Universidade do Minho
Escola de Engenharia

Pedro Miguel da Silva Araújo

Development of naturalness improvements of avatar appearance and behaviour for application in television

Janeiro de 2010

Universidade do Minho
Escola de Engenharia

Pedro Miguel da Silva Araújo

Development of naturalness improvements of avatar appearance and behaviour for application in television

Tese de Mestrado
Mestrado de Informática

Trabalho efectuado sob a orientação de
Prof. Doutor Adérito Fernandes Marcos
Ingeniero David Oyarzun Laura

Janeiro de 2010

ACKNOWLEDGMENTS

First of all, I would like to express my gratitude to my family and friends for all the given support during the work done abroad, for being there and for their love.

Thanks to Professor Adérito Marcos the scientific counseling and guidance through all the work.

Also, I would like to show my gratitude to Amalia Ortiz and David Oyarzun for the continuous encouragement and for taking the time to evaluate my thesis work. Thanks to Aitor Arrieta for his collaboration by assisting me during the motion capture sessions, and to help me to prepare the demonstrations. Furthermore, my gratitude goes to all other VicomTech members for their support and friendship.

Finally, I express my warm gratitude to Filipe Gonçalves and Ricardo Marques for the inspiring discussions, encouragement and close friendship.

RESUMO

Desenvolvimento de melhoramentos na naturalidade da aparência e comportamento de avatares

Os avatares são definidos como personagens virtuais que visam representar ou imitar uma pessoa real. O número de aplicações destes é tão limitado quanto as actividades humanas o são. Os avatares podem ser aplicados em áreas tão diversificadas como jogos de computador, actores virtuais, professores virtuais, ou atletas para simulação em desporto. Um cenário onde a aplicação de avatares poderá ter sucesso é a televisão. No entanto, ainda é uma área de investigação em aberto, e há muitos melhoramentos a fazer de forma a obter uma aplicação adequada e de simples utilização.

O trabalho apresentado nesta tese tem como objectivo de melhorar a naturalidade de avatares aplicados em televisão. Nesse contexto, a naturalidade da aparência e movimento dos avatares têm uma influência determinante na sua credibilidade. O avatar vai interagir com um humano num cenário real, utilizando técnicas de realidade mista, por isso é mais vulnerável à detecção de uma aparência ou comportamento incaracterísticos. Abordamos este problema segundo duas perspectivas: melhorar o aspecto e o comportamento. Relativamente ao comportamento do avatar, melhorámos os seus movimentos utilizando um sistema de captura de movimentos para gerar animações. Estas animações sustentam um módulo baseado em regras, que se encarrega de seleccionar a animação mais adequada para cada momento, dotando assim o avatar de um movimento realista. Por outro lado, o sistema deve permitir uma alteração fácil e rápida da aparência do avatar, aparência esta que deverá ser muito realista. Os avatares que se destinam a televisão, deverão ser criados por modeladores profissionais. Desta forma, um modelo será provavelmente complexo e necessita de ser carregado em tempo real com a rapidez que a televisão exige. Com esse objectivo, desenvolvemos um software que permite a exportação de qualquer tipo de modelo criado na ferramenta comercial de modelação Maya, uma das ferramentas mais populares entre os modeladores, para um ficheiro legível pela plataforma que suporta o avatar. Assim, modelos altamente complexos e realistas, criados por profissionais na área, podem ser facilmente carregados.

ABSTRACT

Development of naturalness improvements of avatar appearance and behaviour for application in television

Avatars are defined as virtual characters that seek to mimic or represent a real person. Their number of applications is as limited as human activities are. Avatars can be applied in the most diverse areas, such as computer games, virtual actors, virtual teachers, or sport simulation athletes. A scenario where applications with avatars would be very successful is the television. However, it is still an open field of investigation and there is a lot of aspects to improve in order to obtain a more suitable and usable application for it.

The work presented in this thesis has the purpose of improve the naturalness of avatars applied in television. Naturalness of appearance and motion of the avatar greatly influences how convincing they are in a television context. The avatar interacts with a human in a real scenario, using mixed reality techniques, therefore is more vulnerable to uncharacteristic behaviour or appearance detection. We have approached this problem from two perspectives: behaviour and appearance improvement. More related with the avatar behaviour, we have improved the movements using a motion capture system to generate animations. These animations feed a rule-based module that selects the most suitable animation in each moment, and so, provides a realistic motion for the avatar. On the other hand, the system must allow an easy and quick change of the avatar appearance, and this appearance must be very realistic. Avatars to be shown in television should be created by expert modeling professionals. In this way a virtual character will be probably a complex model and needs to be loaded in real time with the swiftness that television demands. For that purpose it has been developed a plugin that is capable to export any model from the commercial tool Maya, one of the most popular tools among modeling professionals, to a file readable by the platform that supports the avatar.

CONTENTS

LIST OF FIGURES

1 INTRODUCTION

Simulate real life scenes with 3D computer graphics has been an open field of investigation in the recent past years. The advances of the computational capacities of technologies, has enabled today's stunning levels of realism in 3D graphics. Among the most exciting developments in computer graphics are the virtual humans, who are becoming more and more popular over the last years. Virtual humans can be used in various situations as in games, films, cultural heritage or web-based situations, the field of applications is huge, and there is still a long way to go to achieve a realism level, where a virtual human behaves and resembles to a person in all his complexity.

Several authors had used different terms when it comes to reference virtual humans. Such as: *anthropomorphic*[1] *agents*, *avatars*, *synthetic actors*, *creatures*, *non-player characters* and *embodied conversational agents* [66]. In this thesis it will be mainly used the term avatar, however the term virtual character also refers the same. There are several other definitions about avatars. Each author gives emphasis on different characteristics of them. For example, Bickmore y Cassell [8] definition is focused on avatar capacity in having a conversation with the user:

Embodied conversation agents are animated anthropomorphic interface agents that are able to engage a user in real-time, multimodal dialog, using verbal and nonverbal behaviours to emulate the experience of human face-to-face interaction, such as speech, gesture, gaze, posture, or intonation.

Bates [6] gives an definition, based on the credibility qualities of the avatar:

A believable character does not mean an honest or reliable character, but one that provides the illusion of life, and thus permits the audience's suspension of disbelief.

Thalmann [74] studied avatars capacity to act in virtual environment, he gives the following definition:

Inside these Virtual Environments, Virtual Humans are a key technology that can provide Virtual presenters, Virtual Guides, Virtual actors, and be used to show how humans should act in various situations.

By last, a more general definition was given by Prendinger and Izhizuka [66]:

Life-like characters are animated agents that emulate human-to-human communication skills and thus allow for more natural interaction styles between humans and computers.

1 An anthropomorphic being is a non-human creature with uniquely human characteristics.

(a) Final Fantasy XI (b) Anthropomorphic avatar (c) Cartoon avatar

Figure 1: Examples of different types of three dimensional avatars [59].

The definition of a virtual human can be blurry. In the common immerse virtual environments, for example, an online game, the avatar can be the representation of the user in the virtual world, and may look like the user in appearance. A simple photograph of a person used as his digital representation in Internet chats or email, can be considered an avatar, however the avatar is destitute of any behaviour and incapable of social interaction. When defining a term, each author places special emphasis in one of the avatar characteristics. In this thesis, that emphasis goes to appearance and behaviour. The appearance is the perceptible physical characteristics and can be in many forms (2D/3D,cartoon-anthropomorphic, etc). Basically is a combination of face and body. The behaviour can be intelligent or not. If the avatar does not have intelligence, it executes only predefined orders, otherwise it can take autonomous decisions and generate is own answers during the interaction [59]. The necessary components that allow the the avatar to make decisions are called agents. In Figure 1 it can be seen some examples of avatars.

The realism level of the avatars depends on the purpose that they are intend to. An avatar can have a cartoon appearance in a computer game, but in simulation, for example, in clinical training they need to be realistic as possible, not only in physical appearance, but capable of interact and respond like a person [41]. Convincing realism is one of the most challenging problems in computer graphics. Human beings have a particular sensitivity in regard to motion and physical appearance of others. Our familiarity with body gestures provides us with a sharp perception of unnaturalness in movements and appearance. Mori [55] proposed the "The uncanny valley", an hypothesis where he states that there is theoretical gap between points of human compassion and empathy. That is, from a cartoon look, for example, until they reach a range of motion and likeness that is recognizable as fully human. Humans tend to feel repulsion and antipathy to "almost human", but not perfect humanoids. We are not easy to fool. In her book, Margaret

Hagen [35] introduced a concept where representational pictures are realistic in some aspects and not in others. Different methods produce different varieties of realism by omitting certain properties of the scene. James Ferwerda [24] has extended this concept to computer graphics and identified three types of realism. They are:

- Physical realism, in which the image provides the same visual stimulation as the scene;
- Photo-realism, the image produces the same visual response as the scene;
- Functional realism, in which the image provides the same visual information.

In physical realism, the image must be an accurate point-by-point representation of the spectral irradiance values at a particular viewpoint in the scene. Physically-based methods can produce accurate representations of reality with many applications (illumination engineering, remote sensing, etc), but the rendering is extremely computationally expensive and in most cases the images are not realizable in most displays. Another approach in order to achieve realism is photo-realism, resulting in an image that is indistinguishable from a photograph or scene. A definition of photo-realism is that the image must be photo-metrically realist. Photometry is the measure of eyes response to the light energy in terms of its perceived brightness. It means that the image must produce the same visual response even though the physical energy from the image and the scene are different. These visual-based algorithms, although are faster than the physical-based ones, are still too slow for interactive applications [24]. The third approach, the functional realism, is widely used in computer graphics. It relies on the assumption that realism is the fidelity of the information the image provides. Many simulators provide synthetic scenes that are not physically accurate or photo-realistic, however, they are functionally realistic because they transmit similar information as the real scene [56]. Here fidelity means that the user can perceive important properties from an image as he could in the real world. For example, in a flying simulator the user can learn skills that he can transfer into the real world [24]. Functional realism is a very important concept in this work. An avatar does not have to be photo-realistic to look visually convincing [5].

As it can be seen, the proper development of virtual humans requires knowledge of different disciplines, such as computational geometry, kinematics, artificial intelligence, computer graphics, and bio-mechanics. It is a very complicated process, and demands a strategy that passes by dividing the problem in several sub-problems. Funge et al. [28] proposed a model with five levels hierarchy, shown in Figure 2.

This work is focused in the three lower layers of the pyramid. The lowest one is the geometric layer and concerns the definition and appearance of the virtual human. In the kinematic layer is where the animation is defined. At the third layer, physical, are applied physical laws to obtain complex animations of different parts of the body, like anatomically correct skin deformations or hair movement [38]. The rest of the layers, behavioural and cognitive, concern the actions taken by the avatar given certain stimuli, being the last where the reasoning and knowledge usage of the avatar is governed.

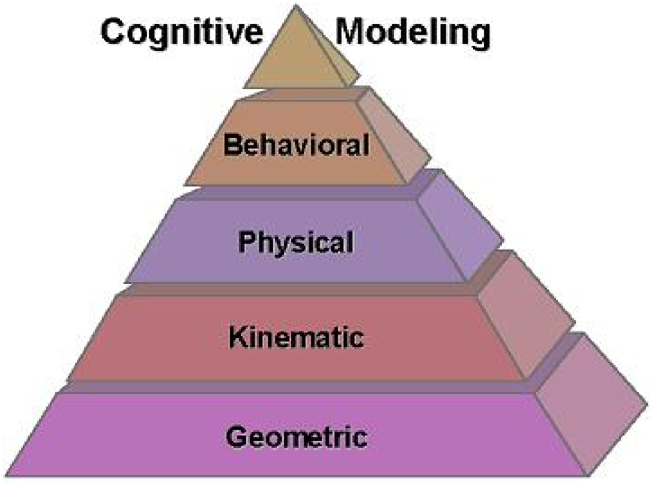


Figure 2: Model proposed by Funge et al.[28].

1.1 MOTIVATION AND OBJECTIVES

There are many application possibilities regarding the avatars. One that sooner or later reflects the developments of technology in general is the television. Among other things, it is focused on the dissemination of the contents to the general public and meanwhile, a unique platform for testing the acceptability of a new development. Thus, it has undergone a major transformation in recent years, including more frequently technological news among the content of its programs. Nowadays it is common to see technology in a sophisticated montages, such as weather report or entertainment television shows. The clearest example of this is founded in the virtual sets for newscast, in which is created all kinds of scenarios giving the impression that the presenter is immersed in them (see Figure 3). This technique consists of first creating a virtual scenario with all the elements that are required. Then, using calibrated cameras and a system for rendering, the final result is achieved by mixing real-time virtual image with the actual image. Furthermore, it is necessary to take into account aspects such as lighting (virtual lights versus real lights) or occlusion (virtual objects in front of real objects) to avoid losing the final realism.

Figure 3: This image shows a virtual studio with a human presenter. He can present and mark various graphics which appear in the virtual environment, for example [19].

However, despite the use of virtual objects and scenes on television is something very common nowadays, it is still hard to find virtual characters embedded in them. Examples of applications that could include interactive avatars may be the competitions and "calling" programs in which a contestant must remain anonymous, but his participation throughout the program and to achieve an effect of continuous presence, he could be represented by an avatar that the same control. Another possibility is use of avatars in entertainment programs, thus allowing an increase in the "fantasy" of the program through the inclusion of characters who really do not exist, or that simulate a real person. The very few current applications of avatars in television lack of usability when it comes to avatar control. These demand time for previous preparation, such as motion capture suit used by an actor that performs the movements that are transformed to movements of the avatar in real-time. This aspect is a problem for daily television programs which are on live or semi-live, where the time to preparation is very short. The time to prepare the avatar should be reduced as possible.

Physical realism as previously described, when applied to avatars presupposes a depth study in the most diverse domains of human investigation, and the capture of human subtle movements is a quite difficult task to accomplish. The pursuit of an high realistic avatar is a very ambitious goal. Therefore, the term naturalness is adopted in this thesis instead of realism, it emphasizes the goal of "illusion of life" and encompasses important concepts such as immersion and presence. Bates [6] explored the idea of believability, he stated that *"a believable character does not mean an honest or reliable character, but one that provides the illusion of life, and thus permits the audience's suspension of disbelief"*. This concept is parallel to functional realism. It is also important to have the notion that the avatars may not have an human appearance, for instance, they can be animals, extraterrestrial, or any other form, nevertheless is expected some naturalness of their movements, they cannot be completely unrecognizable in terms of human behaviour.

This work is a research in avatars naturalness for application in television. It is part of a bigger project that aims the inclusion of avatars in real scenarios using mixed reality techniques, interacting with humans in real time. The avatar is animated in real time and must be also of high quality, to have a consistent mixture between the real and the virtual image. However the mixing between the real world and virtual reality is not a matter of this work, but the quality of that mix in terms of naturalness of the avatar in the real scenario. To achieve this objective, the following partial goals were defined:

- Improve the appearance of the avatar, making him more realistic and coherent with the television contents. For obtaining that, tools that make easier the loading of models created by professional animators will be created.

- Reduce the time needed for setup the application. In television programs, over all in on live daily programs, reduce as much as possible the preparation time is very important. And as less specific devices are needed for the execution as better for television producers.
- Improve the avatar behaviour. It must be natural depending on the avatar current state. This work will focus mainly on the avatar standing posture. Naturalness relies in expected behaviours, and is not expectable that an avatar in television remains static, when it does not receive any order. Even when humans are standing still, they perform some characteristic movements, such as weight transfer from a leg to the other, random walk, etc. In the process of creating an avatar, this issue must be taken into account, otherwise the avatar will look stiff.

1.2 METHODOLOGY

For dealing with the three partial goals defined in last section, this work present these approaches:

APPEARANCE

- The user must have the possibility to choose among different avatar appearances. For this purpose, it was developed a plug-in that exports an avatar model from Maya to a file, with a structure that can be read and loaded to the platform where the avatar will be. Maya is one of the most popular commercial tools for modeling professionals, so the development of the plugin will allow avatar users to load any high quality character modeled with this tool. Moreover, accessories which consist in complements to the avatar physical appearance, such as clothes could be loaded too. It will be used graphical libraries that allow the use of realistic materials for those accessories, making them more realistic and attractive.

Reduction of setup timing

- For reducing the time needed for setup the system, the application will use standard devices like microphones, keyboards and joysticks as input and it does not make use of complex motion capture systems in real time or chrome systems. All of the movements will be prerecorded and launched in real time following a rule-based system or by commands specified by the real actor.

Behaviour

- For this purpose, some animations related with the state of waiting will be obtained, following some rules that define the human standing position. The rules are defined by experts in the field of human behaviour and body language. Then, the animations will be launched by means of a rule-based AI module that evaluates duration of the waiting time and selects the most suitable motion to execute.

This work is part of a bigger project called PUPPET and partially supported by the Basque Government. The final user of this project is a Basque television producer company. In this way, the validation of this work will consist on its integration with the rest of modules of the project and having a final session with representatives of this Basque company, including real actors, producers and directives. The objective will be to check that the system fulfills the requisites of easy-to-use, realism and naturalness of behaviours and direct load of models created by professionals.

1.3 Thesis structure

In this chapter it was introduced the problem that is being treated, and the methods used in order to solve it. The section starts with an introductory text that summarizes and contextualizes the problem, then divides into three subsections: motivation and objectives, methods and techniques and thesis structure. In the motivation and objectives is explained the focus and relevance of the problem, and the sequence of objectives intended to achieve. For that is necessary to adopt methods of investigation and implementation, those are summarized in the methods and techniques subsection.

The second chapter is about avatars, their characteristics and requirements to be effective from a human perspective. Particularly the case of avatars in television, giving also a historic overview. The next chapter deals with realistic human appearance, the natural human posture and the way to capture it with motion capture. The last one it will be described from a technological and historical point of view.

In the fourth chapter is described the implementation, development of the plugin, motion capture session description, and the work results discussion. Finally, in the sixth chapter, some conclusions are exposed, and proposed recommendations for future developments.

2 AVATARS IN TELEVISION

Nowadays, it is common the inclusion of virtual characters in the most diverse applications. Virtual characters are gaining acceptance and are increasingly being used as brand communication devices, information agents, e-learning tutors, or as virtual presenters. They can present in very different contexts, such as demonstrating products [45] or guide and attract users through web-pages [2]. Their presence increases the naturalness and eases human-computer interaction. The use of virtual characters can be successfully extended to other kind of applications witch require real-time and high level of interactivity, live television programs are a great example [61].

The current maturity of animation technologies, allows the improvement of the modeling and production of television content scenario. There are real benefits in using avatars in television. McCulloch [53] identified has the most obvious, the capability of the avatars in operating twenty four hours per day, seven days per week. They can access information instantaneously, holding multiple conversations at the same time in different languages.

In this chapter, it will be presented an overview of the application of avatars in television programs. The relevant work in the field will be reviewed, regarding the requirements fulfilled by an effective avatar.

2.1 HISTORY OF AVATARS IN TELEVISION

The first computer generated television presenter was Max Headroom, and made his debut in 1985. This music program was a sort of virtual character, but only his head was visible. However, was not computer generated. In its place, the technique that was used was to combine a prosthetic suit with video post effects. Max was able to interact with their guests in real time. Max is actually an actor made up and filmed to look like a simulation [42].

It was in 1988 the first television appearance of a virtual character generated and animated by computer in real time. A virtual called puppet Waldo C. Graphic [76], who participated in the program "The Jim Henson Hour", interacting with the rest of puppets (real) that appeared in it. Waldo movements were monitored in real time through an external interface operated by a person. In the very same year, a computer puppet called Mike the Talking Head appears. A single puppeteer controlled all the facial expressions, lip-synching, and head rotations in real time [67].

(a) **Mike the talking head** (b) **Controller for Mike** (c) **Waldo C. Graphic**

Figure 4: Examples of avatars in television.

In 1991, Medialab, Paris, started weekly production of Mat the Ghost. Mat appeared daily on French cable TV (Canal Plus) for almost four years. A team of three puppeteers generated the entire week's animation (five animation minutes) [73].

A famous animated virtual character in real time that appeared in a live television program, was the cat Ratz [15] in 1994 in the BBC children's "Live & Kicking". Moreover, that same year saw the emergence of the television program "Space Ghost Coast to Coast" [47], a talk show presented by virtual characters in a virtual scene. The interview was previously recorded with a real person, and then the production work was to adapt the entire virtual world as well as the different characters to the interview, giving the effect of a conversation in real time. In the subsequent years virtual characters continued to appear in different television programs, such as those developed for the Korean network KBS. On the one hand Pang Pang, a virtual frog in charge of presenting a children's program, and on the other hand Aliang, who played the role of news reporter in the elections for the Korean National Assembly [70]. Both were controlled in real time by a player using an external motion capture system (see Figure 5).

Figure 5: Pang Pang (Left) animated in real time by an actor (Right) using motion capture [70].

In the summer of 2002 has debuted in TV, the presenter known as virtual Maddy in "Tomorrow's World" from the BBC. Maddy was able to interact with the other program presenters (real) and with the viewers. The main peculiarity of this character was that, unlike most virtual characters, his movements were controlled by an outside person, neither had been previously recorded. Instead, Maddy used sophisticated speech synthesis, speech recognition, synthesis, animation and AI, allowing it to interact in a realistic way and in real time [53].

There are other examples of virtual characters that have appeared in recent years, such as Federico, a representation of the journalist Federico Jimenez Losantos that appears in "The Medium" by La Sexta. The Bonsai, which participates in the program "Club Super 3" on TV3. Tuixén, a weather virtual presenter that also appears on TV3 [22].

Figure 6: Drew Forsythe dressed with the motion capture suit. On the monitor is David Tench.

One of the most important milestones in this area took place on August 17, 2006, when it was unveiled the "David Tench Tonight" [18]. This television show was presented by an avatar in a fully integrated real set and did live interviews with various celebrities. The company responsible for creating and animating the virtual presenter was Anima Logic, who were based system captures the Vicon MX40 motion to allow the presenter to interact in virtual real time with the guest. The actor Drew Forsythe was responsible for giving life to David Tench thanks to 30 sensors located throughout his body and another 38 who were in charge of capturing the movements of the face (see Figure 6).

Sam is a virtual weatherman (see Figure 7a), like in the usual style of its human homologues, presents weather forecasts the TV way. It is multi-platform, and can be seen in television, web and mobile phones. The most famous example of the application of an virtual character as a presenter is Ananova. She was a virtual newsreader, in April 2000 marked a turning point in the public perception of digital characters [53]. She was a success and drew attentions from the public, who visited her site to see her beauty, however she can only mimic a few nonverbal expressions. In Figure 7b it can be seen the attractive newscaster.

(a) Meteo Sam [68].

(b) Ananova [20].

Figure 7: Famous virtual presenters.

2.2 EFFECTIVE AVATARS

An effective avatar can be defined according to certain perspectives. Experts from the most diverse specialties, such as computer science, engineering, animation, sport or psychology, give different meanings to the term virtual human and how it can be effective, in other words, if the avatar can transmit the intended message to the target audience with naturalness. Badler et al. [4] presented a list of features that are a significant contribution to the creation of a virtual human. He suggests that ideally, the motions presented by a virtual human will be biomechanically valid, he should move or respond as a human does. The avatars should exist, act and react within a 3D virtual environment, like real humans do. Other important feature highlighted by the author, is the *"human-like"* appearance. Badler et al. recognized that concerning to our perception of acceptable behaviour, models appearance has an important role.

The Thalman gave a large contribution in the field of avatars, in their book *"An Overview of Virtual Humans"* surveyed the most important aspects in the simulation of a virtual human. At the same time, they recognize that lot of research effort is still needed, "the twenty-first century may not be sufficient for a true imitation of the real complex human and her/his behaviour" [52]. The Thalman's work is focused in different aspects:

- *The good representation of face and bodies*, human modeling is the starting point in the creation of virtual humans. But body deformations and believable face emotions are still hard to obtain.
- *Flexible motion control*, flexibility here means that the animation system must to provide the user with the motion control tools, that allow the user to translate his wishes from is own language.

- *A high level behaviour*, this is related with intelligence and autonomy. Autonomous avatars should be able to conduct themselves within a virtual world. Based on perceived information, the behavioural mechanism of the avatar will determine the actions he will perform. The should look spontaneous and and unpredictable, part of the illusion of life relies on these characteristics. Without it, an avatar would just look like a robot.

- *Emotional behaviour*, emotion is the affective aspect of the cousnciousness, it is very important for the avatar believability to be able to express emotions, being a facial expression, a gesture or a specific behaviour.

- *A realistic appearance*, besides a good visual quality in the face or body, the hair skin and clothes also need to look realistic. They are part of human appearance features.

- *Interact with the virtual world*, the interaction between an avatar and an virtual world must be done in a natural way. Avatars must perform according the object witch they interact with, based in information's about the same, emotional state of the avatar, etc.

- *Interact with the real world*, the real people must be aware of the actions of the avatars, that is achieved with the use of interfaces, the major problem is the avatars perceive the behaviour of real people.

From a psychological perspective, Alessi and Huang [1] focused the rules in the emotional domain. They suggested that virtual humans should be social, emotionally expressive, and interactive. They focused the emotional status of the viewers. Alessi and Huang identified some aspects in the avatar development, such as cultural, educational, psychosocial, cognitive, emotional, and that give an appropriate response, which would potentially include speech, facial, and body emotional expression. Other authors followed this emotional approach concerning avatars, Thórisson and Cassel [75] compared different systems, with and without emotional interaction, and concluded that behaviours relating to the process of dialog, increases user's acceptance and without them will result in less believable characters. In a interesting experience, Pertaub et al. [64] observed that virtual characters with a certain level of visual quality, displaying subtle behaviours similar to humans in a certain mental state, users that interact with the virtual characters will respond in a manner consistent as they were in the physical world. The experience was to put some persons in a empty room, doing presentations to an audience composed by avatars, and conclude that the audience response is affected by the behaviour of that audience even though they know it to be virtual.

These studies have shown that is not sufficient to increase the photo-realistic quality of the avatars, in order the enhance their believability and acceptance by humans, their emotional and body language takes an important role in the communication. The naturalness of the avatar depends of all domains that involve humans.

2.3 EFFECTIVE AVATARS IN TELEVISION

And what about avatars in television? Are the concerns about avatars the same as the above ones? Do they need specific requisites? Shao et al. [69] identified two main aspects in the creation of an avatar with the goal of creating the most lovely and suitable one. From the perspective of this author, these are requirements for the audience to find the avatar attractive and well-recognized. The technical aspect is related with the above mentioned, a realistic appearance and communication skills related with verbal and body expression. Shao et al. in their literature, implicitly criticized engineers, stating that they neglect the cultural meaning of an avatar with these characteristics, justifying by their attachment to the development of avatars to computer software or games. They give a lot of importance the social aspect, cultural identification and aesthetic values. These two factors, social and technical, are related to each other, and the latter one is essential to the success of the avatar, because the virtual character is also part of the culture.

Hipler and Martino [36] analyzed the mannerism of a real news presenter, to build a virtual presenter based on real behaviour patterns. They extracted some characteristics that allowed the creation of a rules set. Those rules included the blinking rate, head nodding, eyes rotation. A real presenter keeps a 'base' position most of the time, therefore all the movements go through an attack and a relaxation phase.

Noma and Badler [57] created a virtual human presenter, identifying the following requirements to make him usable:

- *Natural motion with presentation skills;* Naturalness is of the most importance in order to build credibility with users. Non-verbal presentation skills such as motion, must carefully modeled with this in mind.
- *Real-time motion generation synchronized with speech;* If he acts on a interactive environment, it should react immediately according with users input. It is preferable that his speech is synchronized with the motion, that is generated in real-time.
- *Proper inputs for representing presentation scenarios;* The inputs form of the virtual presenter, such as speech text and embodied commands, should enable designers to represent presentation scenarios without its detailed description. For example, speech texts in

an academic environment are completely different from those for weather reports. It must provide the user with the controllability to adequate the inputs according with the scenarios.

Another interesting application of avatars as television presenters, was the one developed by [61] et al.. They used a virtual character based on a cartoon know by the audience form movies. It was focused to be implemented in a interactive real-time environment. The virtual character is controlled by actors (see Figure 8) using techniques of real time phoneme recognition and lips synchronization, and real time corporal animation.

Figure 8: Betizu, a virtual presenter prototype [61].

They established two main goals within this project:

- *High quality of the realism of the virtual presenter;* the animation must be very realistic in order to allow a natural association between the movies cartoon and the virtual character. To achieve this high realism, are used low level API's and low cost animation algorithms allowing thus the animate complex virtual characters in real-time.
- *Real-time interaction in a non intrusive way;* the system must be flexible, the user must be able to use any kind of input devices (e.g., a joystick).

Virtual characters are a intriguing fusion between art and computer science. In the last few years there has been advancements focused on create believable and attractive characters. Using advanced voice recognition systems, animation, etc., it is now possible to create characters that can bring a positive addition to television programs. The above example, Betizu [61], was tested in a Basque Country television, in a program aim to children, and it was a success that was broadcasted daily.

The role of virtual characters can be very important in enhance the mass communication, by *"well representing the information as well as* *improving the general publicity*" [69]. The author shown that the use of avatars in communication could increase the sense of social presence. This is a central concept in media technology. It can be defined as the perceptual illusion in which the user of a medium experiences sensations of being present in an environment [13]. People can experience in two separeted environments, the physical environment in which they are present, and the environment presented via the medium. A conceptualization of presence as an element of realism was presented by Lombard and Ditton [50], concerning the degree which a medium can reproduce accurate representations of objects, events, and people. Social presence occurs when users feel that a form, behaviour, or sensory experience indicates the presence of another intelligence [13]. Avatars, for example, can impersonate historical figures assuming his appearance and narrate a historical event, making it more appealing and fun.

2.4 SUMMARY

In this chapter it was given an overview of the most known applications of avatars in television. It were covered some aspects of an effective avatar, one that look natural, individualizing to the television context. The applications reviewed are not flexible in terms of the avatar appearance, in this work we intend to contribute with a software that allows a quick a easy appearance change of the avatar, according with the context of the television program.

3 REALISTIC HUMAN SIMULATION AND BODY MOTION CAPTURE

One of the most interesting challenges in computer graphics is the simulation of human body motion. A fundamental form of communication among humans it is body language. The more an avatar appearance resembles with a human, more it will be acceptable. For example, a person's mood can be judged uniquely from body language. So, how to get a virtual character to perform motions as a human does? It is quite difficult, humans can perform complex movements that can range from highly dynamic to extremely subtle, such as a back flip or shack hands. Human movement is very complex, it depends of many variables, for example, psychological state and surround environment. The subtlety and spontaneity of human gestures is very difficult to reproduce with virtual characters. Gutierrez et al. [34] identified two main approaches to this problem: the biology (bio-mechanics studies) and the non-biology based (more computer graphics oriented) approach. The biology based approach has been used in the research of bio-mechanical engineering, robotics and neurophysiology, to clarify the mechanisms of human walking, and modeling the muscular actuation systems of animals and human beings [34]. It was stated before on this thesis that realism is seen from a functional point of view, therefore an computer graphics approach is more adequate.

The interest in this field of research has increased in the recent years, and the pursuit of the answer to this question has brought amazing developments. This due partly to the fact that the technologies involved in the process have evolved, meaning the sensors used to capture the motion and the computers capacity to take advantage of them [52]. As it was stated before, the naturalness of an avatar relies a lot in the corporal animation. It is an critical aspect.

The lifelike simulation of human beings it is a field of study that suffered dramatic changes in the past decade, the human representation is no longer a stick figure or simple polyhedral models, it has evolved to high realistic volumetric representations. Every detail of human appearance is taken into account: hair, clothes, skin, behaviour, etc. Hair color, variable thickness, dynamics and all the characteristics that are familiar to humans must be simulated in a realistic fashion, the polygonal hair representation used two decades ago, is no longer accepted. The same problem is applied to cloth. Garment is very important in the beauty and social appearance of an individual. Their simulation and animation are the result of the intersection of different technologies,involving physically-based mechanical simulation in order to adequately reproduce the shape and motion of the clothes over the virtual body [52]. Virtual humans are more and more intelligent, nowadays, they are capable of making decisions and have memory. If they populate a virtual world, they are capabale of interact among each other; cooperate between them, in order to perform tasks [21] and avoid obstacles. Virtual characters are capable of transmitting emotions, Ortony et al. [60] proposed an emotional model where they define 22 emotions based on positive or negative reactions. Ekman defined six basic facial expressions, they are joy, sadness, anger, surprise, fear, and disgust. These are used by many researchers in facial expression analysis , and recognized as universal.

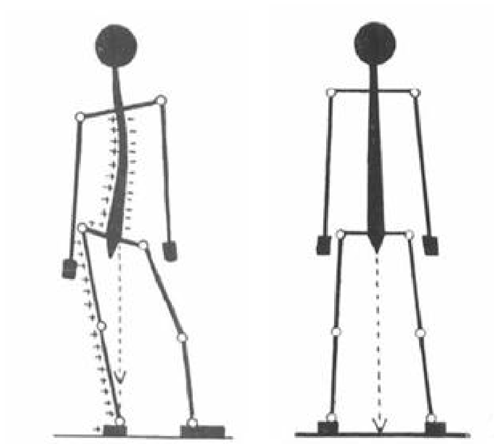
Human motion is an huge field of research, due to it inherent complexity. The everyday movements of people are amazingly fluid yet demanding to reproduce, with actions driven not just mechanically by muscles and bones but also cognitively by beliefs and intentions [4]. On this thesis a particular human pose will be investigated, the standing posture. This particular pose is characterized by an upright posture, supported only by the feet. What is the contribution of improving such a subtle motion? It was stated before that humans have a sharped sensitivity concerning human motion, therefore even small details can look unnatural. The avatar cannot be static when their motions are not solicited.

Although being extremely complex to reproduce the randomness of movements performed by humans when standing or waiting, some models where developed, traducing mathematically the most erratic and chaotic moves. In this chapter the literature concerning these models will be reviewed, but the approach of this thesis will not be to reproduce the accuracy of such models. As it was stated before, the naturalness improvements will be made from a functional realism perspective. With this in mind, the process of recording the moves will not be directed to the imitation of such a complicated models, instead it will be the search for express emotions with body language. In the context of television this assumption gains importance , the content needs to be atractive and an static avatar in the monitor is not acceptable.

This chapter deals with the issues related with the capture and animation of human motion. In first place it will be described the human posture of waiting and the issues that are related to it. The following sections deal with the capture and animation of that posture, starting by the animation techniques, then the motion capture technique that deserves special attention on this thesis.

3.1 Standing posture characteristics

Contrary to what we might think, the stang posture is not a static pose, in fact it is characterized by a varied number of movements, and is a very complex human behaviour. This issue has being dealt by several investigators. During quiet standing the body sways back and forth in a slow fashion, interrupted by fast and gross postural changes [27]. A model to describe the sway has been proposed by several authors, claiming that it is impossible to stand still like a rigid pillar without external support [9]. Collins and De Luca have shown that postural sway is indistinguishable from correlated noise and that it can be modeled as a system of bounded, correlated random walks [14]. A simpler model that can represent this behaviour, can be the non-trivial physical model, the inverted pendulum. Winter et al. [77] demonstrated that this model is valid. Several authors researched with the aim to understand the reflex and sensory information that human use to maintain standing posture [44, 26]. People employ diverse strategies to deal with the disturbances that occur during a quiet stance, these can be divided in ankle strategies and hip strategies [23]. The majority of the researchers in this field of study seem to agree that humans use ankle strategies in response to small disturbances, and hip strategies in larger disturbances.



(a) The asymmetrical attitude

(b) The symmetrical attitude

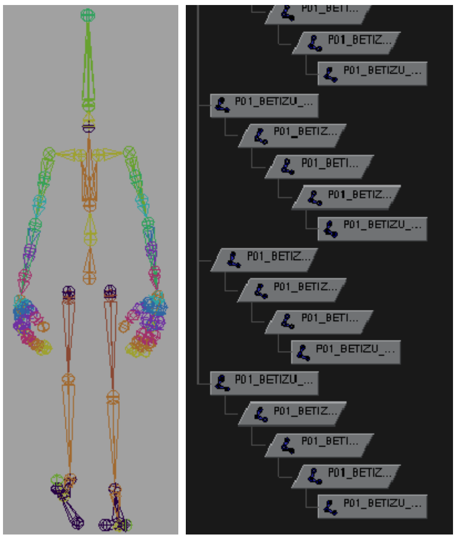
Figure 9: Standing posture main attitudes [71].

Smith [71] found that the attitudes which are standing, fall into two categories. The first of that category is the symmetric attitude (see Figure 9b), and it is characterized by an equal distribution of weight by the feet, both firmly planted on the ground and legs slightly apart. The shoulders are level. The second category is the asymmetrical attitude (see Figure 9a), it principal characteristic is that body weight is supported almost entirely by just one foot, whereas the other carries very little weight, it is only to maintain balance. It can be either the right or left foot. One of the shoulders is lower than the other, depending on the side of the body that bears the weight. For example, in Figure 9a it can be seen that the right shoulder is lower than the left, the right leg supports the weight. The line of the center of gravity must pass by the foot that supports the weight, in order to maintain the comfort and balance. Another important conclusion by Smith, is that he found in the subjects observed, that one of the two asymmetrical attitudes occurred approximately four times as often as the symmetrical attitude.

Freitas et al. identified psychological and physiological factors that explain these postural changes, it are thought to be performed in order to diminish the discomfort caused by increase of tension, mental stress, and reduction of motivation and concentration [27]. Although the physiological factors (increase of venous pooling in the lower extremities, occlusion of blood flow, vertigo, muscular fatigue and increased joint pressure) are important in the science of posturology[1], in this thesis, we assume that these factors do not contribute substantially to the naturalness of the avatar.

3.2 Representing Human Figures

In the representation of human motion, it is a common practice to use skeleton structure in order to define the moving parts of the figure . Human models can be multi-layered or rely only in two layers, where the skeleton is underneath the skin of a human body model [4]. In Figure 10 can be seen a model formed by a skeleton (left) and the skin (right). Other model can be defined the shape of the body with geometric primitives such as ellipsoids and spheres, denominated of volumetric models .



(a) Joints and bones. (b) Hierarchical structure.

Figure 10: Hierarchical skeleton structure.

1 Posturology is defined has the science of human balance in every physiological condition. It involves different disciplines such as bio-mechanics, neurology, anthropology and other.

Leclercq et al. [48] developed a method for generate a smooth mesh from combination of triangle meshes with geometric primitives. The multi-layered models, in the recent works are inspired in the major anatomical layers of the human body, such as the muscle and fat layer.

3.2.1 *The skeleton*

In the past, some application such as 3D games, did the animation by exporting all the animated mesh frame by frame from a modeling tool, however this process demands a lot of memory usage, especially in complex animations. This happened because all the mesh (all vertexes and normals) needed to be stored in memory in order to be animated in real-time. The solution for this problem can be the use of a skeleton inside the mesh. Therefore, instead of animating, it is animated the skeleton within. This makes the animation less memory consuming.

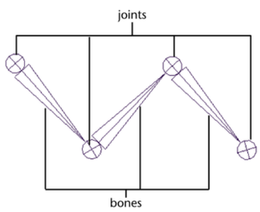


Figure 11: Joints and bones.

The best case scenario would be the avatar to have bio-mechanically valid movements, at least it needs to move in a "human-like" fashion. For that, the skeletal structure must resemble but need not to copy that of humans. Skeletons are hierarchical, articulated structures that allow pose and animate models. A skeleton provides a deformable model with the same underlying structure as the human skeleton gives the human body. The skeleton is typically composed of joints and bones that form joint chains. A joint is the area where two bones are attached for the purpose of motion of body parts.

3.2.2 *Skinning*

Skinning is one type of surface deformer. Every vertex is tied to at least one joint through an influence. An influence stores the vertex, the joint index and a weight which specify how much influence the joint has over the vertex. A vertex can have several influences but only one for each joint. The sum of all the weights in a vertex influences should always be 1. This is process is called skinning because the mesh can be seen as a virtual skin over the skeleton.

"There is no particular mystery in animation... it's really very simple, and like anything that is simple, it is about the hardest thing in the world to do." Bill Tytlaat the Walt Disney Studio, June 28, 1937.

Animation can be defined as *"making the things move"*, to give life and emotional expression to inanimate objects. It can be anything from something as simple as paper and pencil to state of the art computers hardware and software tools. With the maturing of 3D computer animation research, a lot of things had changed since the early computer animations based on traditional animation techniques. Nowadays, the resources are more devoted to image rendering than to animation [46]. In their remarkable book *"Disney Animation: The Illusion of Life"* [40], Thomas and Jonhston enunciated the 12 principles of animation (squash and stretch, anticipation, staging, straight ahead action and pose to pose, follow through an overlapping action, slow in and slow out, arcs, secondary action, timing, exaggeration, solid drawing, appeal). These principles are equally applicable to both traditional forms of character animation as well as computer animation, the artist must have them in mind to create natural characters. However, commercial programs such as Poser, already apply these principles at the time of animation creation, but a good designer must keep them in mind.

In this section it will be presented the most prominent techniques in computer animation. The most relevant in this thesis, motion capture, is described in a dedicated section.

3.3.1 *Keyframe Animation*

Keyframing is one of the methods used to produce animation. Is a popular technique in which the animator explicitly specifies the kinematics by supplying keyframe values whose "in-between" frames are interpolated by the computer [51]. The underlying concept of keyframe is to decompose complicated motions into shorter sections, where the extremes are the important images, the more meaningful ones. This technique was used by Walt Disney since the primordial times of animation [39]. However, this technique demands skillful animators to specify the intermediate poses, and the interpolation can result in invalid motions.

3.3.2 *Inverse and forward kinematics*

With this approach the problem of human motion is stated as: "Given one desired position of a part of the body, what must be the angles of the body joints?" In inverse kinematics, there is the need of specified points, called end-effectors, and on the links are assigned target positions. To

Figure 12: Keyframe sequence. Different postures specified at different time points [16].

solve the Inverse Kinematics problem, it must be found settings for the joint angles so that the resulting configuration of the body places each end effector at its target position [10]. With inverse kinematics the user avoids explicit control over all joint values due to the high level of abstraction for specifying postures that this technique gives.

It is hard to achieve realistic results, having in consideration the complexity of human motion. Early studies were made by several researchers in order to predict the appropriate movements. Paul [63] has researched manipulators control in the robotics field in 1982 using inverse kinematics. Badler et al. [4] provided an interactive system that allows interactive controlling the posture of the character by specifying various geometric constraints such as end effector position and orientation. The system repeatedly evaluates the kinematic constraints and executes the behaviours.

In contrast to inverse kinematics, with forward kinematics the objective is to specify all joint motions in order to determine the position of the end-effector. For example, to move a hand to some location, several arm joints must be rotated to reach the location. However, is no easy to estimate the exact angles of the joints, in order to achieve a certain end-effector position.

3.3.3 *Dynamics*

If the purpose of an animator is a faithful representation of reality in physical terms, let's take, for example, the human locomotion, it will need to use physics-based techniques. These allow high realistic and physically plausible animations. Dynamics takes into account several factors such as friction, gravity and external forces that may act in the system, it relies in the laws of physics, it is distinct from purely geometric approaches such as kinematics. However, the computational cost is enormous, one of the challenges is to develop these techniques with low computational times.

Similar to kinematics, in dynamics there are two approaches, forward and inverse-dynamics. With the first, the objects are controlled

Figure 13: A dynamic "virtual stuntman" plunging and rolling[23].

by manipulating forces and torques in order to obtain trajectories and velocities [30]. In inverse dynamics, on the other hand, the goal is to determine the forces and torques that are required to produce motions. Faloutsos et al. [23] detected a lack of movements contemplated by physically based methods, and proposed a method to generate a variety of human-like motor behaviours. The method uses controllers to define pre-conditions, post-conditions, and expected performance that characterizes a specific motion.

3.3.4 *Parametric animation*

In general, parametric animation consists in merge two or more animations to create a new one. The characteristic parameters for motion are specified and interpolated. With parametric animation is possible to an animator create a small set of animations, and construct a lot more from the initial ones. Byun and Badler [11] proposed a parametrization for a facial animation system, their system applies knowledge in facial physiology and heuristic methods in manipulating animation data.

3.4 CAPTURING THE MOTION

The most popular technique for retrieving human motion is known as motion capture. A realistic motion can be animated by manually specifying a sequence of character poses, such as the keyframing technique explained above. However, it demands a big investment of time by an skilled artist, just to get a few seconds of quality animation. Motion capture is an suitable answer for this problem, it can ease and speed the animation process, it is a very fast and accurate way to bring human motion into a computer animation. Though motion capture has obvious advantages, it can only capture a limited range of movements, it is bounded to human capacity. A movement that does not follow the laws of physics generally cannot be represented, which happens very often in movies and games with some characters with special abilities. Human movement has variations depending of the most diverse reasons (psychological and physical conditions, environment characteristics, etc). The walk motion, for example, is not a repetitive behaviour. This characteristic is very difficult to reproduce with motion capture, it demands an huge amount of different moves recording. In the particular case of this thesis, the recorded motions are relatively simple, the "waiting" state. It is perfectly feasible to use motion capture to retrieve them.

3.4.1 *Background*

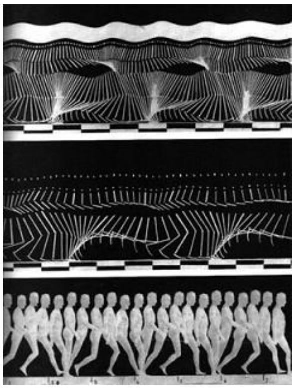
Motion capture (Mocap) is defined by Liverman [49] as the process of obtaining and recording a three-dimensional representation of live action or event, by capturing an object's position and/or orientation in physical space. The information captured can be the simple position of the body in space or as complex as the deformations of the face and muscle masses [72].

Mocap for computer character animation is a relatively new, it begun in late 1970's. However, the attempts of capture human movements in order to get more realistic animations had begun long before. In the late 1800's, Eadweard Muybridge invented the zoopraxiscope. A device for displaying motion pictures previously captured with a photographic camera. The zoopraxiscope projected images from rotating glass disks in rapid succession to give the impression of motion.

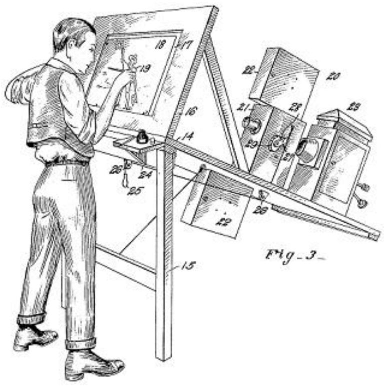
Muybridge along with Etienne-Jules Marey were the first persons to study human motion, photographing and record the movement [37]. In Figure 14a it can be seen a capture of a rapid sequence of images. With several cameras focused on people movement, Marey dissected time and space, revealing the dynamic rhythms of the human body in motion.

To get more convincing motion for the human characters, Max Fleischer invented a technique called rotoscoping. Using a rotoscope (see Figure 14b) the artist draws on a transparent easel, onto which the movie projector is throwing an image of a single film frame [17]. Therefore, rotoscoping can be thought of as a primitive form or precursor to motion capture, where the motion is captured by hand. This technique has been used ever since. Disney's movie Snow White and the seven dwarfs (1937) is an example of an effective utilization of this technique.

It was on the late 1970's that laboratories were beginning to use computers to analyze human motion. In the early 1980's, Tom Calvert [12] attached potentiometers to a body and used the output to drive computer animated figures for choreographic studies and clinical assessment of movement abnormalities. To track knee flexion, for instance, they strapped a sort of exoskeleton to each leg, positioning a potentiometer alongside each knee so as to bend in concert with the knee. The analog output was then converted to a digital form and fed to the computer animation system. Their animation system used the motion capture apparatus together with Labanotation[2] specifications to fully specify character motion [72].



(a) Etienne-Jules Marey, human motion photographic study, 1886



(b) Rotoscope

Figure 14: First motion capture techniques.

Later on, the optical tracking systems have started to appear. Optical trackers typically use small markers attached to the body (either flashing LEDs or small reflecting dots) and a series of two or more cameras focused on the performance space. A combination of special hardware and software pick out the markers in each camera's visual field and, by comparing the images, calculate the three-dimensional position of each marker through time [72]. It is a technology with several limitations, markers can be occluded by the body preventing the reading, there is a limit of positions per second that the markers can capture, and if one marker is too close to another, it is very difficult to differentiate due to the low resolution of the cameras. In 1983 Ginsberg and Maxwell [31] at MIT, presented the Graphical Marionette. Using an optical motion capture system called Op-Eye (that relied on sequenced LEDs), they wired a body suit with LEDs on the body joints (knee, elbow, etc) and another anatomic relevant parts. The capture was made with two cameras with special photo-detectors that returned the 2D position of the LEDs within their field of view. Subsequently, a computer uses the position information from the two cameras to obtain a 3D world coordinate for each LED. The system used this information to drive a stick figure[3] for immediate feedback, and stored the sequence of points for later rendering of a more detailed character.

In 1989, Dozo was created by Jeff Kleiser and Diana Walczak. It is a non real time animation of a woman dancing in front of a microphone singing a song. Motion capture was used to generate a more realistic motion for Dozo. They have used an optical motion capture system and

2 Labanotation is a system of analyzing and recording of human movement. It was invented by Rudolf von Laban (1879-1958) an important figure in European modern dance [33].

3 A stick figure is a very simple type of drawing made of lines and dots, in this case of a human.

Figure 15: Super Mario character controlled in real time using Face Waldo.

multiple cameras to triangulate the images of small pieces of reflective tape placed on the body. The output is the 3D trajectory of each reflector in the space. As was described for the Graphical Marionette, one of the problems with this kind of system is tracking points as they are occluded from the cameras [72]. For Dozo, this had to be done as a very time-consuming post-process. Luckily, some newer systems are beginning to do this in software, significantly speeding up the motion capture process. Around 1992, SimGraphics developed a facial tracking system they called a "face waldo", see Figure 15. Using mechanical sensors attached to the chin, lips, cheeks, and eyebrows, and electromagnetic sensors on the supporting helmet structure, they could track the most important motions of the face and map them in real-time onto computer puppets [72]. The interface is perfectly natural, since the actor could manipulate all the facial expressions of a character by just miming the facial expression himself. In SIGGRAPH '93 Acclaim presented a realistic and complex two-character animation done entirely with motion capture. The system was similar to the ones used for the Graphical Marionette and Doz, an optical motion tracking system, but with the improvement of being able to track up to 100 points simultaneously in real-time.

In the recent years, motion capture techniques have continued to be refined, including new hardware with more possibilities and new software to use the hardware. Recently, this technique has been commonly used in several movies. For example, Polar Express was the first film to successfully utilize facial motion capture for an entire computer generated animation movie [7]. Tom Hanks face was digitalized using a facial motion capture system, with several markers attached to his face and a system that deforms face geometry based on the output data of the muscle and marker movement combined, it was possible to achieve a realistic face. Facial motion capture was a technique used in the Benjamin Button movie [25], when the main character is old, his face is completely generated by advanced face motion capture.

There are many fields of application of this technique of animation besides the cinema, highlighting once again the entertainment as the most used. As previously mentioned, the process can be divided into two stages: the motion capture and transmission of such movement to virtual character. Regarding the first part of the process, there are numerous techniques for capturing the movements. These techniques range from traditional to modern optical systems and magnetic mechanical systems [29].

3.4.2 *Types of motion capture systems*

The available motion capture systems can be categorized into three main groups: the optical systems, magnetic systems, and mechanical systems [43]. Each type has strong and weak points. This section is an overview of those three categories.

3.4.2.1 *Optical systems*

The actor wears a suit with markers attached, or they are directly attached to his body. These markers can be either passive or active. In the case of the passive, the markers are made of a reflective material and their size and shape can be diverse, depending on camera resolutions and captured subject (for example, to capture facial motion is used small markers).

Figure 16: Optical motion capture system used in the film industry. At the left is the actor Andy Serkis wearing the motion capture suit, and in the left is the result of the capture, the creature Gollum (Lord of the Rings) [62].

The cameras in this system are equipped with light-emitting diodes (LEDs), the lights that are emitted by the LEDs are reflected by the markers and captured again. In a active system, the markers are LEDs. They emit their own light to the cameras, where a system activate one marker at the time, eliminating the need to identify each one. All LEDs can be illuminated at once, in this case each one is modulated with different frequency or amplitude. The 3D position of the markers is calculated with a minimum of two cameras.

This system has several advantages. Optical data is accurate, the capture volume can be larger than most other systems and the capture subject can move freely. Hover is more expensive than other types of motion capture systems, markers can be occluded resulting in loss of data and lighting needs to be controlled. It has a limited real-time visual feedback, only stick figures.

3.4.2.2 *Magnetic systems*

These systems are also called of electromagnetic. It is composed of several sensors placed on the captured subject that measure the the spacial relationship to a magnetic transmitter. The sensors output their rotations and translations, thus there is no need for post-processing to compute the rotations. In terms of occlusions these kind of systems are more tolerant than the optical ones, only metallic objects interfere with the output quality. However the electronic equipment also an interfere also. Another great advantage is their price, less expensive than optical systems.

(a) Magnetic system [58]. (b) Mechanical exo-skeleton[54].

Figure 17: Two motion capture systems.

3.4.2.3 *Mechanical systems*

These systems directly measure joint angles of an articulated device used by the captured subject. It is similar to an exo-skeleton. It is free of magnetic or electrical interferences, occlusions and relatively inexpensive. It has a great drawback, these systems are not capable of measuring global translations. For example, if the subject jumps vertically, the data output will be the same as if he stayed on the ground, that data does not follow the jump, because the joint angles remains the same. The system is limited to the measure of joints movements. To overcome this issue, magnetic sensors are added to the system.The equipment restricts subject movements, due the volume and break ability of the exp-skeleton.

3.5 SUMMARY

This chapter covers methods to achieve an realistic motion for avatars. It focus in a particular behaviour, the standing posture. Aiming a believable and natural avatar, in this work we have used a motion capture system to create animations to this posture. We have consulted some work in the field of posturology in order to better comprehend the main characteristics of the standing posture. We also investigated several motion capture paradigms and their pros and cons.

4 IMPLEMENTATION AND RESULTS

In the introductory chapter was presented the problem in question: to improve the naturalness of an avatar in a television environment. This work is part of a project that aims the inclusion of avatars in real scenarios, a more detailed description of the project can be consulted in appendix A. The present chapter it is dedicated to the implementation details, and how the proposed objectives were achieved. At first, it will be explained the development process of the plugin, as well as an overview of the architecture behind the same. Then, it will be shown the process of capturing the human motion and how it was applied in the avatar.

4.1 PLUGIN DESCRIPTION

The plugin is loaded in runtime, and is integrated in Maya graphical interface as an external program that can be easily invoked, allowing an easy and familiar usage for Maya users such as artists, with a button inside of Maya that they simply press and can export data from Maya to a format compatible with their purposes. Figure 18 represents the pipeline of the plugin. When is invoked in Maya it creates a file with information that is used by the application that loads the avatar.

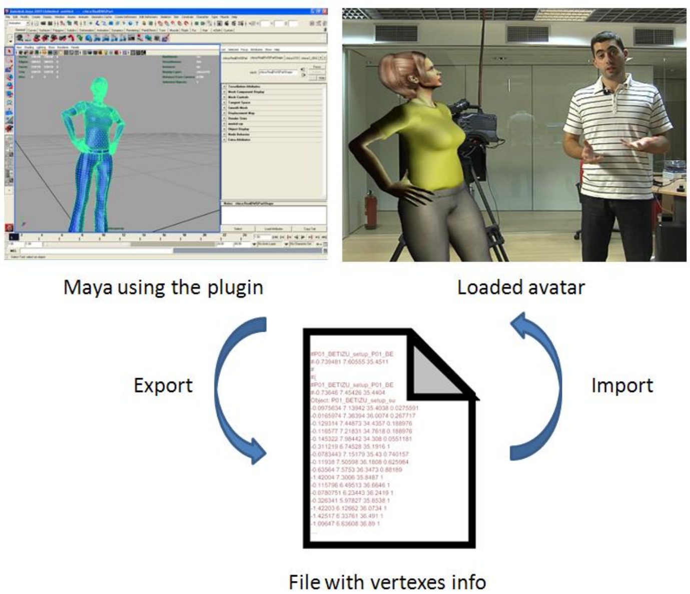


Figure 18: Plugin import/export system example.

Figure 19 shows Maya interface. In the figure can be seen, marked by the red squares, the addition of the *"Exporter"* button in the panel and subsequent pop-up window with several options. The interface allows the invocation of the functions implemented in the plugin.

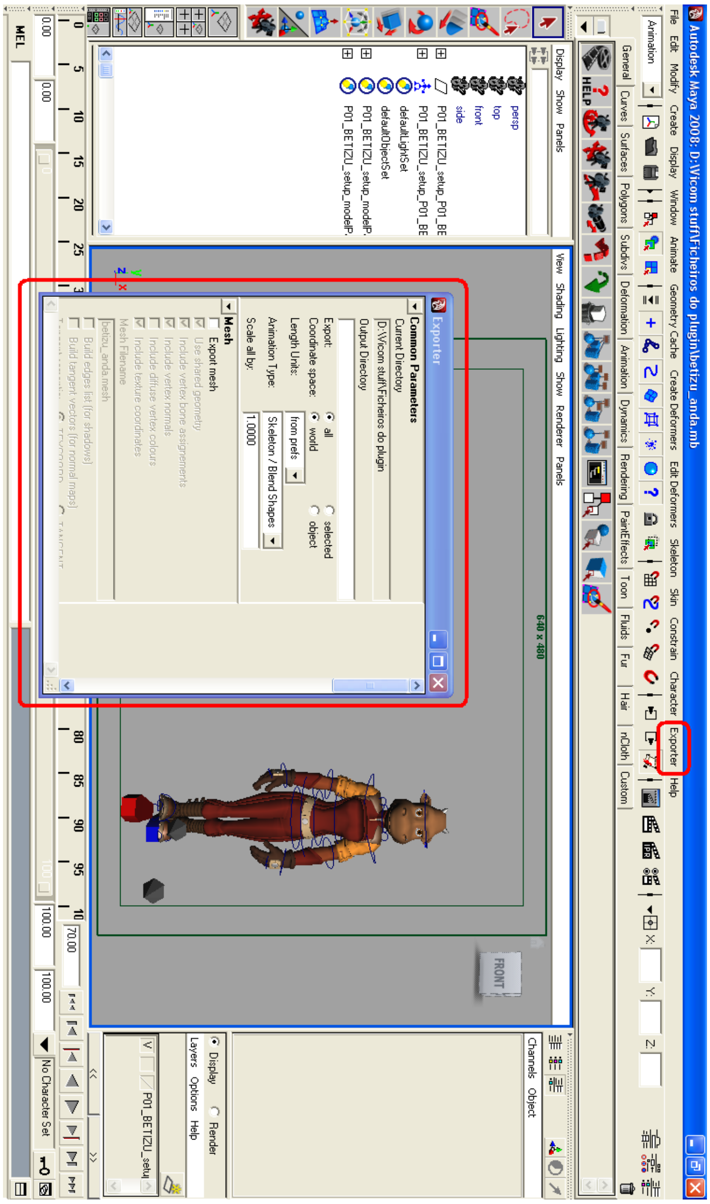


Figure 19: Plugin interface integrated in Maya GUI.

Maya supports many different export formats such as VRML, OBJ and many more, but none of these fitted within the needs of the project.

The plugin output is a file containing the information of how a joint rotation affects the vertexes of the mesh. This file serves the need for realism in the deformation of the mesh. The rotations of the joints could provoke breaks in the mesh, called cracks. The rigid parts of the body model are subdivided in virtual sections, depending on the nearness from a joint in a hierarchical way. In Figure 20 it is shown a visual representation of the weight values for each vertex. A weight is assigned to each vertex, and it is rotated depending on this weight. The weights are assigned during the skinning process.

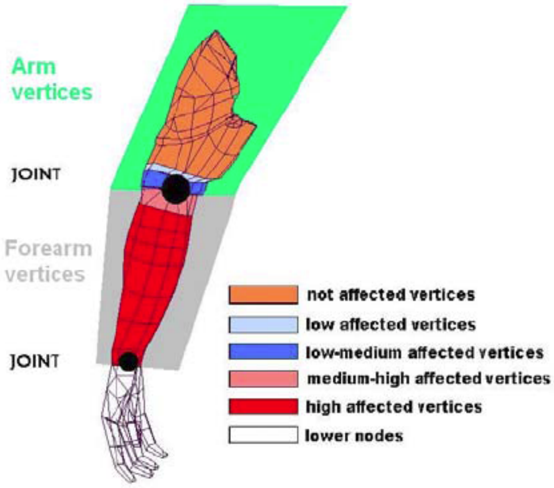


Figure 20: Vertexes affected by the movement of the elbow [61].

The data in the file, must be organized in a structure that reflects the hierarchical nature of the skeleton. A file of this kind cotains the following information:

- Hierarchical structure of the exported model skeleton;
- Joint name, position and list of influenced vertexes;
- Vertex coordinates;
- Joint influence over the vertex;

Each vertex can be affected for more that one joint, therefore can be repeated in the file. One fragment of the resulting file is shown below.

```
#Joint name
#X Y Z (joint position)
#
#{
#Child joint name
#X Y Z (child joint position)
Object: Object name
x y z weight
#
#}
```

The file header is the root joint of the skeleton, all the other joints are children of that joint. To each joint are associated objects, witch are the virtual sections mentioned before. These objects are subdivisions of the mesh, and they contain one or more vertexes affected by the joints rotation disposed in a list.

Typically, artists perform their work in Maya using the graphical interface. It includes menus, buttons, dialog boxes, etc. This interface eases the task of creation in Maya, but this tool allows the accomplishment of the same tasks through two programming interfaces, MEL and C++ API.

MEL is an acronym for Maya Embedded Language [32], and it is especially designed for Maya, it has a simple structure and syntax, making it more accessible for inexperienced programmers. It is an interpreted language, witch is a drawback, because it runs slower than a C++ program, since this one must be compiled from source code to generate actual native machine instructions, and it runs very fast, making the export a quick process. Using the C++ API, user can create plugins and write software that allows you a tighter level of integration with the core Maya package. The API is an interface that consists of a series of C++ class libraries. The plugin is a C++ program that uses and extends the basic Maya classes [32].

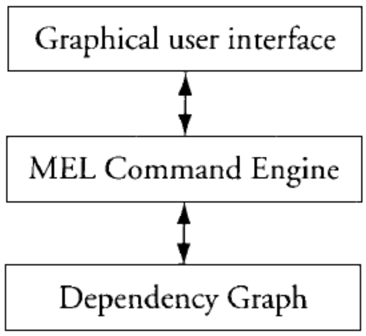


Figure 21: Maya System [32].

The Maya system can be decomposed in three main components, that can be seen in Figure 21. The common user interacts with the Maya Graphical User Interface (GUI). At any operation taken by the user (selection of items, animate and move objects, etc), Maya issues MEL commands that are sent to the Command Engine, where they are interpreted then executed. The majority of the MEL commands operate on the Dependency Graph (DG). The reason for this is that the *"Dependency Graph can be intuitively thought of as the entire scene"* [32]. The DG is really the heart and brains of Maya.

The DG is the core of Maya. Information in the DG is stored in objects called nodes. Nodes have properties designated by attributes, that store the configurable characteristics of each node. Similar types of attributes can be connected together, letting data flow from one node to another. When using the C++ API, it is possible to create custom Dependency Graph nodes. These can be integrated directly into Maya and work seamlessly with all other nodes.

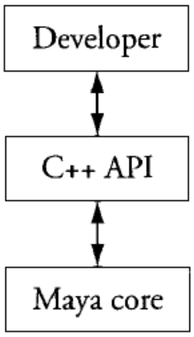


Figure 22: Maya programming interface [32].

It is not possible the direct manipulation of Maya objects and data structures, with the C++ API. What happens, in fact, is that the programmer uses a layer above the actual Maya core. The diagram in Figure 22 shows the different programming layers and how they communicate with each other. The API gives access to the core through interfaces, but at no time gives direct access to Maya core.

The information of the vertex weights is stored in a structure called skincluster. It is a node created during the weight attribution process. The purpose of the skincluster is to store a weight per influence object for each component of each geometry that is deformed. The strategy taken to write the plugin, was to create a structure that hold information relating vertexes, objects and joint relations. By accessing Maya nodes was possible to retrieve all the data needed to create the file. The first step to retrieve the vertex weights, after all meshes and vertexes loaded, is getting the skeleton. The skeleton needs to be in the "bind-pose". This is the pose of the mesh object and the skeleton at the moment the skeleton is bound to the mesh object and before any deformations begin to occur.

```
//This method receives as argument the skincluster retrieved
// with a Maya API method, correspondent to some mesh
MStatus Skeleton::load(MFnSkinCluster* pSkinCluster,ParamList&
    params){

        //Calculate the number of influenced objects
        int numInfluenceObjs = pSkinCluster->influenceObjects(
            influenceDags,&stat);
        (...)
        for (int i=0; i<numInfluenceObjs; i++){
                //Reset to the bind-pose
                MGlobal::executeCommand("dagPose -r -g -bp",true);
                (...)
                //Invoke the method to load the joints
                loadJoint(rootDag,NULL,params,pSkinCluster);
        }
        (...)
}
```

With the skeleton loaded, it is possible to get the joints:

```
MStatus Skeleton::loadJoint(MDagPath& jointDag,...,MFnSkinCluster*
    pSkinCluster){
        //Found a joint
        if (dagPath.hasFn(MFn::kJoint)){

                MFnIkJoint jointFn(dagPath);
                joint_s newJoint;
                // Get all information: matrixes, translation,
                    scale, position...
                        (...)
                //Vector that contains all the joints
                m_joints.push_back(newJoint);
        }

        // Load child joints
        for (i=0; i<jointDag.childCount();i++){
                MObject child;
                child = jointDag.child(i);
                MDagPath childDag = jointDag;
                childDag.push(child);
                //Recursive search for the rest of the joints
        loadJoint(childDag,...,pSkinCluster);
        }
        return MS::kSuccess;
}
```

With all the joints retrieved, it is now possible to fill the following structure that holds the information with the relation between joints, and their influences in the vertexes of a mesh.

```
typedef struct objindarray {

                std::string objName;
                //index of the vertex
                std::vector<double *> vertInd;
                std::vector<float> weight;
        }objIndArray;

// structure to hold joint info
typedef struct jointTag{
        MString name;
        int id;
        (...)
        std::vector<objIndArray> objectsArr;
} joint;
```

All the needed information is now retrieved. The file can be built.

```
MStatus OgreExporter::writeSkeleton(int number, ofstream* ofile) {

        (...)

        for(int i=0; i < skel.at(number).objectsArr.size(); i++){
                //print the name of the object who vertexes belong
                *ofile << "Object: " << skel.at(number).objectsArr.
                    at(i).objName.c_str() << "\n";
                for(int j=0; j < skel.at(number).objectsArr.at(i).
                    vertInd.size(); j++){
                        double* vertex = skel.at(number).objectsArr
                            .at(i).vertInd.at(j);
                        //print vertex x y z values
                        *ofile << vertex[0] << (...)
                }
        }
        //Write the children info
        (...)
        return MS::kSuccess;
}
```

This file contains the information needed to deform a mesh according with the joint movements. Together with geometry information it is possible to quickly load an avatar. This allow character appearance modifications, in other words, the ability to change the character.

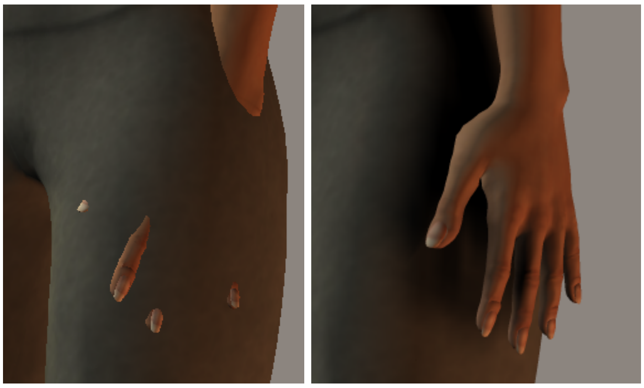
4.3 MOTION CAPTURE SESSION

For the motion capture session, it was used the Animazoo IGS-190 [3], a gyroscopic inertial motion capture system. In order to record a person's motions with this system, a set of markers need to be placed on the body. It is a magnetic system, uses the sensors placed on the body to measure the magnetic field generated by a transmitter source. The sensors and source are connected by cables to an electronic control unit that correlates their retrieved locations within the field. The electronic control units are networked with a host computer that uses a software driver to represent these positions and rotations in 3D space. To solve the angles for the various body joints, it is used Inverse kinematics.

The output of the motion capture system is a BVH[1] format file, standard representation of movements in the animation of human body structures. This file is used to animate the avatars, applying it in the avatar underlying skeleton. It has the advantage of being independent of the 3D model geometry where is going to be used. However, this

1 BVH stands for Biovision Hierarchy, this format was developed by Biovision, a motion capture services company, in order to represent character animation.

representation is not free of problems, since the final result needs some corrections.



(a) Wrong pose. (b) Correct pose.

Figure 23: Avatar pose correction.

In the above figure it can be seen a situation where corrections are needed. The left hand collides with the left leg (see Figure 23a), using a commercial tool, Poser [65], it is possible to correct the hand position. In the right image (Figure 23b) the hand is over the hip without collision.

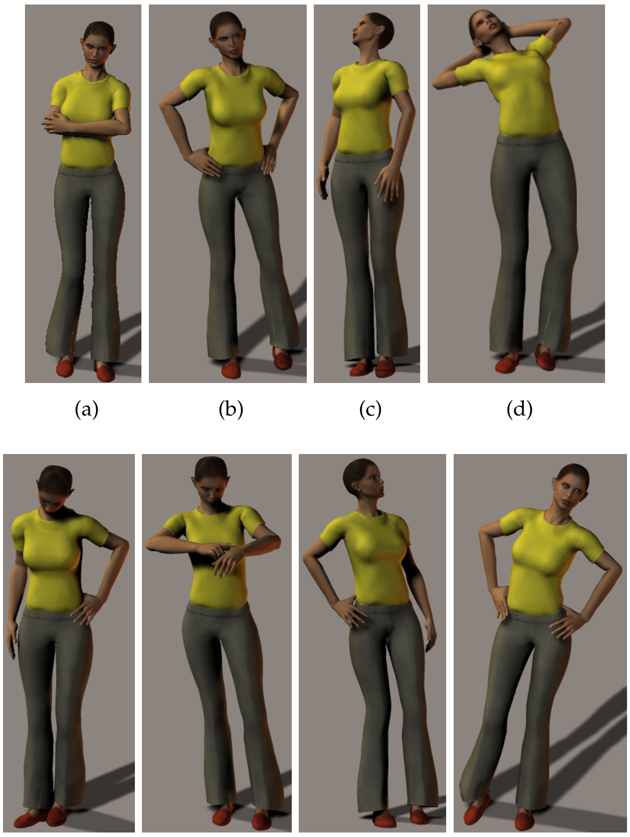
4.3.1 *Recorded motions*

For this session, eight different states of waiting were recorded (see Figure 24). These positions represent different states of mind through corporal language. Anxiety, impatience or tension can be perceived in the avatar motions and pose. The character becomes very expressive and natural. Based on the studied human movements in this position, he reproduced some of those in the most natural way possible. In order to reproduce them, motions are the input to a rule-based AI module that decides, depending on the user interactions and system events, the most suitable movement to reproduce.

4.4 SUMMARY

In this chapter, we have presented the implementation of the plugin that allows the exportation of a model from Maya to a specific format. This format specifies the joint influences to each vertex of the mesh, each with an associated weighting value. The vertexes are grouped by the object where they belong.

The motion of the avatar is done through the animation of the skeleton, generated by motion capture. This technique allows a more realistic human character animation. For this work it were recorded "waiting state" behaviours, applying them to the avatar when his standing to achieve a natural and dynamic behaviour from him.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 24: Recorded positions.

5 CONCLUSIONS

In the previous chapters, it was discussed the realism improvements of avatars, improving their naturalness, thus increasing their believability. The field of virtual humans research is vast, many different topics were addressed. Along with the issues related with realism perception, the techniques were also covered, and how to improve the avatar in the intended terms. In this chapter it is given an overview of the improvements proposed, ending with the limitations of the approach.

5.1 CONTRIBUTIONS

It was defined as final objective, the naturalness improvement of avatars as television presenters, and three partial goals were defined to achieve that objective. This section gives a detailed look on those goals that we have defined, and presents the validation stage that justifies success of the work.

APPEARANCE Previous applications of avatars in television contents lack of flexibility concerning avatar appearance. This work has contributed with a software that allows easy loading of models created by professional animators and direct appearance modification of the avatars. This software is a plugin for the commercial modeling tool Maya. Therefore, professional artist can develop avatars with highly realistic appearance and export them with a simple button press in Maya interface. The plugin was developed in C++, using Maya API. It was integrated in Maya interface as an external program, it enables the users to export a modeled avatar to a file with information about the weight influences on the vertexes caused by deformations of the avatar mesh. This development is also related with the realism of the virtual world where the avatar will be inserted, it is not natural/realistic the in a world with several characters, all of them possesses the same look. It must be possible to choose among different avatar appearances. This software allows the user to pick-up an avatar according with his appearance preferences. This user-defined virtual characters increase the sense of presence, since the user could identify himself more easily if he has the chance to choose the avatar that represents him in a television program.

reduction of setup timing

The integration of the software created in a system that receives inputs from common devices like microphones, keyboards and mice, allows television producers to reduce the time needed for setup the application. This makes easier the use of the system in live or semi-live daily television programs.

behaviour Regarding a natural behaviour of the avatar, when he stands on the waiting position, it was used a motion capture system to record some poses, after some research in human standing position rules. Television demands dynamics and attractiveness, the avatar could not be still while not receive any commands. The recorded movements applied to the avatar improved his naturalness, since now it does not stand static, it performs these pre-defined motions. The avatar will stand side-by-side with a human presenter, therefore it is a situation where the anatomical and behavioral differences between both will be more prominent. This work improve the naturalness needed for this kind of human-avatar interaction situations.

validation The success of this work is measured through its integration in the final PUPPET project system and the agreement of television producers about the fulfilling of the initial objectives and its suitability for creating on-live television programs. So, the validation of the work has been done by means of a session where the television producer company has been able to test the system. Concretely, in this session actors, producers and directives have been present. Actors have tested the usability of the system and producers and directives have checked the suitability of the system for on-live programs. The tests that have been done are:

- Loading of different characters. There has been tested the loading in the system of a virtual character created with Poser program and imported with Maya. Then, the developed plugin has been used for exporting the model to the PUPPET platform. On the other hand, a model created by professional animators using Maya has been loaded in the PUPPET system. Both of them have been loaded successfully and some prerecorded animations have been reproduced in both of them successfully.
- Usability. Professional actors have tested the use of the PUPPET system with the presented work integrated in. They agree that the system is very easy to use and intuitive. Moreover, they do not need extra time for preparing the avatar before the on-live show.

- Waiting behaviour through a rule-based system. The waiting behaviour has been tested successfully. The avatar really creates the illusion of being a human-being interacting with users.

5.2 LIMITATIONS AND FUTURE WORK

The major limitation on this work is the subjective nature on the evaluation of a virtual character naturalness. It was argued in the introduction that it was not the purpose of this work to imitate humans at the most small detail, instead is to provide the illusion of life. To make possible to recognize human characteristics to anthropomorphic models. The use of motion capture is a step on that direction, although the motions care of more precision. The avatar moves as a human does, but it still raises some doubts on the naturalness of the same. The use of a more precise motion capture system could be a solution.

BIBLIOGRAPHY

[1] Norman E. Alessi and Milton P. Huang. Evolution of the virtual human: From term to potential application in psychiatry. *CyberPsychology & Behavior*, 3(3):321–326, 2000. (Cited on page 13.)

[2] Elisabeth André, Thomas Rist, and Jochen Müller. Webpersona: A life-like presentation agent for educational applications on the world-wide web. *In Proc. of the workshop "Intelligent Educational Systems on the World Wide Web", 8th World Conference of the AIED Society, Kobe, Japan*, pages 11–22, August 1997. (Cited on page 9.)

[3] Animazoo. Igs-190. http://www.animazoo.com. accessed october 2009. (Cited on page 37.)

[4] Norman I. Badler, Cary B. Phillips, and Bonnie Lynn Webber. *Simulating humans: computer graphics animation and control*. Oxford University Press, Inc., New York, NY, USA, 1993. (Cited on pages 12, 18, 20, and 23.)

[5] D. Ballin, M. Lawson, M. A. Lumkin, and J. Osborne. Personal virtual humans — inhabiting the talkzone and beyond. *BT Technology Journal*, 20(1):115–129, 2002. (Cited on page 3.)

[6] Joseph Bates. The role of emotion in believable agents. *Commun. ACM*, 37(7):122–125, 1994. (Cited on pages 1 and 5.)

[7] David Bennett. The faces of "the polar express". page 6, 2005. (Cited on page 27.)

[8] Timothy Bickmore and Justine Cassell. Relational agents: a model and implementation of building user trust. pages 396–403, 2001. (Cited on page 1.)

[9] Frank G. Borg. Random walk and balancing. 2004. (Cited on page 19.)

[10] Samuel R. Buss. Introduction to inverse kinematics with jacobian transpose, pseudoinverse and damped least squares methods. *IEEE Journal of Robotics and Automation*, 2004. (Cited on page 23.)

[11] Meeran Byun and Norman I. Badler. Facemote: qualitative parametric modifiers for facial animations. *SCA '02: Proceedings of the 2002 ACM SIGGRAPH/Eurographics symposium on Computer animation*, pages 65–71, 2002. (Cited on page 24.)

[12] T.W. Calvert, J. Chapman, and A. Patla. Aspects of the kinematic simulation of human movement. *Computer Graphics and Applications, IEEE*, 2:41–50, November 1982. (Cited on page 25.)

[13] Yung Kyun Choi, Gordon E. Miracle, and Frank Biocca. The effects of anthropomorphic agents on advertising effectiveness and the mediating role of presence. *Journal of Interactive Advertising*, 2, 2001. (Cited on page 16.)

[14] J. J. Collins and C. J. De Luca. Random walking during quiet standing. *Phys. Rev. Lett.*, 73(5)(5):764–767, Aug 1994. (Cited on page 19.)

[15] QuantumWorks Corporation. Ratz, 1994. (Cited on page 10.)

[16] Nicolas Courty. Kinematic methods for computer animation. http://perception.inrialpes.fr/. accessed september 2006., 2007. (Cited on pages xi and 23.)

[17] Donald Crafton. *Before Mickey: the animated film, 1898-1928*. The MIT Press, 1982. (Cited on page 25.)

[18] Z. Deng, J. P. Lewis, and U. Neumann. Synthesizing speech animation by learning compact speech co-articulation models. pages 19–25, 2005. (Cited on page 11.)

[19] Dak Dillon. Msnbc unveils new 3d studio, November 2008. (Cited on page 5.)

[20] Juan C. Dürsteler. Ananova. http://www.infovis.net/printmag.php?num=181&lang=2. accessed september 2009, 2006. (Cited on page 12.)

[21] Claudia Esteves, Gustavo Arechavaleta, Julien Pettré, and Jean-Paul Laumond. Animation planning for virtual characters cooperation. *ACM Trans. Graph.*, 25(2):319–339, 2006. (Cited on page 18.)

[22] Tony Ezzat, Gadi Geiger, and Tomaso Poggio. Trainable videorealistic speech animation. pages 388–398, 2002. (Cited on page 11.)

[23] Petros Faloutsos, Michiel van de Panne, and Demetri Terzopoulos. The virtual stuntman: Dynamic characters with a repertoire of autonomous motor skills. *Computers and Graphics*, 25(6):933–953, 2001. (Cited on pages xi, 19, and 24.)

[24] James A. Ferwerda. Three varieties of realism in computer graphics. pages 290–297, 2003. (Cited on page 3.)

[25] David Fincher. The curious case of benjamin button. Paramount Pictures, Warner Bros., 2008. Adapted from a story written by F. Scott Fitzgerald. (Cited on page 27.)

[26] R. C Fitzpatrick, J. L. Taylor, and D. I. McCloskey. Ankle stiffness of standing humans in response to imperceptible perturbation: reflex and task-dependent components. *Journal of Physiology*, 454:533–547, 1992. (Cited on page 19.)

[27] Sandra M. S. F. Freitas, Silvana A. Wieczorek, Paulo H. Marchetti, and Marcos Duarte. Age-related changes in human postural control of prolonged standing. *Gait & Posture*, 22:322–330, 2005. (Cited on pages 19 and 20.)

[28] John Funge, Xiaoyuan Tu, and Demetri Terzopoulos. Cognitive modeling: Knowledge, reasoning and planning for intelligent characters. *Proceedings of SIGGRAPH '99*, pages 29–38, 1999. (Cited on pages xi, 3, and 4.)

[29] Margaret S. Geroch. Motion capture for the rest of us. *J. Comput. Small Coll.*, 19(3):157–164, 2004. (Cited on page 28.)

[30] Thanh Giang, Robert Mooney, Christopher Peters, and Carol O'Sullivan. Real-time character animation techniques. 2000. (Cited on page 24.)

[31] Carol M Ginsberg and Delle Maxwell. "graphical marionette". pages 303–310, 1986. (Cited on page 26.)

[32] David Gould. *Complete Maya Programming: An Extensive Guide to MEL and C++ API*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2003. (Cited on pages xi, 34, and 35.)

[33] Christian Griesbeck. Introduction to labanotation, 1996. (Cited on page 26.)

[34] M. Gutierrez, F. Vexo, and D. Thalmann. Reflex movements for a virtual human: a biology inspired approach. 3025:525 – 534, 2004. Virtual Reality Lab. (VRlab), Swiss Fed. Inst. of Technol. (EPFL), Lausanne, Switzerland. (Cited on page 17.)

[35] Margaret A. Hagen. *Varieties of realism: geometries of representational art*. Cambridge University Press, 1986. (Cited on page 3.)

[36] Denise Hippler and José Mario De Martino. Virtual presenter. *IHC 2008 VIII Simposio Sobre Fatores Humanos em Sistemas Computacionais*, pages 296–299, 2008. (Cited on page 14.)

[37] William I. Homer and Aaron Scharf. Concerning muybridge, marey, and seurat. *The Burlington Magazine,*, 104:391–393, September 1962. (Cited on page 25.)

[38] Lucio Ieronutti and Luca Chittaro. A virtual human architecture that integrates kinematic, physical and behavioral aspects to control h-anim characters. *Web3D '05: Proceedings of the tenth international conference on 3D Web technology*, pages 75–83, 2005. (Cited on page 4.)

[39] Mohd Izani, Ahmad Rafi Eshaq, Aishah Razak, and Norzaiha Norhan. A study of practical approach of using motion capture and keyframe animation techniques. pages 52–55, 2004. (Cited on page 22.)

[40] Ollie Johnston and Frank Thomas. *Disney Animation: The Illusion of Life*. Abbeville Press, 1981. (Cited on page 22.)

[41] Patrick Kenny, Thomas Parsons, Jonathan Gratch, and Albert Rizzo. Virtual humans for assisted health care. pages 1–4, 2008. (Cited on page 2.)

[42] Barry King. The burden of max headroom. *John Logie Baird Centre and Oxford University Press*, 30(1-2):122–139, 1989. (Cited on page 9.)

[43] Midori Kitagawa and Brian Windsor. *MoCap for Artists: Workflow and Techniques for Motion Capture*. 2008. (Cited on page 28.)

[44] Herman van der Kooij, Bart Koopman, Ron Jacobs, Thomas Mergner, and Henk Grootenboer. Quantification of sensory information in human balance control. *20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 5:2393–2396, 1998. (Cited on page 19.)

[45] Brigitte Krenn. The neca project: Net environments for embodied emotional conversational agents project note. *Künstliche Intelligenz Themenheft Embodied Conversational Agents*, 17(4), 2003. (Cited on page 9.)

[46] John Lasseter. Principles of traditional animation applied to 3d computer animation. *SIGGRAPH '87: Proceedings of the 14th annual conference on Computer graphics and interactive techniques*, pages 35–44, 1987. (Cited on page 22.)

[47] Mike Lazzo. Space ghost coast to coast, 1994. (Cited on page 10.)

[48] Antoine Leclercq, S. Akkouche, and E. Galin. Mixing triangle meshes and implicit surfaces in character animation. *Proceedings of the Eurographic workshop on Computer animation and simulation*, pages 37–47, 2001. (Cited on page 21.)

[49] Matthew Liverman. *The animator's motion capture guide: organizing, managing, and editing*. Charles River Media, 2004. (Cited on page 25.)

[50] Matthew Lombard and Theresa Ditton. At the heart of it all: The concept of presence. *Journal of Computer Mediated Communication*, 3(2), 1997. (Cited on page 16.)

[51] Nadia Magnenat Thalmann and Daniel Thalmann. Computer animation. *ACM Comput. Surv.*, 28(1):161–163, 1996. (Cited on page 22.)

[52] Nadia Magnenat Thalmann and Daniel Thalmann. *Handbook of Virtual Humans*. John Wiley & Sons, 2004. (Cited on pages 12, 17, and 18.)

[53] Laurie McCulloch. Digital broadcasters. *Computer Graphics World*, 25, October 2002. (Cited on pages 9 and 11.)

[54] Metamotion, September 2009. (Cited on page 29.)

[55] Masahiro Mori. The uncanny valley. *Energy*, 7(4):pp. 33–35, 1970. (Cited on page 2.)

[56] Nick Mourkoussis, Katerina Mania, Tom Troscianko, and Rycharde Hawkes. Assessing functional realism. page 105, 2005. (Cited on page 3.)

[57] Tsukasa Noma and Norman I. Badler. A virtual human presenter. *Proceedings of IJCAI 97 Workshop on Animated Interface Agents, Nagoya, Japan*, pages 45–51, August 25 1997. (Cited on page 14.)

[58] Mathew Orman. High definition motion capture. http://tyrell-innovations-usa.com. accessed september 2009, 2009. (Cited on page 29.)

[59] Amalia Ortiz. *Avatares para la interacción emocional*. PhD thesis, Universidad del País Vasco, Donostia - San Sebastián, Spain, 2008. (Cited on pages xi and 2.)

[60] Andrew Ortony, Gerald L. Clore, and Allan Collins. *The cognitive structure of emotions*. Cambridge University Press, 1988. (Cited on page 18.)

[61] David Oyarzun, Maider Lehr, Amalia Ortiz, Maria del Puy Carretero, Alejandro Ugarte, Karmelo Vivanco, and Alejandro García-Alonso. Using virtual characters as tv presenters. *Lecture Notes in Computer Science*, 4469/2007:225–236, 2007. (Cited on pages xi, 9, 15, and 33.)

[62] Ken P. The man behind lord of the rings' gollum. http://movies.ign.com/articles/384/384094p1.html. accessed september 2009., 2003. (Cited on pages xi and 28.)

[63] Richard P. Paul. *Robot Manipulators: Mathematics, Programming, and Control*. MIT Press, Cambridge, MA, USA, 1982. (Cited on page 23.)

[64] David-Paul Pertaub, Mel Slater, and Chris Barker. An experiment on public speaking anxiety in response to three different types of virtual audience. *Presence: Teleoperators and Virtual Environments*, 11(1):68–78, 2002. (Cited on page 13.)

[65] Poser. http://www.smithmicro.com. accessed october 2009. (Cited on page 38.)

[66] Helmut Prendinger and Mitsuru Ishizuka. *Life-Like Characters: Tools, Affective Functions, and Applications (Cognitive Technologies)*. SpringerVerlag, 2004. (Cited on page 1.)

[67] Barbara Robertson. Mike, the talking head. *Computer Graphics World 11*, 7:15–17, July 1988. (Cited on page 10.)

[68] Meteo Sam. http://www.activamultimedia.com. accessed september 2009., 2009. (Cited on page 12.)

[69] Peiren Shao, Weimin Liao, and Zhigeng Pan. Adopting virtual characters in virtual systems from the perspective of communication studies. *Transactions on Edutainment II*, 5660/2009:70–89, 2009. (Cited on pages 14 and 16.)

[70] Hyun Joon Shin, Jehee Lee, Sung Yong Shin, and Michael Gleicher. Computer puppetry: An importance-based approach. *ACM Trans. Graph.*, 20(2):67–94, 2001. (Cited on pages xi and 10.)

[71] J. W Smith. The act of standing. *Acta Orthopaedica*, 23:2:159 – 168, 1953. (Cited on pages xi and 19.)

[72] David J. Sturman. A brief history of motion capture for computer character animation. *SIGGRAPH 94, Character Motion Systems*, 1994. (Cited on pages 25, 26, and 27.)

[73] David J. Sturman. Computer puppetry. *Computer Graphics and Applications, IEEE*, 18:38–45, 1998. (Cited on page 10.)

[74] D Thalmann. Virtual humans in virtual environments: A new view of multimedia applications. *Intern. Workshop on Synthetic - Natural Hybrid Coding and Three Dimensional Imaging*, pages 3–7, 1997. (Cited on page 1.)

[75] Kristinn R Thórisson and Justine Cassell. Why put an agent in a body: The importance of communicative feedback in human-humanoid dialogue. *Proc. Conf. Lifelike Computer Characters*, pages 44–45, October 1996. (Cited on page 13.)

[76] Graham Walters. The story of waldo c. graphic. *3D Character Animation by Computer, ACM Siggraph 89 Course Notes, ACM Press, New York*, pages 65–79, 1989. (Cited on page 9.)

[77] David A. Winter, Aftab E. Patla, Milad Ishac, and William H. Gage. Motor mechanisms of balance during quiet standing. *Journal of Electromyography and Kinesiology*, 13(1):49 – 56, 2003. (Cited on page 19.)

A

SYSTEM ARCHITECTURE

This document describes the real-time animation system where this work is integrated. The system is constituted by several modules, and the developed plugin is an input of the animation engine, the system core. The different components of the system can be grouped in three different categories:

- Interface; is constituted by input devices and commercial tools such as microphone and motion capture system, that generate outputs in a format readable by the animation engine. This module includes audio analysis in order to covert audio input in streams of phonemes, later used in the facial animation engine of the animation module. The developed plugin is included in this module.
- Animation engine; the animation is done in real time, therefore must be efficient. The synchronization module synchronizes corporal a facial animation engines in order to obtain coordination between the speech and the facial movements.The low level animation engine possesses fast and efficient algorithms, that make use of graphic libraries that support the real time calculation of the animation of geometric models with large computational weight.
- Augmented reality; the avatar will be inserted in a real environment and interact with humans. In this module, it must be created a sense of avatar presence, regardless of changes or camera zoom, so that the ongoing television program is not negatively affected by the presence of the avatar.

Figure 25 shows the architecture of the system described above.

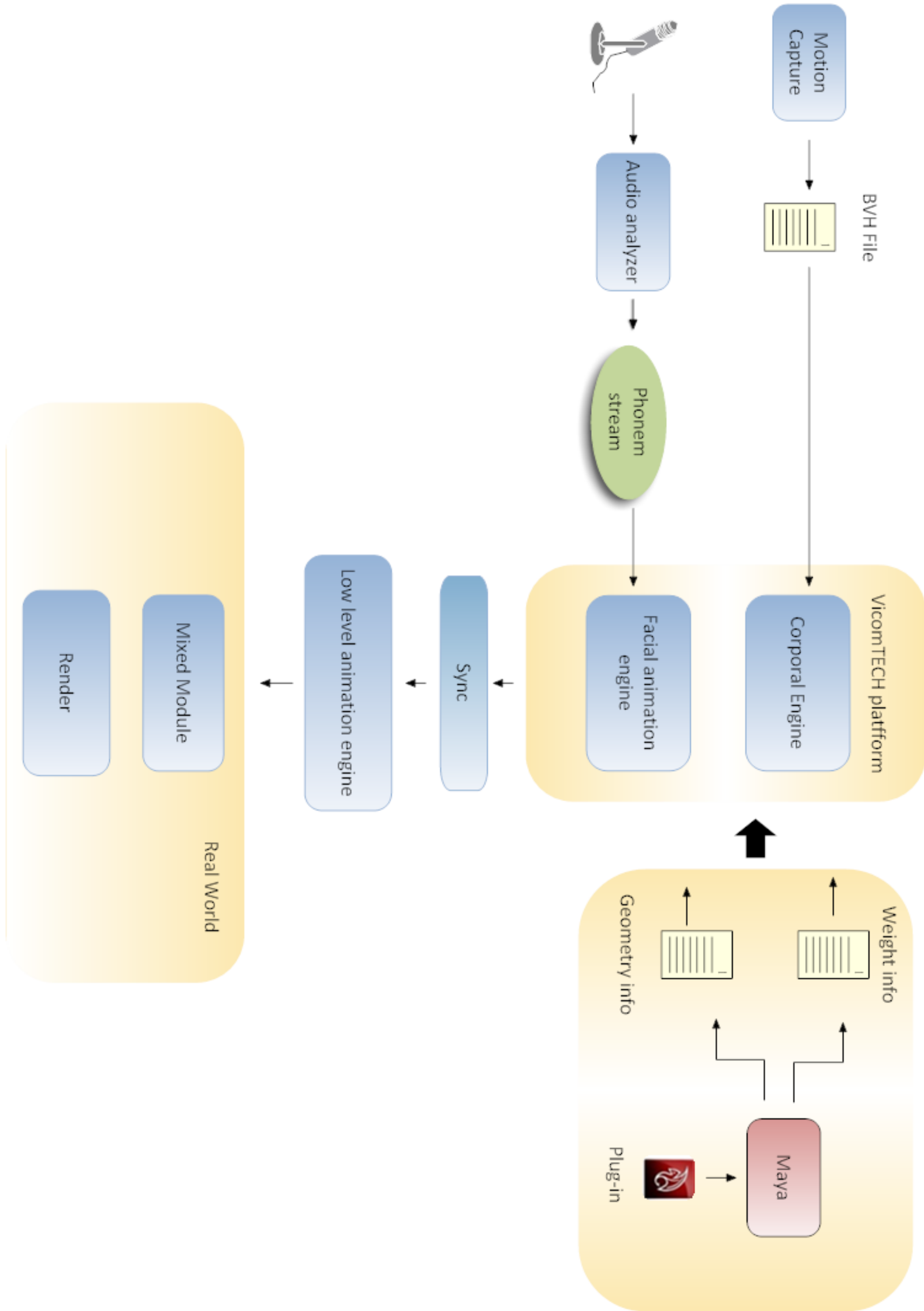


Figure 25: System architecture.