A COMPARISON OF INTERPOLATION METHODS FOR VIRTUAL CHARACTER UPPER BODY ANIMATION

by

Xingyu Lei

A Thesis

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Master of Science

Department of Computer Graphics Technology

West Lafayette, Indiana

December 2020

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Prof. Nicoletta Adamo, Chair

Department of Computer Graphics Technology

Dr. Bedrich Benes

Department of Computer Graphics Technology

Dr. Tim McGraw

Department of Computer Graphics Technology

Approved by:

Prof. Nicoletta Adamo

To my caring and supportive parents

and to all my loved ones

ACKNOWLEDGMENTS

I would like to express my most sincere gratitude to my advisor and committee chair, Prof. Nicoletta Adamo. She accepted me into graduate school, gave me the opportunity to participate in her research and provided me with mentorship on my academic decisions. I could not have imagined a better advisor for my graduate study.

My heartfelt thanks go to the other committee members, Dr. Bedrich Benes, and Dr. Tim McGraw, for their insights and supervision. They also taught me the fundamental aspects of computer graphics, which inspired my vision of this thesis. Without their help, this thesis would not have been possible.

I want to pay special tribute to my consultant, Ashish, and director, Dr. Arman Sabbaghi, at Purdue Statistical Consulting Service. They contributed in a significant way to the experimental design and data analysis of my thesis.

To Junzhe, my classmate and long-time friend—I appreciate your feedback. To Xiao—I appreciate your mental support. To all my other family and friends—thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The realistic animation of virtual characters can enhance user experience. Motion-editing methods such as keyframing and motion capture are effective for pre-determined animations but are incapable of real-time generation. Algorithm-based dynamic simulation and machine learning-based motion synthesis are procedural but too complex. This thesis explores an approach known as animation interpolation, which benefits from the strengths of both types of methods. Animation interpolation generates full animation sequences by assembling pre-defined motion primitives or key poses in real-time.

The purpose of this thesis is to evaluate the naturalness of character animation in three common interpolation methods: linear Euler interpolation, spherical linear quaternion interpolation, and spherical spline quaternion interpolation. Many researchers have studied the mathematical equations, motion curves, and velocity graphs of these algorithms. This thesis focuses on the perceptual evaluation and the implementation of expressive upper body character animation.

During the experimental studies, 97 participants watched 12 animation clips of a character performing four different upper body motions using three interpolation methods. The motions were based on McNeill's classification of body gestures (beat gesture, deictic gesture, iconic gesture, and metaphoric gesture). After viewing each clip, the participants rated the naturalness on a 5-point Likert scale. The results showed that animations generated using spherical spline quaternion interpolation were perceived as significantly more natural than those generated from the other two interpolation methods.

1. INTRODUCTION

Using a virtual character in human-computer interaction scenarios can enhance user experience, according to Matsiola et al. (2005). The ability to express and react to emotions is an essential element of human social interaction. Therefore, it is crucial that virtual characters also have the ability to express and react to emotions (Torre et al., 2019).

The following scenario often occurs in certain genres of video games. When a player with weapon charges at a non-player character, the non-player-character should appear to feel threatened and change from being passive to being aggressive. This change is often achieved by activating a ready-to-attack animation sequence. This example shows that virtual characters require two features: realistic and natural movement (the ability to express emotion) and real-time response (the ability to react). The example also showcases the common game mechanism known as state machine-based animation blending.

1.1 Problem Statement

Although the technology behind games and films is developing rapidly, creating believable character animation in real-time remains challenging. Keyframe animation is a traditional way to generate animation by defining poses at crucial moments. However, the creation, assembly, and control of the animation require great effort and expertise (Heloir & Kipp, 2009). Even a few seconds of animation requires hours of fine-tuning, and the final quality of the animation depends on the animator’s skill (Johansen, 2019).

Motion capture converts a real actor’s performance to virtual character animation. This approach has become more affordable and accurate, with improvements in optical hardware and motion sensors. However, all of the actor’s movements need to be pre-defined and cannot reflect real-time interaction. Today, animation applications increasingly require on-the-fly adaption (Van Welbergen et al., 2014). For example, animation for real-time interactive media cannot be planned (Horswill, 2009). Finally, dynamic simulation methods generate physically correct body reactions; however, even with high-performance computers, full-body dynamics are still very complex to handle (Mezger et al., 2005).

Animation interpolation is a semi-procedural solution. It generates realistic movement algorithmically by interpolating using a limited number of pre-recorded animation segments (Mezger et al., 2005). Popular interpolation algorithms have been well studied in terms of computation complexity, interpolation curve paths, and velocity graphs. Few studies, however, have examined the human perception of different interpolation methods in character animation (Mezger et al., 2005). This dearth of research motivated the present work. The objective of this thesis is to study the effects of different interpolation methods on the perception of naturalness in animated characters' upper body movements. The two standards for evaluating the quality of virtual character animation are the evaluation of naturalness and the objective measurement of real-time response (Wang et al., 2015).

1.2 **Significance**

In the video game industry, as the number of games increases every year, it is expensive for independent developers or artists to create animations using keyframing and motion-capture techniques (Horswill, 2009). Small to medium-sized game studios would benefit from using a procedural way of identifying motion segments and applying interpolation (Johansen, 2019). For animation or visual effects (VFX) studios, using a faster and more natural interpolation could enable artists to generate animation prototypes more quickly, saving resources and allowing more time for creative decisions.

Most importantly, animation interpolation or blending is a common game logic used in many human-computer interaction applications. Finding a more natural solution for blending between animations helps users interact with virtual characters and provides them with a more immersive experience. Furthermore, this semi-procedural approach makes the process of creating animations easier, allowing people with little knowledge of computer graphics to create their own animated stories.

1.3 **Purpose Statement and Research Questions**

The purpose of this study is to compare and evaluate, in terms of naturalness, three methods of animation interpolation: linear Euler interpolation, spherical linear quaternion interpolation, and spherical spline quaternion interpolation. The animations involve the upper body movement of 3D characters.

This thesis aims to answer the following research question: Which of the three interpolation methods (linear Euler interpolation, spherical linear quaternion interpolation, or spherical spline quaternion interpolation) creates the most natural animations as perceived by viewers (with or without computer animation experience)?

This thesis has the following assumptions: The participants were honest during the experiment and answered the survey based on their knowledge and experience, rather than based on what they assumed the researchers would like them to answer.

1.4 **Definition**

Character animation is the specialized area of animation that involves bringing to life a virtual character (either a digital representation or a hand-drawn figure). In this study, "character" refers to a creature with human-like features.

Naturalness is the state or quality of minimum artificial processing. In the context of character animation, naturalness is how realistic a character appears and behaves compared to a real human. Naturalness can also refer to how believable or acceptable a character is.

A skeletal system is a hierarchical set of joints connected by bones. The transformation of the joints drives the skinned surface mesh to produce movements.

Rotation is a type of transformation that describes an object's circular movement around a center (or a point).

Interpolation is a method of estimating intermediate values from a limited number of independent variables. In the context of computer animation, interpolation usually refers to inserting in-between frames of an animation. Given two key poses, interpolation will generate the in-between motion from one pose to the next.

Quaternion is a mathematical term for a number system that extends the complex numbers. It has practical use in applied mathematics when calculating 3D rotations. It was first described in 1843 by the mathematician William Rowan Hamilton (Dam et al., 1998), using the form:

$$q = s + ix + jy + kz \quad s, x, y, z \in \mathbb{R} \text{ where } i^2 = j^2 = k^2 = ijk = -1.$$

Euler angles are three components x, y, and z, used to describe the orientation of a rigid body with respect to the 3D coordinate system.

Rotation Matrix is a linear algebra term describing a 4×4 rectangular array used to perform a rotation in Euclidean space. The rotation of three elemental Euler angles can be represented by multiplying the corresponding rotation matrix.

1.5 Limitations and Delimitations

This thesis has the following limitations:

1. The character model has a unique visual design (art style, texture, level of detail, etc.), which may influence the viewer's judgment.
2. The skeletal joint placement and weight distribution in the character's skinning define the resulting mesh's deformation behavior, which affects the quality of the animation.
3. The quality of the actor's performance during motion capture directly affects the naturalness of the animation.
4. The quality of motion retargeting is affected by the character model's resemblance to the actor in terms of body proportions.

This thesis has the following delimitations:

1. This study focuses on animation in 3D space rather than 2D space.
2. This study evaluates the interpolation of rotation on the character's skeletal joints, equivalent to human bones. Other transformations, such as translation and scale, are not within the scope of this study.
3. This study tests three common interpolation methods: linear Euler interpolation, spherical linear quaternion interpolation, and spherical spline quaternion interpolation.
4. This study uses a character that adheres to the Humanoid Animation Standard (2006).
5. This study focuses on animation performance with an emphasis on upper body movement.

1.6 Deliverables

The outcomes of this study are:

1. A Unity-based system that generates interpolation animations for any given two motion primitives (segments). This system can use any of three interpolation methods and can adjust blending time.
2. An evaluation report of all the data collected from the participants in the interpolation evaluation experiment.

1.7 Summary

This chapter starts by introducing the use of character animation in human-computer interaction. The two essential aspects for such virtual characters are the ability to express and react, which can be translated to performing natural animation in real-time. However, the problem is that motion editing techniques for creating animation, such as keyframing and motion capture technology, although they generate natural movement, need to pre-determine each motion. Algorithmic approaches for creating animation such as dynamic simulation or machine-learning-based motion synthesis are procedural but far too complex.

Animation interpolation, a common game mechanism used in state machine-based animation blending, is a semi-procedural solution for creating natural animation in real-time. By evaluating and comparing the naturalness of three commonly used interpolation methods (linear Euler interpolation, spherical linear quaternion interpolation, and spherical spline quaternion interpolation), this thesis wishes to answer which of the method is most suitable for character animation. This study potentially could save time and resources for professional studios and independent content creators. It can also benefit the general public by enhancing the day-to-day human-computer interaction experience and by giving them easier access to create their own animated stories. Finally, this chapter also covers definitions for important terminologies, the limitation and delimitation for this thesis.

2. LITERATURE REVIEW

This chapter presents previous research in three areas: animation interpolation methods, perception of naturalness in character animation, and expressive body gestures. Section 2.1 analyzes the methods for the interpolation of rotation that can be applied to character animation. Section 2.2 discusses how a viewer perceives human interaction and how they evaluate the quality of animation. Section 2.3 discusses the types of expressive body gestures that affect human-computer interaction.

2.1 Animation Interpolation Methods

Interpolation of Rotation

Dam et al. (1998) introduced two rotation modalities: rotation represented by Euler angles or rotation matrices and rotation represented by quaternions. Their research compared these modalities in terms of strengths and weaknesses: interpolation using Euler angles are commonly used by artists and supported in content creation software; the 4×4 homogeneous matrix can represent not only rotation but also other transformations such as translation, projection, and scale. However, Euler angles or rotation matrices treat rotation independently along three axes, and the rotation order affects the interpolated curve. Moreover, the independent rotation can cause gimbal lock, as shown in Figure 2.1 (Fielding, 2007), a common artifact of Euler rotation in which the rotation loses one degree of freedom and almost always needs to be avoided. Using quaternions produce predictable results, as the rotation is performed along one axis. A quaternion is data-compact, storing only four numbers instead of nine. It is computationally efficient and has a simple composition for sequential rotations. The disadvantage is that quaternions have a complicated mathematical model that is not intuitive to everyday users.

According to Smith (2013), quaternions play a significant part in skeletal animation, both in rigid transformations and geometric skinning, such as dual quaternion linear blending. Furthermore, in most 3D editing software, interpolation is often achieved with quaternion-vector pairs, where the rotation is stored as a quaternion and the translation is stored as a vector.

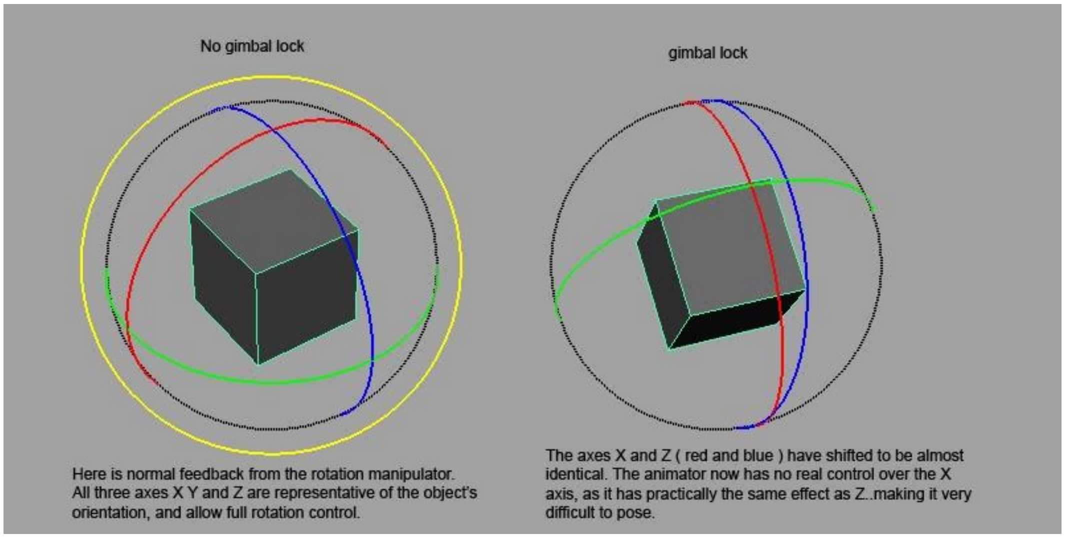


Figure 2.1. Illustration of gimbal lock caused by Euler rotation

Mathematical Model for the Three Interpolation Methods

Linear Euler interpolation (LinEuler) is an interpolation between two tuples of Euler angles. Let a point in 3D space be represented as $P = [x, y, z] \in \mathbb{R}^3$; then a vector from the world's origin to that point can be represented as $v = (x, y, z) \in \mathbb{R}^3$. Consider a unit sphere, where a vector from the center to one point on the surface is represented as $v_0 = (x_0, y_0, z_0) \in \mathbb{R}^3$ and another as $v_1 = (x_1, y_1, z_1) \in \mathbb{R}^3$. Linear Euler interpolation between these two vectors can be represented as follows, with an interpolation parameter $h \in [0, 1]$:

$$LinEuler(v_0, v_1, h) = v_0(1 - h) + v_1 h \tag{2.1}$$

A quaternion consists of a scaler part and a vector part: $q = [s, (x, y, z)] \quad s, x, y, z \in \mathbb{R}$, also written as $q = s + ix + jy + kz \quad s, x, y, z \in \mathbb{R}$ where $i^2 = j^2 = k^2 = ijk = -1$. Spherical linear quaternion interpolation (Slerp) was first introduced by Ken Shoemake (1985). It interpolates rotation along the shortest path on a unit sphere at a constant velocity, which causes a sudden change of angular direction when performing a series of rotations, making keyframes visible. Let H be sets of quaternions, where $p, q \in$ H, $cos(\Omega) = p \cdot q$, and interpolation parameter $h \in [0, 1]$, Slerp can be represented using the formula:

$$Slerp(p, q, h) = \frac{p\, sin\big((1 - h)\Omega\big) + q\, sin(h\,\Omega)}{sin(\Omega)} \tag{2.2}$$

The spherical spline quaternion interpolation (spherical and quadrangle, or Squad) is the spherical cubic equivalent of the Bezier curve. The Bezier curve is defined as a third-order curve where two auxiliary points define the tangent for each control point. A special case occurs when the tangents coincide in the control points. This results in differentiability in the control points (Dam et al., 1998). Shoemake (1987) presented Squad and then proved the continuous differentiability of Squad at control points.

$$Squad(q_i, q_{i+1}, s_i, s_{i+1}, h) = Slerp(Slerp(q_i, q_{i+1}, h), Slerp(s_i, s_{i+1}, h), 2h(1-h)) \quad (2.3)$$

where $h \in [0, 1]$ and $s_i = q_i \exp(-(\log({q_i}^{-1} q_{i+1}) + \log({q_i}^{-1} q_{i-1}))\,/4)$.

The three algorithms generate entirely different interpolation paths (Figure 2.2). Work is continuing to develop more advanced mathematical designs to create smoother quaternion spline curves (Geier, 2020). Dam et al. (1998) proposed a more elaborate method called spherical interpolation using numerical gradient descent (Spring). Barr et al. (1992) proposed a method that takes into account the initial and final angular velocity. The algorithm applied angular velocity constraints, which generated a smooth and consistent rotation path on a unit sphere. The path generated can pass through all control points.

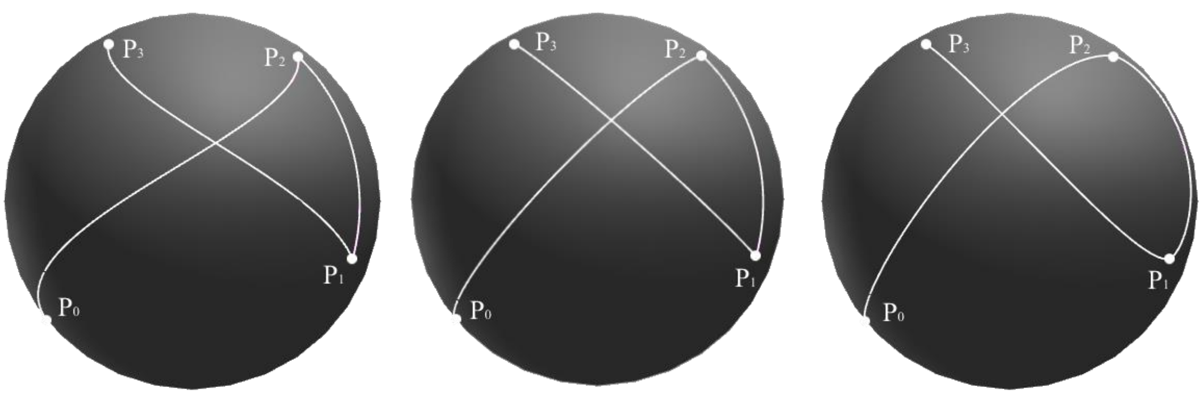


Figure 2.2. Interpolation curve on a unit sphere using LinEuler (left), Slerp (middle), and Squad (right), with no rotation along the twist axis/normal.

Interpolation Between Poses in Character Animation

Traditional hand-drawn animations are based on sequences of manual drawings. The scenes in the storyboard that define important movements are known as *keyframes* and are often drawn by experienced artists. Assistant artists with less experience, known as in-betweeners, fill the gaps between keyframes to create smooth transitions. This method, known as pose-to-pose animation, is one of the most common techniques of creating animations (Figure 2.3; Vail, 2017). In-betweeners can be replaced by computers (Dam et al., 1998), using an estimation method called interpolation. Therefore, pose-to-pose animation has excellent applications in not only traditional 2D animation but also in modern 3D animation (Lasseter, 1987).

It is critical to identify which poses can represent the key moments of the motion. *Key pose* represents a "signature" motion that is unique and extreme (So & Baciu, 2005). Artists usually have a keen sense of identifying key poses from experience. Algorithms are also effective in extracting critical motions. So and Baciu (2005) measured the difference of poses in directional movements. The poses were ranked, with key poses having a more considerable difference and neutral poses having less difference.

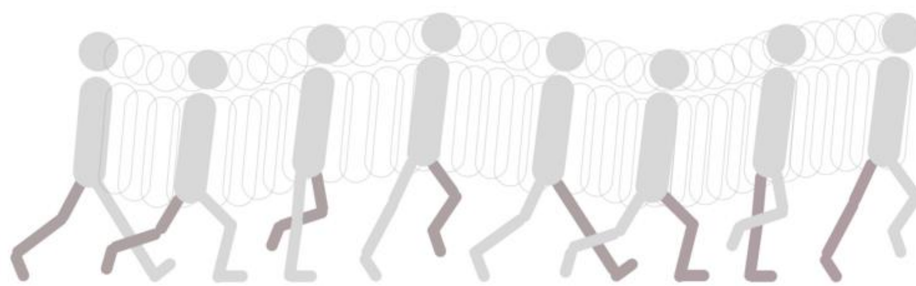


Figure 2.3. Illustration of pose-to-pose animation

The algorithm of the interpolation defines the smoothness of the transition between any two segments. In the context of 3D character animation, a sense of motion is usually created by animating a character's major skeletal joints, which drive the character's 3D mesh. The interpolation of translation has been well studied in flat 3D Euclidean space. Character animation, however, is largely achieved through joint rotation, and rotation lies in non-Euclidean space. Interpolating smoothly between rotations will become more important in the computer graphics field (Barr et al., 1992).

The interpolation path on a unit sphere can be translated into the orientation of a joint rotating from its base. Bloom et al. (2004) identify three key properties of the mechanical analysis of interpolation: the path, angular velocity, and commutativity (the order of interpolation in sequential rotation does not affect the result). Similarly, Wang et al. (2015) suggest dividing breaking down animation interpolation into three parts: blending time (total time spent traveling between two segments), path (interpolation curve of a joint), and velocity (angular velocity of a joint). These properties are similar to Dam's analysis of the interpolation path and velocity graph of a 3D unit sphere (1998).

In reality, any rigid human body is affected by internal and external forces, causing joints to rotate at a non-constant speed. Qiang and Xingfen (2012) demonstrated animation sequences of a character jumping using the same path curve but different velocity curves, one with constant speed and another with initial acceleration and final deceleration. Compared to the one with constant speed, the animation with acceleration and deceleration was more physically correct. As stated in Disney's 12 principles of animation, slow in and slow out are preferred over constant speed to avoid robot-like movement (Thomas & Johnston, 1981).

To examine the blending time factor, Wang et al. (2015) proposed an automatic variable-timing transition method that adapts to simple or complex transitions in 3D character animation. They presented a framework for generating transition animations using different interpolation methods based on different motion categories. Their system automatically adjusted to an appropriate blend time according to the offset orientation of major bones.

2.2 **Perception of Naturalness in Animation**

Even smooth interpolation with angular velocity constraints that are physically correct may appear unnatural or mechanical. The human body is an active system, which contains muscles or motors, providing internal forces that affect motion (Hodgins et al., 1992). A perfectly smooth motion, therefore, may appear "robot-like." According to Mezger et al. (2005), the study of interpolation methods for character animation "needs to be addressed by combining computer graphics and perception research" (p. 1). Moreover, they suggest that "psychophysical measures seemed to be more sensitive and appropriate for quantifying slight quality differences between animation techniques than the tested physical criterion" (p. 8).

Interpolation of rotation is often analyzed at isolated joints rather than the full-body, but full-body involvement is crucial for the naturalness of motion (Jansen & Van Welbergen, 2009). Even though user studies are invaluable for measuring naturalness in character motion, most naturalness metrics do not take into account human observation (Van Welbergen et al., 2010).

Perception of Human Motion

According to Blake and Shiffrar (2007), both the form and the motion greatly influence the perception of human action. Motor learning, which is the viewers' actions and experiences, can also affect their perception of other people's actions. Other aspects, such as social constraints and neural mechanisms, also play an essential part.

Etemad et al. (2016) studied the relative significance of different themes in motion perception. The two sets of themes they studied represent different features of human motion: primary themes specifying actions and secondary themes specifying styles or characteristics. These themes are similar to the concepts of motion and form, respectively.

People tend to make accurate judgments of simple, one-dimensional motion that can be represented as particle motion like a free-falling object. However, they tend to make inaccurate judgments of complex, multi-dimensional motions that cannot be treated as particle motion (Proffitt & Gilden, 1989). According to Vicovaro et al. (2014), people tend to rely on heuristic strategies (a conscious or sub-conscious simplification strategy to save perceptual and cognitive resources) rather than perceptual judgment. Vicovaro et al. also proposed that, given more perceptual information, viewers are more likely to evaluate perceptually rather than heuristically. In the present study, therefore, perceptual information will be displayed along with the stimuli during the experiment.

Visual Perception of Virtual Characters

In a 3D animation, the visual representation of the character can affect how viewers perceive the animation. The "uncanny valley" hypothesis describes how an object's resemblance to a human affects the viewers' emotions. Studies have found that robots designed to be human-like give viewers an eerie feeling (Mori, 1970). The same prediction also applies to 3D computer animation.

In the early stage of the study of visual perception in psychology, experiments were conducted in which the motions from a human were compared with a point-light animation using the same motions. (Point-light animation involves attaching lights to an actor's major joints; the actor then moves against a contrasting background). Viewers can distinguish between human walking, running, dancing, and more, with only 10 to 12 moving points (Johansson, 1973). With the advancement of computer animation, similar studies have been conducted on virtual models. A character with a fully detailed polygon mesh is perceived as more sensitive by viewers compared to a stick figure character (Hodgins et al., 1998).

In character animation interpolation, viewers' sensitivity to error directly determines whether they perceive the motion as natural or artificial. Human traits in non-human objects known as anthropomorphism have been studied using virtual characters (Chaminade et al., 2007). The more anthropomorphic the characters are, the more likely viewers are to report their motion as artificial. This study was supported by an fMRI examination of the brain's mentalizing region. Furthermore, a study was conducted to compare viewers' sensitivity to the motion error between a ball and a human. The results showed that errors caused by local motion change have little impact on the user's sensitivity. However, higher error sensitives are shown when displayed global changes in human characters (Reitsma et al., 2008).

Another factor to consider is the viewer's level of experience with animation. Those who are familiar with animation are very likely to spot errors in human motion; those who are naïve in regard to computer graphics are less sensitive to artifacts (Mezger et al., 2005). This study focuses on both audiences: those who are novices in animation and those who are industry professionals in areas such as animation, VFX, video games, and film.

Another aspect to consider is gender. Both the gender of the viewer and the gender of the character can influence human-computer interaction. A study by Kramer et al. (2016) found a difference in relation to learning when viewers interact with teaching agents of the same or opposite sex. For neutral motions such as walking or conversational gestures, viewers' judgments of male or female characters are similar; when certain emotions are involved, however, such as sadness or anger, gender bias appears more prominent (Zibrek et al., 2015).

Character selection must be carefully considered to ensure it reflects the perception of natural motion (Reitsma et al., 2008). When viewers evaluate the same motion in different characters, their sensitivities may vary between the characters (Hodgins et al., 1998). This thesis used an average complexity, gender-neutral robot character similar to that used in the study conducted by Chaminade et al. (2007) to balance sensitivity, anthropomorphism, and the gender factor.

Evaluation of Naturalness

Jansen and Van Welbergen (2009) explored the methodologies for evaluating the motion of human-like animation. In their experiment, various sources of motion were used, from original motion-captured data, to the balance model, to no motion at all. They proposed three evaluation methods for naturalness based on signal detection theory (Macmillan & Creelman, 2005):

- Two-alternative forced choice (2AFC): Participants compare two animated clips side by side. They must choose the one that looks more natural.
- Yes/No: Participants view one clip at a time and say whether they think it is generated by computational algorithms or real human performance. This is somewhat similar to the motion Turing test where participants judge whether the motion is “machine-like” or “human-like” (Van Welbergen et al., 2010).
- Rating: Participants view one clip at a time and rate the naturalness of the animation based on a scale from “Very unnatural” to “Very natural.”

The results showed that the rating method could provide more valuable information on the naturalness of individual motion than the other two methods. The Yes/No method and 2AFC are more efficient than the rating method and are suitable for discriminating clips. (Jansen & Van Welbergen, 2009). Hyde et al. (2014) used a similar rating method to evaluate the naturalness of a character’s facial expressions when exaggerated or damped.

The three test paradigms experiment proposed by Jansen and Welbergen (2009) used a sample of 29 voluntary participants. Luo et al. (2009) performed a study on naturalness in procedural animation using 21 subjects. Lin et al. (2009) explored the relationship between style parameters and an animation’s expressiveness with 34 participants. All the studies discussed above used the statistical analysis method of analysis of variance (ANOVA). Given the limitations of time and resources in the present study, all three methods could not be tested. Instead, this thesis used the rating method, which is the most informative evaluation method.

2.3 Expressive Upper Body Gestures

Upper Body Motion in Communication

In addition to verbal communication, non-verbal behaviors such as facial expression, eye gaze, and body posture play an essential part in human social interactions (Mehrabian, 1981). The previous section suggested that, when evaluating the motion of interpolated animation, the character model's visual representation must be carefully considered. In this thesis, facial expression and eyes are not included when evaluating the character's non-verbal behavior. Etemad et al. (2016) explored the two major areas of humanoid animation motion segments: face motion and body motion. They suggested that the "face and body can perform independently in terms of actions but not emotions and expressions" (p. 65). Similarly, other research into human and robot interaction has focused on understanding the effects of upper body gestures in humans and robots' reactions using upper body gestures (Ramiro et al., 2013; McColl et al., 2011; Xiao et al., 2014).

In humans, the upper body is more expressive than the lower body. Ogawa et al. (2017) analyzed the relationship between emotions and upper body poses. By using an evaluation grid method, they found that emotions such as surprise, fear, anger, and joy were associated with certain postural features. They also conducted a follow-up study with multiple regression analysis to evaluate the quantitative relationship between these emotions and the upper body poses. The results from both studies suggest that the character's expressiveness is closely connected to the motion of its upper body.

Gunes and Piccardi (2007) provided an expressive body gesture reference table based on research from Coulson (2004) and Burgoon et al. (2005). In their table, the emotion categories and correlated upper body gestures are as follows. 1. Anxiety: hands close. 2. Anger: body extended; hands low on the waist. 3. Disgust: body leaning back; hands touching the neck or head. 4. Fear: body contracted; body leaning back; hands high, trying to cover body parts. 5. Happiness: body extended; hands kept high. 6. Uncertainty: shoulder shrug.

Body Gesture Type

There is an infinite amount of freedom for upper body gestures. However, only a limited amount of motion can be tested. McNeill (1992) classifies four common types of body gestures:

- *Metaphoric gestures* are gestures that create literal, physical form to abstract ideas.
- *Iconic gestures* are gestures that represent an object's attributes, relationships, or actions, such as indicating the size of an object.
- *Deictic gestures* are gestures that connect speech to another object, such as pointing.
- *Beat gestures* are gestures that convey the rhythm of speech.

This thesis used a combination of McNeill's gesture types and Gunes and Piccardi's body gestures reference table to test body animation. By using a range of body movements that fall into all the categories of body gestures, this thesis ensured that the stimuli covered enough variety of body motion. Figure 2.4 shows the four gestures selected for this thesis: right arm throwing an object (metaphoric gesture); both arms moving outward, showing the size of an object (iconic gesture); right arm pointing to the sky (deictic gesture); both arms moving forwards in parallel (beat gesture).

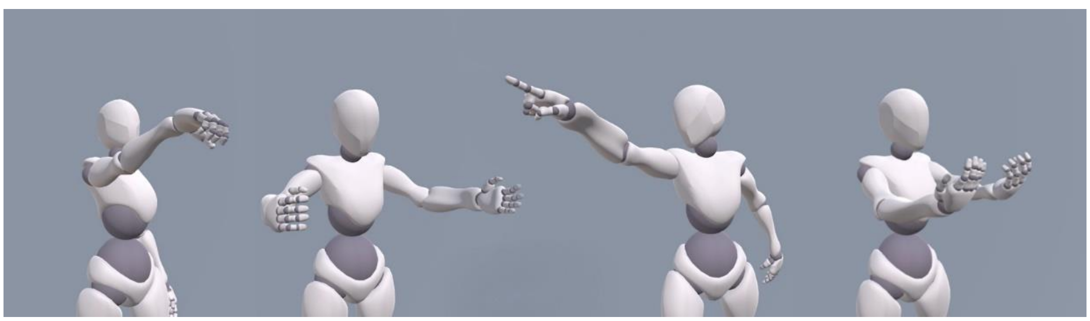


Figure 2.4. Examples of four types of gestures
(from left to right: metaphoric, iconic, deictic, beat)

2.4 **Summary**

This literature review started by presenting research on rotation interpolation. Rotation can be divided into two modalities: rotation in Euclidean space, represented by Euler angles or rotation matrices, and rotation represented by quaternions. Next, the literature review introduced the mathematical equations for the three interpolation algorithms: LinEuler, Slerp, and Squad (Dam et al., 1998). The following part connected the interpolation of rotation to character animation. When interpolation is applied to joint orientation, it drives the character's skinned mesh, which in turn creates body motion. Wang (2015) proposed an animation system in which interpolation is divided into three components: blending time, curve, and velocity. Each component is treated separately but significantly affects the overall motion of interpolation.

The next section described how people perceive both the form and the motion of human action. Visual perception is greatly influenced by how the character is represented, including its level of detail, its level of anthropomorphism, and its gender. With anthropomorphism, viewers perceive motion as more unnatural (Chaminade et al., 2007). With complex mesh, viewers are more sensitive to errors in motion (Reitsma et al., 2008). Jansen and Van Welbergen (2009) proposed three evaluation paradigms based on signal-detection theory (Macmillan & Creelman, 2005) to evaluate how viewers perceive the naturalness of a virtual human. The method 2AFC is the most efficient in discriminating between motions, followed by the Yes/No method. However, the rating method provides the most information about naturalness. Due to the limitations of time and resources, this thesis used only the rating method for the evaluation.

Human motion in non-verbal communication is divided into face motion and body motion. Face and body can perform independently to convey emotion. Next, using an evaluation grid method and a follow-up study involving multiple regression analysis, Ogawa et al. (2017) found that emotion is closely associated with upper-body gestures. Gunes and Piccardi (2006) provided a reference table of different upper body gestures that express different emotions. To ensure that the experiment covered a sufficient variety of motions, four body gestures adhering to McNeill's (1992) classification were used: iconic gesture, deictic gesture, metaphoric gesture, and beat gesture.

3. METHODOLOGY

This thesis aimed to examine the effects of three interpolation methods, namely linear Euler interpolation (LinEuler; Equation 2.1), spherical linear quaternion interpolation (Slerp; Equation 2.2), and spherical spline quaternion interpolation (Squad; Equation 2.3), on the perceived naturalness of upper body character animation. These interpolation algorithms were applied to four different upper body motion sequences.

3.1 Variables and Hypothesis

This experiment included two independent variables: the interpolation method and body gesture type. The interpolation method had three levels, and the body gesture type had four levels, forming a 3×4 within-participants factorial design. Each participant evaluated all 12 animation clips.

Table 3.1. Levels for interpolation methods

Factor 1	Level 1	Level 2	Level 3
Interpolation method	LinEuler	Slerp	Squad

Table 3.2. Levels for body gesture types

Factor 2	Level 1	Level 2	Level 3	Level 4
Body gesture type	Beat	Deictic	Iconic	Metaphoric

The dependent variable of this experiment was the measurement of perceived naturalness for each of the animation clips using the rating evaluation method. During the experiment, animations were viewed and rated as $r \in \{1, 2, 3, 4, 5\}$ based on a 5-point Likert scale (1 = *not natural at all*, 2 = *somewhat unnatural*, 3 = *neither natural nor unnatural* or *neutral*, 4 = *somewhat natural*, 5 = *very natural).* For every gesture type and every interpolation method, a rating value was recorded as the perceived naturalness. The statistical analysis took into account two other factors: the gender of the participants (male or female) and their level of experience in animation (novice or experienced).

The null hypothesis was that all the interpolation methods would be given the same naturalness ratings, meaning that all three interpolation methods have equal effects on the viewer's perception of naturalness in character animation. The alternative hypothesis was that at least one of the interpolation methods would be given a different rating; that is, the viewer perceives naturalness differently depending on the interpolation method, and some methods are better or worse than others.

$$H_{null}: r_{LinEuler} = r_{Slerp} = r_{Squad}$$

$$H_{alt}: not\ all\ r\ values\ are\ equal$$

3.2 Participants and Sampling Design

This study was targeted at the general public, who have little experience in animation. However, Mezger (2005) raised the point that viewers' level of experience with animation affects their perception of naturalness. Those experienced in animation can easily spot animation artifacts, whereas those who have less experience are more tolerant of errors. This study also sought to determine whether animation professionals perceive naturalness differently from less experienced viewers.

Two perceptual surveys (Section 3.4) were designed and hosted on Qualtrics and distributed on the Prolific platform and through the University's Computer Graphics Technology department. Prolific, a participant-recruitment website, allows surveys to be posted and shared with a specific participant pool. For this study, the platform was instructed to recruit participants located in the United States with a personal computer as their access device, as the survey was composed in English and designed for non-mobile devices. The survey was also distributed via email to the students and faculty members of the University's Computer Graphics Technology department to ensure a sufficient number of participants had animation experience. The participants recruited from Prolific received a fixed compensation of $1 for completing the survey; those recruited from the department were given the option to enter a prize draw for two Amazon gift cards worth $40.

To determine an appropriate sample size for this thesis, a pilot study was first run with a small sample of participants. After collecting the results from the pilot study, a power analysis determined the appropriate sample size for the main study. This power analysis is discussed in Section 4.1.

3.3 Experiment Materials

Stimuli

The stimuli were 12 animation clips showing gestures ranging from 1.6 seconds to 2.4 seconds in duration. Animations in the same gesture category have the same frame range (start pose, end pose, and duration), with different in-between animations generated by different interpolation methods (shown in Figure 3.1). Animations in the same interpolation method category use the same interpolation algorithm but on different pre-recorded motion segments.

There were four original motion-captured recordings, each corresponding to one type of gesture. These original recordings each provided the motion primitives, which served as the source for the three algorithm-generated interpolation clips. They were later re-processed to generate the 12 stimuli clips.

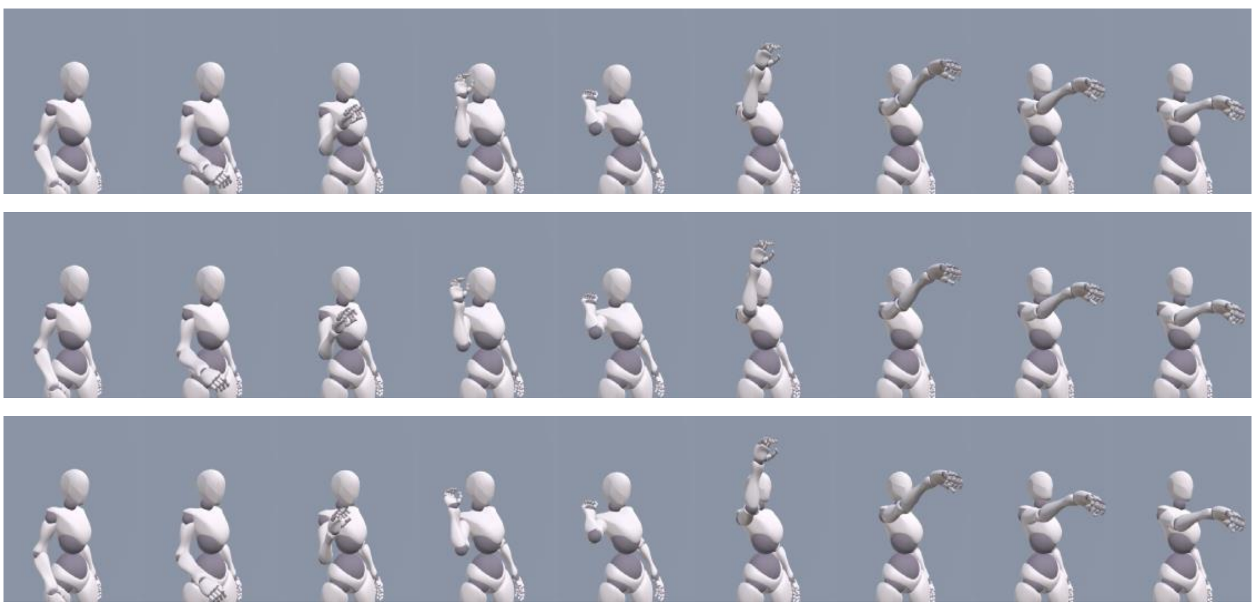


Figure 3.1. X Bot sampled animation of metaphoric gesture (from top to bottom: LinEuler, Slerp, and Squad)

System Overview

This thesis used X Bot, a 3D model of a rigged (ready to animate) humanoid robot model as the target character from Mixamo (n.d.). The character was gender-neutral, with no facial features. It had human-like fingers, posed only to fit the gesture's context and not animated. The character consisted of 19 major joints adheres to the Humanoid Animation Standard (2006).

The system for creating the stimuli can be divided into two major stages: motion primitive preparation and interpolation implementation (Figure 3.2). In the first stage, original motion data are gathered from a real actor's performance. In the second stage, Maya (3D animation software) and Unity (real-time game engine) implement the three interpolation methods.

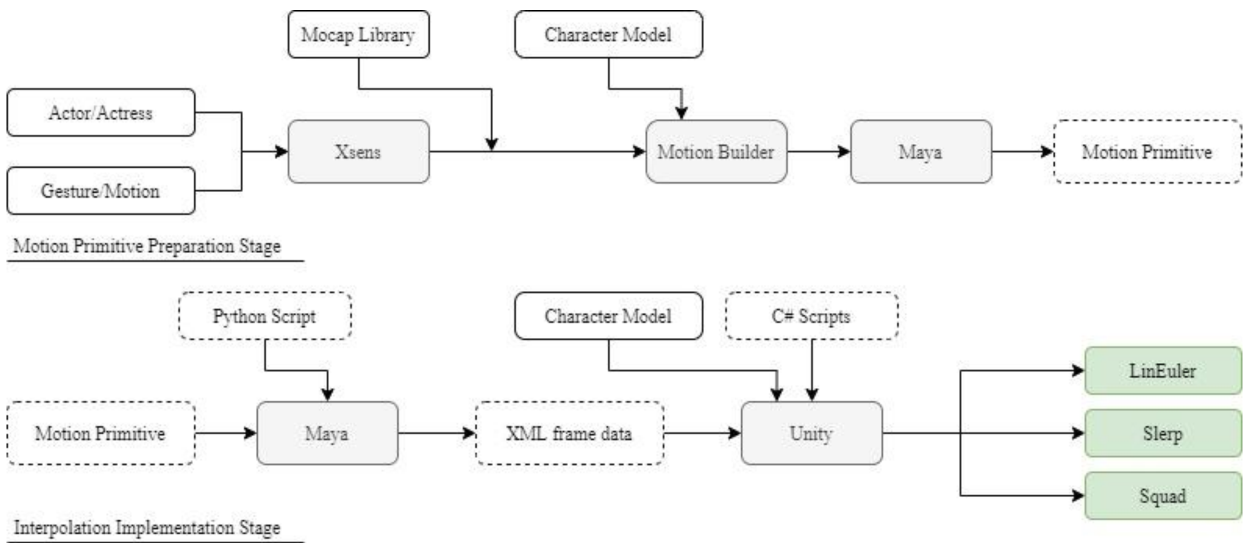


Figure 3.2. System overview

The sensor-based motion-capture system Awinda by Xsens is used to capture real actors' performances. Motion-capture data from Mixamo, an online motion-capture data library, are also used as a supplement. The recordings obtained from Xsens are .fbx files consisting of raw motion data from 17 sensors, which are then remapped to the X Bot character through MotionBuilder's (motion capture software) retarget features. A comparison of the X Bot and Xsens skeletal systems is shown in Figure 3.3. The retarget phase requires some manual adjustments to ensure the X Bot character's motion is correct. The motion data obtained from Mixamo are ready to use without modification, as Mixamo uses X Bot as the default character.

The experiment only needs a small amount of data; thus, the data are imported to Maya and clipped to an appropriate length, concluding the motion primitive preparation stage.

In the next stage, the data are extracted via a custom Python script in Maya. This script extracts all the rotation values from the root joint along its hierarchy for selected frames. It outputs XML files consisting of orientation data for selected keyframes on selected joints. This XML data, along with the X Bot character, are then passed to Unity, which performs the interpolation of rotation on individual joints. For each gesture, custom C# scripts generate three output interpolations: LinEuler, Slerp, and Squad. As a result, the system produces 12 stimuli clips under the same lighting and camera conditions.

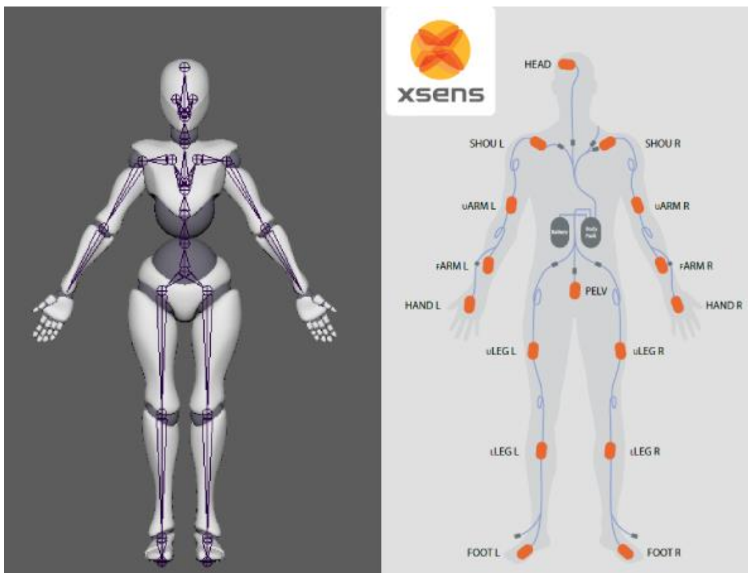


Figure 3.3. Comparison of X Bot joint position and Xsens sensor position

Motion Extraction and Conversion

In the interpolation implementation stage, the motion data are passed between two software applications: Maya and Unity. Although Maya can export .fbx files with preserved motion data and directly import them into Unity, the animation in Unity is encapsulated and hard to modify with desired interpolation algorithms.

Furthermore, the system requires only a small number of keyframes for interpolation. Thus, it is ideal for extracting only the necessary key poses. A custom Python script, “keyframe extractor," was used in Maya to format the rotation values for selected joints of selected frames into an XML file.

```
Int[] frames
String rootJnt

For jnt in Hierarchy(rootJnt):
	For frame in frames:
		WriteXML(jnt, frame, rotationX, rotationY, rotationZ)
```

The final step in the data exchange between Maya and Unity is the rotation data conversion. Maya uses a right-handed vector space with XYZ default rotation order. In contrast, Unity uses a left-handed vector space with ZXY rotation order. Additionally, Unity represents rotation internally using quaternions (Unity Script Reference, 2019). Thus, the Euler angles from Maya need to be converted to the corresponding quaternions (Unity Forum, 2009).

```
Vector3    input
Quaternion output

flipped = Vector3(input.x, -input.y, -input.z)
Quaternion_x = Convert(flipped.x, vector_right)
Quaternion_y = Convert(flipped.y, vector_up)
Quaternion_z = Convert(flipped.z, vector_forward)

output = Quaternion_z*Quaternion_y*Quaternion_x
```

Interpolation Implementation

The implementation of LinEuler interpolation (Equation 2.1) treats rotation as Euler angles, represented as a "Vector3" data type (an ordered list of three values). Unity has a built-in Lerp function, which can linearly interpolate between two vectors with an interpolation parameter.

```
Vector3 from, to
Float interp

LinEuler_result = from*interp+to*(1−interp)
```

The implementation of Slerp interpolation (Equation 2.2) treats rotation as quaternions, represented as a "Quaternion" data type (a data structure in Unity with an ordered list of four values). The built-in Slerp function in Unity supports the functionality to interpolate two quaternions along with the shortest distance.

```
Quaternion from, to
Float interp

dot_result = Dot_Product(from, to)
theta = Acos(dot_result)

Slerp_result = (from*Sin((1-interp)*theta)+to*Sin(interp*theta))/Sin(theta)
```

As stated in Equation 2.3, the Squad is a third-order curve based on Slerp, which requires two additional auxiliary points before and after the control points. There is no built-in Squad function in Unity for sequential rotations, so Squad interpolation was accomplished using a custom Squad package (Myklebust, 2015) and some math extension functions to generate joint rotation values.

```
Quaternion before, from, to, after
Float interp

Define Intermediate(Quaternion a, Quaternion b, Quaternion c)
       value = b*Exp(-0.25*(log(a/b)+log(c/b)))

tangent_from = Intermediate(before, from, to)
tangent_to   = Intermediate(from, to, after)

Squad_from   = Slerp(from, to, interp)
Squad_to     = Slerp(tangent_from, tangent_to, interp)
Squad_result = Slerp(Squad_from, Squad_to, 2*interp*(1-interp))
```

3.4 **Survey Procedures**

The study was divided into two parts: the pilot study and the main study. The pilot study determined the appropriate sample size for the main study. The feedback from the pilot study was used to find more effective ways to present the animated clips and improve the survey design. Both studies were hosted on Qualtrics, a survey website.

The Pilot Study

In the pilot study survey (Appendix D), the participants were first presented with an overview of the survey, including the purpose, the tasks they would be asked to carry out, an estimation of the duration, and other notes. Next, they provided their demographic information, including age, gender, race, highest completed education, and experience in animation or computer graphics. The participants were treated equally regardless of their responses. The only required answer was their level of experience in animation; all other demographic information was optional for the pilot study.

The main part of the survey involved watching and evaluating different animation clips. This part consisted of 12 animation blocks divided into four gesture groups. The gesture groups were presented in random order, and the participants acknowledged a brief description upon starting each gesture group. Within each group, three animation clips of that gesture using different interpolation methods were presented in random order. Each clip (a looped full-body animation) was shown along with a descriptive text (e.g., “the character is throwing an object”). After viewing each clip, the participants were asked to rate the clip on a 5-point Likert scale (from 1, “not natural at all,” to 5, “very natural”) based on the question: How natural is the animation? A hidden timer was used to track the time spent on each page. The survey ended after the participant had rated all 12 clips. They were then encouraged to leave anonymous feedback to improve the survey.

The Main Study

Based on the feedback from the pilot study, slight adjustments were made to the survey design of the main study (Appendix E), including an improved descriptive text for each gesture; a new camera angle, framing only the upper part of the character; and a counter-balanced random order instead of a completely random order for the presentation of animations with different interpolation methods.

As in the pilot study, the participants started the experiment by reading the survey introduction and providing their demographic information, including age, gender, race, highest completed education, and animation or computer graphics experience level. The participants were treated equally, regardless of their answers. Every demographic question required an answer, and these answers were used to determine whether the participants were representative of the target population.

The next stage of the main study involved viewing and evaluating the four gestures using the rating method. There was a total of 12 animation blocks divided into four gesture groups. All four groups were presented in random order, and the participants acknowledged a brief description upon starting each group. Each group included three animated clips, each of which used a different interpolation method. Each animation was looped and shown along with a descriptive text. The participants were instructed to rate the clip on a 5-point Likert scale (from 1, “not natural at all,” to 5, “very natural”) before proceeding to the next clip. A hidden timer was used to track the time spent on each page. The responses were recorded, and the survey ended after the participant had evaluated all 12 clips.

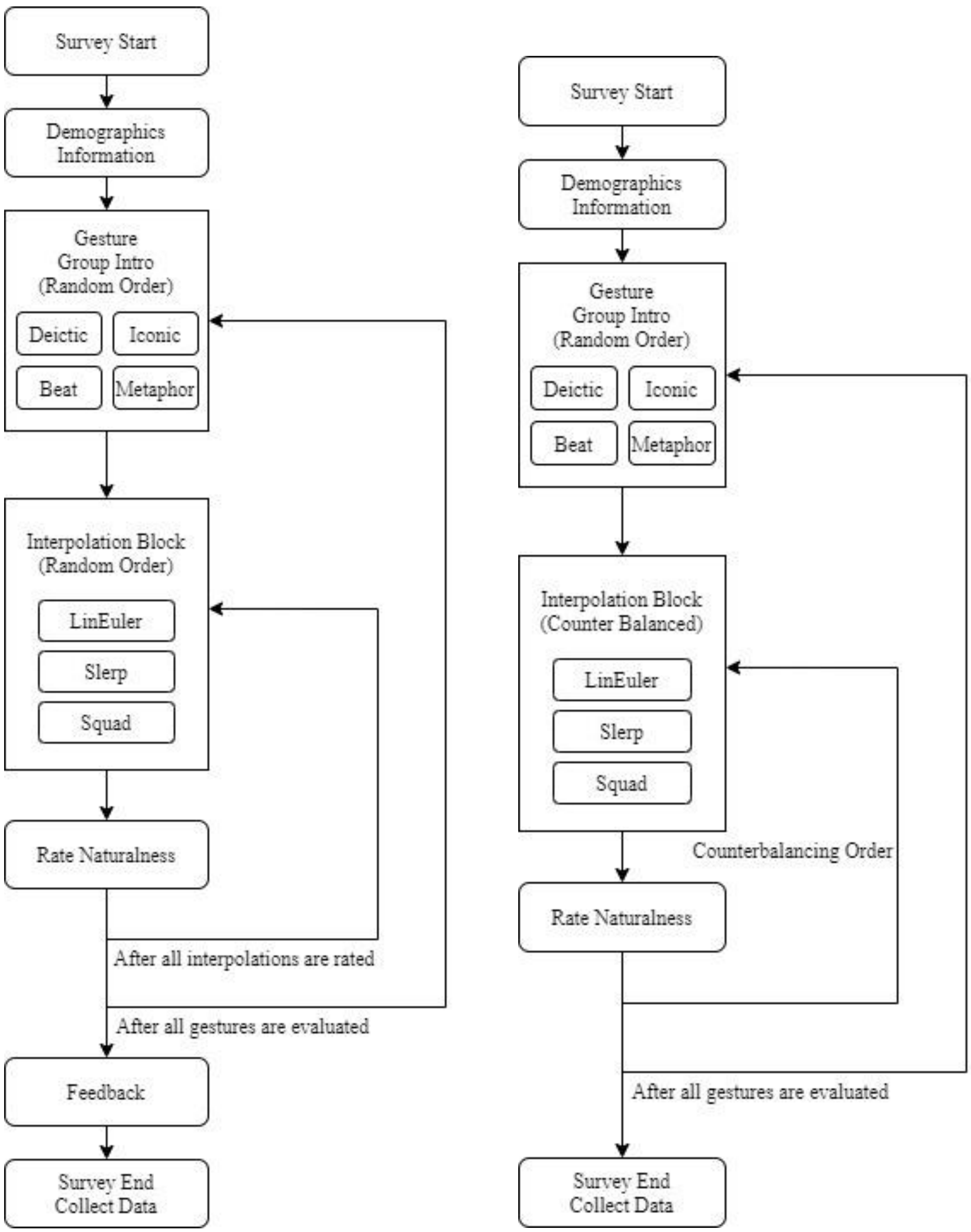


Figure 3.4. Flow chart of the experimental steps
(pilot study on the left and the main study on the right)

3.5 Analysis Strategy

All the participants' responses were recorded automatically by Qualtrics after each individual completed the entire study. Not all the data, however, qualified for analysis. This thesis used two sets of rules to filter out participants who were potentially rushing through the survey by giving patterned or random responses. The first filter excluded the results of participants who gave every clip the same rating. The second filter excluded faster than normal responses, taking into account both the entire time spent on the survey and the time spent on individual questions. The precise cutoff time was determined from the pilot study.

Before the actual data analysis, the sample size was set for the main study. A power analysis was performed based on the data collected from the pilot study. After consultation with the statistical department, the power level was set at 80%.

The participants' evaluation data for the statistical analysis were based on the Likert scale $r \in \{1, 2, 3, 4, 5\}$. The analysis followed a three-step procedure. In the first step, the mean rating for each interpolation method is calculated using linear regression, with LinEuler as the baseline. In the second step, a one-way analysis of variance (ANOVA) test is conducted. Its outcome p-value either rejects or fails to reject the hypothesis that all the interpolation methods would be given an equal rating. If the result rejects the null hypothesis (i.e., there is at least one group different from the other two), a third step is required, in which Tukey's honest significant difference test (Tukey's HSD) is performed for post-comparison. The outcome adjusted p-value from Tukey's HSD identifies which pairs of groups are different. The result further indicates which method is different compared to the other in terms of perceived naturalness.

3.6 Summary

This chapter first introduced the methodology design with the hypothesis for examination and the research method. This experiment used a 3×4 within-participants design, where every participant was assigned to all options of the three interpolation methods (LinEuler, Slerp, and Squad) for each of the four gesture types (beat, deictic, iconic, and metaphoric). The participants were recruited either through Prolific or from among the students and faculty members of the department. This recruitment method ensured that the study population both reflected the general public and included experienced viewers.

The stimuli were 12 animated clips. The clips were created by converting .fbx motion data to motion primitives. The interpolation implementation stage used the Unity game engine to apply the interpolation algorithm and render the output animation. The experimental procedure was divided into two stages: the pilot study and the main study. The pilot study was used to estimate the appropriate sample size and improve the survey design for the main study. Lastly, the analysis strategy consisted of three steps: a linear regression to calculate the mean rating for each interpolation method, a one-way ANOVA to either reject or fail to reject the hypothesis that all the methods would be given an equal rating. If the result rejects the null hypothesis, Tukey's HSD test is used to identify which methods are different.

4. DATA COLLECTION AND ANALYSIS

4.1 The Pilot Study Overview

Data Collection

The pilot study was conducted to estimate the proper sample size and to improve how animation clips were presented to the viewers. Twenty responses were collected from privately invited links. After looking at the length for total completion time and the length for individual clips submission time, a minimum time filter was set. If participants took less than 120 seconds to complete the entire survey or spent less than four seconds to evaluate any individual clip, the result was filtered out. A total of 18 responses were used for further analysis.

Data Analysis

First, a linear regression model was applied to estimate the mean of the three interpolation methods. The linear regression model used LinEuler as the baseline with the intercept value at 2.2500; Compared to the other two methods, Slerp had a value of -0.1806 and Squad a value of 1.1111. As a result, LinEuler had a mean rating of 2.2500, Slerp had a mean rating of 2.0694, and Squad had a mean rating of 3.3611 (Table 4.1).

After calculating each interpolation method's means, a power analysis for ANOVA was performed to determine the sample size. In this thesis, a power level of 80% was chosen to compute the main study's minimal number of samples. After running 100,000 simulations, a minimal of 48 subjects were required for a power level of around 80%.

The third step was to perform a one-way ANOVA test for either reject or failed to reject the hypothesis. The p-value obtained was 1.99e-12, far less than 0.05 (significant level of 95%). Therefore, there was significant evidence to reject the null hypothesis that all interpolation methods have equal ratings.

Now that the ANOVA test had a significant p-value indicates that some group means were different, an additional post-ANOVA comparison, the Tukey HSD, needed to be conducted to identify which pairs of groups were different. The adjusted p-values were generated from three pairs of groups (Slerp-LinEuler, Squad-LinEuler, Squad-Slerp) using a confidence level of 95%.

Based on the adjusted p-value, the Squad-LinEuler pair and Squad-Slerp pair had a value less than 0.05, rejecting the null hypothesis that the two groups had identical ratings. Slerp-LinEuler pair had a p-value greater than 0.05, which failed to reject the null hypothesis that the two groups had equal ratings. Based on the p-value and difference value, Squad had a noticeable more positive rating of 1.1111 compared to LinEuler and 1.2917 compared to Slerp.

Table 4.1. The mean table for LinEuler, Slerp, and Squad in the pilot study

	LinEuler	**Slerp**	**Squad**
Mean	2.2500	2.0694	3.3611

4.2 The Main Study Overview

Data Collection

The main study used two distribution platforms: Prolific and the University department's email list. Fifty responses were collected from Prolific, and 59 responses were collected from the University. After applying a filter of a minimum of 120 seconds for the entire survey and 4 seconds for each clip, 97 responses (42 from Prolific and 55 from the email list) were used for the analysis. The number of participants was far greater than the minimum of 48, which gave this study a power level of approximately 98%.

For the analysis, a linear regression was performed first and then a one-way ANOVA test. If the ANOVA test yielded a p-value less than 0.05, a post-Tukey HSD test was used to identify the pairwise difference between groups.

Demographics

The survey also collected demographic information about the participants, including gender, race, education level, and animation experience, which helped to determine whether the sample was representative. Gender and animation experience were further analyzed by performing the three-step data analysis procedures. Instead of selecting from all the demographics categories, the participant could select “Prefer not to answer” to exclude themselves from the analysis.

Based on the data summary below, the sample was approximately half male and half female, with only 5% of participants reported as “Other.” Most of the participants (over 85%) selected either White or Asian as their race. As for education level, 46% reported having a high school degree or lower, and 54% reported having at least a bachelor’s degree.

Table 4.2. Gender distribution table

Gender	**Count**	**Percentage**
Male	43	44.33%
Female	49	50.52%
Other	5	5.15%

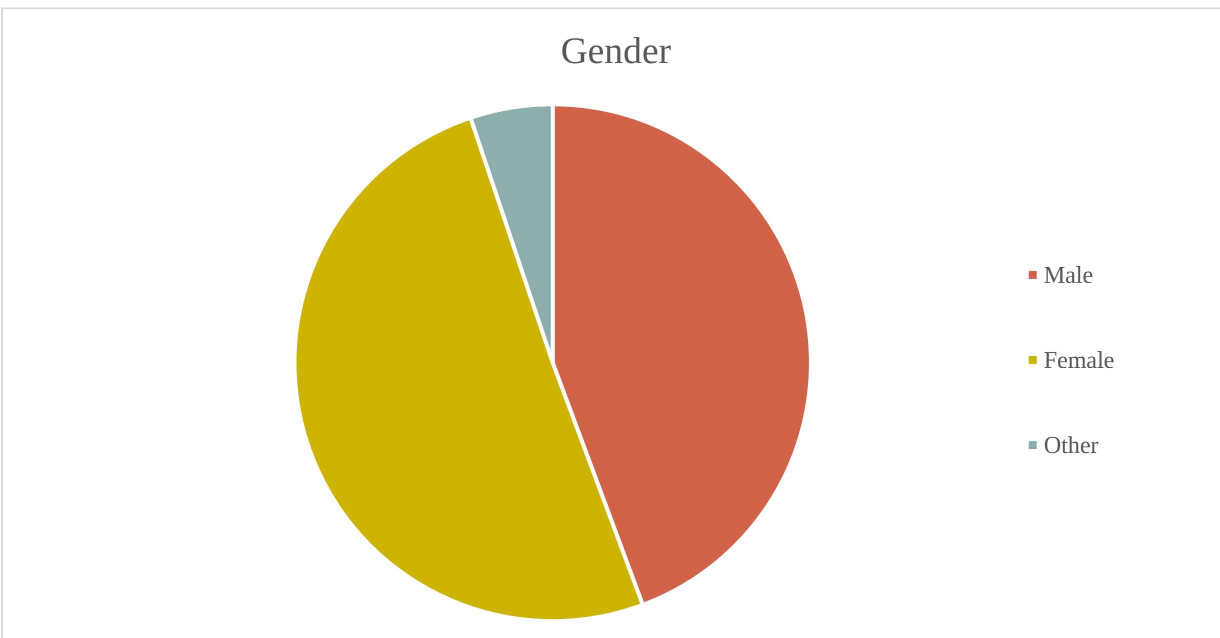


Figure 4.1. Pie chart for gender distribution

Table 4.3. Race distribution table

Race	Count	Percentage
White	59	60.82%
Black or African American	3	3.09%
American Indian or Alaska Native	0	0.00%
Asian	24	24.74%
Native Hawaiian or Pacific Islander	1	1.03%
Other	10	10.31%

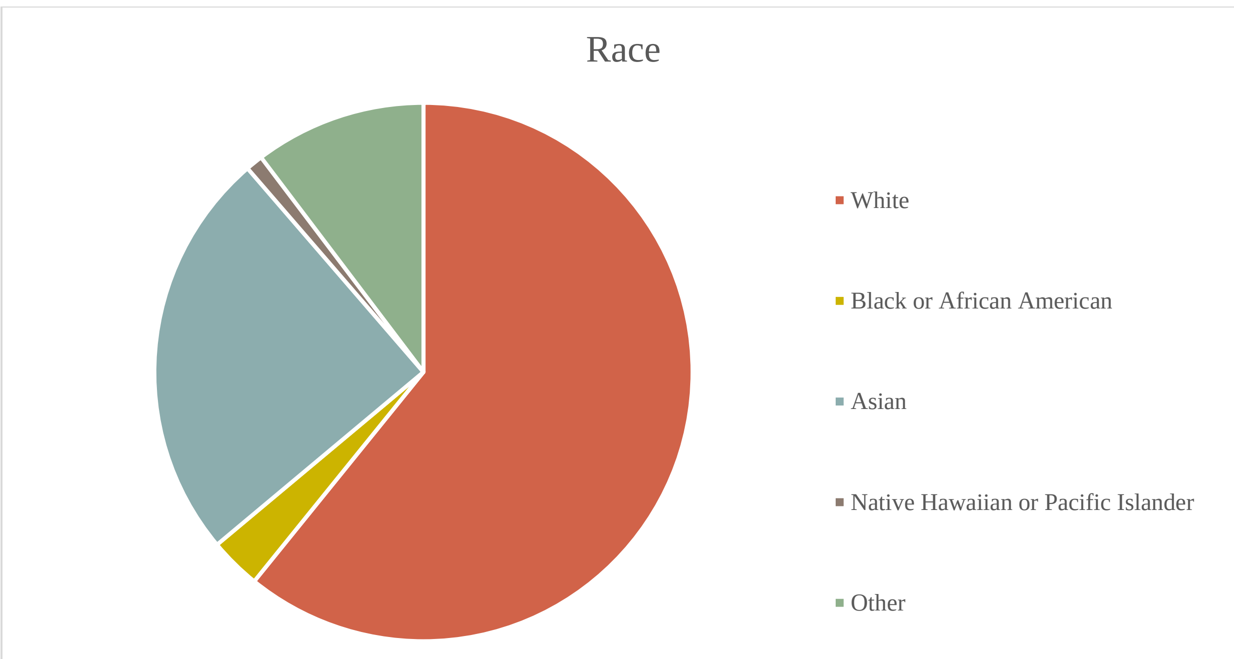


Figure 4.2. Pie chart for race distribution

Table 4.4. Education distribution table

Education	Count	Percentage
High school degree or lower	45	46.39%
Associate degree	9	9.28%
Bachelor’s degree	34	35.05%
Graduate degree or higher	9	9.28%

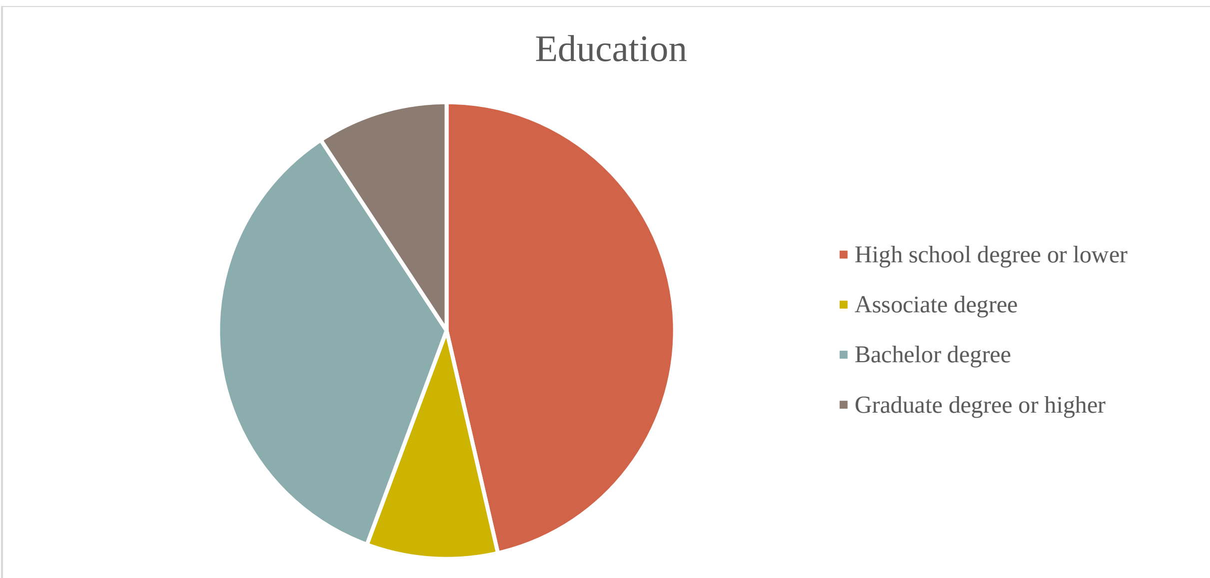


Figure 4.3. Pie chart for education distribution

Table 4.5. Experience in animation distribution table

Animation Experience	Count	Percentage
Novice	72	74.23%
Profession, Major in Animation	25	25.77%

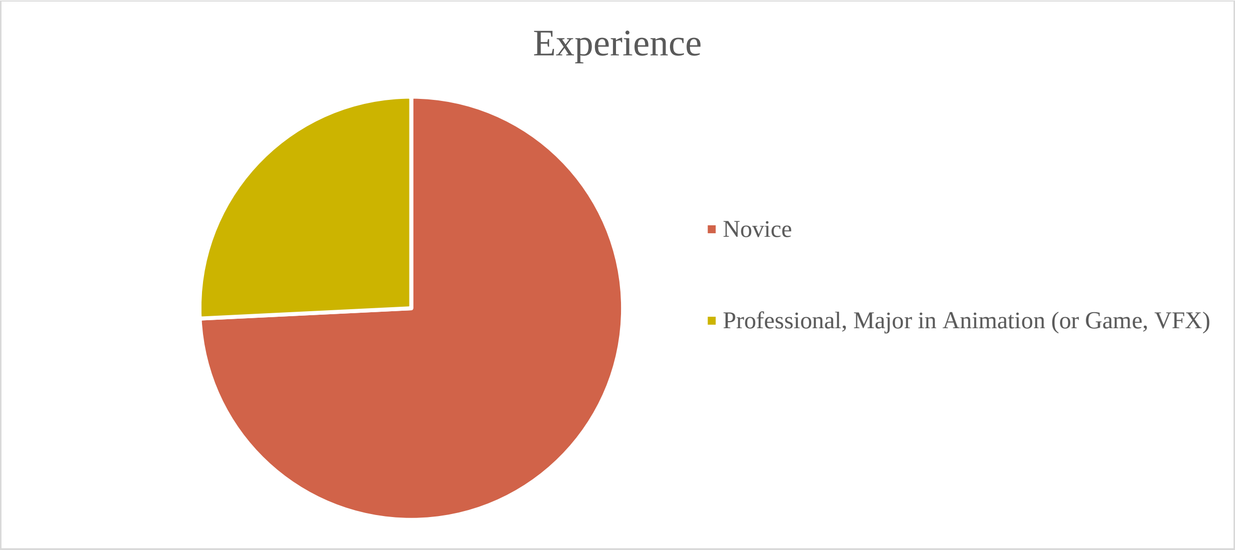


Figure 4.4. Pie chart for experience in animation distribution

Combined Analysis

The linear regression model used LinEuler as the baseline with the intercept value at 2.6598. In comparison, Slerp had a value of 0.0412, and Squad had a value of 0.8351. The mean rating for LinEuler was 2.6598; Slerp, 2.7010; and Squad, 3.4948 (**Error! Not a valid bookmark self-reference.**).

Next, the one-way ANOVA test yielded a p-value equal to 2e-16, much less than 0.05. There was therefore sufficient evidence to reject the null hypothesis that all the interpolation methods would be given equal ratings.

After the ANOVA test resulted in a significant p-value, indicating that some group means were different, Tukey's HSD post hoc test was used to identify which pairs of groups were different. The adjusted p-values were generated from three pairs of groups (Slerp-LinEuler, Squad-LinEuler, and Squad-Slerp) using a confidence level of 95%.

The adjusted p-value from the Tukey HSD matched the result from the pilot study, with the Squad-LinEuler pair and the Squad-Slerp pair's p-values less than 0.05, and so the hypothesis that the two groups had identical ratings was rejected. The Slerp-LinEuler pair with a p-value greater than 0.05 which failed to reject the hypothesis that the two groups had equal ratings. Based on the p-value and difference value, Squad had a more positive rating of 0.8350 compared to LinEuler and 0.7938 compared to Slerp.

Table 4.6. Mean table for the combined main study

	LinEuler	Slerp	Squad
Mean	2.6598	2.7010	3.4948

Data Analysis on Prolific Platform

The main study contained results from two platforms: Prolific and Department email list. The latter was biased toward animation-major students; therefore, this section repeated the analysis solely on data collected from Prolific, with a total of 42 subjects.

Starting with the linear regression model using LinEuler as the baseline, this study got an intercept value at 2.8333; Comparing it to the other two methods, Slerp had a value of 0.0595 and Squad a value of 0.5595. The linear regression result showed that LinEuler had a mean of 2.8333, Slerp a mean of 2.8929, and Squad a mean of 3.3929 (Table 4.7). At first glance, even though the differences between interpolation were perceived less overall, the rankings for mean value was aligned with the combined data.

Next, this study performed a one-way ANOVA test with a result of 1.1e-6 for the p-value, which was still significantly lesser than 0.05. Therefore, there was significant evidence to reject the null hypothesis that all interpolation methods had equal ratings. The Tukey HSD post hoc test was used to identify which pairs of groups were different with a confidence level of 95%.

The result from the Tukey HSD had Squad-LinEuler pair and Squad-Slerp pair less than 0.05, rejecting the hypothesis that these two groups were identical. As for the Slerp-LinEuler pair, its p-value was greater than 0.05, which failed to reject the hypothesis that the two groups had equal ratings. These data further aligned with the combined data, with less difference value. In short, Squad had a more positive rating of 0.5595 compared to LinEuler and 0.5000 compared to Slerp.

Table 4.7. Mean table for Prolific subjects in the main study

	LinEuler	**Slerp**	**Squad**
Mean	2.8333	2.8929	3.3929

4.3 Analysis of Experience Level in Animation

This section presented two same analysis procedures on two datasets separated from the combined study. These two datasets were divided based on the participant's experience level in animation. There was a total of 72 subjects for the novice group, and for the experienced group (Major or Professional in animation, VFX, or games), there were a total of 25 subjects. This section explored how the viewer's level of experience would play a role in evaluating the naturalness of different interpolations. The result showed that viewers with more experience tend to give lower naturalness ratings but could distinguish between clips significantly better.

LinEuler continued to be the baseline for linear regression. The novice group's intercept value was at 2.7778, while for the experienced group, the intercept was at 2.3200. Comparing it with the other two interpolations, the novice group had a difference value of 0.0347 for Slerp and a difference value of 0.6910 for Squad. The experienced group had a difference value of 0.0600 for Slerp and Squad, a difference value of 1.2500.

The sample mean for LinEuler in the novice group was 2.7778, and for the experienced group, 2.3200. The sample mean for Slerp in the novice group was 2.8125, while the experienced group had a value of 2.3800. Finally, the sample means for Squad in the novice group was 3.4688, and for the experienced group, 3.5700 (Table 4.8).

Next, a one-way ANOVA test was performed with a significant level at 95%. The result for both groups had a p-value of 2e-16, which was significantly lesser than 0.05. Therefore, for either the novice group viewers or the experienced group viewers, there was significant evidence to reject the null hypothesis that all interpolation methods had equal ratings. To identify which pairs of groups were different, a Tukey HSD post hoc test with a confidence level of 95% was further needed.

The Tukey HSD test results for the two groups showed that the Squad-LinEuler pair and Squad-Slerp pair were less than 0.05, rejecting the hypothesis that they were identical. As for the

Slerp-LinEuler pair, both groups' p-values were greater than 0.05, which failed to reject the hypothesis that the two groups have equal ratings. The result indicated that the Squad interpolation was different from the other two methods: for the novice group, Squad had a more positive rating of 0.6910 compared to LinEuler and 0.6562 compared to Slerp. For the experienced group, Squad had a more positive rating of 1.2500 compared to LinEuler and 1.1900 compared to Slerp.

When comparing the two groups together, the viewers with more experience gave lower ratings on both LinEuler and Slerp than the novice group. However, they gave higher ratings for Squad compared to viewers in the novice group. This was aligned with the literature review that participants with animation experience perceived more differences between interpolation methods.

Table 4.8. Mean table for different levels of animation experience in the main study

	Novice	**Experienced**
LinEuler	2.7778	2.3200
Slerp	2.8125	2.3800
Squad	3.4688	3.5700

4.4 Analysis of Different Gesture Type

This study used four different types (beat, deictic, iconic, and metaphoric) of gestures for character animation to cover various upper body motion. Some gestures showed more differences in naturalness compared to the others. This section analyzed and compared these types of gestures using the same procedure. Since each participant rated animated clips for all gesture types, a sample size of 97 was used across this analysis.

The first thing was to generate the linear regression model and get the mean values for different gesture types of different interpolation methods (Table 4.9).

For LinEuler interpolation, the beat gesture had a mean value at 2.5464, deictic gesture at 3.0825; iconic gesture at 2.5155; and metaphoric gesture at 2.4948.

For Slerp interpolation, the beat gesture had a mean value at 2.5567, deictic gesture at 3.0412; iconic gesture at 2.6082; and metaphoric gesture at 2.5979.

For Squad interpolation, the beat gesture had a mean value at 3.3814, deictic gesture at 3.6289; iconic gesture at 3.6701; and metaphoric gesture at 3.2989.

In the second step, this study performed one-way ANOVA tests with a significant level at 95%. The results for the p-value were all significantly lesser than 0.05. Therefore, there was significant evidence to reject the null hypothesis that all interpolation methods had equal ratings for all gesture groups. Then, the Tukey HSD post hoc tests to identify which pairs of groups were different with a confidence level of 95% was performed.

The Tukey HSD showed all Squad-LinEuler pair and Squad-Slerp pairs had p-values less than 0.05, rejecting the hypothesis that these two groups were identical in all gesture groups. As for the Slerp-LinEuler pair, all of its p-values were greater than 0.05, which failed to reject the hypothesis that the two groups had equal ratings in all gesture groups. The result indicated that the Squad interpolation was different from the other two methods for all four gesture types.

For beat gesture, Squad had a more positive rating of 0.8351 compared to LinEuler and 0.8247 compared to Slerp; For deictic gesture, Squad had a more positive rating of 0.5464 compared to LinEuler and 0.5876 compared to Slerp; For iconic gesture, Squad had a more positive rating of 1.1546 compared to LinEuler and 1.0619 compared to Slerp; Lastly, for the metaphoric gesture, Squad had a more positive rating of 0.8041 compared to LinEuler and 0.7010 compared to Slerp.

When looking at the results for all four gesture types together, the iconic gesture showed the most difference in naturalness between interpolation methods among all four groups. The deictic gesture showed the least difference in naturalness rating. However, the deictic group also had the highest average naturalness ratings, which meant the gesture could be potentially more challenging to evaluate than other gestures.

Table 4.9. Mean table for different gesture types in the main study

	Beat	**Deictic**	**Iconic**	**Metaphoric**
LinEuler	2.5464	3.0825	2.5155	2.4948
Slerp	2.5567	3.0412	2.6082	2.5979
Squad	3.3814	3.6289	3.6701	3.2989

4.5 **Analysis Based on Gender**

This section presented two same analysis procedures on two datasets separated from the combined study. These two datasets were divided based on the participant's gender. For a total of 97 subjects, the majority of them reported as either male or female. There were 43 subjects for the male group, and for the female group, a total of 49 subjects. This section was interested in comparing how the viewer's gender would affect the evaluation of naturalness in different interpolations. Based on the analysis, male participants had slight differences compared to female participants when evaluating different interpolations.

The analysis first started by using LinEuler as the baseline for linear regression. The intercept value for the male group was at 2.6512, while for the female group, the intercept was at 2.6888. Comparing it to the other two interpolations, the male group's Slerp had a difference value of -0.0930 and Squad a difference value of 0.9361. The female group's Slerp had a difference value of 0.1531 and Squad a difference value of 0.7194.

The sample mean for LinEuler in the male group was 2.6512, and the female group 2.6888. The sample mean for Slerp in the male group was 2.5581, while the female group had a value of 2.8418. Finally, the sample means for Squad in the male group was 3.5872, and for the female group, 3.4082 (Table 4.10).

Next, this section performed the one-way ANOVA test with a significant level at 95%. The male group had a p-value of 2e-16 and for the female group a p-value of 1.13e-11, which were both significantly lesser than 0.05. Therefore, for both male and female viewers, there was significant evidence to reject the null hypothesis that all interpolation methods had equal ratings. To identify which pairs of groups were different, this section further performed a Tukey HSD post hoc test with a confidence level of 95%.

The Tukey HSD test for the two groups showed that the Squad-LinEuler pair and Squad-Slerp pair's p-values were less than 0.05, rejecting the hypothesis that they were identical. As for the Slerp-LinEuler pair, both groups' p-values were greater than 0.05, which fails to reject the hypothesis that the two groups had equal ratings. The result indicated that the Squad interpolation was different from the other two methods: for the male group, Squad had a more positive rating of 0.9361 compared to LinEuler and 1.0291 compared to Slerp. For the female group, Squad had a more positive rating of 0.7194 compared to LinEuler and 0.5663 compared to Slerp.

Table 4.10. Mean table for different gender group in the main study

	Male	Female
LinEuler	2.6512	2.6888
Slerp	2.5581	2.8418
Squad	3.5872	3.4082

4.6 Summary

This chapter started by introducing the two studies conducted for this thesis: the pilot study and the main study. For both studies, the same three-step analysis was performed, including linear regression, a one-way ANOVA test, and a Tukey HSD post-comparison test. A power analysis was conducted on the pilot study to determine the appropriate sample size for the main study, with a power level above 80%.

The pilot study, main study, and the isolated Prolific study all found sufficient evidence to reject the null hypothesis that all three interpolation methods (LinEuler, Slerp, and Squad) would be given equal naturalness ratings. Squad interpolation was different from the other two interpolation methods, with noticeably more positive ratings. The analysis further separated the dataset between each gesture. All four gesture groups (beat, deictic, iconic, and metaphoric) also yielded the same results as all of the previous studies. The deictic gesture group was given higher ratings than the other three gesture groups, but it exhibited the least difference between interpolation methods. The iconic gesture group, on the other hand, exhibited the most difference between interpolation methods.

This thesis also explored how gender and animation experience affected the evaluation by splitting the dataset into corresponding subsets. For viewers experienced in animation and novices, the results provided sufficient evidence to reject the null hypothesis that the naturalness of the animations produced by all three interpolation methods would be perceived equally. The results also showed that the animations using Squad interpolation were given more positive ratings than the animations using the other two methods. The analysis found that viewers with more animation experience gave lower ratings on average but gave a wider range of ratings for different interpolation methods. No significant difference was found between the male participants and the female participants when evaluating non-emotional speech gestures.

5. CONCLUSION

5.1 Findings

The findings from the data analysis provide sufficient evidence to reject the hypothesis that LinEuler, Slerp, and Squad interpolation would be given the same naturalness ratings. After dividing the dataset into subsets and analyzing those subsets, it was found that upper body animation generated by Squad interpolation was perceived as significantly more natural than that generated by LinEuler or Slerp. This conclusion holds not only for audiences with different levels of expertise in animation (novice or professional) and different gender groups (male or female), but also for different gesture types (beat, deictic, iconic, or metaphoric).

Although LinEuler and Slerp use entirely different rotation models and interpolation calculations, there was insufficient evidence to reject the hypothesis that LinEuler and Slerp interpolation would be given the same naturalness rating. These two interpolation methods, therefore, had the same effect on the viewer's perception of naturalness in character animation. This finding was consistent across all the data subsets (i.e., different levels of animation expertise, different gender, and different gesture types).

Based on the 5-point Likert scale (from 1, "not natural at all," to 5, "very natural"), the average naturalness rating for animations generated using Squad interpolation was between 3 ("neutral") and 4 ("somewhat natural"). In contrast, animations generated using LinEuler and Slerp were rated between 2 ("somewhat unnatural") and 3 ("neutral"). Squad interpolation is superior to the other two linear models due to its algorithm, which generates a smoother path and continuous angular velocity. Perceptual studies have shown it to produce the most natural animations, which further aligns with the result of the mechanical analysis in the literature review.

The participants who identified as majors or professionals in animation (VFX or games) rated LinEuler and Slerp much lower than the participants who identified as novices in animation. However, compared to the notice group, they rated Squad higher in naturalness. This study found that experienced viewers could distinguish interpolation methods more clearly than inexperienced viewers. The literature review also found that experienced viewers are less tolerant of and more sensitive to errors in character motion than inexperienced viewers. This thesis also found that there was no significant difference between male and female participants' evaluations of naturalness, which is consistent with the findings presented in the literature review.

The usage of McNeill's classification of body gestures provides a broad range of upper body motions. One of the analyses was examined the naturalness of different gesture types. The iconic gesture was rated the most different between interpolations, and the deictic gesture showed the least difference. For all three interpolation methods, however, the deictic gesture received the highest average naturalness ratings, which meant that viewers were less likely to distinguish between interpolation methods in the deictic gesture group. The reason for this result might be better acting during the motion capture phase or better key pose sampling results.

5.2 Limitations and Future Work

For this thesis, a comprehensive perceptual experiment was conducted with consistent results across all analysis scenarios: animation interpolation using Squad is significantly more natural than animation interpolation using the other two methods. Given the range of possible upper body motions, however, more gesture types could be tested to further confirm the conclusion of this thesis. Moreover, due to the time and resource limitations of this thesis, only the rating method was used to evaluate the naturalness of the interpolation methods. The other two test paradigms proposed by Jansen and Welbergen (2009), 2AFC and the Yes/No method, are also worth using. Post-experiment qualitative questions could also help explain why Squad is perceived as more natural than the other two methods. The feedback from the pilot study suggested that camera angles also play a considerable part in viewers' perception of motion. It is vital to ensure that the participants have a clear view of the animation and that the motions are not hidden.

One of the essential properties of interpolation algorithms is the control points. In character animation, control points are translated by extracting key poses. This thesis used the researcher's judgment based on experience to identify the key poses of a motion-captured clip. With the advancement of motion identification, it is possible to algorithmically determine the key pose to interpolate. This method can potentially be used for other purposes such as frame reduction. Furthermore, the number of control points and the frame interval between each point also play a considerable part in the animation's smoothness. Although there are usually no limits to the number of keyframes that animation or VFX companies can add, in a real-time application, the control points need to be wisely planned. When the motion range exceeds 180 degrees between two orientations, using quaternions as rotation modality will interpolate along the shortest path, resulting in the direction of rotation being flipped. Applying more rules can avoid such an extreme case.

Lastly, the algorithms themselves have some limitations. All three methods interpolate along a perfectly planned path, which is humanly impossible. Human motion, although natural, is imperfect. Furthermore, as animation interpolation is performed on isolated joints, the rotation starts and ends at the same time with no variation. They also fail to handle the collision of body parts, which could result in 3D model clipping (or intersection). Many more improvements could be implemented. For example, adding random noise could avoid perfectly smooth curves; Interpolation with timed offsets based on joint hierarchy could add secondary motion; specifying rotation constraints for limbs could reduce the chance of model clipping.

5.3 Implications and Contribution

The findings of this thesis could benefit both real-time and non-real-time applications. The state-machine mechanism is commonly used in games and interactive applications. The more natural behavior of the virtual character, in turn, enhances user experience. The implementation of a more natural animation can also reduce stress for developers: By saving time on repetitive tasks and allocating more resources for creative decisions, a more natural animation interpolation enables independent developers to create content more quickly and larger studios to generate animation prototypes more quickly. Furthermore, this semi-procedural and simple approach for creating animation sequences allows people who have little professional knowledge to create their own animated stories.

The mechanical analysis suggested that Squad interpolation results in constant angular velocity and angular direction. Therefore, it appears to be the most smooth and natural of the three interpolation methods. The 12 principles of animation (Thomas & Johnston, 1981) states that most objects follow an arc when moving. The findings from this perceptual study agree with that principle. Moreover, this thesis has created a basis for further research into the evaluation of animation techniques.

REFERENCES

Barr, A. H., Currin, B., Gabriel, S., & Hughes, J. F. (1992). Smooth Interpolation of Orientations with Angular Velocity Constraints using Quaternions. *Computer Graphics*, 8.

Blake, R., & Shiffrar, M. (2007). Perception of Human Motion. *Annual Review of Psychology*, *58*(1), 47–73. https://doi.org/10.1146/annurev.psych.57.102904.190152

Bloom, C., Blow, J., Software, B.-A., & Muratori, C. (2004). *Errors and Omissions in Marc Alexa's "Linear Combination of Transformations."* 5.

Burgoon, J. K., Jensen, M. L., Meservy, T. O., Kruse, J., & Nunamaker, Jr, J. F. (2005). Augmenting human identification of emotional states in video. *Proceedings of the International Conference on Intelligent Data Analysis*. https://www.yumpu.com/en/document/read/11978926/augmenting-human-identification-of-emotional-states-in-video

Chaminade, T., Hodgins, J., & Kawato, M. (2007). Anthropomorphism influences perception of computer-animated characters' actions. *Social Cognitive and Affective Neuroscience*, *2*(3), 206–216. https://doi.org/10.1093/scan/nsm017

Coulson, M. (2004). Attributing Emotion to Static Body Postures: Recognition Accuracy, Confusions, and Viewpoint Dependence. *Journal of Nonverbal Behavior*, *28*(2), 117–139. https://doi.org/10.1023/B:JONB.0000023655.25550.be

Dam, E. B., Koch, M., & Lillholm, M. (1998). *Quaternions, Interpolation and Animation*. 103.

Etemad, S. A., Arya, A., Parush, A., & DiPaola, S. (2016). Perceptual validity in animation of human motion: Perceptual validity in animation of human motion. *Computer Animation and Virtual Worlds*, *27*(1), 58–71. https://doi.org/10.1002/cav.1631

Fielding, C. (2007, April). *Quick Trick: Gimbal Lock... Just Ignore It! (With a little help from Maya)*. http://fliponline.blogspot.com/2007/04/quick-trick-gimbal-lock-just-ignore-it.html

Geier, M. (2020). *Quanterion-nursery*. https://github.com/mgeier/quaternion-nursery

Gunes, H., & Piccardi, M. (2007). Bi-modal emotion recognition from expressive face and body gestures. *Journal of Network and Computer Applications*, *30*(4), 1334–1345. https://doi.org/10.1016/j.jnca.2006.09.007

Heloir, A., & Kipp, M. (n.d.). *EMBR: A Realtime Animation Engine for Interactive Embodied Agents*. 2.

Hodgins, J. K, Sweeney, P. K., & Lawrence, D. G. (1992). Generating Natural-looking Motion for Computer Animation. *Proceedings of the Conference on Graphics Interface'92*, 265–272.

Hodgins, J. K., O'Brien, J. F., & Tumblin, J. (1998). Perception of human motion with different geometric models. *IEEE Transactions on Visualization and Computer Graphics*, *4*(4), 307–316. https://doi.org/10.1109/2945.765325

Horswill, I. D. (2009). Lightweight Procedural Animation With Believable Physical Interactions. *IEEE Transactions on Computational Intelligence and AI in Games*, *1*(1), 39–49. https://doi.org/10.1109/TCIAIG.2009.2019631

Hyde, J., Carter, E. J., Kiesler, S., & Hodgins, J. K. (2014). *Assessing Naturalness and Emotional Intensity: A Perceptual Study of Animated Facial Motion*. 8.

Jansen, S. E. M., & Van Welbergen, H. (2009). Methodologies for the User Evaluation of the Motion of Virtual Humans. In Z. Ruttkay, M. Kipp, A. Nijholt, & H. H. Vilhjálmsson (Eds.), *Intelligent Virtual Agents* (Vol. 5773, pp. 125–131). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-04380-2_16

Johansen, R. S. (2019). *Automated Semi-Procedural Animation for Character Locomotion*. 114.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, *14*(2), 201–211. https://doi.org/10.3758/BF03212378

Ken, S. (1985). Animating Rotation with Quaternion Curves. *SIGGRAPH 1985*.

Ken, S. (1987). Quaternion calculus and fast animation. *SIGGRAPH Course Notes* 10.

Lasseter, J. (1987). Principles of Traditional Animation applied to 3D Computer Animation. *ACM Computer Graphics*, *21*(4), 35–44.

Lin, Y.-H., Liu, C.-Y., Lee, H.-W., Huang, S.-L., & Li, T.-Y. (2009). Evaluating Emotive Character Animations Created with Procedural Animation. In Z. Ruttkay, M. Kipp, A. Nijholt, & H. H. Vilhjálmsson (Eds.), *Intelligent Virtual Agents* (Vol. 5773, pp. 308–315). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-04380-2_33

Luo, P., Kipp, M., & Neff, M. (2009). Augmenting Gesture Animation with Motion Capture Data to Provide Full-Body Engagement. In Z. Ruttkay, M. Kipp, A. Nijholt, & H. H. Vilhjálmsson (Eds.), *Intelligent Virtual Agents* (Vol. 5773, pp. 405–417). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-04380-2_44

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide (2nd ed.)*. Lawrence Erlbaum Associates Publishers.

Marcos-Ramiro, A., Pizarro-Perez, D., Marron-Romera, M., Nguyen, L., & Gatica-Perez, D. (2013). Body communicative cue extraction for conversational analysis. *2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, 1–8. https://doi.org/10.1109/FG.2013.6553741

Matsiola, M., Dimoulas, C., Veglis, A., & Kalliris, G. (2005). *Augmenting user interaction experience through embedded multimodal media agents in social networking environments*.

McColl, D., Zhang, Z., & Nejat, G. (2011). Human Body Pose Interpretation and Classification for Social Human-Robot Interaction. *International Journal of Social Robotics*, *3*(3), 313–332. https://doi.org/10.1007/s12369-011-0099-6

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought.* University of Chicago press.

Mehrabian, A. (1981). *Silent Messages: Implicit Communication of Emotions and Attitudes*.

Mezger, J., Ilg, W., & Giese, M. A. (2005). Trajectory synthesis by hierarchical spatio-temporal correspondence: Comparison of different methods. *Proceedings of the 2nd Symposium on Appied Perception in Graphics and Visualization - APGV '05*, 25. https://doi.org/10.1145/1080402.1080406

Mori, M. (1970). The Uncanny Valley. *The Uncanny Valley*, *7(4)*, 33–35.

Myklebust, V. (2015, September). *Spherical Spline Quaternions For Dummies*. https://www.gamasutra.com/blogs/VegardMyklebust/20150911/253461/Spherical_Spline_Quaternions_For_Dummies.php

Ogawa, M., Lee, S., & Okada, Y. (2017). Analyzing Relationship between Upper Body Poses and Emotions Using Multiple Regression Analysis. *Hong Kong*, 4.

Proffitt, D. R., & Gilden, D. L. (1989). Understanding natural dynamics. *Journal of Experimental Psychology: Human Percetpion and Performance*, *15*(2), 384–393.

Qiang, W., & Xingfen, W. (2012). Study of interpolation algorithm based on velocity function. *2012 5th International Congress on Image and Signal Processing*, 761–765. https://doi.org/10.1109/CISP.2012.6469773

Reitsma, P. S. A., Andrews, J., & Pollard, N. S. (2008). Effect of Character Animacy and Preparatory Motion on Perceptual Magnitude of Errors in Ballistic Motion. *Computer Graphics Forum*, *27*(2), 201–210. https://doi.org/10.1111/j.1467-8659.2008.01117.x

Smith, M. (2013). *Applications of Dual Quaternions in Three Dimensional Transformation and Interpolation*. 36.

So, C. K. F., & Baciu, G. (2005). Entropy-based motion extraction for motion capture animation. *Computer Animation and Virtual Worlds*, *16*(3–4), 225–235. https://doi.org/10.1002/cav.107

Thomas, F., & Johnston, O. (1981). *Disney Animation: The Illusion of Life*. Abbeville Press.

Unity Forum: Right Hand to Left Handed Conversions. (2009, April). https://forum.unity.com/threads/right-hand-to-left-handed-conversions.80679/

Unity Script Reference: Quaternion. (2019). https://docs.unity3d.com/ScriptReference/Quaternion.html

Vail, N. (2017, May). *Understanding Linear Interpolation in UI Animation*. https://www.freecodecamp.org/news/understanding-linear-interpolation-in-ui-animations-74701eb9957c/

Van Welbergen, H., Van Basten, B. J. H., Egges, A., Ruttkay, Zs. M., & Overmars, M. H. (2010). Real Time Animation of Virtual Humans: A Trade-off Between Naturalness and Control. *Computer Graphics Forum*, *29*(8), 2530–2554. https://doi.org/10.1111/j.1467-8659.2010.01822.x

Van Welbergen, Herwin, Yaghoubzadeh, R., & Kopp, S. (2014). AsapRealizer 2.0: The Next Steps in Fluent Behavior Realization for ECAs. In T. Bickmore, S. Marsella, & C. Sidner (Eds.), *Intelligent Virtual Agents* (Vol. 8637, pp. 449–462). Springer International Publishing. https://doi.org/10.1007/978-3-319-09767-1_56

Vicovaro, M., Hoyet, L., Burigana, L., & O'sullivan, C. (2014). Perceptual Evaluation of Motion Editing for Realistic Throwing Animations. *ACM Transactions on Applied Perception*, *11*(2), 1–23. https://doi.org/10.1145/2617916

Wang, Y., Lang, F., Wang, Z., & Xu, B. (2015). Automatic Variable-timing Animation Transition based on Hierarchical Interpolation Method: *Proceedings of the 10th International Conference on Computer Graphics Theory and Applications*, 309–316. https://doi.org/10.5220/0005264703090316

Web3D Consortium. (2006). *ISO/IEC 19774—Humanoid Animation*. https://www.web3d.org/documents/specifications/19774/V1.0/index.html

X Bot. (n.d.). Mixamo Characters. https://www.mixamo.com/#/?page=3&type=Character

Xiao, Y., Zhang, Z., Beck, A., Yuan, J., & Thalmann, D. (2014). Human–Robot Interaction by Understanding Upper Body Gestures. *Presence: Teleoperators and Virtual Environments*, *23*(2), 133–154. https://doi.org/10.1162/PRES_a_00176

Zibrek, K., Hoyet, L., Ruhland, K., & Mcdonnell, R. (2015). Exploring the Effect of Motion Type and Emotions on the Perception of Gender in Virtual Humans. *ACM Transactions on Applied Perception*, *12*(3), 1–20. https://doi.org/10.1145/2767130

APPENDIX A. PILOT STUDY DATA

Table A.1. Clip combination and abbreviation

	Iconic Gesture	Deictic Gesture	Metaphoric Gesture	Beat Gesture
Linear Euler Interpolation	L-I	L-D	L-M	L-B
Spherical Linear Quaternion Interpolation	SL-I	SL-D	SL-M	SL-B
Spherical Spline Quaternion Interpolation	SQ-I	SQ-D	SQ-M	SQ-B

Table A.2. Pilot study data with 18 participants

Gender	**Experience**	**L-B**	**SL-B**	**SQ-B**	**L-D**	**SL-D**	**SQ-D**	**L-I**	**SL-I**	**SQ-I**	**L-M**	**SL-M**	**SQ-M**
Female	Novice	3	3	3	1	1	2	2	1	3	2	3	4
Female	Professional	2	3	4	3	2	3	4	2	5	3	2	5
Male	Professional	1	1	4	2	1	4	1	1	4	2	1	3
Male	Professional	2	3	4	2	1	4	4	3	5	2	2	4
Male	Professional	2	3	4	2	3	3	4	2	5	2	2	2
Male	Novice	1	1	3	1	1	1	2	1	4	5	5	4
Male	Professional	2	3	4	3	2	4	4	2	4	4	2	5
Male	Professional	3	2	2	2	2	3	2	4	3	2	1	4
Female	Novice	2	4	5	1	1	2	5	3	5	4	2	5
Other	Novice	2	2	3	2	1	3	1	2	3	1	2	4
Male	Professional	3	2	4	1	2	4	2	1	5	3	3	5
Other	Novice	2	2	5	1	1	1	3	3	3	1	1	1

Male	Novice	1	1	2	1	1	2	1	1	3	1	1	1
Female	Novice	3	4	4	3	2	5	3	3	3	3	3	4
Male	Novice	3	3	4	4	2	3	2	3	4	3	4	4
Male	Professional	1	1	2	1	1	2	1	1	2	1	1	2
Male	Novice	2	2	3	1	1	1	3	4	4	3	3	3
Male	Novice	2	2	2	2	2	2	3	2	3	3	3	3

APPENDIX B. MAIN STUDY DATA

Table B.1. Clip combination and abbreviation

	Iconic Gesture	Deictic Gesture	Metaphoric Gesture	Beat Gesture
Linear Euler Interpolation	L-I	L-D	L-M	L-B
Spherical Linear Quaternion Interpolation	SL-I	SL-D	SL-M	SL-B
Spherical Spline Quaternion Interpolation	SQ-I	SQ-D	SQ-M	SQ-B

Table B.2. Main study data with 97 participants

Gender	**Experience**	**L-B**	**SL-B**	**SQ-B**	**L-D**	**SL-D**	**SQ-D**	**L-I**	**SL-I**	**SQ-I**	**L-M**	**SL-M**	**SQ-M**
Male	Novice	2	3	4	4	2	3	3	2	4	3	2	4
Male	Professional	3	3	4	3	3	4	3	3	3	3	3	3
Male	Novice	5	5	5	5	4	5	3	3	4	4	3	3
Female	Professional	1	2	1	3	1	4	1	1	3	1	1	4
Male	Novice	3	3	4	3	2	4	4	3	5	2	3	3
Male	Professional	2	2	4	4	4	5	1	2	5	2	2	3
Female	Novice	4	1	3	3	2	3	1	2	3	3	4	4
Male	Novice	4	3	4	5	4	5	4	3	4	4	3	5
Male	Professional	4	2	4	1	1	2	2	2	4	1	2	3
Female	Professional	3	3	5	4	4	5	3	4	4	4	5	3
Male	Professional	2	4	5	4	4	5	1	2	4	2	4	2

Female	Professional	2	3	4	3	3	4	2	3	5	2	3	4
Female	Novice	3	2	4	3	4	3	2	3	4	3	2	3
Male	Novice	1	3	4	1	3	5	1	2	5	1	2	4
Female	Novice	3	2	3	4	3	4	2	3	2	4	4	5
Female	Novice	3	4	5	4	5	4	4	2	4	3	3	2
Male	Novice	2	3	4	3	3	4	3	1	4	2	2	3
Male	Professional	1	1	3	2	2	3	2	1	4	2	1	4
Female	Professional	2	2	3	3	2	3	3	2	4	4	4	4
Female	Professional	2	1	4	4	3	4	1	3	4	1	1	3
Female	Novice	2	2	2	4	4	4	3	2	4	4	4	5
Female	Novice	3	1	3	3	2	4	3	2	4	3	1	2
Female	Novice	3	3	4	2	3	2	3	4	2	2	3	4
Male	Novice	3	2	3	2	3	4	3	3	5	3	1	4
Male	Professional	2	3	4	3	3	3	2	2	3	3	3	3
Male	Professional	2	3	3	4	3	3	2	2	2	4	5	4
Female	Novice	2	2	1	3	1	2	1	3	5	2	4	5
Male	Novice	2	2	3	4	2	3	2	3	4	2	2	3
Male	Professional	3	2	4	2	2	3	3	2	3	3	2	4
Female	Novice	2	2	4	2	4	3	3	2	4	2	3	4
Male	Novice	2	2	5	2	3	3	2	2	4	3	3	4
Female	Professional	1	1	2	3	2	4	1	1	4	2	2	4
Male	Novice	3	4	4	3	4	4	3	2	3	2	2	3
Female	Novice	1	3	2	5	4	3	2	3	4	4	2	5
Male	Novice	2	3	3	3	2	4	2	2	4	2	1	3
Female	Novice	2	2	4	3	2	3	2	1	3	2	4	3

Female	Novice	1	2	2	2	2	3	1	2	4	1	2	4
Female	Novice	2	2	3	3	3	4	1	2	4	1	1	3
Female	Novice	3	3	4	2	3	2	4	4	4	4	5	4
Male	Professional	2	2	3	2	2	3	4	3	3	3	3	3
Male	Professional	3	2	4	3	3	5	3	2	4	1	2	2
Female	Professional	2	3	4	3	4	5	3	4	4	3	2	3
Other	Novice	1	1	2	2	2	3	2	2	3	2	3	3
Male	Professional	3	2	3	4	4	5	4	3	4	4	5	5
Female	Professional	1	1	4	2	3	5	1	1	4	3	3	3
Other	Novice	1	1	2	2	3	4	1	2	3	2	2	3
Male	Professional	1	1	2	1	2	2	1	2	3	2	2	2
Female	Novice	1	2	3	4	3	4	4	4	3	3	2	4
Female	Novice	3	3	2	5	4	4	3	4	4	2	3	3
Female	Novice	3	2	4	3	4	5	2	3	4	2	3	4
Male	Professional	1	1	2	2	1	3	1	1	3	1	1	2
Male	Novice	2	1	3	5	2	5	2	4	2	3	3	4
Female	Novice	2	5	4	3	2	2	4	3	4	2	2	1
Male	Professional	2	4	3	1	3	3	3	2	4	4	1	3
Male	Professional	3	4	5	2	2	5	1	1	5	2	3	5
Male	Novice	2	2	2	4	4	4	3	3	4	1	2	2
Female	Novice	2	3	4	3	3	4	2	2	4	1	1	2
Female	Novice	2	3	2	4	4	3	2	1	3	4	3	4
Female	Novice	1	2	4	3	4	3	2	5	5	3	4	2
Female	Novice	4	4	3	3	3	3	1	5	4	3	1	4
Male	Novice	4	4	4	3	4	3	3	3	4	3	3	4

Other	Novice	4	4	4	4	3	5	2	1	3	2	2	4
Female	Novice	5	5	4	5	5	5	3	4	4	3	4	4
Female	Novice	3	4	4	4	5	5	3	4	5	2	1	2
Female	Novice	2	1	2	4	2	3	3	3	4	2	3	4
Female	Novice	3	4	4	3	4	4	4	4	5	3	3	4
Male	Novice	3	4	4	2	3	2	5	2	5	1	2	2
Male	Professional	2	1	4	2	3	5	1	1	4	1	3	4
Male	Novice	3	4	4	3	4	3	1	2	4	2	1	1
Female	Novice	3	4	1	5	4	4	4	4	3	4	3	2
Male	Novice	3	3	5	3	4	4	1	2	2	4	2	4
Male	Novice	5	2	4	4	4	5	3	4	5	4	3	3
Female	Novice	3	2	4	2	3	3	3	4	3	2	3	2
Other	Novice	4	2	3	2	4	3	3	5	4	4	2	5
Female	Novice	3	3	4	4	4	3	2	2	4	1	2	3
Male	Novice	1	3	3	4	4	5	3	2	3	1	2	2
Male	Novice	1	1	4	4	3	5	4	5	5	3	3	4
Female	Novice	5	3	4	2	4	3	3	4	4	2	4	3
Female	Novice	1	3	3	3	4	2	2	2	3	3	3	4
Male	Novice	2	2	2	4	3	3	3	2	3	2	3	2
Female	Novice	1	3	3	2	3	4	3	1	4	1	3	4
Female	Novice	2	2	1	2	4	3	2	3	2	1	2	1
Female	Novice	3	4	3	4	4	4	4	4	4	4	3	3
Female	Novice	3	3	4	2	3	3	3	2	2	2	1	3
Male	Novice	4	1	3	2	1	3	4	2	5	3	2	1
Male	Novice	5	5	4	3	4	4	4	4	4	3	3	4

Female	Novice	3	3	3	4	4	4	3	3	3	4	4	4
Male	Professional	4	2	3	1	1	1	2	1	2	3	4	4
Female	Novice	4	2	3	2	2	3	3	4	2	2	3	4
Other	Novice	2	3	4	4	3	5	2	3	4	3	3	4
Female	Novice	4	2	3	5	2	4	2	2	1	1	4	2
Female	Novice	2	2	4	3	2	3	3	2	3	1	1	3
Male	Novice	3	1	2	3	3	2	2	3	3	4	4	4
Female	Novice	4	3	4	4	4	3	3	3	2	2	1	2
Male	Novice	2	2	3	1	2	3	3	2	4	2	2	3
Male	Novice	3	3	4	4	2	5	3	2	4	4	3	5
Female	Novice	3	3	4	4	4	3	4	4	4	2	2	2

APPENDIX C. ANALYSIS RESULT

Table C.1. Linear regression table for the pilot study

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept) LinEuler	2.2500	0.1263	17.810	<2e-16
(Method) Slerp	-0.1806	0.1787	-1.011	0.313
(Method) Squad	1.1111	0.1787	6.219	2.61e-09

Table C.2. ANOVA table for the pilot study

	Df	Sum Sq	Mean Sq	F value	Pr(>|F|)
Method	2	70.45	35.23	30.66	1.99e-12
Residuals	213	244.76	1.15		

Table C.3. Tukey HSD table for the pilot study

pair	diff	lwr	upr	p adj
Slerp-LinEuler	-0.1805556	-0.6022368	0.2411256	0.5709836
Squad-LinEuler	1.1111111	0.6894299	1.5327923	0.0000000
Squad-Slerp	1.2916667	0.8699855	1.7133479	0.0000000

Table C.4. Linear regression table for the combined main study

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept) LinEuler	2.65979	0.05241	50.749	<2e-16
(Method) Slerp	0.04124	0.07412	0.556	0.578
(Method) Squad	0.83505	0.07412	11.266	<2e-16

Table C.5. ANOVA table for the combined main study

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Method	2	171.9	85.95	80.64	<2e-16
Residuals	1161	1237.4	1.07		

Table C.6. Tukey HSD table for the combined main study

pair	diff	lwr	upr	p adj
Slerp-LinEuler	0.04123711	-0.1327027	0.2151769	0.8432303
Squad-LinEuler	0.83505155	0.6611117	1.0089913	0.0000000
Squad-Slerp	0.79381443	0.6198746	0.9677542	0.0000000

Table C.7. Linear regression table for the Prolific study

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept) LinEuler	2.83333	0.08182	34.628	<2e-16
(Method) Slerp	0.05952	0.11571	0.514	0.607
(Method) Squad	0.55952	0.11571	4.835	1.77e-06

Table C.8. ANOVA table for the Prolific study

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Method	2	31.7	15.865	14.11	1.1e-06
Residuals	501	563.5	1.125		

Table C.9. Tukey HSD table for the Prolific study

pair	diff	lwr	upr	p adj
Slerp-LinEuler	0.05952381	-0.2124807	0.3315283	0.8643576
Squad-LinEuler	0.55952381	0.2875193	0.8315283	0.0000053
Squad-Slerp	0.50000000	0.2279955	0.7720045	0.0000556

Table C.10. Linear regression table for the novice subjects

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept) LinEuler	2.77778	0.06009	46.227	<2e-16
(Method) Slerp	0.03472	0.08498	0.409	0.683
(Method) Squad	0.69097	0.08498	8.131	1.48e-15

Table C.11. ANOVA table for the novice subjects

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Method	2	87.3	43.65	41.97	<2e-16
Residuals	861	895.4	1.04		

Table C.12. Tukey HSD table for the novice subjects

pair	diff	lwr	upr	p adj
Slerp-LinEuler	0.03472222	-0.1647921	0.2342365	0.912105
Squad-LinEuler	0.69097222	0.4914579	0.8904865	0.000000
Squad-Slerp	0.65625000	0.4567357	0.8557643	0.000000

Table C.13. Linear regression table for the experienced subjects

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept) LinEuler	2.3200	0.1025	22.642	<2e-16
(Method) Slerp	0.0600	0.1449	0.414	0.679
(Method) Squad	1.2500	0.1449	8.626	3.89e-16

Table C.14. ANOVA table for the experienced subjects

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Method	2	99.41	49.70	47.34	<2e-16
Residuals	297	311.83	1.05		

Table C.15. Tukey HSD table for experienced subjects

pair	**diff**	**lwr**	**upr**	**p adj**
Slerp-LinEuler	0.06	-0.281337	0.401337	0.9098685
Squad-LinEuler	1.25	0.908663	1.591337	0.0000000
Squad-Slerp	1.19	0.848663	1.531337	0.0000000

Table C.16. Linear regression table for the beat gesture

	Estimate	**Std. Error**	**t value**	**Pr(>|t|)**
(Intercept) LinEuler	2.54639	0.10515	24.217	<2e-16
(Method) Slerp	0.01031	0.14870	0.069	0.945
(Method) Squad	0.83505	0.14870	5.616	4.61e-08

Table C.17. ANOVA table for the beat gesture

	Df	**Sum Sq**	**Mean Sq**	**F value**	**Pr(>F)**
Method	2	44.54	22.271	20.77	3.76e-09
Residuals	288	308.87	1.072		

Table C.18. Tukey HSD table for the beat gesture

pair	**diff**	**lwr**	**upr**	**p adj**
Slerp-LinEuler	0.01030928	-0.3400181	0.3606366	0.9973536
Squad-LinEuler	0.83505155	0.4847242	1.1853789	0.0000001
Squad-Slerp	0.82474227	0.4744149	1.1750696	0.0000002

Table C.19. Linear regression table for the deictic gesture

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept) LinEuler	3.08247	0.10220	30.160	<2e-16
(Method) Slerp	-0.04124	0.14454	-0.285	0.77562
(Method) Squad	0.54639	0.14454	3.780	0.00019

Table C.20. ANOVA table for the deictic gesture

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Method	2	20.87	10.436	10.3	4.78e-05
Residuals	288	291.81	1.013		

Table C.21. Tukey HSD table for the deictic gesture

pair	diff	lwr	upr	p adj
Slerp-LinEuler	-0.04123711	-0.3817569	0.2992827	0.9561301
Squad-LinEuler	0.54639175	0.2058720	0.8869116	0.0005571
Squad-Slerp	0.58762887	0.2471091	0.9281487	0.0001825

Table C.22. Linear regression table for the iconic gesture

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept) LinEuler	2.51546	0.10061	25.002	<2e-16
(Method) Slerp	0.09278	0.14229	0.652	0.515
(Method) Squad	1.15464	0.14229	8.115	1.42e-14

Table C.23. ANOVA table for the iconic gesture

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Method	2	79.84	39.92	40.66	2.8e-16
Residuals	288	282.78	0.98		

Table C.24. Tukey HSD table for the iconic gesture

pair	diff	lwr	upr	p adj
Slerp-LinEuler	0.09278351	-0.2424258	0.4279928	0.7913306
Squad-LinEuler	1.15463918	0.8194299	1.4898484	0.0000000
Squad-Slerp	1.06185567	0.7266464	1.3970649	0.0000000

Table C.25. Linear regression table for the metaphoric gesture

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept) LinEuler	2.4948	0.1046	23.842	<2e-16
(Method) Slerp	0.1031	0.1480	0.697	0.487
(Method) Squad	0.8041	0.1480	5.434	1.18e-07

Table C.26. ANOVA table for the metaphoric gesture

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Method	2	37.14	18.570	17.48	6.82e-08
Residuals	288	305.90	1.062		

Table C.27. Tukey HSD table for the metaphoric gesture

pair	diff	lwr	upr	p adj
Slerp-LinEuler	0.1030928	-0.2455467	0.4517322	0.7656414
Squad-LinEuler	0.8041237	0.4554843	1.1527632	0.0000004
Squad-Slerp	0.7010309	0.3523915	1.0496704	0.0000102

Table C.28. Linear regression table for the male subjects

	Estimate	**Std.Error**	**t value**	**Pr(>|t|)**
(Intercept) LinEuler	2.65116	0.07846	33.790	<2e-16
(Method) Slerp	-0.09302	0.11096	-0.838	0.402
(Method) Squad	0.93605	0.11096	8.436	3.36e-16

Table C. 29. ANOVA table for the male subjects

	Df	**Sum Sq**	**Mean Sq**	**F value**	**Pr(>F)**
Method	2	111.4	55.72	52.63	<2e-16
Residuals	513	543.2	1.06		

Table C.30. Tukey HSD table for the male subjects

pair	**diff**	**lwr**	**upr**	**p adj**
Slerp-LinEuler	-0.09302326	-0.3538371	0.1677906	0.6793416
Squad-LinEuler	0.93604651	0.6752327	1.1968603	0.0000000
Squad-Slerp	1.02906977	0.7682559	1.2898836	0.0000000

Table C.31. Linear regression table for the female subjects

	Estimate	**Std. Error**	**t value**	**Pr(>|t|)**
(Intercept) LinEuler	2.68878	0.07386	36.405	<2e-16
(Method) Slerp	0.15306	0.10445	1.465	0.143
(Method) Squad	0.71939	0.10445	6.887	1.47e-11

Table C.32. ANOVA table for the female subjects

	Df	**Sum Sq**	**Mean Sq**	**F value**	**Pr(>F)**
Method	2	56.3	28.148	26.33	1.13e-11
Residuals	585	625.5	1.069		

Table C.33. Tukey HSD table for the female subjects

Pair	**diff**	**lwr**	**upr**	**p adj**
Slerp-LinEuler	0.1530612	-0.09236406	0.3984865	0.3084079
Squad-LinEuler	0.7193878	0.47396247	0.9648130	0.0000000
Squad-Slerp	0.5663265	0.32090124	0.8117518	0.0000003

APPENDIX D. PILOT STUDY SURVEY FORMAT

PURDUE
UNIVERSITY

Description

What is the purpose of this study?

The purpose of this study is to **evaluate the naturalness of a set of animated clips** generated using 3 different interpolation methods

What will you do?

You will view 12 short animated clips: **4 sets of 3 different interpolation types**

You will be asked to rate the naturalness from **not natural at all** to **very natural**

Hong long is the survey?

7 to 8 minutes

NOTE:

Please take this survey with your computer (ideally) under good internet
Please take at least 20 seconds to watch each clip, clips are looped
Pay particular attention to the speed and pauses of the arm movement

→

What is your age?

- 17 or younger
- 18-59
- 60 or older

What is your gender?

- Male
- Female
- Other

What is your race?

- White
- Black or African American
- American Indian or Alaska Native
- Asian
- Native Hawaiian or Pacific Islander
- Other

What is your highest completed education?

- High school degree or lower
- Associate degree
- Bachelor degree
- Graduate degree or higher

What is your experience in Animation?

- Novice
- Professional, Major in Animation (or Game, VFX)

← →

In the following section, you will see **3** clips of a character **throwing an object**.

Please watch it carefully for at least **20** seconds (clips are looped)

Pay attention to the change of speed on the arm motion.

→

[The character is throwing something]

How natural is the animation?

Not natural at all		Neutral		Very natural
1	2	3	4	5
○	○	○	○	○

→

[The character is throwing something]

How natural is the animation?

Not natural at all		Neutral		Very natural
1	2	3	4	5
○	○	○	○	○

→

[The character is throwing something]

How natural is the animation?

Not natural at all		Neutral		Very natural
1	2	3	4	5
○	○	○	○	○

→

[The character is throwing something]

How natural is the animation?

Not natural at all		Neutral		Very natural
1	2	3	4	5

→

In the following section, you will see **3** clips of a character **pointing to a direction.**

Please watch it carefully for at least **20** seconds (clips are looped)

Pay attention to the change of speed on the arm motion.

→

[The character is pointing]

How natural is the animation?

Not natural at all		Neutral		Very natural
1	2	3	4	5
○	○	○	○	○

→

[The character is pointing]

How natural is the animation?

Not natural at all		Neutral		Very natural
1	2	3	4	5
○	○	○	○	○

→

[The character is pointing]

How natural is the animation?

Not natural at all		Neutral		Very natural
1	2	3	4	5

→

In the following section, you will see **3** clips of a character **moving both arms forward**.

Please watch it carefully for at least **20** seconds (clips are looped)

Pay attention to the change of speed on the arm motion.

→

[The character is moving both hands forward]

How natural is the animation?

Not natural at all		Neutral		Very natural
1	2	3	4	5
○	○	○	○	○

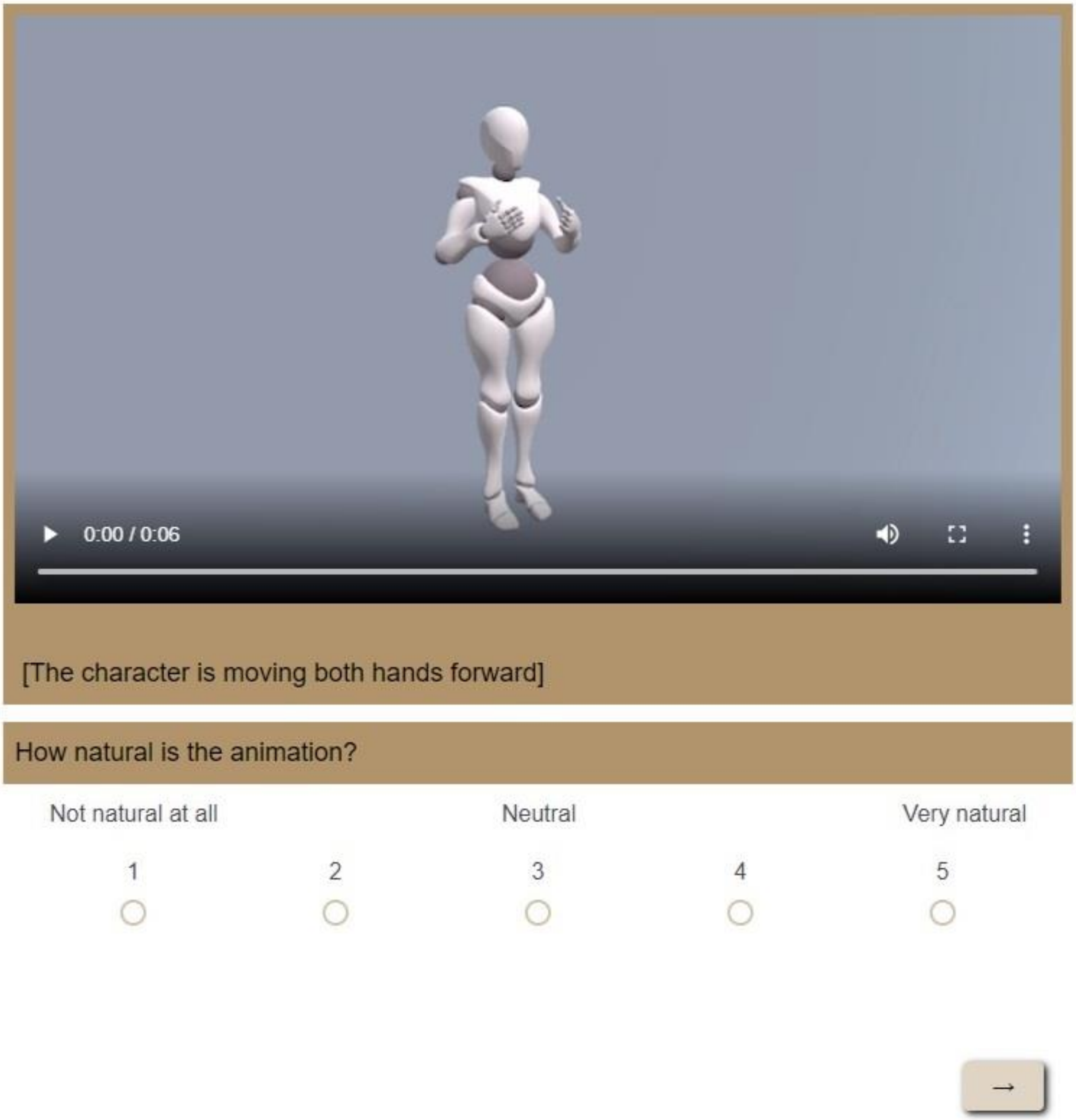
→

[The character is moving both hands forward]

How natural is the animation?

Not natural at all		Neutral		Very natural
1	2	3	4	5
○	○	○	○	○

→



[The character is moving both hands forward]

How natural is the animation?

Not natural at all		Neutral		Very natural
1	2	3	4	5
○	○	○	○	○

→

In the following section, you will see **3** clips of a character **showing the width of an object.**

Please watch it carefully for at least **20** seconds (clips are looped)

Pay attention to the change of speed on the arm motion.

→

[The character is showing the width of an object]

How natural is the animation?

Not natural at all		Neutral		Very natural
1	2	3	4	5
○	○	○	○	○

→

[The character is showing the width of an object]

How natural is the animation?

Not natural at all		Neutral		Very natural
1	2	3	4	5
○	○	○	○	○

→

[The character is showing the width of an object]

How natural is the animation?

Not natural at all		Neutral		Very natural
1	2	3	4	5
○	○	○	○	○

→

APPENDIX E. MAIN STUDY SURVEY FORMAT

PURDUE
UNIVERSITY

What is the purpose of this study?

Evaluate the naturalness of character animation generated using 3 different interpolation methods

What will you do?

1. Watch 12 short animated clips: **4 sets of 3 different interpolation types**
2. Rate their naturalness from **not natural at all** to **very natural**

(Naturalness: how realistic a character behaves compared to a real human)

Hong long is the survey?

4 to 8 minutes

NOTE:

-Please take this survey with your computer under good internet connection

-Please take at least 10 seconds to watch each clip

-Pay close attention to the speed and pauses of the arm movement

→

What is your age?

- 17 or younger
- 18 or older
- Prefer not to provide

What is your gender?

- Male
- Female
- Other
- Prefer not to provide

What is your race?

- White
- Black or African American
- American Indian or Alaska Native
- Asian
- Native Hawaiian or Pacific Islander
- Other
- Prefer not to provide

What is the highest level of education you completed?

- High school degree or lower
- Associate degree
- Bachelor degree
- Graduate degree or higher
- Prefer not to provide

What is your experience in Animation?

- Novice
- Professional, Major in Animation (or Game, VFX)

← →

In this section, you will see **3** clips in which a character performs a **outward gesture** (showing the width of an object)

The clips are different, so please watch each clip carefully for at least **10** seconds. Pay close attention to the arm motion

→

[Outward gesture]

How natural is the animation?

Not natural at all				Very natural
1	2	3	4	5
○	○	○	○	○

← →

[Outward gesture]

How natural is the animation?

Not natural at all				Very natural
1	2	3	4	5
○	○	○	○	○

←

→

[Outward gesture]

How natural is the animation?

Not natural at all				Very natural
1	2	3	4	5
○	○	○	○	○

← →

In this section, you will see **3** clips in which a character performs a **throwing gesture** (using one arm to throw an object)

The clips are different, so please watch each clip carefully for at least **10** seconds. Pay close attention to the arm motion

→

[Throwing gesture]

How natural is the animation?

Not natural at all				Very natural
1	2	3	4	5
○	○	○	○	○

← →

[Throwing gesture]

How natural is the animation?

Not natural at all				Very natural
1	2	3	4	5
○	○	○	○	○

← →

[Throwing gesture]

How natural is the animation?

Not natural at all				Very natural
1	2	3	4	5
○	○	○	○	○

← →

In this section, you will see **3** clips in which a character performs a **beat gesture** (moving both arms parallel forward to emphasize information during speech)

The clips are different, so please watch each clip carefully for at least **10** seconds. Pay close attention to the arm motion

→

[Beat gesture]

How natural is the animation?

Not natural at all				Very natural
1	2	3	4	5
○	○	○	○	○

← →

[Beat gesture]

How natural is the animation?

Not natural at all				Very natural
1	2	3	4	5
○	○	○	○	○

← →

[Beat gesture]

How natural is the animation?

Not natural at all				Very natural
1	2	3	4	5
○	○	○	○	○

← →

In this section, you will see **3** clips in which a character performs a **pointing gesture** (moving one arm to point towards a direction)

The clips are different, so please watch each clip carefully for at least **10** seconds. Pay close attention to the arm motion

→

[Pointing gesture]

How natural is the animation?

Not natural at all				Very natural
1	2	3	4	5
○	○	○	○	○

← →

[Pointing gesture]

How natural is the animation?

Not natural at all

Very natural

1 2 3 4 5

← →

[Pointing gesture]

How natural is the animation?

Not natural at all | Very natural

1	2	3	4	5
○	○	○	○	○

← →