Animating emotions in ECA's for interactive applications

Lars Zilmer

lzilme08@student.aau.dk

Aalborg University Copenhagen, spring 2012

Abstract: Within interactive applications, emotions expressed through animations are often overlooked and when animating for a mobile platform, certain constraints have to be considered, such as a limited amount of polygons and bones in the model. This master's thesis, aims to investigate how to apply emotional expressions through body language by the use of animation to an embodied conversational agent (ECA), used for an interactive application on a mobile device. The research suggests the three most suited emotions to test as being: angry, sad and happy. Furthermore a list of emotional expressions to implement in the animations was identified, and a test was then conducted to validate the findings, by asking the participants how they perceived the emotional states of the ECA. The results indicate that it is possible to implement emotions through body language to an ECA, when constrained by a mobile platform.

Keywords: Animation, 3D, Embodied Conversational Agent, ECA, Low Poly, Mobile Platform, Body Language, Emotions

TABLE OF CONTENTS

I. INTRODUCTION AND MOTIVATION

Animation was previously limited to 2D as drawings on paper, and was most often used as entertainment in the form of cartoons for kids. But since Pixar released the first feature-length computer animated movie Toy Story in 1995, the concept of 3D animation within computer graphics has become increasingly popular, and today this type of animation is seen in for example advertising, education and movie productions.

Within these media forms, both animation, audio as well as the story are often used to express the emotional state of the characters, however when looking at interactive applications such as computer games, this emotional state is often expressed mainly by the story and the audio, through the voice actors and the background music setting the mood. The characters are then left looking stiff with no real emotional expression.

As this low level of emotional expressiveness seems to be the general standard in many interactive applications, it raises an interesting issue of how a desired emotional expression can be achieved on an animated 3D character in an interactive application.

Therefore this master's thesis aims to investigate how to apply emotional expressions by the use of animation to a 3D character used in an interactive application, which leads to the initial problem statement.

1. Initial problem statement

How can a 3D character be animated to express emotions in an interactive application?

II. PRELIMINARY ANALYSIS

1. Embodied conversational agents

Within interactive applications the term embodied conversational agent (ECA) is used to describe a graphical representation of an animated character on the screen [1]. *“ECAs are virtual embodied representations of humans that communicate multimodally with the user (or other agents) through voice, facial expression, gaze, gesture, and body movement.”* [21, p. 1]. Throughout this study only the visual modality will be considered, as the focus will be on animation.

2. Level of detail

When designing the model of the ECA a choice has to be made about the level of detail, this choice should be decided based on the target platform of the animation. The level of detail addresses the amount of triangles (or polygons) in a mesh. A high amount of polygons (high poly) in a model is used mostly

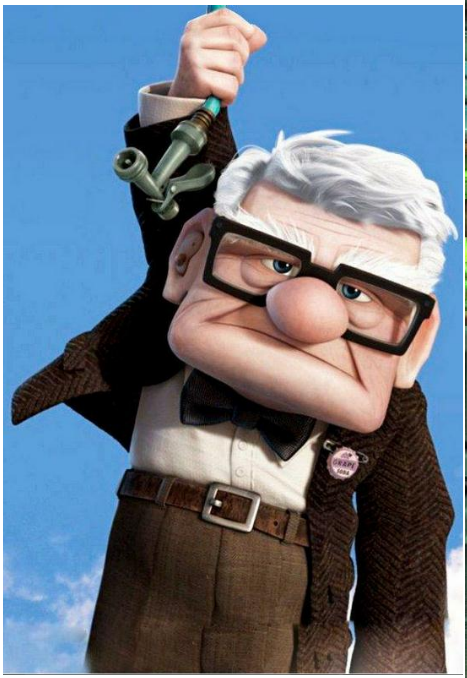

Figure 2 – Example of high poly, from the Pixar movie Up.

Figure 1 – Example of low poly, from the game Shadowgun.

for pre-rendered animations in movies and cartoons, whereas a low amount of polygons (low poly) is better for real-time animation in interactive applications since the device has to be able to process the amount of polygons.

When addressing the level of detail of the model, the skeletal setup has to be considered as well, as more bones in the skeleton require more calculations. This is not a problem for pre-rendered animations, however for real-time animation it has to be considered.

Some choices for target platforms for low poly models are:

Platform	Polygon count	Bone limit
Desktop platform (Windows, OS X etc.)	1500 - 4000	No suggested limit
Game console (Play Station3, Xbox 360 etc.)	5000 - 7000	No suggested limit
Mobile device (iPhone, Android etc.)	300 - 1500	30

Table 1 – Recommended polygon count, and bone limit from [9]

The main differences between the mentioned target platforms in relation to the poly count are the processing power and the graphics engine. It is not possible to give an exact value as to the number of polygons these devices can handle as a number of factors weigh in. The values given in the table above should only be considered as an approximation.

In the case of games, the role of the ECA and the genre of the game also affect the target polygon count. For example in a third person shooter game the hero should have a higher polygon count than the bad guys, as the hero will be in the center of the screen at all times [9]. Furthermore characters in a shooter game should have a higher poly count than characters of a strategy game, as the strategy game usually have a lot more characters on the screen at a time and the camera will be zoomed out, making it less relevant to have a lot of details in the models.

A key aspect when designing an interactive application is the frame-rate, the number of frames drawn per second. Besides the poly count, factors such as shading, script complexity, texturing, and scene complexity all affect the frame-rate. With this in mind it seems necessary to investigate just how to reduce the poly count on an ECA model, without affecting the overall look of the animation. Since the mobile device has the lowest performance this seems like the obvious choice of target platform to test on for the purpose of this project.

When the ECA has been developed, an interactive application that the ECA can be integrated with should be found or created. For this purpose this study will be created in collaboration with another master's thesis project at Aalborg University aiming at creating an interactive application. The application will then function as an environment for the ECA.

3. State of the art

The amount of applications for mobile devices containing 3D animation is increasing. However many of these are very limited in terms of expressive animations, the best example at the moment within this field is The Sims. Therefore it seems interesting to investigate how this application uses animation to express the state of the characters, as it might offer some inspiration.

The Sims for Android/iPhone

The Sims is a game with the purpose of taking care of a 3D character, by fulfilling the character's needs, such as sleep, hunger etc. illustrated by bars on the graphical user interface. The game starts out

Figure 3 – Screenshot from The Sims

with the creation of the character. The player selects the character's sex, look and personality traits, before creating and furnishing a home for the character.

Figure 4 – Screenshot from the character creation screen in The Sims

The game uses animation to inform the player about the needs of the character, when any one of the bars mentioned before gets too low, an animation is used to give an indication of what the character is in need of. For example if the character is tired he will start yawning and stretching, and walks slower with the arms hanging and feet dragging behind. These animations are often very exaggerated. Waving the arms high in the air and doing backflips are some examples of exaggerated animations used in the game. Sounds and visual effects are also used in combination with the animations to guide and inform the player.

The character model is very low poly, which is easy to see on the character creation screen since the camera is so close to the character, however during gameplay this is not an issue.

The camera is positioned above the scene at a slight angle, with rotation, translation and zooming functionalities.

4. Final problem statement

How can we overcome the constraints imposed by mobile platforms when expressing emotions in animated embodied conversational agents?

III. ANALYSIS

This chapter aims to explore ways of expressing emotions through body language in animations. The research is carried out by first looking at how to express emotions using body language. Before looking into the concept of animation, to get an understanding of how animations are generally used in interactive applications, and an understanding of some of the fundamental aspects of animation, by looking at principle and techniques used within animation. These findings will then form the theoretical framework, which will provide a way to answer the problem formulation.

1. Emotions in body language

Due to the earlier discussed constraints imposed by the mobile platform, namely the amount of polygons and bones that the ECA model should contain, it might be problematic to express emotions through facial animation, as a large amount of bones would be required to be able to control the mouth, eyes, eyelids, eyebrows and so on. Therefore it is important to investigate how emotions can be expressed by the use of body language.

When it comes to understanding the different poses and motions of body language it's important to take the whole body in to consideration and not just focus on one specific gesture, since many expressions can have multiple meanings. For example crossing the arms over the chest can be an indication of disagreement, resistance, guardedness or discomfort but is also a very comfortable pose for many people [13] [23]. *"Count to three. That is, refrain from assuming that any single gesture has a particular meaning until you see two corroborating gestures that reinforce that same meaning."* [13, p. 15] Furthermore, culture plays a role in the understanding of certain gestures.

Paul Ekman's research on facial expressions suggests that six universally identifiable emotions exist, these being: happiness, sadness, anger, fear, surprise and disgust [15]. Mark Coulson used this research as inspiration to investigate emotions in body language in static images. His research showed that the three most recognizable emotions are anger, sadness and happiness [16]. Therefore it has been chosen to focus on these three emotions for this study. The following section will provide a list of different poses and expressions that could help in achieving the desired emotional states for the animations. The list is based on findings from *The Nonverbal Advantage* [13], *I Can Read You Like a Book* [14] and *What Every BODY is Saying* [23].

Angry

- Chin pushed out.
- Leaning forward with a closed body posture can signal hostility.
- Hands on the hips, is one of the most common gestures people use to communicate an aggressive, super confident, or independent attitude.
- Kicking the ground, or different objects.
- Flailing arms.
- Clenched fists. [13] [14] [23]

Sad

- Bowed head / Look down.
- Stooped shoulders.
- Slumped body posture.
- Slow walking. [13] [14] [23]

Happy

- Head high and straight.
- Chest puffed out.
- An erect body posture is usually a sign of an upbeat mood.
- An open body position is perceived more positively than a closed body position.
- Arms relaxed at the sides of the body, is generally a sign of openness and accessibility.
- Evenly paced gait.
- Bouncy steps (Light on the feet).
- Arms swinging at the side in an easy and relaxed manner, when walking. [13] [14] [23]

2. Animation in Games

As the target application of the ECA can be thought of as a kind of game, it seems interesting to investigate how animation is generally used in games.

The most common animations used in games are idle, *"Idle animations are what you see when the character is inactive and waiting for input."* [12, p. 176], and walk/run animations. The point of the idle animation is to keep the character looking alive and to not suddenly become a static statue [10]. These animations are made as loops, meaning that the first and last frame of the animations have to be identical. This is to make the animations look seamless while playing in the game [11] [18].

When going from one animation to another, the current frame of the animation playing is faded to the first frame of the target animation over a short period of time, to give a smooth transition [18].

When creating an animation loop it's often a good idea to pick an uneven number of frames for the animation, as this will provide a center point on the timeline [11]. This becomes obvious when looking at for example a walk cycle.

A walk cycle is animated as if the character is walking on a treadmill [18]. Meaning that the character is walking in one spot without actually moving, instead the feet are sliding along the floor at a constant rate. This rate then determines how fast the character should move in the game, which is controlled through code.

The walk cycle consists of two steps, each containing 5 poses [11], where the first and last poses are mirrors of each other, resulting in the first and last poses of the complete cycle being identical. The center pose of the complete cycle can be considered to be the last pose of the first step, as well as the first pose in the second step. As mentioned a step consists of 5 poses, starting with the contact pose when both feet are in contact with the ground. When the back foot is lifted, so that the weight of the body is on the front foot, the body moves down, this is called the down pose. The next pose is called the pass pose. Here the lifted foot passes the other foot, which is getting stretched out. The grounded foot then lifts the body up into the up pose, by only touching the ground with the toes, at the same time the body is leaning forward and the lifted foot is moved in front to catch the body as it becomes out of balance. The step then ends with the lifted foot touching down on the ground again, in another contact pose, catching the body. This process is then repeated for the other foot, giving a complete cycle [11] [18].

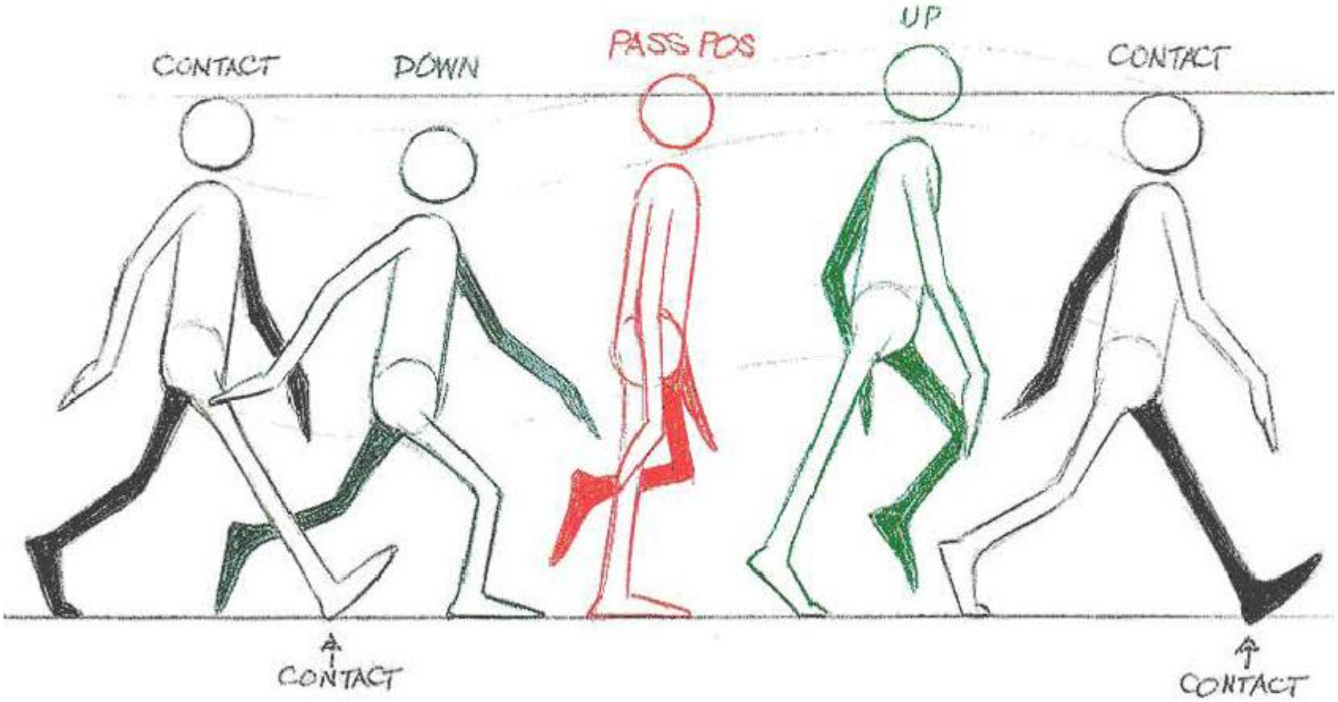

Figure 5 – The different poses in a walk cycle from [8].

Emotions can then be layered on top of the cycle by changing the amount of frames used up or down, resulting in a slower or faster walk respectively. Or adjusting how much different parts of the body are moving.

As mentioned the first and last frame of an animation loop should be identical, for this reason the last frame has to be removed before adding it to the game, as it would otherwise give an unsmooth transition when looping the animation.

3. The twelve principles of animation

The following section will look at Disney's twelve principles of animation to get a better idea of some of the important aspects and things to consider while animating. The section is mainly based on *The Illusion of Life* [2] and *Principles of traditional animation applied to 3D computer animation* [3]. The principles are listed below, followed by a brief description of each.

1. Squash and Stretch
2. Anticipation
3. Staging
4. Straight Ahead Action and Pose to Pose
5. Follow Through and Overlapping Action
6. Slow In and Slow Out
7. Arcs
8. Secondary Action
9. Timing
10. Exaggeration
11. Solid Drawing
12. Appeal

Squash and Stretch

Squash and stretch defines the rigidity and mass of an object. The less rigid an object is the more distorted its shape will become during an action. A chair and a table will not distort at all, whereas a soft ball thrown at a surface will squash down and have its sides stretched out, as seen on the picture below. *"The most important rule to squash and stretch is that, no matter how squashed or stretched out a particular object gets, its volume remains constant."* [3, p. 36]. During squash and stretch the shape of the objects don't necessarily have to be distorted. A person jumping will squash by bending down in the knees and then stretch out when initiating the jump [2] [3]. Referring back to the walk cycle this is also seen in the up and down poses.

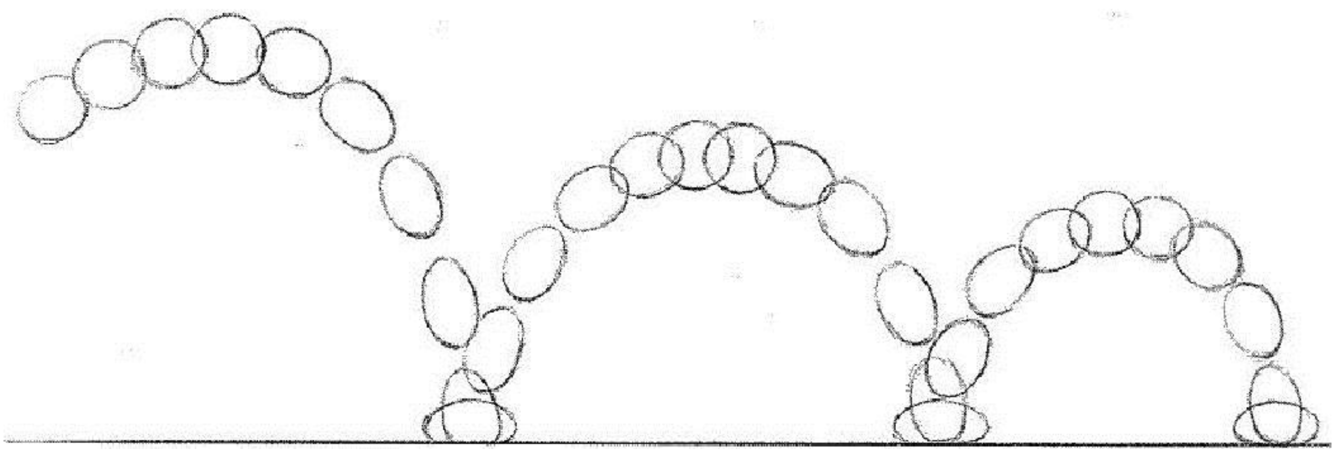

Figure 6 – Example of squash and stretch with a bouncing ball from [8].

Anticipation

According to the animator Bill Tytla there are only three things in animation, anticipation, action and reaction [8], which Charlie Chaplin formulated as.

1. *"Tell 'em what you're going to do."*
2. *"Do it."*
3. *"Tell 'em that you've done it."* [8, p. 273]

To keep the audience of an animation from getting confused when an action is carried out, that action has to be anticipated by the audience. There are multiple aspects to anticipation, when kicking a ball the foot has to be pulled back before kicking. Anticipation can also be used as an indication of weight, if the character is picking up an object, bending way down before starting the lift will indicate that the object is very heavy, whereas if the character just walks over to the object and picks it up without any anticipation it will seem very light [2] [3]. In an interactive application this seems less relevant as the actions of an ECA are usually reactions to the user input.

Staging

"Staging is the presentation of an idea so that it is completely and unmistakably clear." [3, p. 38]. Having the right camera angle, mood, and expression that best suit the situation, to make sure the audience doesn't miss the idea. The important aspects of the scene should be in focus, if it's important for the audience to see a specific facial expression, it's not a good idea to show it in a long shot, as the audience might miss it [2] [3].
Dealing with an interactive application, you do not have full control over the camera and what the user is looking at. In this case staging can be considered more of a guide to what is important in the animation, for example if the user is unable to zoom in completely on the ECA, there is no point in animating detailed facial expressions, and instead the overall body language should be used to express the situation.

Straight Ahead Action and Pose to Pose

This principle deals with the way the animator approaches a scene. In straight ahead action the animator starts from the first frame and just working towards the end improvising along the way. The advantage with this approach is that it has a nice flow, but it's lacking structure. In pose to pose all the key-poses are laid down first, then the in-betweens are added in the end. This approach has structure but is lacking the flow. By combining these approaches you

get the flow of the straight ahead action and the structure of the pose to pose [2] [3].
The way to utilize this principle within key-frame computer animation is to go through the scene focusing on one part of the body at a time, starting with the root of the joint hierarchy. The body is keyed before the arms, the arms before the hands and so on. *"Instead of animating one complete pose to another, one transformation is animated at time, starting with the trunk of the hierarchical tree structure, working transformation by transformation down the branches to the end."* [3, p. 40].

Follow Through and Overlapping Action

Follow through deals with the fact that nothing in nature comes to a complete stop all at once. If a person is running and then suddenly stops, loose clothes and appendages will keep moving for a bit before settling.
Overlapping action means that one action should not be finished completely before starting the next action. They should overlap and seem as one continuous action. *"It is not necessary for an animator to take a character to one point, complete that action completely, and then turn to the following action as if he had never given it a thought until after completing the first action. When a character knows what he is going to do he doesn't have to stop before each individual action and think to do it. He has it planned in advance in his mind."* [3, p. 40].
Another part of this principle is called moving hold. When the character is supposed to hold a pose for some time, the character should not be stuck in the same static position for the given time. Instead the pose is pushed a bit extra for the duration, to keep the character in motion, as having consecutive still frames would ruin the illusion of life of the character [4].

Slow In and Slow Out

This principle deals with the speed which objects are travelling at the start and end of a motion. Most non mechanical objects don't travel at a constant speed. Instead they slowly accelerate at the beginning of the motion and slowly decelerate at the end of the motion [2]. Looking at the arms in a walk cycle, in both ends of the motion their movement will slow down and almost pause before continuing in the opposite direction.

Arcs

Objects moving in a completely linear motion will have a stiff and mechanical feel. In nature most movements have an arcing motion [2]. An obvious example of this is the swinging motion of the arms when walking. The arms will rotate around the shoulder joint creating an arc.

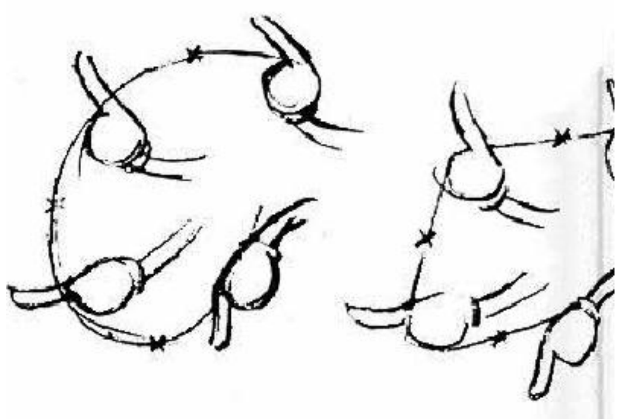

Figure 7 – Left side shows an arcing motion, right side shows a linear motion from [2]

Secondary Action

A secondary action is used to fortify the primary action. It should not dominate the primary action, but merely add an extra bit of realism to the action. An example could be a sad person whipping a tear away, the sad expression of the person would then be the primary action, and the whipping of the tear the secondary action. The secondary action would not work by itself, but in this case adds a bit of feeling to the scene [1] [3].
Within a game this principle can be used to give a bit more life to an animation, by having small variations in the animation, such as a tapping foot or the head turning to look around [12]. Additionally this could be used together with the findings from the emotions in body language chapter, to further enhance the expressions.

Timing

Timing is an essential part of any animation, and is controlled by the amount of frames between any two given key-poses [5]. Few frames results in a fast motion and many frames will give a slower motion. Having the correct timing is important, as the action could otherwise be misinterpreted. If one person is supposed to gently pad another, having a too fast motion would result in the pad looking like a slap, ruining the idea of the scene. However it's difficult to give a real definition of what good timing is. Something that works in one situation might not work in another. "*The only real criterion for timing is: if it works effectively on the screen it is good, if it doesn't, it isn't.*" [5, p. 12].

Exaggeration

Exaggeration is pretty straight forward, though should not necessarily be considered as distorting shapes or objects or making an action unrealistic. It should be used to push an action to be more convincing, and to make a bigger contact with people [2]. *"If a character was to be sad, make him sadder; bright, make him brighter; worried, more worried; wild, make him wilder."* [2, p. 65]. In many cartoons, such as Droopy and Tom & Jerry exaggeration is used in extreme ways to emphasize feelings and actions.

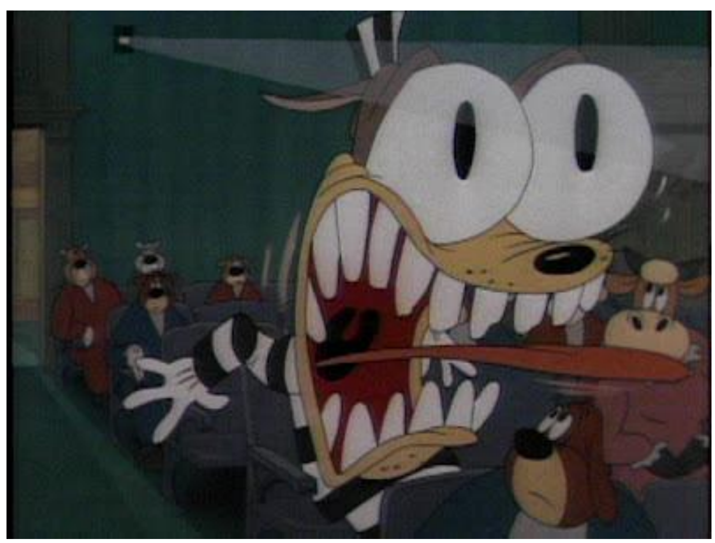

Figure 8 – Extreme example of exaggeration. A scene from Droopy.

Since a scene is composed of many different elements, such as objects, actions, sounds etc. it's important to keep the exaggeration balanced among these components, as having just one element exaggerated in an otherwise normal looking scene, will make the exaggeration stand out and look unrealistic [3]. As with the secondary actions this principle can be used to further enhance the emotional expressions of the ECA.

Solid Drawing

Solid drawing in traditional animation refers to the quality of the drawings. *"Does your drawing have weight, depth and balance?"* [2, p. 67]. Within 3D computer animation, this is still relevant, as it refers to things like lighting, texturing and rendering. Frank Thomas and Ollie Johnston describe another aspect to this principle in [2, p. 67] as "twins". This means that one side of the body is completely mirrored over to the other side, and no asymmetry is present in pose. This should be avoided, as it is not natural [2].

Appeal

Appeal is related to the style and design of anything on the screen, the characters, the actions, the poses etc. *"The word appeal is often misrepresented to suggest cuddly bunnies and soft kittens. It doesn't; it means anything that a person likes to see."* [3, p. 42]

All of the above principles were designed with cartoon animation in mind. However they are still relevant when animating for interactive applications, as most of the principles deal with things to consider while animating.

Two of the principles, namely secondary actions and exaggeration can be used to design different intensities of the animated emotions, as it was discovered that secondary actions can give an extra bit of feeling to an emotion, and exaggeration can emphasize the emotions in extreme ways.

4. Animation techniques

The two most used techniques for animation within computer graphics are motion capture and key-frame animation. These two techniques are usually used for very different purposes. Motion capture offers the possibility to have an actor play out the role of the character, by recording his movements and applying them directly to the character. This means that very expressive and realistic movements can be created in relatively short time. However more exaggerated motions cannot be achieved using this technique, as the actor has to be able to perform the action. The technique is often used within live action movies and games. Another downside to this type of animation is if the bone structure of the character is too simplistic, data might be lost in the process. For this reason, it seems obvious that motion capture should be disregarded for this study. With key-frame animation on the other hand, the animator is the actor [17]. Using this technique the movements can be as exaggerated and unrealistic as possible as only the animator's imagination limits what can be achieved. This also makes this type of animation the preferred choice for making cartoons. It also makes it the better choice for this study, as it provides more control over the animated poses and expressions.

An important part of animation is the preparation process before anything is animated. According to the 50/50 rule 50% of animation is preparation and the other 50% is actual animation [18]. Before animating anything, it's important to know exactly how the scenario of the animation is going to play out. This is usually done using sketches and storyboards. However since this is a game, the animations are not going to play in one continuous sequence. Instead a table should be made, listing the different animation clips and loops needed for the application.

A way to get more natural and convincing movement in the animated character when using key-frame animation is by utilizing the technique called rotoscoping. This is done by animating according to a reference video of an actor acting out the actions that the character is supposed to do [19]. The reference video allows for the action to be studied frame by frame, making it easier to get the details of the whole body into the animation, although it's a good idea to not copy the movements exactly. If the character doesn't resemble the actor it can give the animation the wrong sense of weight and balance. Furthermore the actors acting skills plays a big role in how natural the movements will be. The best way to utilize rotoscoping is to only use the reference video for the key poses, to get the overall motion and timing, and then improvise to get the last level of details [6].

Before the character can be animated it has to be rigged. Rigging refers to setting up a skeleton for the model, and adding controls for the skeleton, so that it can be animated. Two different ways exists to manipulate the skeleton, Forward Kinematics (FK) and Inverse Kinematics (IK). *"Forward Kinematics (FK) and Inverse Kinematics (IK) are two manipulation techniques used in animation to control the bending of joints in a chain."* [10, p. 139] FK is the simplest way to manipulate the skeleton, by rotating the joints. Animating using FK is done by starting at the root joint and rotate each joint down the skeleton chain. IK constrains the joints to only bend in a single direction, which is controlled by a pole vector. An IK handle positioned on the last joint in the chain, is translated, causing the joints higher up in the chain to rotate [20].

As FK gives the possibility to rotate the joints individually, it makes it easier to create swinging motions, such as the arms swinging down the side of the body when walking. It is however difficult to target the end of the limb in 3D space, to for example have the feet stick to the ground, or grab an object with hand. For these cases IK is the preferred choice, as only the IK handle has to be moved, which saves a lot of time.

IV. THEORETICAL FRAMEWORK

Based on the knowledge gained from the previously reviewed literature, ways of expressing emotions in animations when constrained by the mobile platform can thereby be proposed.

As the target platform for the ECA was chosen to be a mobile platform, the model has to be low poly with a simple skeleton setup. Ways of lowering the poly count of the model and keeping the amount of bones in the rig low will be discussed in the design chapter.

Before turning to the design of the model and the rig, it's important to know which animations have to be implemented. This way, all unnecessary details can be eliminated and only the required bones will be added to the rig. It has been decided that three different emotions will be used for this application, these being:

- Angry
- Sad
- Happy

From the analysis it was discovered that the principles of secondary actions and exaggeration could be used to create different intensities of emotions, three different intensities of each emotion can then be proposed, to see which way is the most efficient to portray emotions, when only using body language.

- The first intensity will focus on poses, with very subtle animations.
- The second intensity will be the same as the first, though with the addition of secondary actions.
- The third intensity will use more exaggerated animations.

Furthermore, to keep it simple only idle and walk animations will be implemented. In table 2 the chosen expressions that will be used for the different intensities of emotions can be seen. These expressions have been decided upon based on the findings from the emotions in body language section.

	Idle	Walk
Angry1	Lean forward, head straight, hands on the hips, in the shape of fists	Lean forward, determined steps, hands in fists
Angry2	Shake fist in front of body	Shake fist in front of body
Angry3	Kick the air	Stomping steps
Sad1	Head hanging, arms hanging, body bent forward	Head hanging, arms hanging, body bent forward, slow walk
Sad2	Hands cover eyes (cry)	Hands cover eyes (cry)
Sad3	On the knees, hands cover eyes (cry)	Same as Sad1, but everything exaggerated a bit more
Happy1	Chest puffed out, straight body, one arm in the side and the other hanging down	Relaxed body, swinging arms, some movement in the head
Happy2	Arms a bit up in a cheering motion	Arms a bit up in a cheering motion
Happy3	Jump with arms up (cheer)	Jumping walk

Table 2 – The different animations

By combining various relevant theories from the reviewed literature, three different ideas of how to express emotions, when taking the constraints imposed by the mobile platform into consideration, have been identified. It should thereby be possible to answer the problem statement by implementing these expressions and by performing a test, see if the test participants are able to perceive the emotions the way they were intended.

V. DESIGN

As earlier mentioned the target application of the ECA would be supplied by another master's thesis project, however since that project has a different focus, only the actual environment, which is in the form of a classroom, will be used. This means that a few basic controls should be created, and a choice has to be made regarding camera settings.

As mentioned, the distance from the camera to the ECA is essential for the amount of detail the ECA model should have. As the focus should be on the body and not the face, the camera will be placed above the ECA though tilted a bit to allow the users to see the body. Furthermore the camera should be placed as close to the ECA as possible while still keeping the face out of focus.

To help in achieving natural looking animations, several different reference videos will be used as inspiration. Mainly videos found on http://www.referencereference.com and http://endlessreference.com, as these websites offer a lot of videos of different people walking and doing other basic movements, which will be very helpful in understanding how the body moves during different actions. Knowing what kind of animations to create the attention can then be turned to designing the ECA. The bone structure has to be kept really simple to not exceed the 30 bone limit suggested for the mobile platform from table 1 in the preliminary analysis. The areas that require the most bones would usually be the head and the hands. But since this study is disregarding facial animation completely, only one bone will be left in the head for basic head rotation. A way to lower the amount of bones for the hands is to setup one bone chain for the thumb, and one bone chain to control the rest of the fingers. This will eliminate individual finger control, but will still give enough control for the hand to go from open hand to a

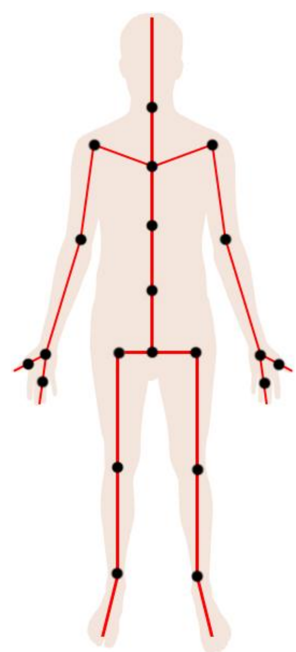

Figure 9 – Bone structure for the ECA.
Red lines are the bones.
Black dots are the joints.

closed fist. And looking at the table of the animations to create, from the theoretical framework, any additional hand gestures are not relevant. With this in mind, it's possible to create a rig containing only 27 bones, as seen on figure 9, which will still provide the desired controls.

As earlier established the poly count of the model should be kept below 1500. A few ways of reducing the amount of polygons will be discussed in the following section.

When creating a round or cylindrical shape, such as arms and legs, it can be sufficient to use a hexagonal shape or even in some cases a rhombus, reducing the poly count substantially.

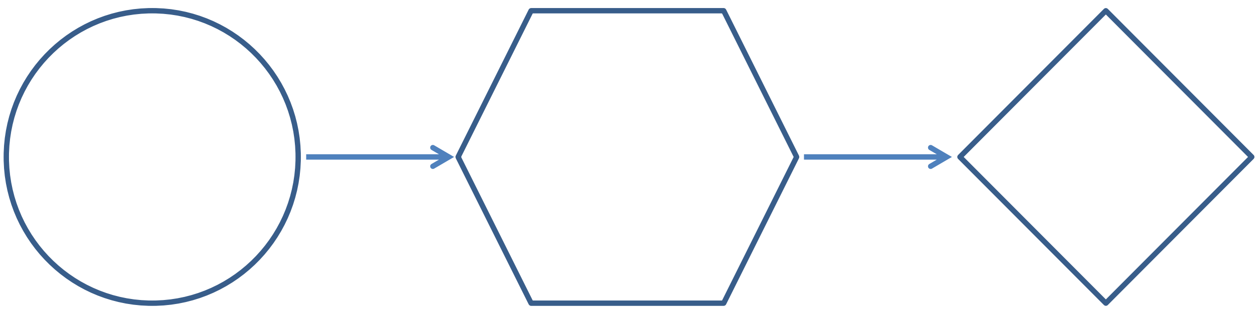

Figure 10 – Choices for optimizing the poly count of a round shape.

To keep the poly count low and not have it affect the animations it seems like a good idea to look into how the vertices should be placed around the joints. Looking at for example the knee joint, the simplest way is to have just a single edge loop around the joint, however as seen on figure 11 this will cause a collapse in the mesh, when the joint is bent. By adding an extra edge loop on both sides of the existing loop the mesh will keep its volume while

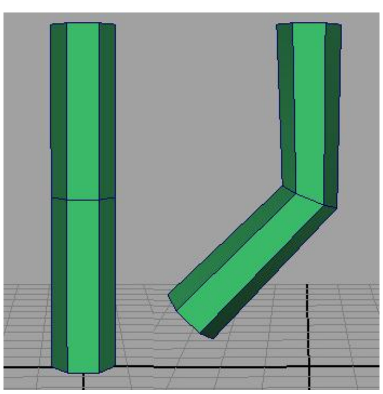

Figure 11 – Single edge loop, collapse

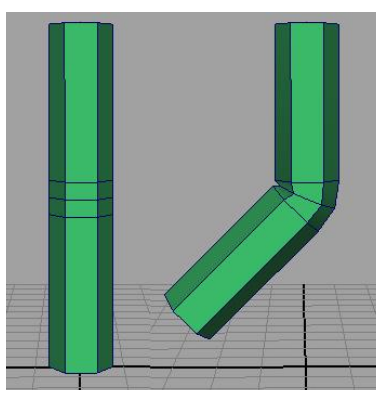

Figure 12 – Triple edge loop, collapse

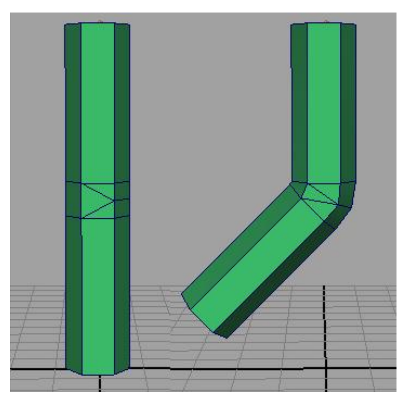

Figure 13 – Triple edge loop simplified, no collapse

bending, however from figure 12 it can be seen that there still is a small collapse on the back side of the joint. This can be further optimized by removing the backside of the middle edge loop, thereby providing a non-collapsing bend using fewer vertices as seen on figure 13. The same approach can be used around several other joints such as the elbow and finger joints.

As discussed the hands will be rigged using only two bone chains each, instead of one for each finger, therefore there is no point in modeling all the fingers, as they cannot move individually from each other. Instead only modeling the thumb and then have the rest of the fingers as one, will save a lot on the poly count. Another area to think about is the head. Since no facial animation will be created, the head will not need a lot of details. Just getting the basic shape of the head, and adding the details through textures should be fine.

Since the camera will be at some distance from the ECA, a poly count of 1000 triangles should be sufficient for the model, as it is not necessary to create a lot of details in the model, although it will still allow for the extra resolution around the joints.

As the controls from The Sims seems very intuitive, the same controls will be implemented in this application, meaning that the character will be controlled by tapping on the floor making the character walk to the tapped location, the camera should be moved by dragging a finger across the screen and rotation of the camera will be done by moving two fingers in a circular motion on the screen. Since the camera distance and height has already been determined, zooming will be disregarded.

VI. IMPLEMENTATION

The program Autodesk Maya was used to create the model and the animations and Photoshop was used to paint the textures for the model. Furthermore the game engine Unity 3D was used to implement the application.

The first step of the implementation was to create the model of ECA. The target poly count was decided to be 1000 triangles.

The hand was modeled from a cube, by only shaping the basic outline of the hand. From here the technique described in the design chapter regarding the vertex placement around the joints, was used for the finger joints, to allow the fingers to bend without the mesh collapsing.

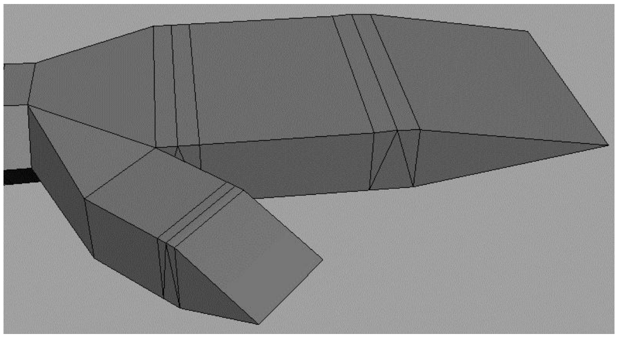

Figure 14 – 76 triangles were used for the hand.

To keep the poly count low for the head, only the basic shape was created, with a few polygons spend to create the nose and mouth. The hair

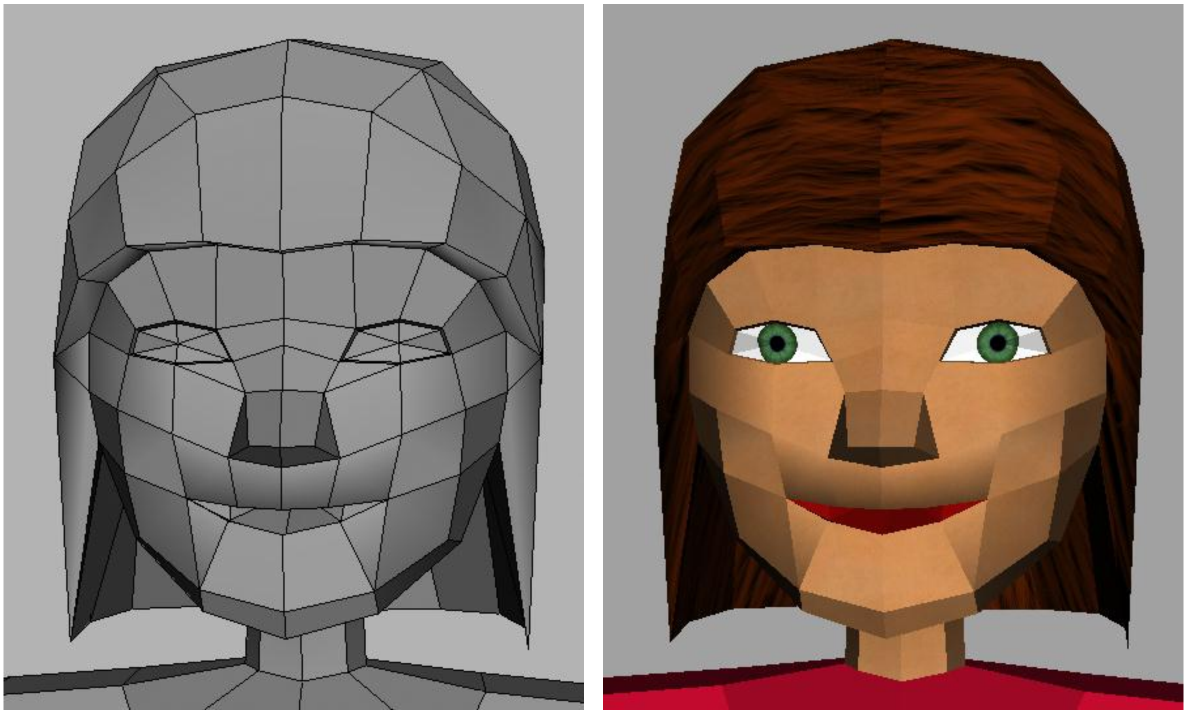

Figure 15 – 336 triangles were used for the head, and additional details such as hair and eyes were done using textures.

was shaped to cover the ears, so no extra polygons had to go into modeling these. Normally the eyes would be modeled as separate object, which allows them to be animated to look around. However since they were not supposed to be animated in this case, a small indentation was made in the model to outline the position and shape of the eyes. And the eyes were then painted on using a texture.

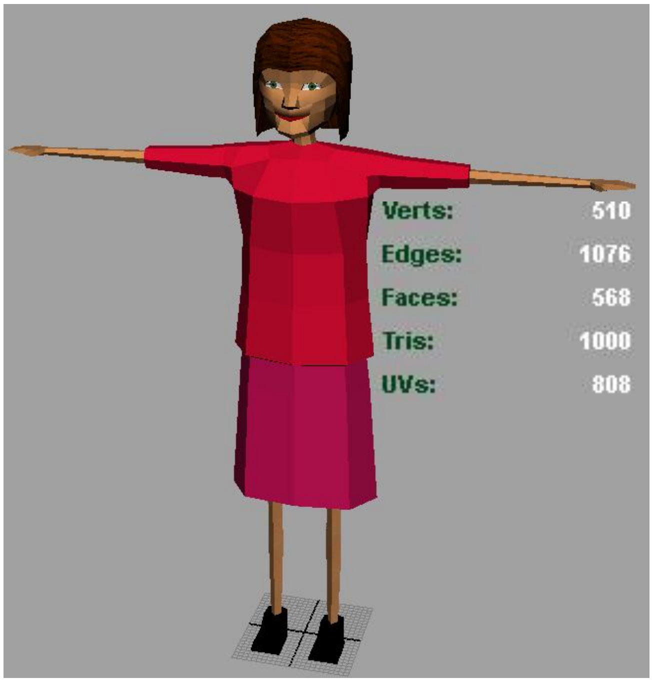

Figure 16 – 1000 triangles were used for the final model.

The model was created in a T-Pose, to make the rigging process easier [7]. By utilizing the different techniques described in the design chapter, the final model ended up with just over the 1000 poly limit. But by going through the mesh and deleting unnecessary polygons, it was possible to get it down to exactly 1000 triangles.

With the model done, the next step was to start the rigging process. All the bones were laid out according to figure 9 of the design chapter. Objects of different shapes and colors were created to control the attributes of the bones and positioned in a way as to indicate which bones they were controlling (see figure 17).

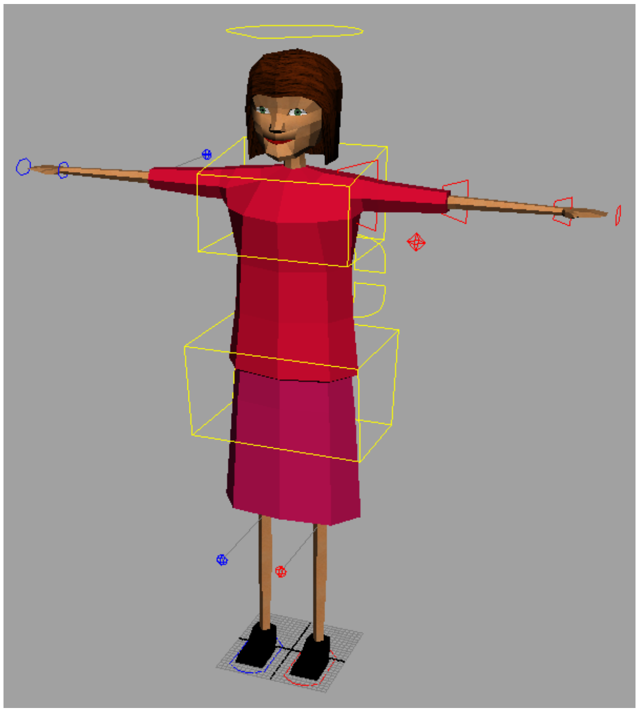
Figure 17 – The final rig with control objects.

The legs were set up with Inverse Kinematics, and for the arms a switch was created to allow the possibility to switch between Forward- and Inverse Kinematics. Figure 17 shows the left arm set to Forward Kinematics, with controls for each individual joint, and the right arm set to Inverse Kinematics, with only a control for the IK handle and for the pole vector controlling the direction of the elbow.

Having finished the rig, the attention can then be turned towards the actual animation. Since nine different variations of the walk cycle had to be produced, a neutral walk was first created, providing a base to build upon. This neutral walk could then be tweaked and adjusted with the different expressions from table 2 of the theoretical framework to reach the different intensities of emotions, thereby speeding up the process of creating the walk cycles, as the different variations of walks shouldn't be created from scratch each time. Additionally a high intensity emotion could then be created from a lower intensity of the same emotion. An example of this can be seen on figures 18-20.

Figure 18 – The 5 poses of the neutral walk cycle.

Figure 19 – The 5 poses of the sad1 walk cycle, created from neutral walk cycle.

Figure 20 – The 5 poses of the sad3 walk cycle, created from sad1 walk cycle.

The idle animations on the other hand were created from the default T-Pose, as this can be considered pretty close to an idle stance.

Each animation was then exported into individual files separate from the model, with the naming convention 'model name'@'animation name'.fbx, in this way when importing the model and animations into Unity 3D the program automatically finds and assigns all of the animations to the model [22]. Another reason for using this method is that it makes it a lot easier to go in and adjust an animation, if something needs to be corrected.

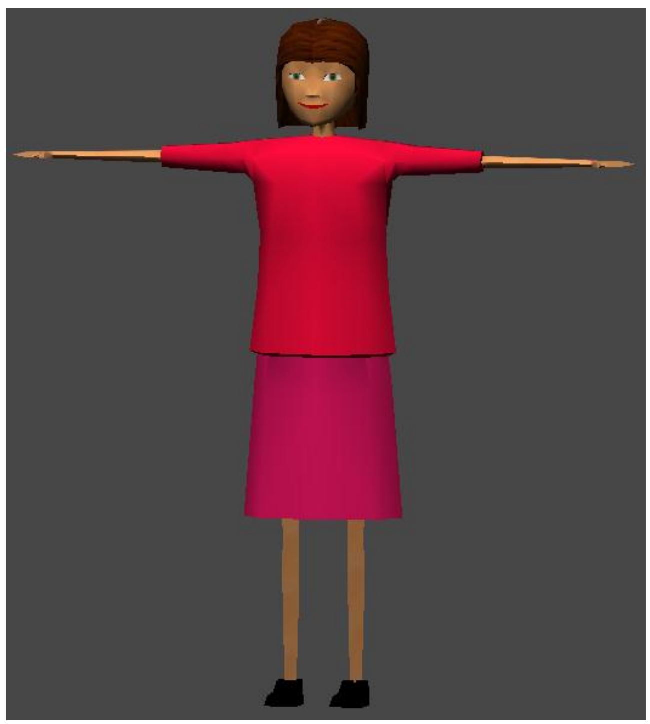

Figure 21 – The model in Unity 3D after smoothing the normals.

To get rid of the rough edges and make the model look smoother, Unity 3D has an option to calculate the normals for the mesh, and applying a smoothing angle to this calculation. The result of this can be seen on figure 21.

The final application was built for an Android phone and the application's installation file can be found on the attached CD, in the folder Android Build.

VII. METHOD OF RESEARCH

To figure out how to express emotions through body language, the different intensity levels should be tested to see if it’s possible to identify the intended emotions.

The three different intensities of emotions were decided to be

1. Subtle animations, focus on poses
2. Subtle animations, focus on poses, secondary actions
3. Exaggerated animations

Since the third intensity level should be the most expressive, it seems like a good idea to start with the lowest intensity. To keep the focus on the animations, no story has been designed.

The test will be carried out like an interview, were the test participants will be asked to think out loud.

The participants will be asked to guess the emotion of the character, and explain the reason for their guess. They will not be told if they have the correct answer or not, as this might influence the rest of the test. Furthermore they will not be given any answers to choose from, as it is important to get the participants own interpretations of the expressions. All the comments will be written down during the test, as these might help getting a better understanding of the results afterwards.

Furthermore a short introduction will be given to the test participants, explaining the controls of the application.

The test can be done in a couple of ways. One way is to show the same emotion three times with the different intensities, for example happy1, happy2 and happy3. This would make it easy to see if the subjects would find it easier or more difficult to identify the emotions as the intensity increases. However there is a chance that they might see a pattern in the order of animations, which could influence the results. Another way of testing would then be to change the emotion as the intensity increases, for example to show angry1, sad2 and happy3, giving it a more random order, which seems like a better choice.

To test all nine intensities of emotions, three different tests with pre decided orders can be made.

- Test 1: angry1, sad2, happy3
- Test 2: sad1, happy2, angry3
- Test 3: happy1, angry2, sad3

Each test participant will then get all of these tests. However the results could end up being biased by the fact that in the later tests the subjects will already have seen the emotions before, therefore the order of the tests should change so that the first person will start from test 1, the second will start from test 2 and so on, giving three different test cases. This will give two ways of interpreting the results, one being to simply look at all the answers given by the subject, the other to only look at the first test presented to the subjects, as in this case they will not previously have seen the emotions. It also means that the three test cases should be tested an equal amount of times.

Figure 22 – Screenshot from the test application.

The test will be carried out at Aalborg University with students from Medialogy.

As each test has three different emotions, three question marks have been added to the application, as shown on figure 22, making it possible to switch between the emotions. The question mark at the top represents the lowest intensity, and the bottom question mark represents the highest intensity.

VIII. RESULTS

In this study the interest was set on figuring out how to express emotions in an ECA, when constrained by a mobile platform.

In order to determine how the different intensities of emotions were interpreted by the participants, the different answers will be counted for each of the different intensities of emotions. When counting the perceived emotions, different degrees of an emotion will be overlooked, meaning that for example glad, happy, euphoric & ecstatic will all be counted as being a happy emotion.

In total 24 test participants were used, 8 for each test case.

The rows show the different emotions tested for, and the columns show the emotions guessed by the test subjects, so for example happy with an intensity of 2 was perceived to be happy by 18 of the 24 subjects.

	Angry	Sad	Happy	Annoyed	Complacent	Neutral
Angry1	10	0	1	0	0	13
Angry2	24	0	0	0	0	0
Angry3	21	0	1	2	0	0
Sad1	0	22	0	0	0	2
Sad2	0	24	0	0	0	0
Sad3	0	24	0	0	0	0
Happy1	1	0	0	1	4	18
Happy2	0	0	18	2	2	2
Happy3	0	0	24	0	0	0

Table 3 – Sum of observations.

To get a better overview of the results, these values will be visualized through a bar chart.

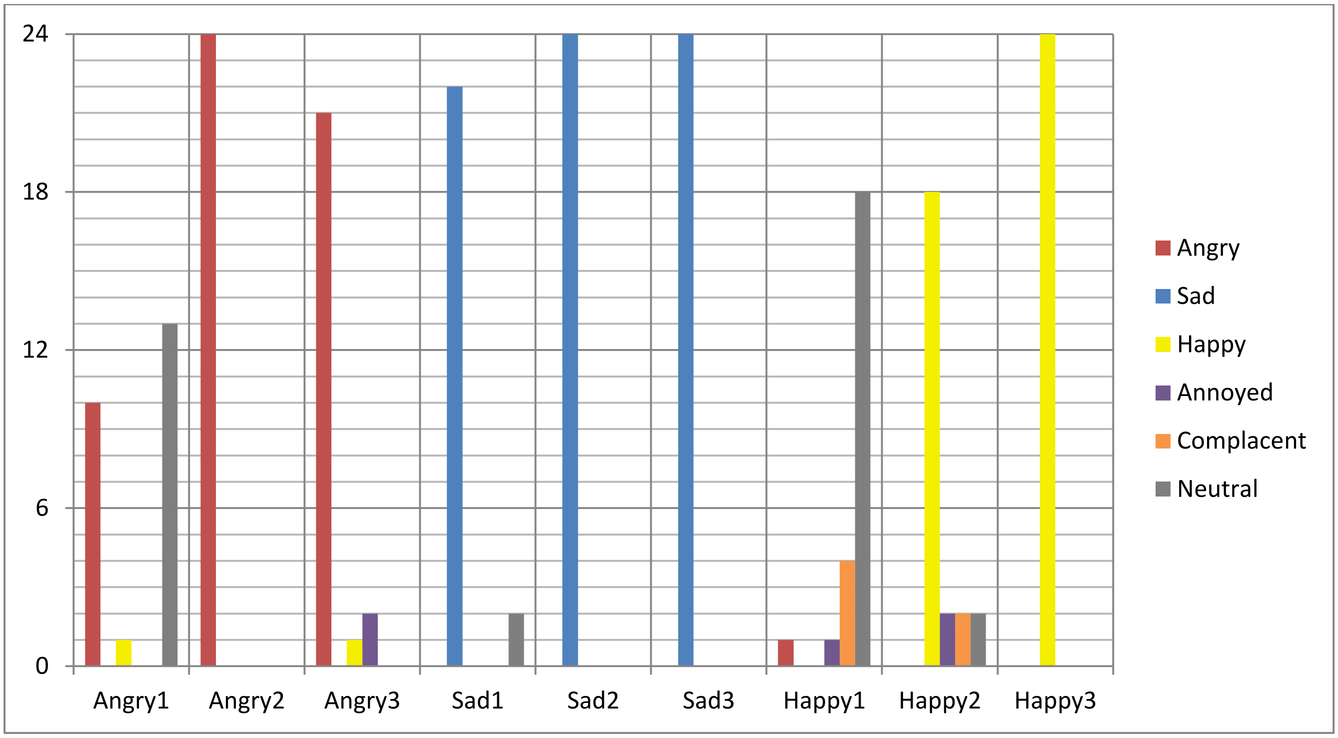

Figure 23 – Results from all the tests

The results show that the third intensity of the different emotions was easy for most of the subjects to identify. The same goes for the second intensity although more subjects could identify the second intensity of the angry emotion, and a small drop can be seen for the happy emotion. The first intensity however stands out for the angry and happy emotions, as less than half of the test participants were able to identify the angry emotion, and no one could see the happy emotion. Instead these two animations were interpreted to be neutral, with no indication of emotions.

As earlier stated some of the results could have been biased by the fact that the subjects had already seen some animations presenting the same emotions of different intensities. Therefore a second graph should be made where only the results, of the first of the tests each subject went through, will be plotted.

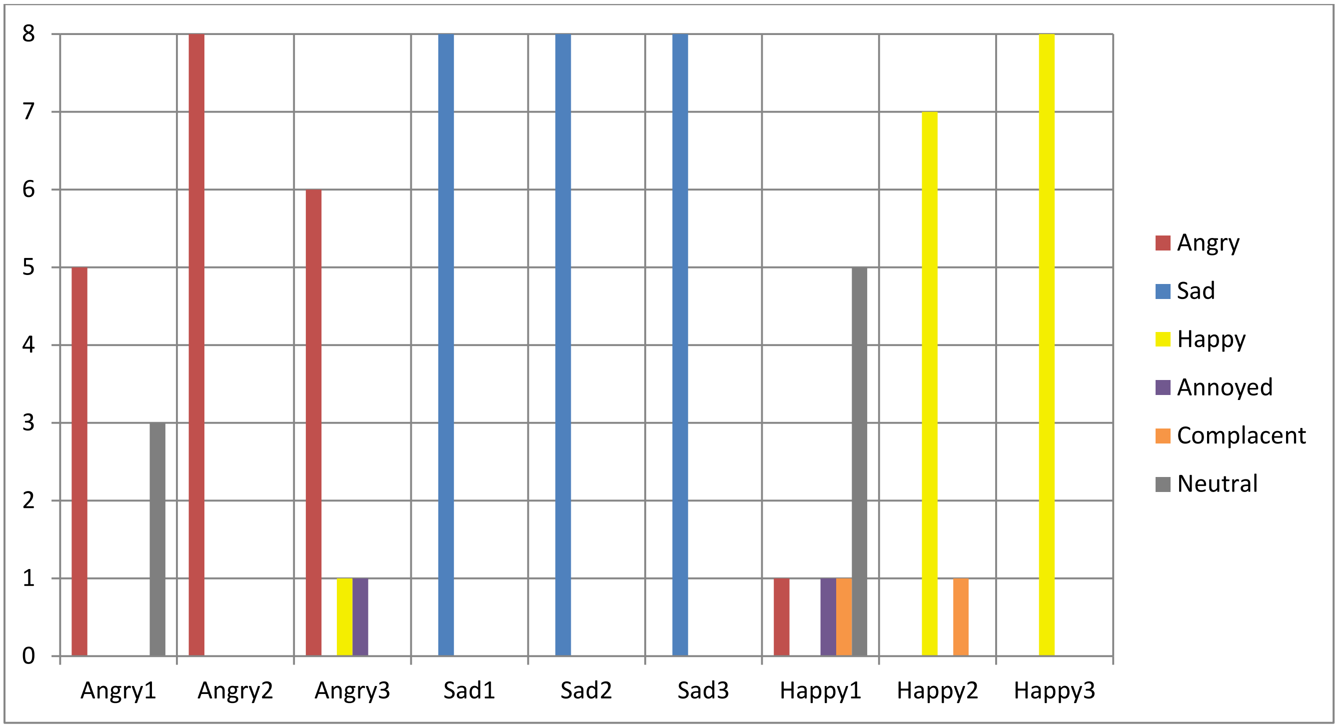

Figure 24 – Results from the first tests only

To better compare the results, the percentage of answers perceived as intended from the two graphs are shown in the below table.

	Angry1	Angry2	Angry3	Sad1	Sad2	Sad3	Happy1	Happy2	Happy3
All	41,67%	100%	87,5%	91,67%	100%	100%	0%	75%	100%
First	62,5%	100%	75%	100%	100%	100%	0%	87,5%	100%

Table 4 – Percentage of answers perceived as intended from the two previous graphs.

The results show a difference in 4 out of the 9 emotions, where angry1 has the biggest change of 20,83 percentage points.

The presented results will be discussed in the following chapter.

IX. DISCUSSION

With the first intensity of the different emotions, most of the test subjects were struggling to come up with any ideas as to what the portrayed emotions could be and these were also the ones where the subjects took the most time in giving an answer. It was observed that many were trying to see the face of the character to get a better indication of the emotion. Even with the first sad emotion, that most of people were able to identify, some were at first saying that the ECA looked tired. Looking at the results, it can be argued that this level of intensity was not expressive enough.

As the second intensity basically was the same as the first with the addition of the secondary actions, it seems obvious that more people would be able to identify the intended emotions. Everyone guessed the angry and sad emotions, which could indicate that this is a good way of expressing the emotions. However some of the test subjects were confused by the arm movement of the happy emotion. A few said that their first instinct was to say that the ECA was happy, but because the cheering motion of the arms was present all the time, it made the ECA look more annoyed. Others suggested that the arm movement made it look like the ECA was telling a story and gesticulating with the arms. When using secondary actions to strengthen the emotions, it should probably be done in a more subtle way. Additionally a deeper research into selecting the correct secondary action might have given an even better result.

The highest intensity of the emotions gave the best overall results, angry was the only one not receiving a perfect score, and again the choice of action seems like the decisive factor. The subjects not guessing the angry emotion, all agreed that the walk was very aggressive, however the kicking motion of the idle animation made them guess something else. This level of intensity in the emotions doesn't seem like the best way of expressing emotions unless for extreme cases, as some of the comments given were:

Happy: *"She just won the lottery."*, *"She is celebrating."*

Angry: *"She is gonna kill someone."*, *"She is walking like she is going to a fight."*

Sad: *“Her world has broken down.”*, *“Someone close to her died.”*

As it was suggested that seeing the same emotions multiple times could have influenced the results, the answers from only the first tests viewed by the subjects were gathered. Since the results show a difference of up to 20,83 percentage points in 4 of the 9 emotional states, it can be argued that it might have had some impact on the results. Since 6 of the 8 participants who saw the lowest intensity of the angry emotion first, were able to identify the intended emotion, and from the total 24 participants this number was 10. This means that from the remaining 16 subjects who had seen a higher intensity of angry before seeing the lowest, only 2 thought it was angry, suggesting that the results were influenced by the subjects having previously seen higher intensities of the angry emotion. However for the remaining 3 emotions with percentage changes, the difference was too small to indicate any bias in the results.

Additionally it can be argued that the level of expressiveness in the chosen emotions was a factor, since almost all of the subjects perceived the intended emotion for all three intensities of sad.

X. CONCLUSION

The primary interest throughout this project was to find a way to express emotions in an embedded conversational agent, when the ECA was to be used in an interactive application on a mobile platform. The problem stated:

"How can we overcome the constraints imposed by mobile platforms when expressing emotions in animated embodied conversational agents?"

In order to answer the problem statement, research within the fields of reading body language and animation had to be conducted, with the purpose of identifying ways to express emotions in animations when taking the constraints imposed by the mobile platform into consideration and to get a fundamental understanding of how to approach animation. The emotions angry, sad and happy were chosen to be used in the research and three different intensities of these emotions were decided to be implemented. A model for the ECA was then designed and implemented to fit the suggested requirements of the mobile platform while still being able to perform the intended emotional expressions. A test of the animations was carried out to see if the intended emotions were perceived by the test subjects. The testing proved that it is possible to achieve the desired emotions by using subtle animations with the addition of secondary actions. The test subjects did perceive emotions in the ECA, although some did not perceive the intended emotions, but through a more extensive research of body language, expressions more specific to the chosen emotions could have been found. By exaggerating the expressions, the emotions can be further enhanced. However this might result in the emotions becoming too extreme.

Therefore it can be concluded, that it is possible to express emotions through body language for an ECA with a limited amount of polygons and bones, and without the use of facial expressions.

XI. FUTURE PERSPECTIVES

As learned from the discussion and conclusion the main issue was with the chosen emotional expressions. Therefore this would be the main aspect to look further into in the later development phase.

Furthermore it could be interesting to add additional non controllable ECAs into the equation, as this could give the emotions an aspect of relations between the ECAs. Additionally an idea could be to widen the collection of chosen emotions, to see if the same principles can be applied to other emotions.

XII. BIBLIOGRAPHY

[1] Petta P., Pelachaud C. & Cowie R., *Emotion-Oriented Systems: The Humaine Handbook,* Springer, 2011

[2] Thomas F. & Ollie J., *The Illusion of Life Disney Animation*, Walt Disney Production, 1981

[3] Lasetter J., *Principles of traditional animation applied to 3D computer animation*, Computer Graphics (Volume 21 Number 4), 1987

[4] Lew J., *Learning 3D Character Animation* , http://www.jefflew.com/l3dca2.html [Video], 2009

[5] Whitaker H. & Halas J., *Timing for Animation*, Focal Press, 2002

[6] Ratner P., *3-D Human Modeling and Animation Second Edition*, Wiley, 2003

[7] McDermott W., *Creating 3D Game Art for the iPhone with Unity,* Focal Press, 2011

[8] Williams R., *The Animators Survival Kit*, Faber & Faber, 2002

[9] *Modeling Characters for Optimal Performance*, http://unity3d.com/support/documentation/Manual/Modeling%20Optimized%20Characters.html [Online], 2011

[10] McKinley M., *The Game Artist's Guide to Maya*, Sybex, 2005

[11] Athias D., Creating Walk Cycles in Maya 2011, http://www.digitaltutors.com/11/training.php?pid=354 [Video], 2010

[12] Steed P., *Animating Real-Time Game Characters,* Charles River Media, 2002

[13] Goman C. K., *The Nonverbal Advantage*, Berrett-Koehler, 2008

[14] Harley G. & Karinch M., *I Can Read You Like a Book*, Career Press, 2007

[15] Ekman P., *Universal Facial Expressions of Emotion*, California Mental Health Research Digest (Volume 8 Number 4), 1970

[16] Coulson M., *Attributing Emotion To Static Body Postures: Recognition Accuracy, Confusions, and Viewpoint Dependence*, Journal of Nonverbal Behavior (Volume 28 Number 2), 2004

[17] Orwall B., *What's the Buzz? `Toy Story' director John Lasseter talks about story and character -- two things that Movie makers often forget these days.*, Wall Street Journal (Eastern Edition) p. R17, 1998 (March 19)

[18] Maestri G., *[digital] Character Animation 2 Volume II: Advanced Techniques*, New Riders, 2002

[19] Lew J., *Animating Power and Weight*, http://www.jefflew.com/apw1.html [Video], 2009

[20] Maraffi C., *Maya Character Creation: Modeling and Animation Controls*, New Riders Publishing, 2003

[21] Mancini M., Hatmann B. & Pelachaud C., *Non-verbal behaviors expressivity and their representation*, 2004

[22] *Character Animation*, http://unity3d.com/support/documentation/Manual/Character-Animation.html, [Online], 2011

[23] Navarro J. & Karlins M., *What Every BODY is Saying: An Ex-FBI Agent's Guide to Speed-Reading People*, William Morrow Paperbacks, 2008