UNIVERSITÄT KOBLENZ · LANDAU
Fachbereich 4: Informatik

LMR Laboratory for Mixed Realities
GFKI Gesellschaft zur Förderung künstlerischer Informatik e.V.
Laboratory for Mixed Realities

Adding Emotion to Expressionless Animations of Virtual Characters

Diplomarbeit

zur Erlangung des Grades eines Diplom-Informatikers
im Studiengang Computervisualistik

vorgelegt von

Taek-Bong Kim

Erstgutachter:	Prof. Dr. Stefan Müller (Institut für Computervisualistik, AG Computergraphik)
Zweitgutachter:	Dipl.-Inform.(FH) Stefan Zingel

Koblenz, im Mai 2007

Erklärung

Ich versichere, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

	Ja	Nein
Mit der Einstellung der Arbeit in die Bibliothek bin ich einverstanden.	☐	☐
Der Veröffentlichung dieser Arbeit im Internet stimme ich zu.	☐	☐

....

(Ort, Datum) (Unterschrift)

Kurzfassung

Zahlreiche Studien belegen, dass menschliche Bewegungen Informationen über den Akteur in sich bergen. Beobachter sind daher in der Lage, Dinge wie Persönlichkeit, Geschlecht und Gefühlslage allein aus Bewegungen von Menschen zu erkennen. Um dem Ziel nach glaubwürdigen und realistischen virtuellen Charakteren näher zu kommen, verbesserte sich in den letzten Jahren vorwiegend das Aussehen der Charaktere. Dank moderner Techniken und einer rapiden Entwicklung der Computer Hardware können heute visuell extrem realistische Charaktere in virtuellen Echtzeitumgebungen dargestellt werden. Trotz ihrer visuellen Qualität werden sie jedoch in interaktiven Umgebungen häufig als *mechanisch* wahrgenommen. Diese Störung der Illusion, einem lebendigen, Menschen ähnlichem Lebewesen gegenüber zu stehen ist in einem mangelndem menschlichen Verhalten des virtuellen Charakters begründet. Daher können ausdrucksvolle Bewegungen, die einen emotionalen Zustand des Charakters vermitteln, dazu verhelfen dem Menschen ähnlichere und daher glaubwürdigere Charaktere zu realisieren. Im Rahmen dieser Diplomarbeit wird die Umsetzbarkeit eines Systems zur automatischen Generierung emotional expressiver Charakter Animationen untersucht. Übliche Techniken zur Erstellung von Animationen sind sehr aufwendig und zeitintensiv. Um alle möglichen Variationen von Bewegungen in einer interaktiven Umgebung zu erstellen kommen solche Ansätze daher nicht in Frage. Um interaktive Charakter zu ermöglichen, welche in der Lage sind ihre Gefühle zum Ausdruck zu bringen, wird daher diese Problematik im Zuge dieser Diplomarbeit behandelt werden. Einschlägige Literatur aus Forschungsgebieten, welche sich mit Emotionen und Bewegungen befassen werden im Rahmen dieser Arbeit untersucht. Eigenschaften, anhand derer Menschen Emotionen in Bewegungen erkennen, werden technisch in einem Animationssystem umgesetzt, um aus neutralen Animationen emotionale Bewegungen zu generieren. Abschliessend werden die erstellten Ergebnisanimationen in Tests ausgewertet in Bezug auf Erkennbarkeit der Emotionen und Qualität der Ergebnisse.

Abstract

Research has shown that people recognize personality, gender, inner states and many other items of information by simply observing human motion. Therefore the expressive human motion seems to be a valuable non-verbal communication channel. On the quest for more believable characters in virtual three dimensional simulations a great amount of visual realism has been achieved during the last decades. However, while interacting with synthetic characters in real-time simulations, often human users still sense an unnatural stiffness. This disturbance in believability is generally caused by a lack of human behavior simulation. Expressive motions, which convey personality and emotional states can be of great help to create more plausible and life-like characters. This thesis explores the feasibility of an automatic generation of emotionally expressive animations from given neutral character motions. Such research is required since common animation methods, such as manual modeling or motion capturing techniques, are too costly to create all possible variations of motions needed for interactive character behavior. To investigate how emotions influence human motion relevant literature from various research fields has been viewed and certain motion rules and features have been extracted. These movement domains were validated in a motion analysis and implemented in a system in an exemplary manner capable of automating the expression of angry, sad and happy states in a virtual character through its body language. Finally, the results were evaluated in user test.

Acknowledgments

Firstly, I would like to thank my advisor professor Stefan Müller for his support and especially for his teachings at the university, which aroused my interest in computer graphics.

I am also grateful to my supervisor Stefan Zingel at the Laboratory for Mixed Realities for constant advice and answers to countless questions. Further I am indebted generally to my colleagues at the LMR, for being a vivid example that professional work and great fun do not necessarily have to contradict.

The Perception Action Cognition Lab at the University of Glasgow deserves special thanks for freely providing just the right motion capture data for my thesis.

Last but not least, I would like to thank my family and friends for mental support, patience and for excusing my months of absence from a normal social life.

Contents

List of Figures

1 Introduction

This first chapter outlines the thesis research. It is exposed that this study is relevant and needed in current research. Overall aims of the research are formulated and a research methodology intended to achieve those aims is presented. Finally, the thesis structure is outlined and a brief summary of each chapter is given.

1.1 Motivation

During the last decade virtual characters evolved to an essential element in many interactive three dimensional graphics simulations. Virtual characters appear in various educational applications. In warfare or rescue training simulations they instruct or assist the user in completing complex tasks. They are used as conversational representatives, called avatars, and in virtual therapy. In computer games synthetic characters play a major role in dragging the user into the story and creating mood within the simulated scenarios. Virtual characters define a human-computer interface, which benefits from the familiar human appearance and acting, which *invoke peoples' automatic responses to the human form and behavior, and thereby achieve a kind of empathic interaction* [VVS06]. Credibility is a key factor, which all these applications try to convey to their characters. Especially the convincing human appearance of such characters is a problem frequently addressed in research. Thus the look of synthetic characters constantly increased in terms of visual quality and realism. Advances in technology, particularly the increase of performance in computer and in graphics hardware, enabled the use of highly realistic virtual characters in real time simulations.

However, in interactive virtual environments these characters, though visually convincing, tend to loose their credibility. As described in [BB04] the importance of the avatars' behavior realism far outweighs visual fidelity in some applications. Researchers believe that expression contributes to a richer interaction [Pic97]. Thus, interpersonal attitudes, communication of emotions and personality traits are important behavioral abilities of virtual characters. Human behavior is believably simulated in areas where no interaction is required and activities are predefined, as in computer generated movies. This human behavior is mainly displayed through facial expression, gestures, bodily movements and other non-verbal communication signals (see Chapter 2) . Traditionally, this is achieved by skilled animators who manually create expressive motion sequences in an off line process. More recently, motion capturing techniques, which arose from human motion analysis, provide artists with natural looking animations captured from real actors. Both techniques are capable of producing realistic and expressive results which convey important inner properties of the characters, such as personality and emotion.

However, to create animations with these techniques is time consuming and labor intensive. Therefore, in interactive environments these techniques are not feasible, due to the vast number of situations and reactions of characters which could possibly occur. New techniques have to be developed, which reproduce emotional expression in virtual humans. Many researchers identified bodily expression of movements as an accurate means by which individuals show their emotional state (see Chapter 2). Relatively few approaches were presented in computer science, which use motions to express emotions in a character. As quoted above, realistic animation sequences of motion can be created. Therefore techniques that modify neutral animations and add emotional expression would be a great advance towards more expressive and hence more human-like, believable characters.

1.2 Research Aims

This thesis investigated cues in human motion in order to generate emotionally expressive animations from neutral motions. A prototype was implemented and its results were evaluated in user tests in respect to perceptability of emotion in the created animations. The overall aims, which guided the research are listed below.

1. Identification of basic features and rules from which emotions are percepted in human motions
2. Validation of findings in an analysis of real human motions
3. Implementation of a prototype system for generation of emotionally enriched animations
4. Evaluation of the system by user tests

1.3 Chapter Overview

- Chapter 2 introduces basic techniques and terms relevant for character animation. All techniques used in the implementation are described.
- Chapter 3 surveys research of human motion and emotion expression. Also related approaches from computer science are introduced.
- Chapter 4 describes concepts which influenced the resulting implementation. Also an overall system overview is given.
- In Chapter 5 the analysis of real movement and the validation of found domains important for emotions perception is described.
- Chapter 6 explains the concrete implementation of a system for emotionally expressive animation generation.

- Chapter 7 describes the performed evaluation of the system, presents results and discusses the findings.
- Chapter 8 concludes the thesis.

2 Basic Principles and Terms

The center of this thesis are animations of virtual three dimensional characters. Therefore, this chapter gives an introductory overview of techniques used in character animation for real time environments. Basic techniques from character setup to animation gathering are presented. Furthermore, specific subjects like orientation in three dimensional space and interpolation are described, since they are closely related to the introduced techniques. Finally, third party libraries and software used is described.

2.1 Interpolation

If a discrete data set is given, interpolation is a method of constructing new data. Often data has been gathered by sampling of experiments. Finding a function which approximately fits these points is called curve fitting. Interpolation is a special case of curve fitting, in which the function goes exactly through all sample points. In character animation interpolation is used to calculate intermediate bone positions and translations. For interpolation between two 3d positions linear interpolation is used in most cases. Here, a point P between two key frames K_0 and K_1 at a time t is determined by:

$$P = K_0.value + ((t - K_0.time)/(K_1.time - K_0.time)) * (K_2.value - K_1.value) \quad (1)$$

This kind of interpolation can be easily calculated and is therefore used by many real time animation systems. But the results of this technique can be very jerky. This is caused by the connecting curve, which goes through all sample points. At transition points the trajectory is not smooth. Mathematically this is described as a $G0$ continuity of the curve. Smoother curves can be achieved if $G1$ continuity is demanded. This results in tangents at each point for both segments (before and after), which have the same direction but not necessarily are of the same magnitude. Consequently, for

$C1$ continuity the tangents have to be in the same direction and poses the same magnitude.

2.2 Orientations in three dimensional space

Orientation is an important attribute of objects in three dimensional space. Especially in character animations a clear and convenient orientation representation of objects is necessary. Different representations for orientation in space exist. The following paragraphs describe the most common methods used in three dimensional space, their advantages and drawbacks.

2.3 Euler Angles

Euler Angles are triplets which are used to describe rotations or relative orientations of rigid bodies in an three-dimensional euclidean space. The main concept of Euler Angles is splitting a complete rotation into three simpler rotations around the axes of the system. Each value of an Euler Angle describes the amount of rotation around a specific axis. Unfortunately there are different definitions of Euler Angles concerning the order in which the rotations are applied. Since rotations are not commutative the order is important. Therefore, when working with Euler Angles the order and axes should always be supplied. Interpolating between Euler Angles is computationally expensive, since numeric integration has to be used. Euler Angles suffer from a serious drawback called *Gimbal Lock*. If during rotation two axes get aligned, one degree of freedom is lost, and therefore one reference gets canceled. By the use of constraints and limitations this effect is avoidable. Euler angles are not intuitive, but unlike other representations comprehensible. Therefore it is mainly used by artists.

2.4 Matrix Rotations

An orientation or rotation in space can conveniently be expressed in a rotation matrix. The interpolation however is quite complex. The main problem that occurs is known as matrix drift, which is produced by finite accuracy. A 3x3 matrix is orthonormal if it satisfies the condition $M * M_t = E$. Its nine components represent the three degrees of freedom of the rotation. Because of the arithmetic drift the orthonormal condition can be violated. This may result in non valid rotation matrices with unexpected results. It is avoided by continually checking the matrix for orthonormality and correctness (e.g. *Gram-Schmidt* algorithm). Despite of being computationally expensive these steps can still lead to imprecise rotations.

2.5 Using Quaternions for Rotation

Another rotation representation are Quaternions. A quaternion consists of two components, a vector defining the rotational axis and a scalar value. For rotation representation, exclusively unit quaternions are used. This guarantees, that multiplication of two quaternions results in another unit quaternion, thus forming a mathematical group. The advantage of quaternions is the low amount of operations needed for a multiplication, contrary to matrix multiplications. Quaternions are easy to parametrize and therefore well suited for interpolation. Since they describe a path on the surface of a four-dimensional hypersphere, quaternions are not intuitive. The value of a quaternion has no correlation to the resulting rotation, which could be predictable by the user.

2.6 Morphing of character meshes

Morphing (or blending) between character meshes is a basic method used in real-time character animation. Multiple copies of the same character geometry are used, each one representing the same character in a different pose. All objects are required to contain the same number of vertices. The animation is created by interpolating the positions of corresponding vertices in the different objects. The result is a smooth animation with no gaps or visible seams in the character model's geometry. Modern graphics processors have the functionality to calculate simple linear interpolations. Therefore, this method has a high performance. The downside of this technique however is it's memory usage. In order to effectively control the animation a lot of key frames (here, copies of the character object) have to be used. That leads to a large amount of data that has to be stored. In addition to this, all object copies that are needed in an animation sequence will have to be kept in memory for at least as long as it takes to finish the animation sequence. Moreover, characters animated with this method are less flexible respecting later modification of animations. The whole character animation has to be predefined off-line and no on-the-fly generation or modification of animation data is possible.

2.7 Hierarchically articulated Objects

Another approach for animating three dimensional characters is the use of hierarchically articulated objects (as described in [Lan06]). In such a hierarchic character organization, each body part of the articulated character is represented as a separate object. These separate objects are stored in a hierarchy and are joined to each other at their pivot points. Each object also references its child objects - attached objects of a lower hierarchic order. The character is setup and displayed by traversing the hierarchic data structure starting from the root, passing down the transformations of parent objects to it's children. There, the parent transformation matrices and the local matrix are concatenated. This allows a more convenient control over movements of the character, than just having a mesh of vertices. Furthermore the memory usage remains small as the vertex and transformation information for each object contained in the model needs to be stored just once. One of the main drawbacks of this technique, however, is the resulting visual quality. Since all objects in the hierarchic character are separate objects, it is inevitable that gaps appear between these objects while the character is animated. It is possible to hide these gaps by overlapping the objects that make up the model, but still the visual quality remains poor, because of disturbing seams.

2.8 Skeletal Animation

Skeletal animation combines the main advantages of the previously described animation techniques. It provides a simple control over the character as well as visually good results. To achieve this, an additional control layer is added to the character. This layer consists of an endoskeleton, which is a hierarchic structure of joints. A skin, which is the vertex mesh representing the shape of the character is attached to and driven by this control structure. This separation between mesh data and hierarchic position information into two distinct data-structures, makes skeletal animation and mesh skinning superior to the previously mentioned techniques.

A bone is simply a transformation matrix, determining the relative offset of the bone to the respective parent bone. All bones of the articulated object together form the skeleton. Explicitly, only a skeleton is animated, using an algorithm similar to that one used for animating a hierarchically articulated object, which in turn implicitly animates the skin. Memory usage for skeletal animation is small, as all skin vertices have to be stored only once. It requires a significantly lower amount of information to be stored, than with the mesh inbetweening method discussed above.

2.9 Characters skinning

Skinning is the process of attaching a mesh to a skeleton. Two main methods exist, which are widely used: rigid and the soft skinning. In a rigid skinned character each of the skin vertices is associated with a single bone. The resulting animated mesh is seamless. But close to joints deformations can result in extreme stretching or inter penetration. These effects are almost invisible when low-polygon objects that do not have a lot of modeled detail are used. In this case, this technique is preferred for it's simple computation. The final position of a vertex v_{final} can be computed as:

$$v_{final} = M_{bone} * v_0 \qquad (2)$$

with v_{orig} being the original vertex positions and M_{bone} being the derrived transformation matrix of the linked bone. An advanced skinning method with better visual results is soft skinning. In a soft skinned character each vertex can be influenced by multiple bones. This is done via a vertex weighting, which is a mechanism to keep track of what bones affect the vertex and to which amount it is influenced by the different bones. Every linked bone contributes a value between 0 and 1 to the weighting of the vertex, which all sum up to 1.

$$v_{final} = \sum_{i=0}^{n} w_i * M_{bone\_i} * v_{orig} \qquad (3)$$

The final vertex position v_final is the sum of all multiplications of the bone weights w_i with the transformation matrix of the according bone M_bone and the original vertex position v_orig. For v vertices and n bones there have to be made $n*v$ multiplications at worst case. This additional computations are usually accepted for better visual results. Since vertex weighting can be computed in a vertex shader on the GPU it is used in most real time applications.

It should be mentioned, that in real human bodies the skin is deformed by the underlying musculature and not by the bones. Techniques which simulate musculature to deform the skin are still subject of ongoing research. As yet, such a technique is still computationally intensive and therefore not common in real time animation.

2.10 Character animation data

Motions of virtual characters have to be availabe as animation data on order to be used in the virtual environment. Basic principles of animation as well as the storage of such data is explained throughout the next paragraphs. Moreover the most common ways of character animation creation are briefly introduced.

2.10.1 Animation

An animation is a series of still images shown in a rapid successive sequence. Each of the still images is just a little different than the one before it and after it. Our eyes interpret these quickly showing images as a fluid motion. This impression starts with about eight images per second. Empirical tests have shown that movements appear fluent with an increasing number of frames per second. However, above 70 fps there is little improvement discerned by most viewers.
Giving a character the illusion of life by animating it, is what is referred to as character animation. The animations are usually created in an off line process. In three dimensional character animation the character's pose is changed slightly in each frame. This creates the impression of a living and moving character. Such character animation sequences are usually created by an animator in an off line process. When displaying

such complex models, like virtual human characters, a lot of information would have to be stored. In the worst case, the position of each vertex for each frame would have to be stored. With a model of thousands of vertices and an estimate frame rate of 30 frames per second this would easily lead to memory insufficiencies on even the most potent machines. Because of that key frames are used in computer animation.

The term key frame was invented in the Walt Disney Studios. A key frame in a character animation is a frame at a certain time position, in which the character is in a particular occurrence, pose or position. In classical animation, these key frames are drawn by a master animator. The junior animators then create the frames between two subsequent key frames. This process is called inbetweening, or in short, tweening.
Similarly, in computer animation, key frames are stored instead of the whole animation, thus saving a lot of memory. Key frames for an animation can come from various sources. They can be created by an artist, generated ith the aid of a program by algorithms or extracted from motion capture data. What is stored in a key frame depends on the character's topology. If the character is made up by nothing else than a mesh, all vertex positions have to be stored for each key frame. But as mentioned earlier, this technique is only used for very simple objects and animation. In most cases, skeletal animation is used. In this case, each key frame stores bone positions and orientations. In that way, less data has to be stored per key frame. The rendering engine just has to update the bone positions. Due to the skinning the mesh is transformed along with the skeleton. The twenning task, which is done by an junior artist in classical animation, is done by the computer in computer animation. Given two subsequent key frames, time dependent positions between the key frames

have to be calculated. This is done by interpolation techniques which are explained in the next chapter.

2.10.2 Manually created animation

Traditionally, character animations are created by artists in an off-line process. By posing the character and storing this data into key frames for a certain time position, complex animations are created. Such animation data is then loaded into the real time application and triggered at a given time. Depending on the creating animator, such animations can reach natural-looking and visually impressive results. However, the animation is predefined and is usually used unaltered.

2.10.3 Motion Capturing

Motion capturing, or *mocap* as it is often abbreviated, describe techniques of measuring an object's position and orientation in physical space and then storing that information in a form usable by a computer. Such captured objects are usually human actors which perform motions and facial expressions. Motion capturing originally has been developed as an analyzing tool for biomechanics research. But recently, it's importance as a source of motion data for computer animation has increased. It is mainly used for video games and cinema productions, where a high amount of realism is required. A great variety of motion capturing systems exist. Inertial systems, mechanical systems, magnetic systems and optical tracking systems with active or passive markers. All have their pros and cons, nevertheless the best quality is achieved by optical tracking and is therefore the choice of most high level productions, where cost is a minor issue.
The principle of all motion capturing techniques is similar. During a mocap session, an actor is put into a suit with markers at predefined positions. The positions (and orientations) of these markers are then sampled multi-

ple times a second. This information is recorded and stored as animation data. This data can be mapped onto virtual character models or skeletons. In that way, very natural animations can be obtained for virtual characters, since the data results directly from actual human movement.

2.11 Physically based animation

Driving a character mainly by physical laws of biomechanics is called a physically based or procedural animation. No key frames exist, which were predefined in an off-line process. All modifications of the skeleton are calculated at real time, dependent on forces which are determined by a physics engine. That allows the definition of highly dynamic characters capable to perform a great variety of motions and to react on different situations. Since forces and other influences which drive human motions are quite complex to simulate, this approach often leads to unnatural and stiff motions. Thus, the quality of the created animations is inferior to the visual results of animation manually created or motion caputred.

2.12 OGRE

OGRE (Object-Oriented Graphics Rendering Engine) is a scene-oriented, flexible 3D engine written in C++ designed to make it easier and more intuitive for developers to produce applications utilizing hardware-accelerated 3D graphics. The class library abstracts all the details of using the underlying system libraries like Direct3D and OpenGL and provides an interface based on world objects and other intuitive classes. (from the official OGRE homepage)

A collection of important features are enlisted below:

- Free under the LGPL license
- object oriented design with a powerful plug-in architecture
- scene graph based

- multi platform support
- high level shader language support (GLSL, Cg, HLSL, asm)
- hardware supported animation engine of weighted multiple bone skinning
- many exporters for the most 3D modeling tools
- strong community support
- Being a rendering engine most of OGRE's features are graphics related. But also math classes for vectors, matrices, quaternions etc. are provided within OGRE as well as other little helper classes.

2.13 Character Studio

Character Studio is a plug-in for 3D Studio Max©. It provides a set of professional tools for animating three dimensional characters. A great number of professional animators as well as hobbyists utilize Character Studio. As described above a concept of separating the model into a geometry and a skeletal control layer is followed. The biped, a specialization of a skeleton, can be quickly created in the modeling environment. By skinning methods this biped skeleton is linked to the mesh. The skeleton can be animated via free-form animation supported by a powerful constraining system, footstep animation, inverse kinematic solver and so on. Moreover motion capture data can be mapped onto the Biped. This allows a flexible character animation process. Blending and sequencing of animation clips is also possible in Character Studio. Finally, crowds of characters can be generated and animated using a system of delegates and procedural behaviors.

3 Affective Human Motion

The synthesis of affect and its perception, these two subjects can not be examined separately. This chapter surveys which impact an emotion has on the motion of humans. Therefore the creation of emotional signals on the sender side, as well as the perception of emotion on the receiver side have to be studied closely. This chapter introduces research results mainly from the field of social psychology, which is devoted to the survey of emotion. Definitions and Terminology are explained and finally approaches and related work from computer science are introduced.

3.1 Emotions

The term emotion is derived from the Latin emovere 'to set in motion'. The initial meaning of physical motion changed to a figurative meaning of mental movement. In psychoanalysis sometimes the term affect is used. In some references the terms emotion and affect are separated strictly [Mas00]. However throughout this thesis the terms emotion and affect are used interchangeably.

Everyone knows what an emotion is; until asked to give a definition. [FR84]

This citation found in most of the emotion literature clearly describes the problem respecting the term emotion. The search for a general and universally accepted definition of the term emotion remains without a result. Due to the complexity of the term definitions vary depending on author, context or research field. Psychology, philosophy, sociology, biology and other research fields are concerned about emotion. Accordingly emotions are studied through a great variety of means, such as physiological changes, such as heart beat or skin conductivity, through neural activity, or through changes in behavior. However, a possible general definition can be found in [wik]:

Emotion, (...) is an intense mental state that arises automatically in the nervous system rather than through conscious effort, and evokes either a positive or negative psychological response.

A philosophical analogy of the term emotion was given by *Descartes* (1649) comparing emotions to body pains. Pains are signals alerting that something important is happening to your body. In the same way emotions are signals alerting that something needs attention in our 'soul', in our thought [OJ96]. Descartes also emphasizes that emotions are phenomena that are not under the control of our mind, but it is possible to recollect our thought to not let the soul be driven completely by passions (*Descartes*).
Another prominent approach to define the term emotion was done by James and Lange. They independently developed a theory where emotions are described as 'awareness of body responses' to an event. Emotions are generated by body changes such as facial animation or posture. This theory brought forward the important role of the body in connection with the phenomena emotion.
Frequently, the term basic emotion appears in literature. Basic emotions are defined as separate emotions, which differ among each other in terms, such as their appraisal, probable behavioral response, physiology and other characteristics. Ekman, Friesen, and Ellsworth identified basic emotions as anger, disgust, fear, joy, sadness and surprise [EF78]. However, an agreed composition of basic emotions can not be found throughout science.

3.2 Non-verbal Communication

A very abstract definition of communication could be the exchange of information. The most dominant type of communication in human society is speech. By speech people communicate needs, ideas, thoughts and other complex information to their environment. However, a great variety of other types of communication exist among humans, which in related research were found to play a crucial role in social interpersonal behavior among humans. These communication channels are summarized under the term non-verbal communication. Non-verbal describes the fact, that all messages which are exchanged during non-verbal communication do not imply speech or written words. It was considered by Argyle that non verbal behavior takes place whenever one person influences another by means of facial expressions, gestures, bodily contact, body posture, spatial behavior, gaze and pupil dilation, clothes, or non-verbal vocalization [Arg79]. While interacting, humans continually display and interpret the other person's non-verbal signals. Hundreds of expressive movements are used every day during social events of the day and their correct use is an essential part of people's social competence and skills. [Arg79] enumerates the following areas, in which non-verbal communication is utilized:

- Support of speech
- Alternative for speech
- Expression of attitude towards other people
- Rituals and formalities
- Expression of emotion

The next paragraphs describe the channels of non-verbal communication, which are tightly connected to the expression of emotions.

3.3 Facial expressions

Facial expressions are used by many animals but are highly developed among primates, and most of all, in humans. Facial expressions are an important channel of nonverbal communication. They result from motion or position of muscles of the face. Facial expressions convey the emotional state of an individual to observers, and thus reveal part of the feeling side of a person's private life. Therefore facial expressions still are substantial among human communication, since they convey such information more naturally and involuntarily than by e.g. speech. The different aspects of the study of human facial expressions vary from computer simulation and analysis to understanding it's role in a nonverbal communication, and the emotional process.
Frequently, Facial Action Coding System (FACS) is used by researches, which was developed by Paul Ekman and Wallace Friesen in 1977 [EF78]. Using this method, researchers measure facial expressions by identifying the muscular activity underlying the changes in facial appearance.

3.4 Gestures and bodily movements

Gestures are movements, which are mainly performed by the hands but also by other parts of the body and which are used in combination with or instead of speech. Gestures are used to express various things such as emotions, thoughts, actions, personality, etc. All gestures can be described as signs with an agreed meaning. The signs themselves as well as their meanings differ among cultures and social environments. One example for this is a nod. This vertical movement of the head is interpreted in most Western countries as a 'yes' or a confirmation. However, in Bulgaria it is the appointed gesture for 'no'. The most prominent researcher in the field of gesture research is Adam Kendon. He has investigated many aspects of gestures, including their role in communication, conventionalization of gesture, integration of gesture and speech, and the evolution of language. Also some bodily movements can be counted to the members of non-verbal communication, if the performance of the motion transmits signals other than pure motion information. In [Vol03] Volpe calls that category of movements 'non-propositional' movements.

Contrary to gestures, these are instead embodied in the direct and natural emotional expressions of body movement based on fundamental elements such as tempo and force that can be combined in a wide range of movement possibilities. Therefore, non-propositional movements do not rely on specific movements, but build on the quality of movements [Vol03].

The ability to decode basic emotions from motion is developed early in life and is not dependent on social and cultural environment [RB98]. It rather has it's origins in the human evolution.

3.5 Posture and stance

A posture is the position of the body or body parts. Three basic human postures exist [Arg79]: Standing, sitting (crouching, kneeing) and recumbency. Each of them having multiple variations depending on the position of the limbs and the bending of the body. The human posture is an important means to express impersonal attitude. Moreover posture is directly connected to the inner emotional state of people. This is either driven by physiological impact of the inner state on the posture or by symbolically agreed and learned meaning. Therefore, like gestures, posture is strongly influenced by the cultural background and environment. A classification of postures can be found in [Bir70].

3.6 Visual perception of human motion

The following overview was made, because finding and research results from communication psychology, motion analysis and other emotion related fields would have to be used as a basis for the implementation of an emotional motion generation system. Therefore appropriate literature was consulted, which would help to model techniques, algorithms and rules to allow the automatic generation of emotional expression in computer character animations. It was discovered that research relevant for the realization of such a problem was relatively rare. While a lot of research regarding the encoding of emotions in facial expressions and consistent solutions exist in the field of computer science, research on body motions and their connection to emotion often was found to be not profound or inapplicable for a practical implementation The following paragraphs describe relevant findings from motion analysis and psychology and introduce existing approaches from computer science.

In 1973 the Swedish psychologist Gunnar Johansson attempted to determine the information, which is carried through human motion. Special point lights were placed on the joints and other significant positions on the body of actors. While performing actions in a dark environment these actors were filmed by a camera. In that way clips were created, in which just white spots on a black background were seen [Joh73]. Human movement was reduced to the motion of a small set of points with all other characteristics of the actor subtracted away. When these limited motion stimuli were shown to the test subjects, it was still possible for them to recognize the movement which was performed. This result dmonstrated that even highly reduced information such as the position of distinctive spots of the body can contain enough information to create a vivid expression of the movement.

Having such an instrument for motion analysis, research of human movement perception with point light displays was continued by other

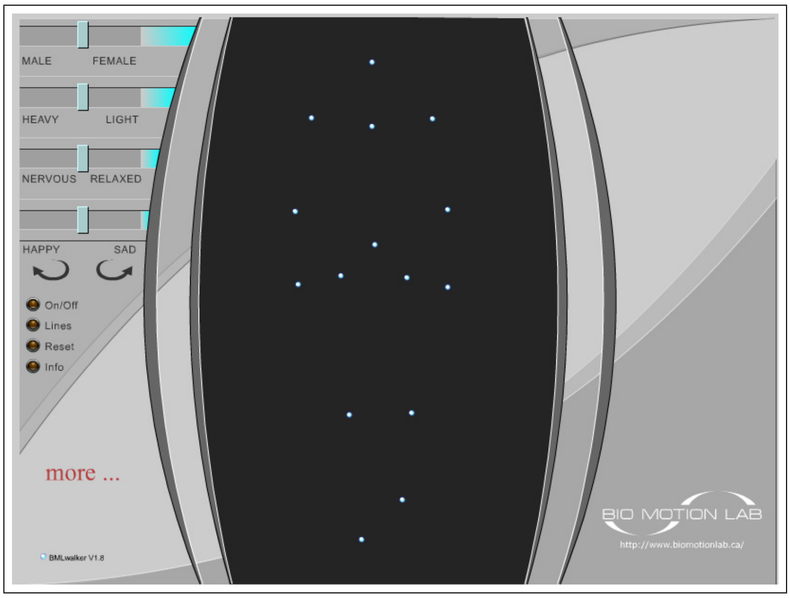

Figure 1: BMLWalker showing a human in point light representation [BML]

researchers. [JCK78] showed that not only the motion itself but also gender of the performing actor could be identified by this motion cues. Ongoing research of e.g. [TB81] demonstrated the recognition of style and personality from such point light displays. These researchers affirmed the assumption that a lot more information is transmitted through human movement than only movement recognition.

The question, whether transmitted information also includes emotion, was studied first in the specialized example of stylized dance by [WD96]. A test was setup to determine to what extent emotional expression recognition was based on this pure motion data and to what amount on other visual cues. Two professional dancers were filmed who tried to convey the emotions fear, anger, grief, joy, surprise and disgust. The subjects were asked to judge which emotion was portraited. Half of the subjects were shown full body movie clips, the other half observed clips on point light displays. 88% of the emotions shown in the full body movie clips and still

63% of the emotions from the point light clips were identified correctly by the subjects. These results, which were significantly above chance, showed that emotions could be expressed through bodily motion.
Recently, a similar test was performed by [Pol04] to explore the impact of emotions to everyday movements. Human movements like knocking on a door or drinking a cup of coffee were recorded in ten basic emotional states (afraid, angry, excited, happy, neutral, relaxed, strong, sad, tired and weak). Then these clips were shown to the test candidates as video clips and as point light representation. Test candidates could actually recognize the emotions with a high probability, where the most certainly identified emotion was anger. It was exhibited that dance, but also normal movements could convey information about an emotional state of the performer.

The concept of body language cues which result in a positive feedback in conversation was named *Immediacy* and *Proxemics* by Mehrabian [AM69]. This concept is based on main cues which include general posture, physical contact, spacial proximity, eye contact and body orientation. For example, it could be detected that frequent head motions resulted in a more positive and friendly impression of an interlocutor while a static head position conveyed aggressiveness. However, these cues can not directly be mapped to concrete emotions. Rather they describe cues for a general positive or negative tendency of feedback.

3.7 Cues for emotion in motion

It has been shown, that human movements seem to carry information of an inner or emotional state of the performing subject. The research also showed that this information can be decoded and therefore identified by other observers to a certain amount. The following paragraphs present literature from respective information which focuses on visible features and cues which act as indicators for emotions in human movements.

The research of DeMeijer was concerned with the attribution of emotions in human motion [DeM89]. In test sessions, various body movements of different actors were videotaped. These movements were exactly classified regarding seven given general dimensions: trunk movement, arm movement, vertical direction, sagittal direction, force, velocity and directness. In an experiment subjects judged these motions in terms of emotional expression. In that way a correspondence between specific movement patterns and particular emotion attributes was examined. Three factors were extracted from the original ratings and interpreted as Rejection-Acceptance, Withdrawal-Approach, and Preparation-Defeatedness as domains in a direct connection to emotions shown in body movements.

[RB98] extensively studied the decoding mechanisms of emotions of children. Six movement features such as 'frequency of upward arm movement, the duration of time arms were kept close to the body, the amount of muscle tension, the duration of time an individual leaned forward, the number of directional changes in face and torso, and the number of tempo changes an individual made in a given action sequence'[RB98] were used to test the ability of recognizing emotions in expressive body movement of the children. It could be shown that these features are strongly involved in the recognition of four basic emotions fear, anger, happiness and grief. Moreover these features were already recognized in early stage of life. This led to the conclusion that these special cues are not dependent on education or social background and have its origin in evolutionary history of men.

The role of velocity as a major factor for emotion discrimination and perception was examined in [Pat01]. In experiments the speeds of affective human arm movements were manipulated. This was achieved by simply changing the duration of movements. Test candidates perceived a modulation of affect by velocity changes. For example, sad movements which were displayed faster were categorized as angry. Such results emphasized the important role of velocity in emotion discrimination especially in modulating the intensity of the perceived affect. 'However finding a direct connection between emotions and stimuli is elusive {...}'[Pat01] and therefore needs further research. Moreover Paterson stated that 'it is clear that there are other properties of the movements that were not controlled by velocity, but that play a role in the discrimination of affect'[Pat01].

Similar features for expressive movements were found in another theory from a different field of research. The Labanotation is a sophisticated language for the description of movements which was originally developed for dance. Nevertheless this theory is also applicable to all other human movements. Developed by the dance teacher and researcher of human movement Rudolph Laban, this technique is used widely in dance and theater as well as sport analysis until now. In the Laban Movement Analysis (LMA) the expression of movements is examined. The main concept which is used to describe expression and therefore the emotional content of a motion is the Effort and Shape model. This model involves the 'dynamic' qualities of the movement and the inner attitude towards using energy. Space, Weight, Time and Flow are the four motion factors Effort comprises of. In LMA these factors are seen as the basic qualities, meaning that they are the minimum units required to describe an observed movement. Each factor is a continuum between two extremes: *Space* between indirect/direct, *Weight* between light/strong, time between sustained/sudden, and *Flow* between free/bound.

3.8 Related work from computer science

Expressing emotions has been studied extensively during the last decades. Unfortunately, most of this research was focused on facial expressions. Many approaches and techniques, especially in connection with the facial coding system, have been introduced, and have led to promising results [PK91][DK96]. On the other hand, the research of expressing emotional states through motion or more generally through body language is sparse but existent. Some researchers in computer science surveyed this problem and suggested methods for adding expressiveness to given neutral motion. The relevant research in this field is presented in this chapter.

Ken Perlin was one of the first researchers who laid effort into the creation of 'remarkably lifelike, responsively animated characters in real time' [Per95]. The created characters, which were called puppets, could be controlled through a set of buttons each representing a primitive action or a discrete emotional state. These primitive actions could be blended into one animation. The contribution of each primitive action into the end result was determined by the use of weights, scalars ranging up from 0.0 to 1.0. To add more naturalness and expressiveness to the animation results, they could additionally be tuned by adding certain biases to the joint motions. The values for these biases were retrieved from a pseudo random noise function which then modified joint angles and amplitudes over time. These time dependent slight modifications were called 'the texture' of motion. This texture was the means for varying the expression of the character. However, controlling the character by applying scalar joints of sine and cosine functions seems to be non-intuitive. Moreover, setting transition times and action weights also requires a certain artistry and skill to achieve convincing results. Naively applied parameters result in unrealistic animations.

Unuma et. al. used Fourier analysis techniques to describe and manipulate human periodic motion in real time [MUT95]. The animation data was interpolated and extrapolated to capture a wide variety of expressions. For instance, various degrees of tiredness could be created by interpolating between a 'normal' and a 'tired' walk. Differences between the coefficients of a Fourier function model for a neutral locomotion and those for emotion-driven locomotion could be quantified. With these results, emotion specific fourier characteristic functions could be generated. These functions were then applied individually or in combination to other neutral locomotions to animation data to produce different variations and expressions. However, the process of generating such Fourier functional models and characteristic functions could be very complex and lengthy.

Another approach was followed by Badler and colleagues at the University of Pennsylvania. They developed EMOTE (Expressive Motion Engine), a 3D character animation system that allows specification of *Effort* and *Shape* parameters to modify independently defined arm and torso movements'[DCB00]. The concepts of *Effort* and *Shap*e were inspired by the work of the researcher and choreographer Rudolf Laban. He developed the Theory of Effort to qualitatively describe human movement. This was done in terms of four main dimensions: *Space*, *Time*, *Weight*, and *Flow* as descibed in section3.7. These effort parameters were determined by the use of the *Laban Movement Analysis* method on emotional gestures. These results were then applied to arm and torso animations to obtain results with enhanced expressiveness.

[KAC96] proposed a method to produce emotional animation from neutral, expressionless motion. His technique was composed of two steps. Firstly, the so called 'emotional transforms' had to be identified. Therefore tracked human actors performed motions in different emotional states, such as angry, sad and neutral. This animations were then analyzed using signal processing techniques. The results showed that mainly two parameters varied noticeably between the different emotional movements: speed (the timing of the animation) and the spatial amplitude (range of the animation). In respect to these parameters emotional transform functions were calculated. These functions were applied to joints positions of the object depending on time. The resulting animation was able to express an emotion.

All introduced techniques are capable of generating more or less expressive movements. However, some techniques require modification of low level parameters for the control of expressions which is not very intuitive. Others are tailored for specific motions like locomotion (Fourier) or limited to parts of the body, like [DCB00] to torso and arm movements. Still others have to be modeled in an off-line process for each different type of expression. Such a way that some artistry or expertise is demanded in order to generate natural, expressive movements.

4 Concept and Design

The previous chapter showed that psychology research and theories from art and humanities seem to identify similar parameters for expressive human movement. Especially attributes like force, dynamics and speed are terms that are mentioned in many theories and seem to influence emotional expression to a certain amount. This chapter discusses the design considerations for an emotional animation system which modifies character movements by such basic attributes. The aim was to implement a prototype for the synthetic creation of emotionally expressive animations from neutral input movements. The system was supposed to be capable of simulating the basic emotions sad, angry and happy. To investigate the effect of these animations and to allow a validation of the system the results were shown to human observers in a final experiment. The evaluation is introduced in Chapter 5.

4.1 Considerations

The implemented system was designed to modify existing neutral human character animations in a way that results conveyed emotional expression. Below are the core design considerations and requirements for such a system.

4.2 From Neutral to Affective

The followed approach of enhancing neutral animation with emotional expressiveness is based on the following concept. The diversity of human motion is enormous. Motions have various purposes and functionalities. Therefore a survey to determine what impact emotions would have on every single existing action would not be possible. To diminish the mass of motions, they can be categorized into groups of actions. Among others this has been done for human locomotion, dance and other subgroups of actions (see Chapter 3). Still the number of subdivisions that would be needed to reasonably cover all possible motions would be large. Hence most of the research, which is concerned with the generation of expressive motions focused on a special motion or limited group of movements. Therefore features in human motion, which are generally valid for all motions effected by a specific emotion would be extremely helpful to find. Such universal features could be applied to 'neutral' animations to change them into more expressive movements. The type of motion, which the animation represents would not have to be categorized first. Based on the results of research concerned with the features of affective motion (introduced in Chapter 3) three abstract domains were defined:

1. Speed

2. Range

3. Posture

Certainly no absolute values could be used, since they would have been dependent on many factors as the performed motion, the performing actor, situation and so on. Therefore relative values had to be found, which describe the relative change of a feature in respect to a neutral state of this motion. Real human movements samples performed in angry, sad and happy states were analyzed in respect to these movement domains. The results were used to estimate reasonable values, which were then used in the implemented system. The analysis described in the next chapter was carried out in order to measure these differences in the defined domains and to evaluate their suitability for the animation system.

4.3 Character Topology

In order to modify characters and their animation, well structured virtual characters were needed, which allow convenient, flexible and full control of the model. Hierarchically structured characters with skeletal animations satisfy these requirements. Such organized models can easily be modified and controlled by their joints. No special attention has to be paid to the surrounding mesh, which tends to become quite complex for detailed characters. This is automated by the use of skinning technologies. Another factor for the employment of skeletal animation was its major role in state of the art real time applications, such as games. In these areas skeletal animation is a de-facto standard real time animation technique. Therefore, most current virtual characters are structured in that way. The final mentioned advantage are the possible use of pre-generated animation data as well as motion capture data. Such data can easily be mapped onto such other skeletons. In that way flexible access to lots of motion libraries and collections is granted.

4.4 Arbitrary character Support

A huge variety of virtual characters exist, all differing in size, structure and other properties. An underlying skeleton is dependent on the logical structure of the character. The length of bones can vary as well as their offset positions and scale. More importantly the topology of the skeleton differs among characters. Some have additional body parts like tails or wings which all need extra bones for control. Other models need more precise control of the mesh and have more joints than others. The application was designed to preferably support all human characters or at least the majority of them. Therefore a general skeleton structure has been defined. This skeleton was supposed to contain just the basic, significant bones, which a human being would have. On the other hand, all bones required for ma-

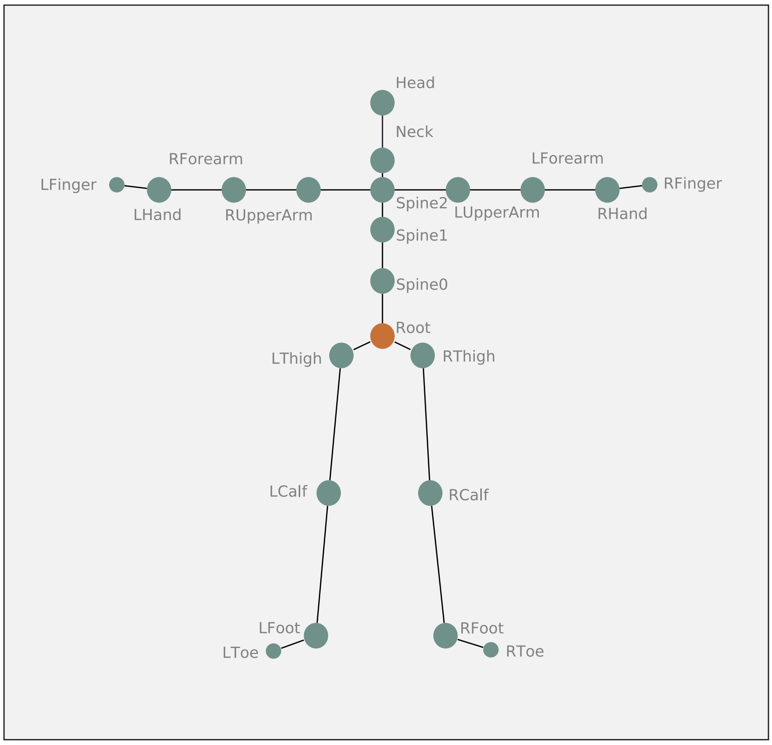

Figure 2: Reference skeleton structure

nipulation had to be provided. The resulting minimal skeleton is depicted in Figure 2. In a configuration process joints of the specific character model would have to be mapped onto the standard skeleton. Arbitrary characters, whose skeleton satisfy the basic requirements of the minimal skeleton could be used. Since all algorithms would be designed to work on such a standard skeleton, the implementation would be independent from actual character objects.

4.5 Modifiers

Each domain, which was examined during the analysis was implemented into a separate modifier class. A modifier is an abstract class, which simply works on existent data and modifies it by certain means. By separating the different aspects in separate modifiers, it was possible to control their execution. Modifiers can work on top of each others results. For evaluation and testing of certain parts, modifiers could be applied or not. This resulted in a clearer structure of the code, which could be debugged and maintained easier. A range modifier analyzed the limb extensions and apex angles of joints. On the basis of the result data of the analysis, it would alter the extreme positions of joints, which would be taken during the animation. The speed modifier would adjust the velocities of an animation. It had to change the overall speed of the movements in an animation evenly. Moreover certain acceleration behaviors typical for a specific emotion had to be simulated. Therefore, non linear velocity changes of the animation were required as well as body part specific modifications. Posture was found to be a strong indicator for an inner state of a human. Therefore a posture modifier would create stereotypical body poses. These body poses would the influence the result animation by blending.

4.6 OGRE3D integration

In order to focus on the research, programming effort was tried to minimize. Therefore the *OGRE3D* was used as an auxiliary library. Mainly the *OGRE* library was utilized to visualize the results. With its existing methods and data structures for different basic animation techniques such as vertex, morph and most importantly skeletal animation it was well suited for this task. Required techniques such as skinning, basic animation blending or interpolation methods like Spherical Linear Interpolation (SLERP) could be used from the library. Moreover the well organized and yet performant data structure for skeleton and joints could be accessed conveniently from inside code. Basic mathematical structures like vectors and quaternions, which were extensively used throughout the implementation, were also provided by the OGRE library.

4.7 Autodesk character studio

For analysis and evaluation, test characters and animation were needed. Character Studio embodies a highly flexible concept for character animation called the biped. It allows interactive creation and animation of skeletal armatures inside 3D Studio Max©. Apart from creating animations, motion scripts can be loaded and combined on a biped skeleton. These scripts can also be combined into sequential or overlapping motion scripts. Moreover, controls for importing motion capture data exist. A great deal of formats for motion capture data, such as *.csm*, *bvh* and the native *.bip* files are supported. Via the integration of character studio into the work flow animations from multiple sources could comfortably be used for analysis and evaluation.

4.8 Scenario

In the following paragraph a scenario of the usage of the application is shown. All use cases of the system are illustrated in Figure 3.

The system had to provide a mechanism to allow the definition of a used character. A character selected by the user. All needed resources as well as an character dependent configuration file are loaded. The file includes all necessary character specific information. Then an animation can be chosen by user input. The different selectable emotion states, in which the animation is to be shown are provided by the graphical user interface. Finally the animation is shown in the state requested by the user.

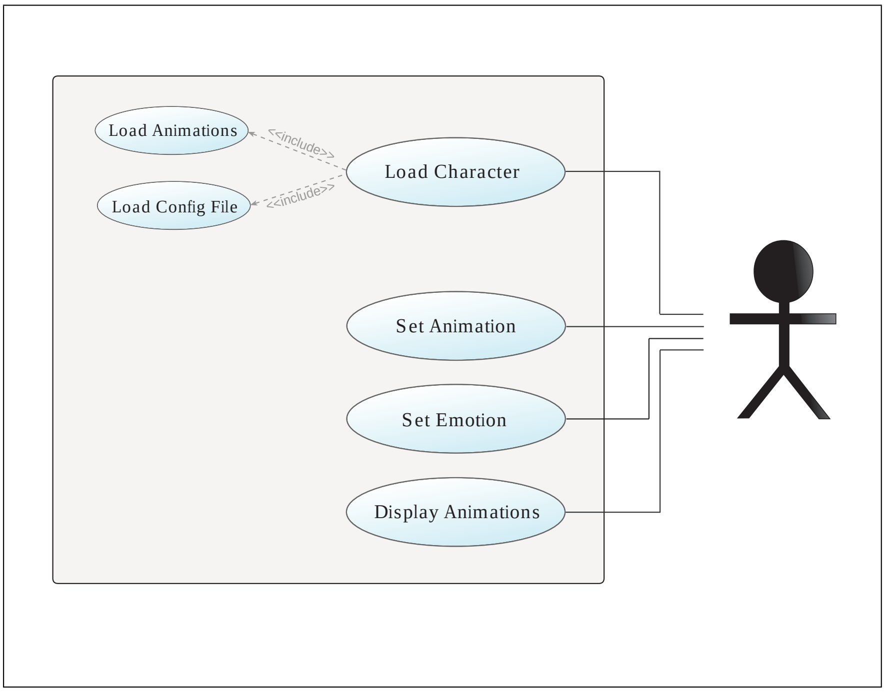

Figure 3: Use-Case diagram of the scenario

5 Analysis

The analysis carried out in the scope of the thesis was supposed to examine different aspects of motion. Firstly, the feature domains speed, range and posture, which were identified in the previous chapters had to be studied closely. The main question to answer was their suitability for emotion representation in the different movements. It had to be shown that the properties of these domains varied significantly depending on emotion.
Secondly, concrete values were extracted from examination for emotions. These were gathered in respect to the body parts and the relative differences to a neutral reference animation. The resulting values were then used in the actual implementation. For natural and authentic data, motion capture sequences of real human movements was used, which is described in the following chapter.

5.1 The Test Data

The test data used throughout the analysis was obtained from the Perception-Action-Cognition (PACo) lab at the University of Glasgow. The research of this group is focused on human movement as a 'rich source of information about the intentions, affect and identity of an individual.' Besides fundamental research in that field, the lab also surveys application domains such as computer animation and artistic expression in music and dance. The lab provides their research motion capture data freely as a motion library in the Internet. The used capture data was available in formats for character studio (csm-files) as well as matlab.
The used full-body movement library was recorded especially for the research of emotional movements. It therefore was well suited for the defined analysis aims. The motion capture data was gathered from thirty test sub-

jects in the age of seventeen to twenty eight. The individuals were recorded while performing walking, knocking, lifting and throwing actions. Each of these actions was performed in four different emotional states.

5.2 Focus

The available data set contained a lot of information. Thirty actors performing four different actions in four different emotional states summed up to 480 unique motions. With each motion sample containing information about time, velocity, acceleration, apex angle etc the observed data had to be limited. Therefore the available sample data was reduced to a small set of actors. Additionally the examined joints were restricted to joint, which represent significant body parts, such as arms, legs, head and so on. Finally, due to the existing sample data, only four emotional states were considered: angry, happy, sad and neutral. This reduced data set could then be surveyed in respect to the feature domains, derived from the results of related research presented in Chapter refChapterAffectiveHumanMotion. In order to extract usable data for the implementation, a reference motion was defined per group of movements. A group consisted of same movements performed in the different emotional states. The neutral action served as the reference sample data. All following comparisons were made in respect to the reference data and describe a delta offset or difference. The findings of the survey are introduced in the following paragraphs.

5.3 Speed

In a first step, the duration of different motion clips were analyzed. The time which was needed to perform a specific action in a neutral state was measured. This values were then compared to the duration of a same action performed emotionally. Since the duration of an action is directly proportional to the velocity, this gave an inference on the velocity, in which an ac-

emotion	**avg.**
sad	0,3
angry	0,2
happy	0,5

Table 1: Average durations for actions performed in happy, angry and sad

tion was performed. Rather than actual velocities, which would be dependent on the performed action and could therefore not be compared among different movements, a relative difference between the neutral and the emotional animation performance was determined. These relative value could be compared among different motions. In that way trends and similarities could be identified. It could be seen that the relative differences of arbitrary motions acted in the same emotional state to their neutral references resembled between each other. For example, actions in a sad state generally took longer to finish than neutral ones. Angry actions however, needed less time in average. Actions acted in a happy state showed more ambiguous results and no clear trend could be found. Figure 1 shows a list of the results averaged over the examined motions.

The results for overall duration of an action already showed significant characteristics for emotions. However, durations are just a rough cue, since they give no information about the actual speed in which the body, or limbs are moving, nor the acceleration behavior. To obtain more meaningful values for the motion evaluation, the acceleration behavior and speed of actions was surveyed. Speed and acceleration was mainly determined for limb joints. It was found, that acceleration of limbs and the whole body give strong hints regarding the emotional state of the actor. Trajectories of acceleration curves from angry motions showed high peaks in both directions, which represents strong acceleration and deceleration values. Contrary the acceleration curves of sad performances showed no significant peaks. The sad actions were performed in a comparatively constant velocity. Happy actions could be identified by an evenly distributed but slightly higher speed than the neutral reference. Figure 4 shows speed curves in the different emotional states of a forearm bone during a throwing action.

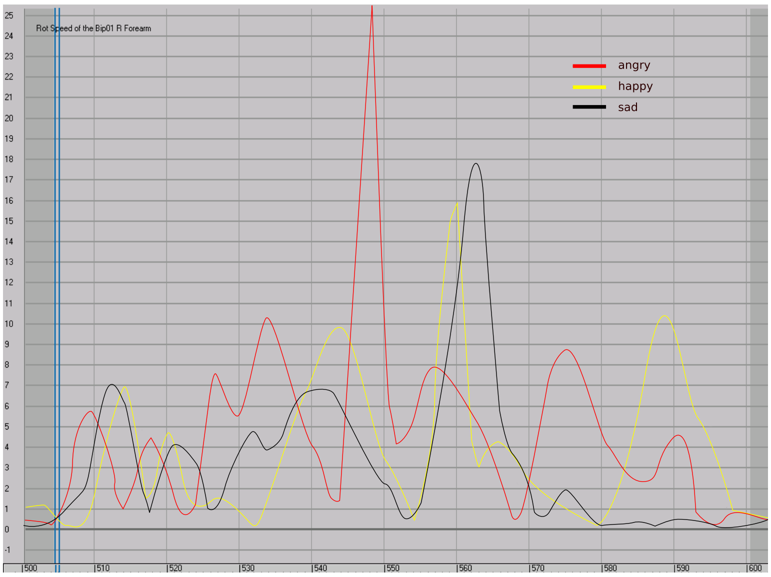

Figure 4: Speed curves of a forearm performing a throw action

It was found that differentiated results could be found, when the speed of actions was examined for single segments of the whole action. These periods were bound by the time of direction changes of a defined main end effector. For example, a throw action could be divided into the following basic periods: hand to object, strike out, throw. These periods were then classified into two groups, periods directing away from the body and periods directing towards the body. Results showed a tendency for faster speed of outgoing movements in angry states. On the other hand no differences could be found regarding sad motion periods.

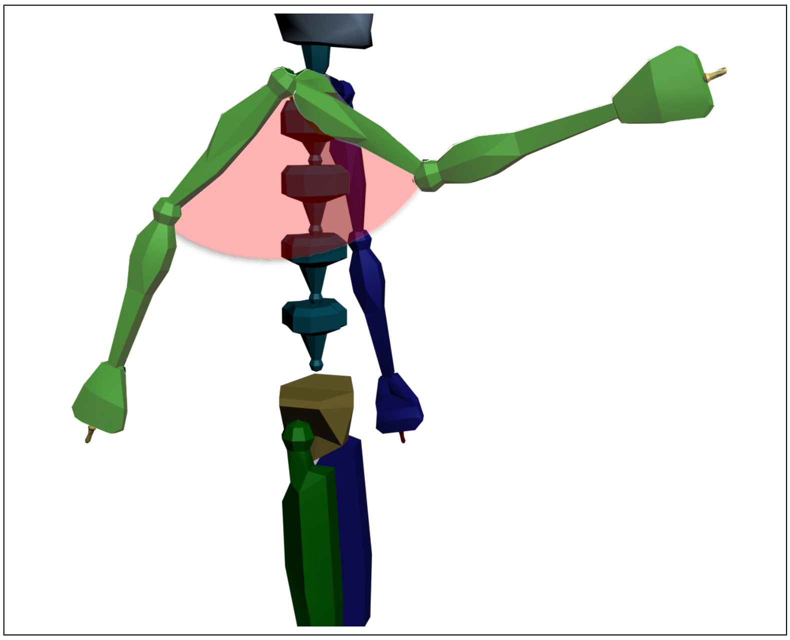

Figure 5: The range of an upper arm motion

5.4 Range

The second observed dimension was the range of a motion. By range the difference between the maximum and minimum apex angle of a joint is meant. This is illustrated for an upper arm in Figure 5. Ranges were measured for segments of a motion, in which a limb joint was moved from one end position to another. These angles were compared among the different motions.

It could be observed that the distinct joints of a body were influenced differently. However groups of joints showed similar behavior. Therefore the joints could be arranged into categories, in analogy to [KAC96]. Figure 6 illustrates these categories.

The emotions had variable impact on these defined categories. The range of joints in *Category A* changed significantly in a happy state. Especially both forearms and hands performed more excessive motions. Angry emotions lead to a stronger movement in the shoulder and upper arm areas as well as the thighs, which belong to the *Category B*. In sad motion all movement ranges were decreased and especially joints of *Category B*. Changes of range in joints in *Cathegory C* were minimal in emotions.

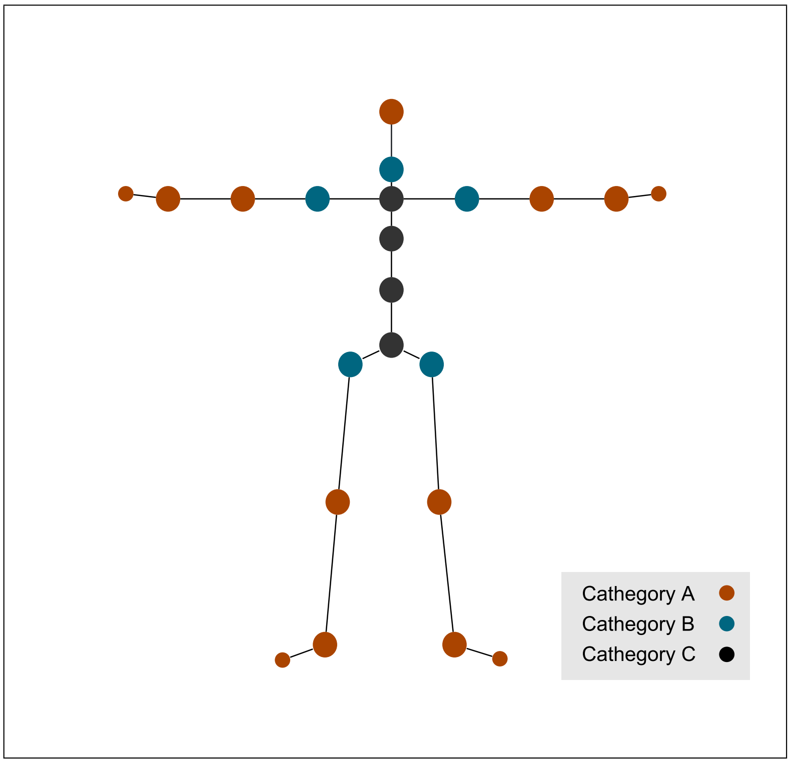

Figure 6: The cathegories of movement range

5.5 Posture

Certain postures have symbolic meanings which are equally interpreted by observers [Arg79]. Most of these communicative body postures are static and not applicable to motions. Some posture aspects however could be observed in the human movements. For example a forward bent head or hanging shoulders are strong indicators for a sad inner state. Such head

and shoulder positions also varied throughout the test data and had an eminent impact on sad movements.
Contrary, the head was held high in happy moods or the chin stretched forward in aggressive, angry stances. These postures are stereotyped often in cartoons, acting and other fields of art. Figure 7 shows such stereotype poses found in the test data.

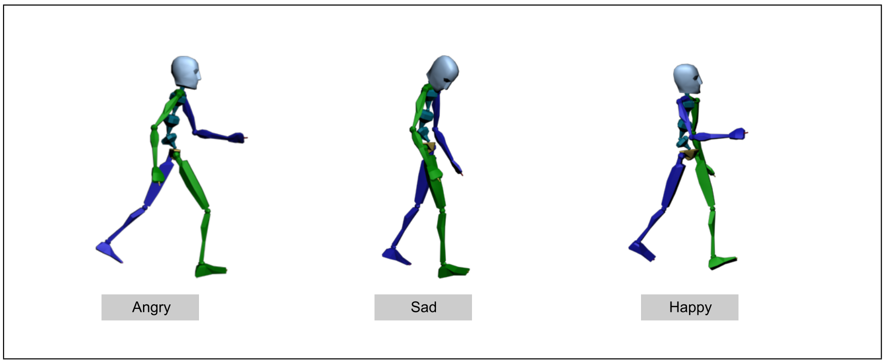

Figure 7: Stereotypical postures found in the test data

5.6 Conclusion

The results of the examination showed observable differences in the focused domains of speed, range and posture. Some emotions seem to possess strong and significant features, which distinguishes them from others. On the other hand, mostly minimal changes in a transition from neutral to happy could be extracted. Hence, it was hard to find a general trend for happy motions, which was equally relevant for all subjects and motion types.

Nevertheless trends could be found in all emotions, which were not dependent on a specific motion and could be observed in the motions of most test actors. The next Chapter describes how these domains were systematically realized and combined into an emotional animation creation system.

6 Implementation

In the scope of the thesis, a proof of concept implementation of an emotional animation modification system has been realized. Given an animated three dimensional character model the system creates emotionally enhanced animations from neutral animations. After a configuration process, this modifications can be applied to every virtual humanoid character, with skeletal animations. This is achieved by the use basic emotion properties, which simulate general changes from neutral to emotional within the limits of predefined motion domains.

The system was implemented in the C++ programming language. The object oriented elements of C++ were used to clarify the structure of the implementation. The dominant role of C++ in relevant industry areas, such as computer game programming, was another factor, which benefited the choice. The development was carried out on a Windows XP©operating system and in the Microsoft's Visual Studio C++ 2005 Express©development environment. The following system hardware configuration was used for testing:

- AMD Athlon©64 Processor 3200+ 2.0Ghz
- 1 Gigabyte RAM
- Nvidia©Geforce©6100 Graphics Board

During the implementation more effort was put into the aspect of clarity than into performance. Therefore emotional animations are automatically generated at system initialization, hence the application of modifiers in real time to achieve interactive changes of the animation is not possible.

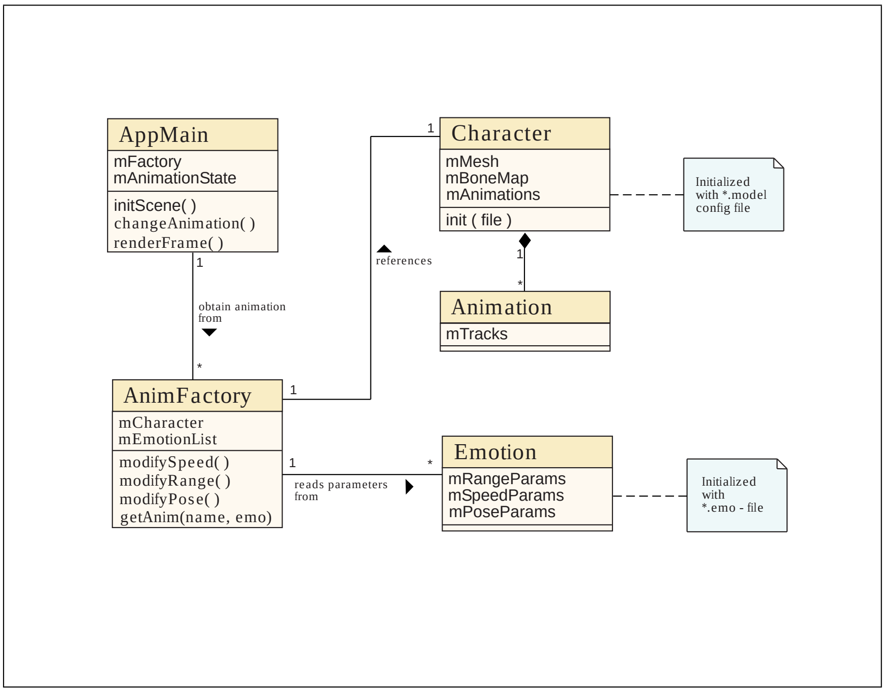

Figure 8: Class diagram of the system

6.1 System Overview

The implemented system generates affective animation variations from a given neutral source animation. The overall class structure of the implemented system is illustrated in Figure 8. A detailed description of the functionality of different classes is given below.

6.2 Animation Class

An animation in OGRE is represented by the Animation class. An animation is an assembly of tracks each referencing a different bone. Each track has one or more key frames, each defining a relative position and orientation of a bone depending on time. At runtime bone positions are updated and interpolated between key frames. Position can be interpolated in two basic modes, linear and spline interpolation. For orientations spherical linear interpolation (SLERP) as well as linear interpolation (LERP) can be used.

6.3 Character Class

The character class was used to represents a virtual character object in the target system. The character was created from a configuration file loaded at runtime. This configuration file provided the resource names for mesh and skeleton files of the character. The file also contained a list of the mapping between character specific joint names and the reference skeleton model, which was used internally. Animations available in the linked skeleton were provided as well as the segmentation into basic periods. This had to be done manually since an automatic generation of periods led to unsatisfying results and was not further researched due to limited time and irrelevance for the defined results. The character class formed the interface of the particular virtual model and the rest of the system. In that way arbitrary models could be used for the system. The following lines show the structure of a model configuration file in xml syntax.

```
<?xml version="1.0"  standalone='no' >
<!-- "model.mesh" model configuraton file -->
<Character>
    <Animations>
        <Animation name="Wave" bone="Bip01_R_Finger0" >
            <Period start="0.0" end="1.372" cat="2" />
            ...
        </Animation>
        ...
    </Animations>
    <Bonemapping>
        <Joint id="Root" name="Bip01" />
        <Joint id="Head" name="Bip01_Head" />
        <Joint id="Neck" name="Bip01_Neck" />
        ...
    </Bonemapping>
</Character>
```

Figure 9: Example model configuration file

6.4 Emotion Class

An emotion class was used to give access to parameters used for a specific emotion. In the final systems three different emotion objects were used representing sad, happy and angry emotions. The parameters values were estimated or averaged based on results of the analysis described in the last chapter. Emotions were loaded from specific files at runtime. Range coefficients for the joint categories, speed values for different categories of motions as well as posture modifications were read from files and used to initialize the emotion objects. Such an abstract emotion class initialized with a configuration file allowed a flexible loading and convenient definition of new emotion usable by the system. An emotion configuration file is shown below.

```
<?xml version="1.0"  standalone='no' >
<!-- Emotion specific data -->
<Emotion name="happy" vCat0="1.0" vCat1="1.3">
    <Bone name="Neck" range="0.3" speed="1.1" yaw="" pitch="-10.0" roll="" />
    <Bone name="Head" range="0.3" speed="1.3" yaw="" pitch="5.0" roll="" />
    ...
</Emotion>
```

Figure 10: Emotion configuration file

6.5 AnimFactory Class

The animation factory class was implemented to generate the results and provide them to the main application. An animation factory administers all information needed for the creation of the modified animations. This includes a character object as well as access to the different emotion objects and their parameters. All available animations are modified by three functions which change basic properties of each motion in order to achieve emotionally expressive animations. The functions, modifyRange(), modifySpeed() and modifyPose() are explained below. The modifyRange method operated directly on the animation data of a character. Therefore each key frame of an animation track was analyzed first. For outer joints key frames which represented an end position of the joint were determined. This was done by a discrete approach where a velocity vector between two key frames was calculated. Directional changes of two velocity vectors, which lay within a certain delta defined time positions of an end position of the joint. The rotation values of these key frames were then enlarged. For that rotation values were transformed into a angle axis representation, in which the angles could easily reduced or enlarged by modifying the angle value and leaving the axis untouched.

Adaption of speed and velocity of animations was performed in the *modifySpeed* method. The analysis showed that velocity and acceleration behavior was dependent on basic periods of an animation. These basic periods were defined by the direction of motion of a main end effector for that period. Therefore, for each period of the animation the respecting end effector velocities were calculated. Due to the relative nature of key frames world positions had to be calculated for the end effector. This was done by a multiplication of parent matrices and and the initial position of the end effector.

$$BonePos_{world} = (M_n * .. * M_3 * M_2 * M_1) * BonePos_{initial} \tag{4}$$

Once velocities of a period for neutral and emotional animations were calculated period velocities for emotional animations had to be normalized. This was done due to possible velocity changes caused by prior range modifications. The normalized velocity could then be modified according to modification values of the respective emotion, as seen in 5, where $Velocity_{period}$ is the velocity of the current period, N_{period} the normalization factor for the period and $CatVelocity_e$ the emotion velocity parameter for the cathegorization of this period.

$$PeriodVelocity_{final} = CatVelocity_e * (N_{period} * Velocity_{period}) \quad (5)$$

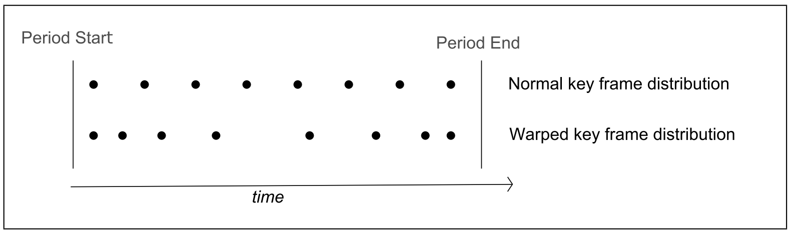

Figure 11: Time warping of key frames

Changes in velocity were done by means of time-warping. Time warping is a technique, where the distribution of time is not seen as fixed, as in the real-world. By the warping of time acceleration behaviour of periods could be modeled. Figure 11 illustrates the main principle of the technique. The advantage of this approach lies in the use of existing data. In that way the motion trajectories of the original animation is kept and not altered to new positions and orientations.

Finally the modifyPose function was run on the animations. Pose modifications were mostly of a constant nature and not dependent on the animation. Nevertheless changing initial orientations of bones could not be done due to the differing pose modifications per emotion. Changes of emotion of a single character would require to hold track of the modifications done to the initial pose. For a more flexible and convenient solutions, emotion specific posture changes were directly calculated into all key frames of the bone tracks. In that way an original initial pose

$Orientation_{joint}$ could be used for all different animations and emotions. In a hierarchical structure like the skeleton, the reorientation of an inner node or joint results in an orientation change of its children. Therefore an inversed matrix $Mod_{emotion}^{-1}$ of the done modification had to be multiplied onto the original orientations of the child joints $Orientation_{child\_Orig}$. In that way their original orientation was maintained. Equation 6 and 7 show the effected node and child equations

$$Orientation_{new} = Orientation_{joint} * Mod_{emotion} \tag{6}$$

$$Orientation_{child\_new} = Orientation_{child\_orig} * Mod_{emotion}^{-1} \tag{7}$$

This convenient modification of orientations could be done due to the use of quaternions. In that way orientations could be simply right handedly multiplied. The affected joints were limited to spine, neck, head and shoulder joints.

6.6 Scenario

A simple runtime scenario of the implemented system is shown in Figure 12. The application firstly initializes an instance of a AnimFactory class. During the *AnimFactory* initialization the character object is created. All required information is read from an according configuration file. Available animations are the modified by the factory. Thus the speed, range and posture modifications are performed. The parameters for each emotion are read from the different emotion objects. The main application uses the factory class to reference animation data. The animations are provided by animation name and requested emotion. The retrieved animation is then displayed by the application.

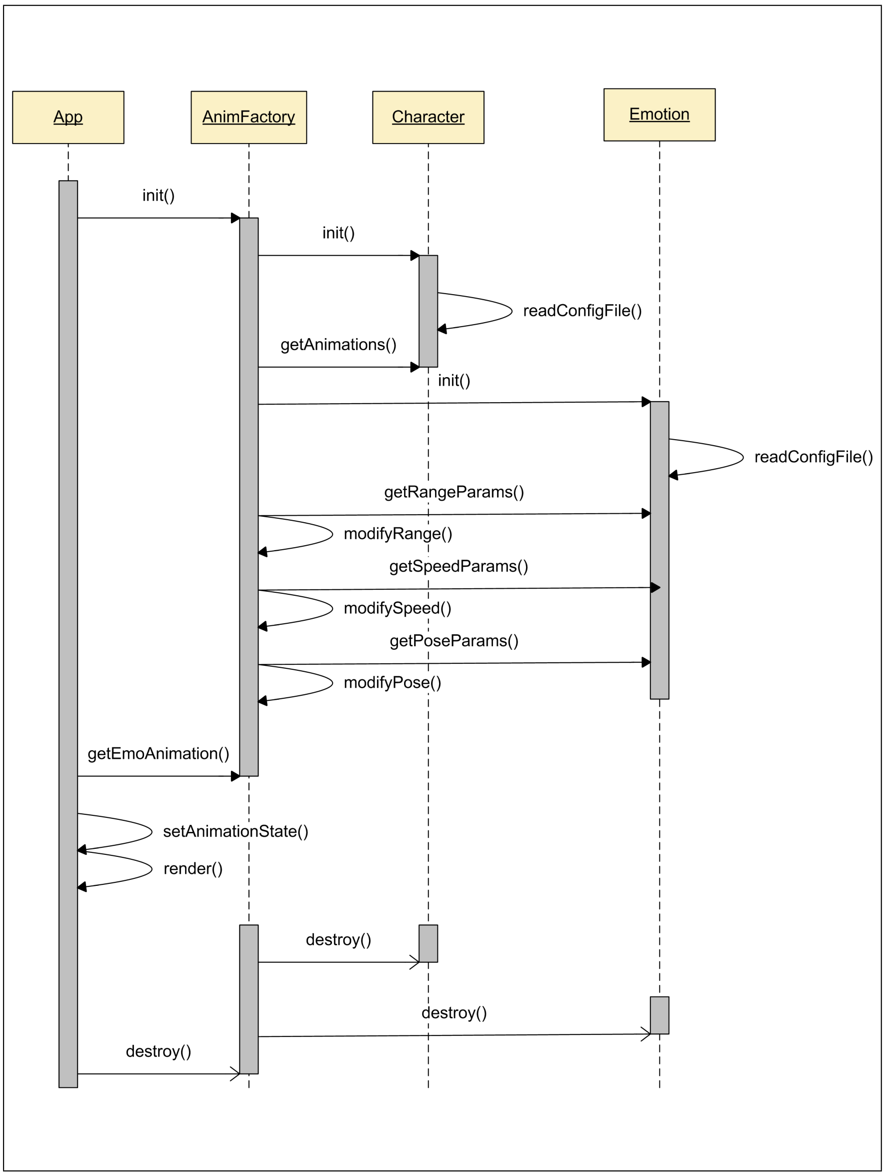

Figure 12: Time warping of key frames

7 Results

The implemented system changes given animations in respect to a selected emotion. Different emotional results of the same animation differ from each other (see Figure 13). In order to evaluate results of the system the animations had to be reviewed and judged respecting their impression by human observers. The chapter describes the performed evaluation of the test system. The test setup and procedure are introduced . Afterwards the results are presented and discussed.

Figure 13: Walk actions modified by the system with angry, sad and happy emotions (from top to bottom)

7.1 Evaluation

Ten different animations expressing no noticeable emotion were used as input data for the test. Additionally, the animations were required not to imply an emotion indirectly, such as hitting or boxing, which are usually performed in an aggressive state. All animations that served as test input data for the evaluation are listed below:

- hand waving
- walking
- pointing
- hammering
- squatting
- athletic exercise
- soccer goal shoot
- football passing
- jogging
- jumping

Each animation was modified by the implemented system with parameter sets for three basic emotions. The used emotions were sad, happy and angry. This resulted in thirty test animations for evaluation. In order to obtain results that could be evaluated a special balloon character model was used throughout the tests. The character's appearance carried no visual information such as facial features, cloth or a characteristic figure. In that way possible cues for emotions were restricted to the motion of the character. Since there exists no accurate method to measure or distinguish emotions, human test candidates were consulted to judge the system. In total fourteen candidates took part in the evaluation. During the evaluation the test animations were displayed to the candidates. In order to evaluate the impression of a single animation and to exclude decisions motivated by comparison or a process of elimination just one animation was shown at a time. Also no neutral animation was shown before, which could have served as a reference. Based on the viewed sample motion, an inner state of the character had to be estimated by the user. Neutral, sad, angry and happy emotions could be selected in the evaluation application. A screen shot of the user interface of the test environment can be seen in Figure 14.

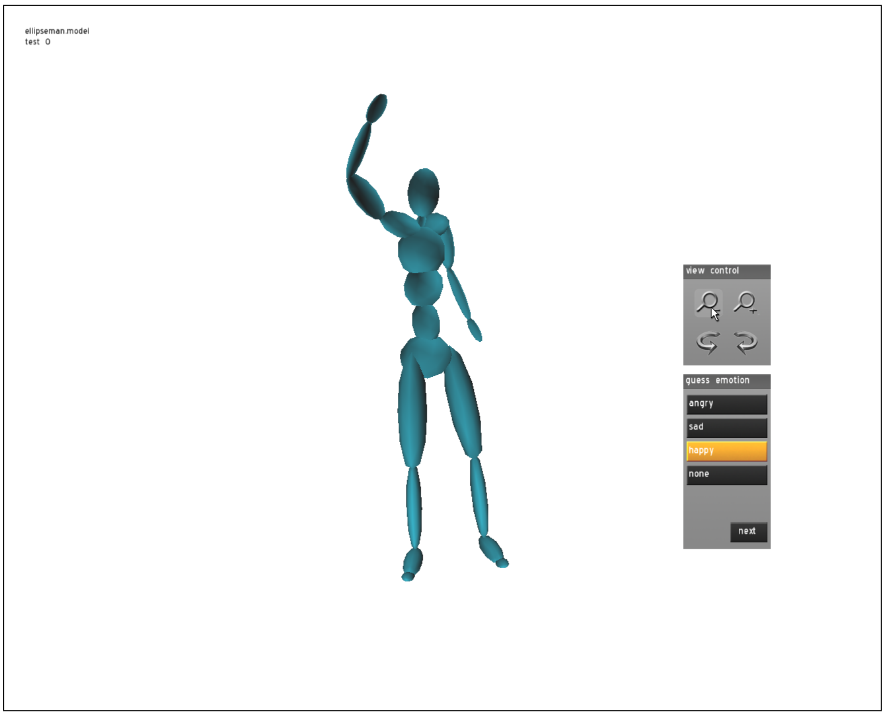

Figure 14: Screenshot of application used for the evaluation of the system

7.2 Results and Discussion

In total, 341 generated emotional motions have been evaluated by users. 218 motions have been identified correctly, which corresponds to 64%. The most steady recognized emotion was 'sad'. Hereby 90% of the sample movements were assigned correctly. The other emotions, angry and happy, were matched less frequently. Angry states were recognized by a probability of 0,55 and happy emotions by 0,46. The results show recognition rates which lie above chance. Especially sad emotions were identified extreme dependently. This conforms with findings of other researches, who tested the impact of emotion to bodily movements. [WD96] who evaluated the recognition of real motion data shown on point light displays reported recognition rates of 63%. Therefore the generated motion data seems to convey a similar emotional content as real motion data. However, significant differences between the results of the used emotions are apparent. This shows that the used motion properties, which were used to modify the animations, are not equally valid for all animations. Figure 15 shows the complete list of results sorted by emotion.

	Number of tests	Correct	Average in %
Total	434	278	64
Angry	154	87	57
Happy	140	65	46
Sad	140	126	90

Figure 15: Correctly matched samples sorted by emotion

In order to estimate the validity of the implemented motion domains for all kinds of motion, the results had to be analyzed dependent on the performed action. An animation dependent view on the results showed great differences in emotion recognition. While all emotional styles of walking could be recognized by a probability of over 0,7, squat actions were averagely matched with a probability of only 0,4. A reason for these dissimilarities could stem from user experiences of real life. A walk is a motion often witnessed in all kinds of variation and different emotional states. Hence an image of different styles of walk already exists in memory of people and is therefore well identified. On the other hand a person who squats angrily is hardly seen in everyday life. The low recognition rates for such rarely experienced combinations of movement and emotion affirm this assumption (i.e. an angry jump (0.28), a happy soccer shoot(0.21) or an angry jogging motion(0,07)).
Another problem was the definition of neutrality respecting animations. The tested animations which served as input data an were defined to be neutral were also tested. Neutrality could was recognized in 40% of all tested animations. This demonstrates that a definition of a neutral state seems to be as difficult as for emotional states. A list of the results in respect to the performed animation is seen in Figure 16.

Often, it was stated by the contestants, that a particular emotion was conveyed to a certain extend, but the realism of the animation was poor. This general problem originated from the approach of the thesis. It was tried to find modifications for different animations, which are detached from the semantics of a particular motion. Since modifications were equally applied to all kinds of movements in a specific emotional state, this could in some cases lead to unnatural results, most apparent in the velocity behavior. Some velocities were perceived as too slow others as too fast

	total	angry	happy	sad
walk	83	79	71	100
goal shoot	73	93	21	86
hammering	71	93	29	93
jogging	52	7	57	93
hand pass	62	50	50	86
pointing	60	57	21	100
exercise	76	50	93	86
jumping	55	29	50	86
squatting	40	14	14	64
waving	64	57	57	79

Figure 16: Correctly matched animations in %

in order to appear realistic. Another visual problem which occurred was the lack of a constraining system. Due to range modifications achieved by forward kinematics the exact position of end effectors could not be retained (see Figure 17). This had serious consequences for animations where a defined position of a body part was demanded. For example, a grabbing motion could miss an aimed object. Such problems could be solved by a constraining systems, where body postures at defined time positions have to be preserved.

While reviewing the results, it became clear that the modification of bodily motions can be used to create a certain kind of expression. The conveyed emotion could be recognized at rates above chance. Still a dependable identification of emotions only by means of bodily motions seems not to be practicable. The relatively low percentage of recognized happy and angry movements confirms this assumption. As mentioned in the survey

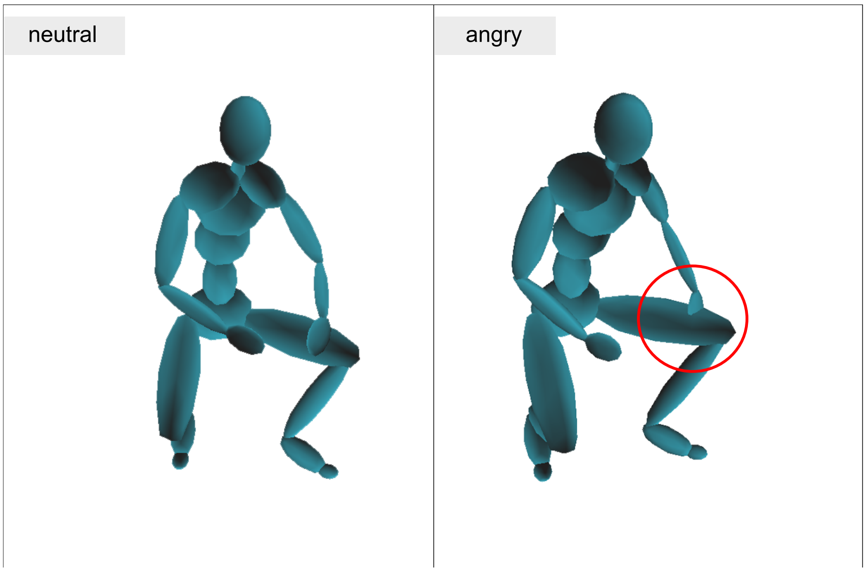

Figure 17: Endposition in original animation (left) and in angry state (right)

of relevant literature the expression of emotion manifests through different channels. Therefore the effect of combined communication channels for emotion expression have to be examined. In future work, a system capable of facial expressions and body motion modification could be implemented and evaluated. Such an approach could be a step towards expressive characters.

Finally, it has to be mentioned that motion of characters are complex. By the use of basic parameters to modify animations a tendency towards an emotional impression can be realized. But believable and expressive character motion is influenced by many other factors which can not be generalized for all possible motions and emotional states. Some motion features are tightly coupled to specific emotions such as specific gestures. Other cues for emotion are dependent on personality, gender, age and other properties of the character. Such features can not be realized with the presented general approach. Nevertheless, this diversity is believed to be necessary to create visually convincing expressive motion results.

8 Conclusion

Expressivity is a key factor to believable virtual characters. In the scope of the thesis, a method has been developed to create emotional animations from neutral motion. This method is based on research from psychology and motion analysis.

Three basic domains were identified to have a vital impact on emotional body movements: range of motion, velocity and body pose. These domains were examined closely from real motion data in order to extract the difference between neutral and emotional movements. In the developed system these parameters were applied to neutral animation data in order to add an emotional expression. An evaluation was carried out to verify the results. Ten different neutral motion sequences were modified by parameter sets extracted from angry, sad and happy motions. These animations were judged by test candidates respecting the emotion which was perceived.

Generally such generated emotions were recognized to a similar amount as emotions from real human motions shown on pointlight displays. The results for the different emotions differed noticably among each other. Sad movements could be identified with most accuracy, which equals result from other researches [WD96].

The system operates on keyframed animation data. Therefore the system allows the reuse of animation data obtained from arbitrary sources. The quality of the resulting animation depends on the input data. It was also found that animations, which showed unusual animation-emotion combinations hardly seen in real life, could not be identified well. Dependent on the input animation the system produces poor results. Due to changes in motion range endeffector positions could not be preserved which resulted in self intersection and other issues. This problem could be adressed by the use of a constraining system which retaines certain key poses.

Summing up it can be argued that the found parameters are capable of enhancing the expressivity of neutral animation data. However, due to the drawbacks and quality issues, they should not be used exclusively to model emotional expressivity of virtual characters. Rather the found parameters should be seen as a basis for future research. An refinement of the found parameters and a deeper analysis of various motion data could lead to better results. However, it seams that different non-verbal communication channles, such as facial expression, gestures and expressive movements have to be combined to achieve a satisfying synthesis of emotion in virtual characters.

References

[AM69] J.T. Friar A. Mehrabian. Encoding attitude by a seated communicator via posture and position cues. In *Journal of Counseling and Clinical Psychology, 33*, pages 330–336, 1969.

[Arg79] M. Argyle. *Koerpersprache und Kommunikation*. Junfermann Verlag - Paderborn, 1979.

[BB04] J.N. Bailenson and J. Blascovich. *Avatars*. Berkshire Publishing Group, September 2004.

[Bir70] R. Birdwhistell. Kinesics and context. 1970.

[BML] Biomotion lab. BMLWalker.

[DCB00] L. Zhao D. Chi, M. Costa and N. Badler. The emote model for efort and shape. In *Proceedings of SIGGRAPH '00*. ACM Press, 2000.

[DeM89] M. DeMeijer. The contribution of general features of body movement to the attribution of emotions. In *Journal of Non-verbal Behavior Vol. 13 No.4*. Springer Netherlands, 1989.

[DK96] et.al. D. Kurlander. Comic chat. ACM press, 1996.

[EF78] P. Ekman and W.V. Friesen. Facial action coding system. 1978.

[FR84] B. Fehr and J.A. Russel. Concept of emotion viewed from a prototype perspective. page 464, 1984. Journal of Experimental Psychology: General 113.

[JCK78] D.R. Proffitt J.E. Cutting and L.T. Kozlowski. A biomechanical invariant for gait perception. In *Journal of Experimental Psychology: Human. Perception and Performance*, pages 357–372, 1978.

[Joh73] G. Johansson. Visual perception of biological motion and a model for its analysis. In *Perception and Psychophysics 14*, pages 201–211, 1973.

[KAC96] A. Bruderlin K. Amaya and T. Calvert. Emotion from motion. 1996.

[Lan06] J. Lander. On creating cool real-time 3d. October 2006. www.gamasutra.com.

[Mas00] R. Masters. Compassionate wrath: Transpersonal approaches to anger. 2000. Journal of Transpersonal Psychology.

[MUT95] K. Anjyo M. Unuma and R. Takeuchi. Fourier principles for emotionbased human gesture animation. In *Proceedings of SIGGRAPH'95, Annual Conference Series*. Addison Wesley, 1995.

[OJ96] K. Oatley and J. Jenkins. Understanding emotions. 1996.

[Pat01] H.M. Paterson. The role of velocity in affect discrimination. 2001.

[Per95] K. Perlin. Real time responsive animation with personality. 1995.

[Pic97] R.W. Picard. *Affective Computing*. MIT Press, September 1997. ISBN 0-262-16170-2.

[PK91] N. Magnenatthalmann D. Thalmann P. Karla, A. Mangili. Smile: A multilayered facial animation system. In *Proceedings IFIP conference on Modelling in Computer Graphics*, pages 189–198. Springer, 1991.

[Pol04] F.E. Pollick. The features people use to recognize human movement style. 2004.

[RB98] J.G. Cunningham R.T. Boone. Children's decoding of emotion in expressive body movement: The development of cue attunement. 1998.

[TB81] T. Buckner T. Beardsworth. The ability to recognize oneself from a video recording of one's movements without seeing one's body. In *Bulletin of the Psychonomic Society 18*, 1981.

[Vol03] G. Volpe. *Computational models of expressive gesture in multimedia systems*. Ph.d. dissertation, University of Genova,Italy, April 2003.

[VVS06] A. Steed E. Tanguy X. Pan C. Loscos V. Vinayagamoorthy, M. Gillies and M. Slater. Building expression into virtual characters. 2006. Eurographics 2006.

[WD96] S.E.G. Lea D. Morgan W.H. Dittrich, T. Troscianko. Perception of emotion from dynamic point-light displays represented in dance. pages 727–738, 1996.

[wik] Emotions - wikipedia. website.