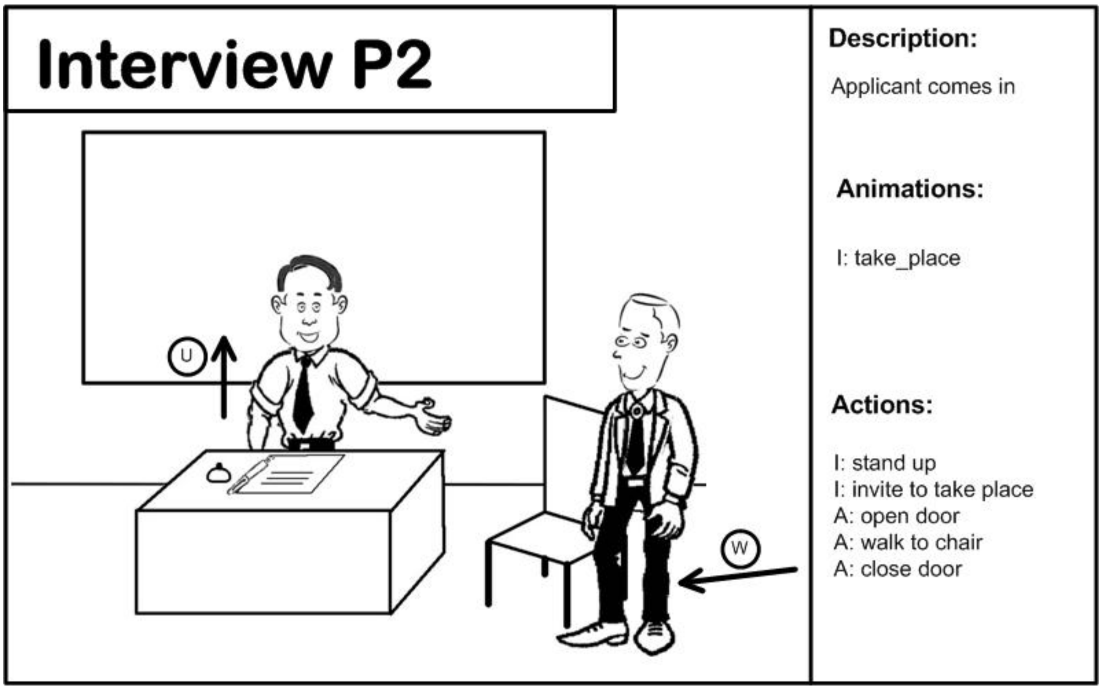
To conclude this thesis, it can be stated that in current virtual worlds interaction is feasible in a high extent, but has a number of limitations regarding the expression of body language and controls. Improvements can be expected with the development and propagation of new user input devices which allow a better tracking of the user's motions and translate them to motions of the virtual characters.

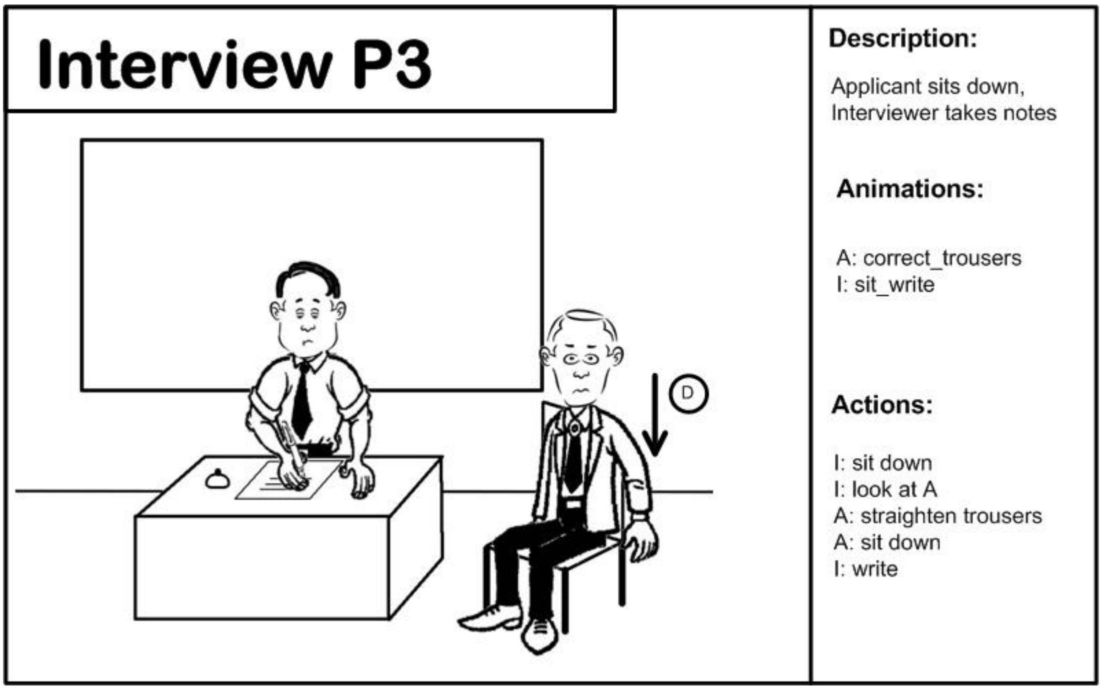
Appendix A

Storyboard

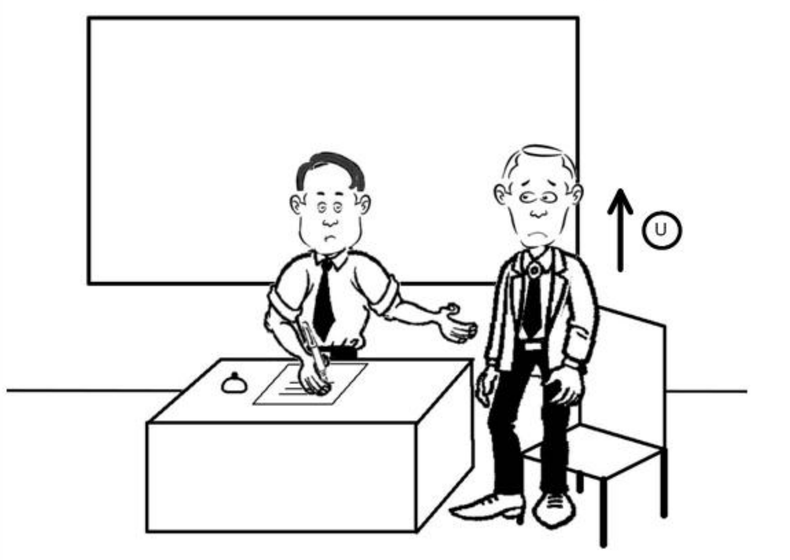





Interview P2
U
W
Description:
Applicant comes in
Animations:
I: take_place
Actions:
I: stand up
I: invite to take place
A: open door
A: walk to chair
A: close door


Interview P3
D
Description:
Applicant sits down,
Interviewer takes notes
Animations:
A: correct_trousers
I: sit_write
Actions:
I: sit down
I: look at A
A: straighten trousers
A: sit down
I: write

Interview P4



Description:

Interviewer tells applicant to stand up again

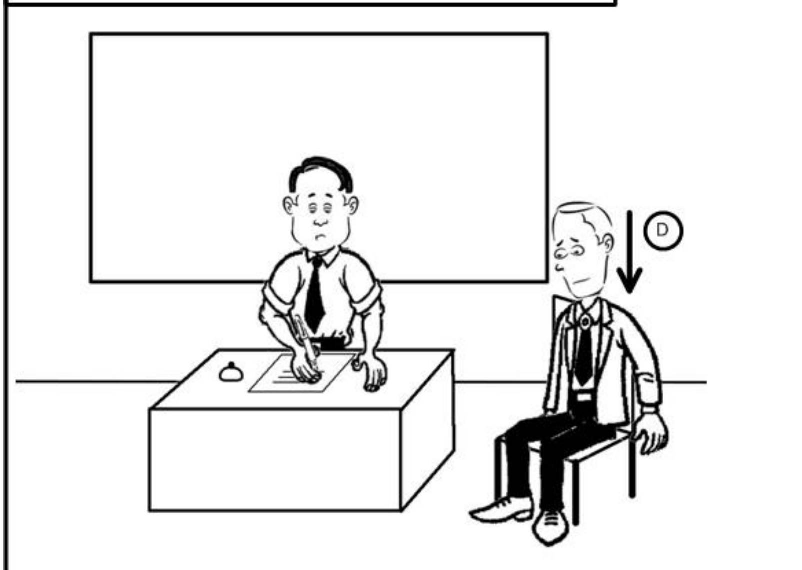
Animations:

I: take_a_seat
A: fix_suit

Actions:

I: show A to stand up
A: stand up
A: fix suit
I: show A to sit down
A: sit down
I: take notes

Interview P5



Description:

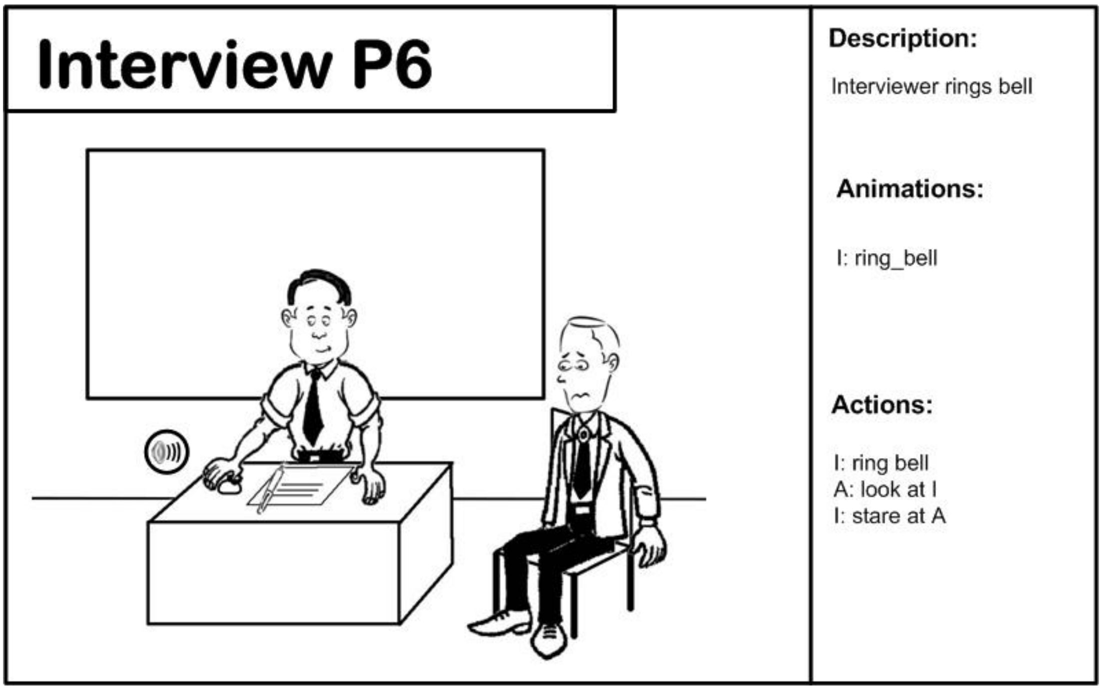
Interviewer confuses applicant with questions

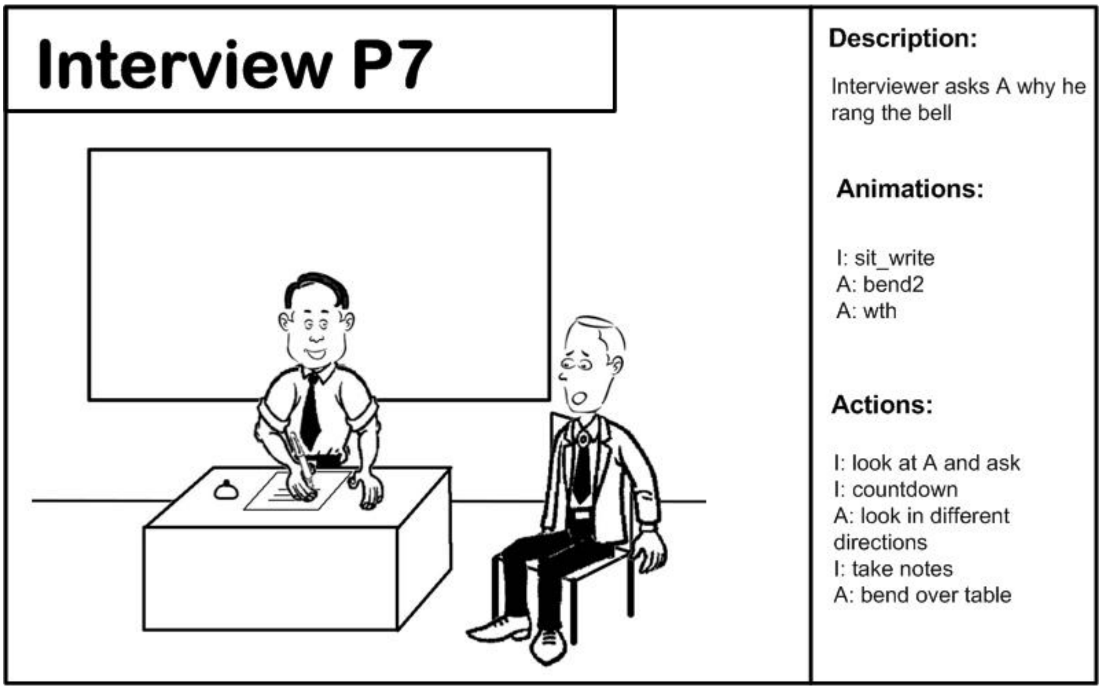
Animations:

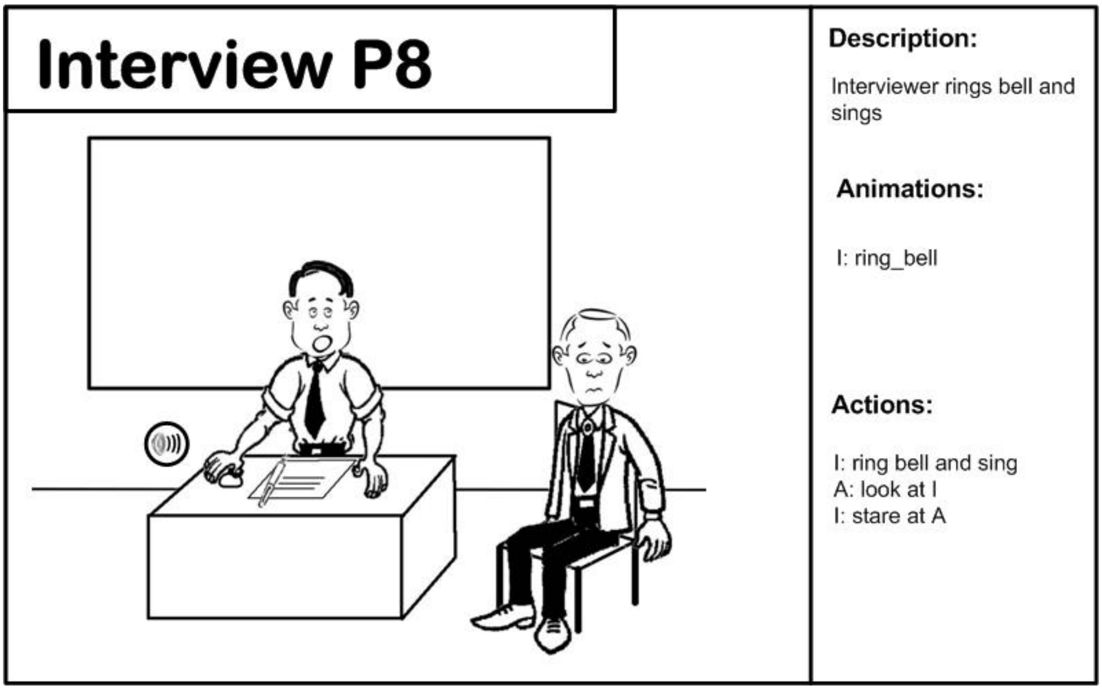
I: sit_write
I: shake_head
A: bend1
I: laugh

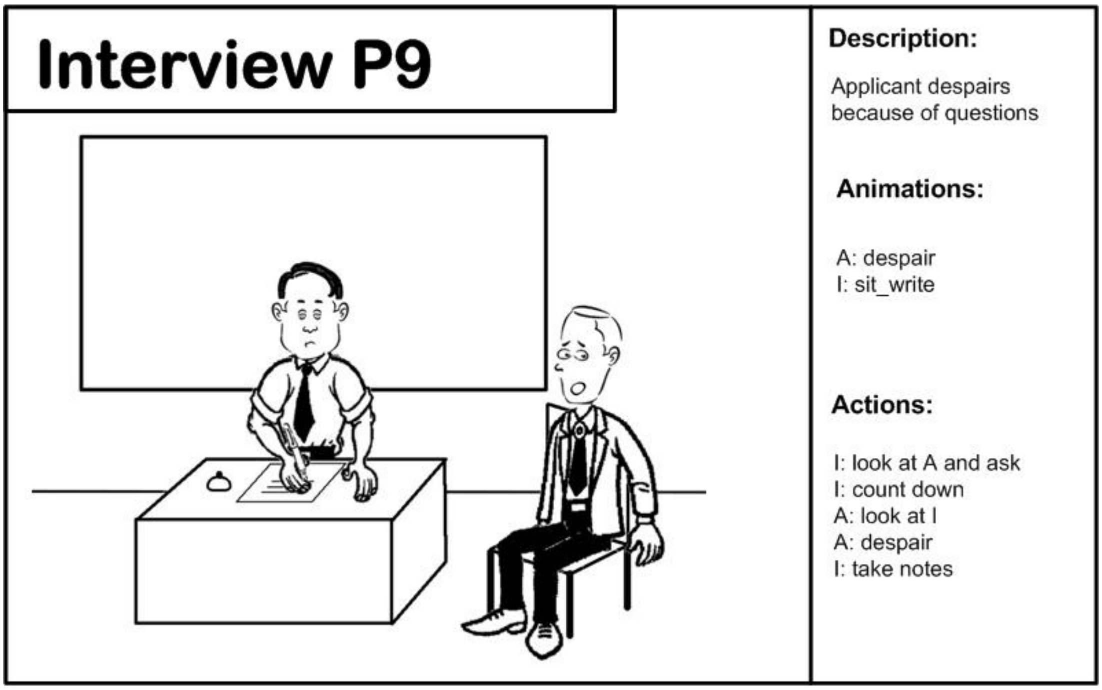
Actions:

I: look at A and ask
A: look at I and respond
I: shake head
I: take notes
A: bend over table
I: laugh


Interview P6
Description:
Interviewer rings bell
Animations:
I: ring_bell
Actions:
I: ring bell
A: look at I
I: stare at A


Interview P7
Description:
Interviewer asks A why he rang the bell
Animations:
I: sit_write
A: bend2
A: wth
Actions:
I: look at A and ask
I: countdown
A: look in different directions
I: take notes
A: bend over table


Interview P8
Description:
Interviewer rings bell and sings
Animations:
I: ring_bell
Actions:
I: ring bell and sing
A: look at I
I: stare at A


Interview P9
Description:
Applicant despairs because of questions
Animations:
A: despair
I: sit_write
Actions:
I: look at A and ask
I: count down
A: look at I
A: despair
I: take notes

Interview P10

Description:

Interviewer makes face and strange noise

Animations:

I: taunt1 or taunt2
A: wth

Actions:

I: make face and look at A
A: look confused around

Interview P11

Description:

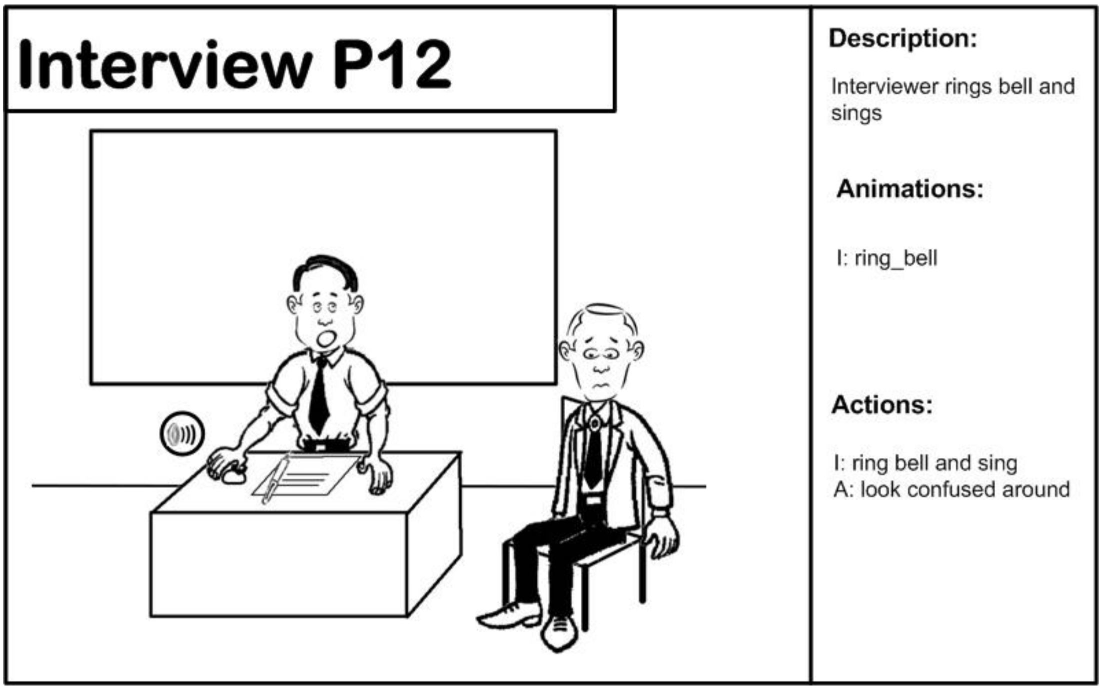
Interviewer asks applicant for name

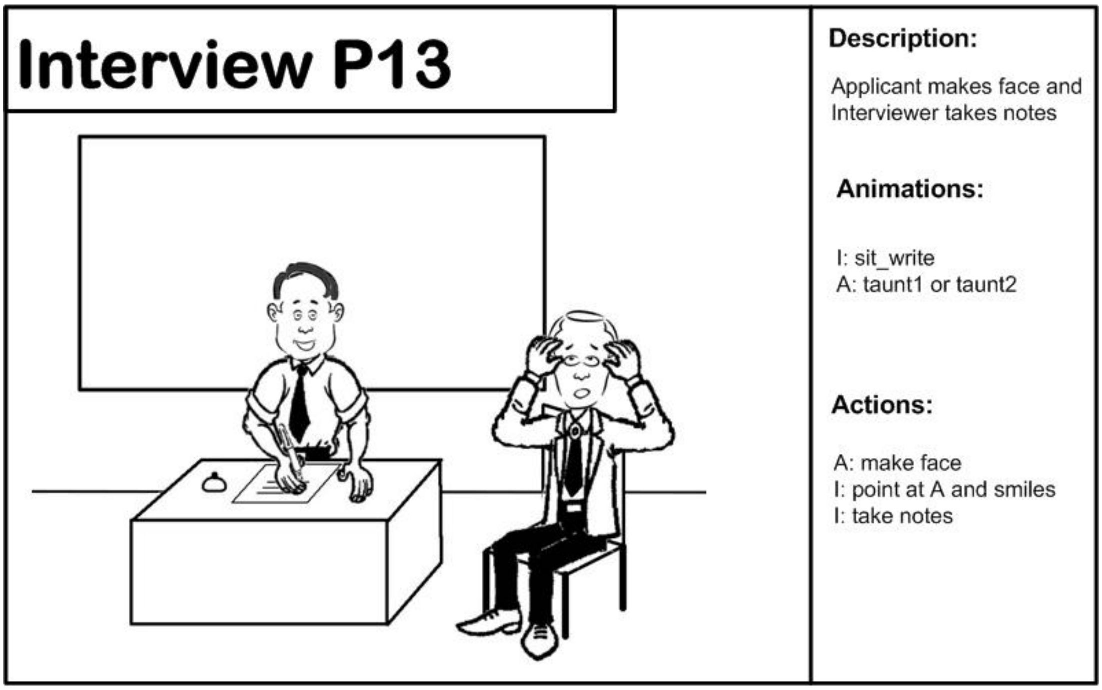
Animations:

I: sit_write
A: bend2

Actions:

I: look at A and ask
A: look at I and respond
I: take notes
A: bend over table
I: stare at A


Interview P12
Description:
Interviewer rings bell and sings
Animations:
I: ring_bell
Actions:
I: ring bell and sing
A: look confused around


Interview P13
Description:
Applicant makes face and Interviewer takes notes
Animations:
I: sit_write
A: taunt1 or taunt2
Actions:
A: make face
I: point at A and smiles
I: take notes

Interview P14

Description:

Interviewer calls evaluators

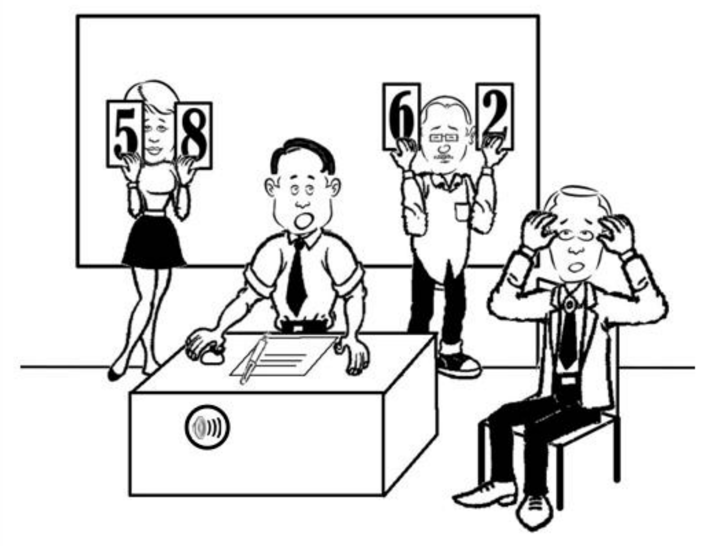
Animations:

none

Actions:

I: call evaluators
E: come in
A: look at E

Interview P15



Description:

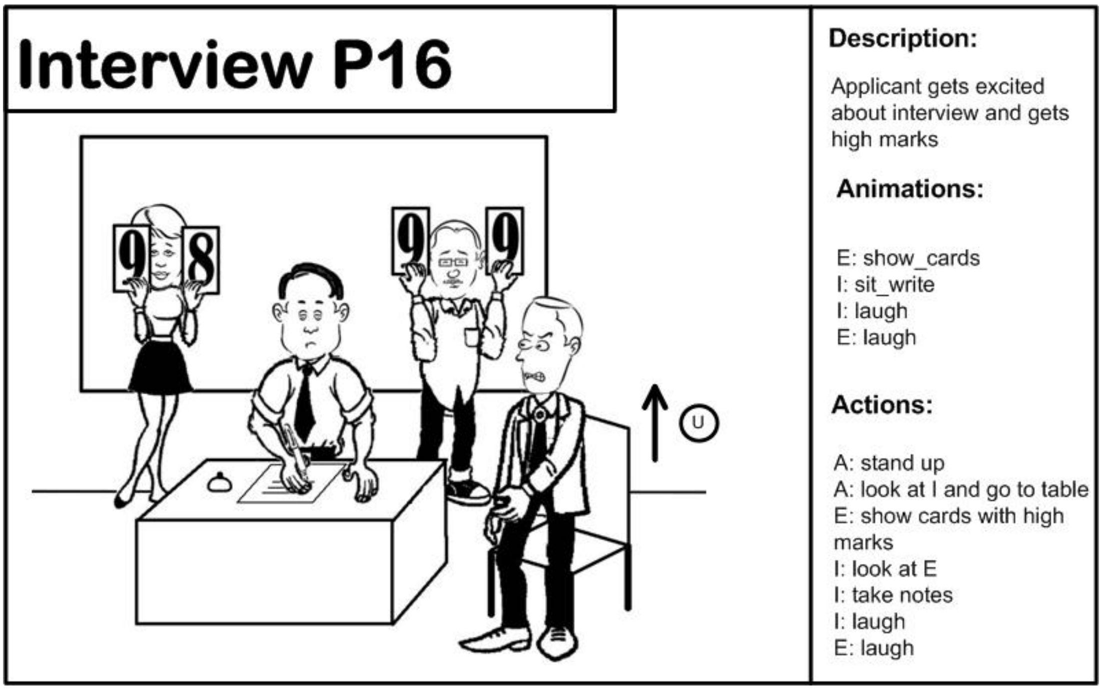
Applicant makes face and gets low marks

Animations:

I: ring_bell
A: taunt1 or taunt2
E: show_cards
I: sit_write

Actions:

I: ring bell and sing
A: make face
E: hold cards with low numbers up
I: look at E
I: take notes
E: put cards away



Interview P16
9
8
9
9
U
Description:
Applicant gets excited about interview and gets high marks
Animations:
E: show_cards
I: sit_write
I: laugh
E: laugh
Actions:
A: stand up
A: look at I and go to table
E: show cards with high marks
I: look at E
I: take notes
I: laugh
E: laugh

Bibliography

M. Argyle. *Bodily Communication.* Methuen, London, 1988.

M. Argyle. *Social Interaction.* Methuen, London, 1969.

M. Argyle and M. Cook. *Gaze and Mutual Gaze.* Cambridge University Press, Cambridge, 1976.

N. I. Badler, C. B. Phillips, and B. L. Webber. *Simulating Humans: Computer Graphics Animation and Control.* Oxford University Press, New York, 1993.

J. N. Bailenson and J. J. Blascovich. Avatars. *Berkshire Encyclopedia of Human-Computer Interaction*, pages 64–68, 2004.

R. H. Bartels, J. Beatty, and B. A. Barsky. *An Introduction to Splines for Use in Computer Graphics and Geometric Modelling*, chapter Hermite and Cubic Spline Interpolation. Morgan Kaufmann, 1998.

R. Bartle. *Designing Virtual Worlds.* New Riders, Indianapolis, 2003.

G. Beattie. *Visible Thought: The New Psychology of Body Language.* Routledge, London, 2003.

R. Belleman. *Interactive Exploration in Virtual Environments.* PhD thesis, Universiteit van Amsterdam, 2003.

R. J. Bloomfield. Worlds for study: Invitation - virtual worlds for studying real-world business (and law, and politics, and sociology, and...). online, 2007. URL `http://ssrn.com/abstract=988984`. accessed 29-Sept-2007.

B. Book. Virtual worlds review. online, 2006. URL `http://www.virtualworldsreview.com`. accessed 28-Oct-2007.

B. Book. Virtual world business brands: Entrepreneurship and identity in massively multiplayer online gaming environments. online, 2005. URL `http://ssrn.com/abstract=736823`. accessed 12-Nov-2007.

B. Brown and M. Bell. CSCW at play: There as a collaborative virtual environment. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*, 2004.

C. Calongne and J. Hiles. Blended realities: A virtual tour of education in Second Life. In *Proceedings of 12th ANNUAL TCC Worldwide Online Conference*, 2007.

E. Castronova. The price of bodies: A hedonic pricing model of avatar attributes in a synthetic world. *Kyklos*, 57(2), 2004.

D. de Nood and J. Attema. Second life, the second life of virtual reality. online, 2006. URL `http://www.epn.net/interrealiteit/EPN-REPORT-The_Second_Life_of_VR.pdf`. accessed 05-Oct-2007.

T. D. DeRose and B. A. Barsky. Geometric continuity, shape parameters, and geometric constructions for catmull-rom splines. *ACM Transactions on Graphics*, 7(1), 1988.

D. Dwyer, K. Barbieri, and H. Doerr. Creating a virtual classroom for interactive education on the web. In *Proceedings of WWW 95, Third International Conference on World Wide Web*, 1995.

S. P. Egen. The history of avatars. online, 2005. URL `http://www.imediaconnection.com/content/6165.asp`. accessed 04-Oct-2007.

P. Ekman, W. Friesen, and P. Ellsworth. *Emotion in the Human Face: Guidelines for Research and an Integration of Findings.* Pergamon, New York, 1972.

Encarta. Interaction. online, 2007. URL `http://encarta.msn.com/dictionary\_1861621732/interaction.html`. accessed 12-Nov-2007.

M. Enderton. The art and practice of 3D computer-generated character animation. *Computer Science Capstone Paper, Macalester College*, 2003.

K. Evert. Theater im virtuell geteilten Raum. In *Proceedings of Transforming Spaces in Darmstadt*, 2002.

M. Fabri and D. Moore. The use of emotionally expressive avatars in collaborative virtual environments. In *Proceedings of the AISB Symposium on Virtual Social Agents*, 2005.

M. Fabri, D. Moore, and D. Hobbs. Expressive agents: Non-verbal communication in collaborative virtual environments. In *Proceedings of Autonomous Agents and Multi-Agent Systems*, 2002.

M. Garau, M. Slater, S. Bee, and M. Sasse. The impact of eye gaze on communication using humanoid avatars. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2001.

M. Gerhard, D. Moore, and D. Hobbs. Continuous presence in collaborative virtual environments: Towards the evaluation of a hybrid avatar-agent model for user representation. In *Proceedings of the Third International Workshop on Intelligent Virtual Agents*, 2001.

M. Gleicher. Animation from observation: Motion capture and motion editing. *SIGGRAPH Comput. Graph.*, 33(4), 2000.

T. Goodliffe. New to SL? here are some ways to make money, virtual to reality. online, 2006. URL `http://www.vtoreality.com/date/2006/`. accessed 02-Nov-2007.

A. Guye-Vuillème, T. Capin, I. Pandzic, N. Thalmann, and D. Thalmann. Non-verbal communication interface for collaborative virtual environments. *Virtual Reality*, 4:49–59, 1999.

L. Herda, P. Fua, R. Plänkers, R. Boulic, and D. Thalmann. Skeleton-based motion capture for robust reconstruction of human motion. In *Computer Animation*, 2000.

J. K. Hodgins, J. O'Brien, and R. Bodenheimer. Computer animation. *Wiley Encyclopedia of Electrical and Electronics Engineering*, 3, 1999.

N. Hogan. Impedance control - an approach to manipulation. *Journal of Dynamic Systems, Measurement and Control*, 107:1–24, 1985.

R. Imbert, A. de Antonio, and J. Segovia. Improving communication in 3D virtual environments by means of task delegation in agents. In *Proceedings of the 5th International Workshop on Cooperative Information Agents*, 2001.

K. Jalkanen. Incarnation of a divine being. online, 2001. URL `http://www.tii.se/tools/project.php?id=4\&language=\_en`. accessed 12-Nov-2007.

D. Jones. I, avatar: Constructions of self and place in Second Life and the technological imagination. online, 2006. URL `http://www.gnovisjournal.org/files/Donald-E-Jones-I-Avatar.pdf`. accessed 12-Nov-2007.

R. Kehl. *Markerless Motion Capture of Complex Human Movements from Multiple Views*. PhD thesis, ETH Zürich, 2005.

G. D. Kessler, L. F. Hodges, and N. Walker. Evaluation of the CyberGlove as a whole-hand input device. *ACM Transactions on Computer-Human Interaction*, 2(4), 1995.

M. L. Knapp. *Nonverbal Communication in Human Interaction*. Holt, Rinehart and Winston, New York, 1978.

D. H. Kochanek and R. H. Bartels. Interpolating splines with local tension, continuity, and bias control. In *Proceedings of the 11th Annual Conference on Computer Graphics and Interactive Techniques*, 1984.

J. Lasseter. Principles of traditional animation applied to 3D computer animation. In *Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques*, 1987.

Linden Labs. Economic statistics. online, 2007. URL `http://secondlife.com/whatis/economy-graphs.php`. accessed 02-Nov-2007.

N. Magnenat-Thalmann and D. Thalmann. *Encyclopedia of Computer Science and Technology*, volume 29, chapter Computer Animation. CRC, 1993.

N. Magnenat-Thalmann and D. Thalmann. *Computer Animation: Theory and Practice*. Springer, New York, 1990.

S. N. Matsuba. VRML dream on-line performance. *SIGGRAPH Comput. Graph.*, 34(2), 2000.

S. N. Matsuba and B. Roehl. “bottom, thou art translated”: the making of VRML dream. *Computer Graphics and Applications, IEEE*, 19(2), 1999.

M. McKeon and S. Wyche. Life across boundaries: Design, identity, and gender in SL. online, 2005. URL `http://www.mattmckeon.com/portfolio/second-life.pdf`. accessed 12-Nov-2007.

M. Meredith and S. Maddock. Motion capture file formats explained. online, 2004. URL `http://www.dcs.shef.ac.uk/intranet/research/resmes/CS0111.pdf`. accessed 21-Oct-2007.

M. Mine. Virtual environment interaction techniques. Technical Report TR95-018, UNC Chapel Hill, 4, 1995.

M. Mine. Moving cows in space: Exploiting proprioception as a framework for virtual envrionment interaction. Technical Report TR97-003, UNC Chapel Hill, 18, 1997.

J. O'Brien. Writing in the body: Gender (re)production in online interaction. *Communities in Cyberspace*, 1999.

C. R. Ondrejka. A piece of place: Modeling the digital on the real in second life. online, 2004. URL `http://ssrn.com/abstract=555883`. accessed 05-Oct-2007.

C. Pinhanez and A. Bobick. It/I: theater with an automatic and reactive computer graphics character. In *ACM SIGGRAPH 98 Conference Abstracts and Applications*, page 302, 1998.

I. Poupyrev, M. Billinghurst, S. Weghorst, and T. Ichikawa. The go-go interaction technique: Non-linear mapping for direct manipulation in virtual reality. In *Proceedings of the ACM Symposium on User Interface Software and Technology*, 1996.

J. Ringard. L'utilisation de la vidéo dans avatars 3D. Master's thesis, Université des Sciences et Technologies de Lille, 2006.

W. Robinett and R. Holloway. Implementation of flying, scaling and grabbing in virtual worlds. In *Proceedings of the 1992 Symposium on Interactive 3D Graphics*, 1992.

R. Schroeder. *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments*. Springer, London, 2002.

M. Slater and M. Usoh. *Artificial Life and Virtual Reality*, chapter Body Centred Interaction in Immersive Virtual Environments. Wiley, 1994.

F. Sparacino, C. Wren, G. Davenport, and A. Pentland. Augmented performance in dance and theater. In *Proceedings of the International Conference on Dance and Technology*, 1999.

N. Stephenson. *Snow Crash*. Bantam, 1992.

A. Vasalou, A. Joinson, and J. Pitt. Constructing my online self: Avatars that increase self-focused attention. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2007.

J. Zhao and N. I. Badler. Inverse kinematics positioning using nonlinear programming for highly articulated figures. *ACM Transactions on Graphics*, 13 (4), 1994.