INFLUENCE OF ENVIRONMENTAL CONTEXT ON AFFECT RECOGNITION OF STYLIZED MOVEMENTS

BY

MADISON HEIMERDINGER

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Mechanical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2017

Urbana, Illinois

Adviser:

Assistant Professor Amy LaViers

Abstract

Modifying the style of movements will be an important component of robotic interaction as more and more robots move into human-facing scenarios where humans are (consciously or unconsciously) constantly monitoring the motion profile of counterparts in order to make judgments about the state of these counterparts. This thesis includes two main contributions: (1) the development of two MATLAB tools that are designed to aid in the creation and simulation of stylized movement trajectories in varied contexts and (2) three user studies that explore the effects of environmental context on a human's perception of stylized movement.

First and foremost, the results from all of the user studies indicate that environmental contexts and stylized walking sequences both impact affect recognition. In the first two studies, participants were asked to categorize stimuli as one of seven affective labels. The results show that the labels were not applied consistently and so it was concluded that the affect of a multi-dimensional stimuli cannot be adequately categorized using a single affective label. In the third study the stimuli were evaluated on multiple scales and classified using ratings of valence and arousal rather than affective labels. The results were used to create a least squares model for the dataset that decomposed the affect ratings of animations to display the compound effects of stylized walking sequences and environmental contexts on affective ratings.

Acknowledgments

First and foremost, I would like to sincerely thank my advisor and mentor, Amy, for all of the guidance and support you have provided me on my research journey. Without your help, I would not have been able to complete all of the goals I hoped to accomplish in graduate school. You challenged me to think differently and utilize tools from disciplines that are not typically viewed as 'engineering related'. Thanks to you, not only do I feel that I am leaving Illinois as a better engineer, but I feel like I am leaving Illinois as a more competent and well rounded individual.

I would also like to thank everyone who I have had the pleasure of working with in the RAD Lab. To Umer and Ishaan, thank you for always being open to discuss questions I had about my research and academics. Your support and willingness to help made it possible for me to successfully complete my research projects. To Catie, the month you spent in the RAD Lab teaching about movement was invaluable. It seems like I found inspiration and new ways to further my research everyday I worked with you. I would also like to Aimy Wissa for providing me with the motivation to attend graduate school.

To Mike, thank you for always supporting me regardless of the circumstance. You have been extremely supportive and have pushed me to preserver even when I felt like there was no light at the end of tunnel. I could not have done this without your love and support. Lastly, to my parents, Daniel Heimerdinger and Evelyn Vance, I can never thank you enough for everything you have done for me. You have both been exceptional engineering role models for me. Without your love, support, and guidance I would have never made it to where I am.

Table of Contents

List of Tables . . . . **vi**

List of Figures . . . . **vii**

Chapter 1 Stylized Movement Studies **1**
1.1 Machine-Driven Style Recognition 2
1.2 Human Perception of Stylized Movements 3

Chapter 2 Frameworks for Quantifying Style and Affect **7**
2.1 Laban/Bartenieff Movement Studies (LBMS) 7
2.2 Affective Domain of Modern Psychology 10

Chapter 3 Methods and Tools for Generating Data-Driven Stimuli **14**
3.1 Generating Baseline Stylized Walking Sequences 14
3.2 Stimuli Creation Process for Thesis User Studies 16
3.3 MATLAB Tool for Viewing Simulations of Data Driven Trajectories in Environments 19
3.4 MATLAB Tool for Creating Simulations of Data Driven Trajectories in Environments 26

Chapter 4 Study 1: Exploring Context-Dependency of Previous Labeling of Affective Gait **34**
4.1 Images Selected to Simulate Environmental Context 34
4.2 Questionnaire Format 36
4.3 Results and Discussion 37

Chapter 5 Study 2: Forced Choice Labeling of Stylized Gait **41**
5.1 Questionnaire Format 41
5.2 Results and Discussion 42

Chapter 6 Study 3: Modeling the Compound Effects of Environment and Gait on Perceived Affect **46**
6.1 Questionnaire Format 46
6.2 Calculating Affective Ratings 49
6.3 Results and Discussion 50

Chapter 7 Towards Understanding the Relationship Between Movement, Context, and Affect Recognition **60**
7.1 Publications 62

Appendix A Read Me for V-SDDTE **63**

Appendix B Read Me for C-SDDTE **70**

Appendix C Screen Captures of User Study Questions **75**

Appendix D Additional Results from User Studies **78**

References **80**

List of Tables

2.1 LBMS Effort Motion Factor Qualities 9

4.1 Affect Selection for Neutral Walking Sequence in Study 1 38
4.2 Affect Ratings of White Background Sequences 38

5.1 Affect Selection for Neutral Walking Sequence in Study 2 43

6.1 Semantic Differential Rating Scales 47
6.2 Study 3 Stimuli Groupings 48
6.3 Study 3 Demographic Statistics 49

List of Figures

2.1 LBMS Basic Effort Actions . . . 10

3.1 Video Stimuli Created for User Studies . . . 17
3.2 V-SDDTE MATLAB Tool . . . 21
3.3 Environmental Context Display in a MATLAB Simulation . . . 23
3.4 3D Environment Incorporated in MATLAB Tool . . . 24
3.5 C-SDDTE MATLAB Tool . . . 27
3.6 Phases of Gait . . . 29
3.7 Baseline Trajectories Available in C-SDDTE . . . 31
3.8 Secondary Feature Panel in C-SDDTE . . . 32

4.1 Images Utilized in User Studies . . . 35
4.2 Circumplex Mapping of Images Utilized in User Studies . . . 35
4.3 Study 1 Results . . . 39

5.1 Total Label Selection Rates in Study 2 . . . 43
5.2 Label Selection Rates With Respect To Walking Sequences in Study 2 . . . 45
5.3 Label Selection Rates With Respect To Environmental Contexts in Study 2 . . . 45

6.1 Circumplex Mapping of Environmental Contexts Ratings from Study 3 and OASIS Reference Images . . . 51
6.2 Circumplex Mapping of Environmental Contexts Ratings from Study 3 and Alternate OASIS Images . . . 53
6.3 Circumplex Mapping of Original and Normalized Environmental Contexts Ratings from Study 3 . . . 54
6.4 Circumplex Mapping of Original and Normalized Stylized Walking Sequence Ratings from Study 3 . . . 55
6.5 Circumplex Mappings of Normalized Animation Ratings Separated by Environment from Study 3 . . . 56

A.1 Subfolder structure for V-SDDTE . . . 66

C.1 Example of Study 1 Question Type 1 and Study 2 Question . . . 75
C.2 Example of Study 1 Question Type 2 . . . 75
C.3 Example of Study 1 Question Type 3 . . . 76
C.4 Example of Study 3 Semantic Differential Scales . . . 76

C.5 Example of Study 3 Human Verification Question 77

D.1 Circumplex Mappings of Normalized Animation Ratings Separated by Walking Sequence from Study 3 78
D.2 Circumplex Mappings of Original Animation Ratings Separated by Environment from Study 3 79
D.3 Circumplex Mappings of Original Animation Ratings Separated by Walking Sequence from Study 3 79

Chapter 1

Stylized Movement Studies

In the coming years, automated systems will interact with human users in increasingly unstructured tasks. These systems, which may manifest as animated avatars or moving machines, need to tap into human movement patterns and conventions in order to operate successfully. An important characteristic of human movement that is considered in HRI studies is the ability of a human to display internal affect (emotions) through stylized movements. Internal human affect is externally manifested through multiple modalities including gestures, postures, and facial expressions. Stylized body movements are nonverbal ways in which affective information about the state of the mover can be recognized and understood by other agents in their environment.

In this thesis, we investigate the relationship between stylized movements, context, and affect recognition. This chapter will present previous work that focussed on stylized movements as well discussion about various applications where stylized movement studies are of particular interest. Section 1.1 discusses previous work in stylized movement recognition and Section 1.2 presents previous work in stylized movement generation. Chapter 2 presents existing frameworks that have been previously used to model subjective qualities, such as style and affect, using quantitative metrics. Chapter 3 presents the methods that were used for creating the stimuli used in the user studies presented in this thesis as well as the two MATLAB tools that were developed. Study 1 investigated environmental context selection rates when users are primed with the intended affect of a stylized walking sequence and is presented in Chapter 4. Study 2 investigated how human affect recognition rates of stylized walking sequences varied based on the environment where they were portrayed and is included in Chapter 5. Finally, the compound effects of environmental contexts and stylized gait on affective ratings that are explored in Study 3 are discussed in Chapter 6. Chapter 7 concludes the thesis with a discussion about how the work presented in this paper helps towards understanding the relationship between movement, context, and affect recognition.

1.1 Machine-Driven Style Recognition

Current research in stylistic movement recognition and generation can be broadly categorized as machine-driven studies or human-driven perception studies. The machine-driven studies often utilize machine learning algorithms to analyze and categorize the style of a movement sequence using information previously obtained from a training set. The machine learning algorithms used for stylized movement recognition are generally *supervised* learning algorithms. The goal of a supervised learning algorithm is to use a training data set to learn a function that maps inputs directly to outputs [1]. Both *generative* and *discriminative* learning algorithms are commonly used for affect recognition algorithms. To help explain the difference, consider x to be the input and y to be an output label. A generative learning model leverages a joint probability $p(x, y)$ to determine the most likely output label while a discriminative learning model would try to learn $p(y|x)$ directly [2].

In [3], Bernhardt and Robinson decomposed movement sequences of a single action into motion primitives and then used a family of binary support vector machines (SVMs) (supervised, discriminative classifier) to classify the style of each motion primitive. The single action sequence was decomposed into segments that had overall motion energy above a certain threshold and then the motion primitives were derived by k-means clustering of the decomposed sequences. The SVMs classified each segmented motion primitive as either neutral, happy, angry, or sad. The algorithm determined the classified the entire movement sequence using majority rule of the motion primitive classifications. Bernhardt and Robinson used this study as a pilot study for [4]. In [4], Bernhardt and Robinson created an algorithm to classify the style of action sequences consisting of knocking, throwing, lifting, and walking movements. A Hidden Markov Model (HMM) (supervised, generative classifier) was first used to segment the connected action sequences into isolated action. The isolated actions were then classified with SVMs. The process was similar to the process performed in [3].

In [5], Etemad and Arya created an artificial neural network (ANN), which they called a radial basis function neural network (RBFNN), for classifying stylized movements. The authors subtract a 'neutral' portion of an action sequence from the original sequence so only the 'stylized' portion of the action sequence remains. The algorithm they developed then decompose the remaining features into Gaussian radial basis functions (RBFs) which are used as input to the RBFNN. The RBFNN is a supervised learning algorithm that classified the input sequences as either happy, sad, energetic, tiered, young, or old.

In general, the overall goal of previous work in machine-driven style recognition studies has been to create algorithms that accurately classified a small set of well defined action inputs as a finite set of qualitative descriptions using kinematic motion parameters. The majority of the machine-driven studies ignore the role of context, operating under the assumption that movement alone can be affective [6].

1.2 Human Perception of Stylized Movements

Human perception of affect communicated through movement and context is of particular interest in the field of animation. Since animations do not face the same physical limitations that are present in real world scenarios, animated character movement and simulated contexts can be exaggerated to make the affective intent of the characters more evident. However, character movements and simulated contexts can be exaggerated past the point of believability and can cause the affective intent of the character to become less clear, as seen in [7]. Disney's Twelve Principles of Animation is one set of guidelines that animators commonly use to increase the believability of the animations [8]. These principles have also been used as guidelines for creating stylized robotic movement [9, 10].

Human perception studies generally consist of stylized movement generation that is validated by means of a user study. In [11], Etemad and Arya proposed a model that used a transformation function to convert the secondary features of an action sequence while maintaining the integrity of the base action. In a separate study, Etemad and Arya presented a modeling approach that utilized Gaussian radial basis functions (RBFs) to transform neutral base action sequences into action sequences with stylized features [7]. The modeling approach developed in [7] will be further described in Section 3.1.

Both [11] and [7] validate the quality of the transformed action sequences with user studies where participants were asked to rate how well a particular style was represented in the generated movement. In these user studies, the participants were not provided with context for where the movement was occurring and were told the intended style of the generated movement prior to evaluation.

1.2.1 Influence of Priming on Perception

In this thesis, the word *priming* will be used to describe individual bias that is associated with decision making. In this thesis, priming will be referred to as either *inherent* or *subliminal*. We consider priming to be inherent when a decision is influenced by preferences developed in the long-term. Inherent priming can be influenced by variables such as cultural upbringing, social bonds, and affluence/social class [12]. This definition follows the notion that past experience effect current perception [13].

The primary form of priming that is examined in this thesis will be referred to as subliminal priming. We consider priming to be subliminal when an individual is prompted with stimuli shortly before making a decision. subliminal priming can be intentional (i.e. designed to sway the decision) or unintentional (i.e. inadvertently influences the decision). Repetitiveness is a form of subliminal priming that has been shown to decrease motivation [14]. In user studies involving human evaluation, the order that questions are asked is often randomized to mitigate the affects of priming caused by repetition [7, 15, 16]. Relativity is another form of subliminal priming that can cause rating scales to become nonuniform. This occurs when a stimulus is rated with respect to the ratings of previous stimuli rather than on an absolute scale [13].

1.2.2 Role of Context in Affect Recognition

The affect experienced by a human viewing a stylized movement is also heavily dependent on the situation where the movement is being experienced or imagined. If no context (environment, observed subject, situation/task, etc.) is provided, there is a significant chance that the observer will misinterpret the affective intent of the movement [6]. It is worth noting that multiple studies in psychology present theories that the components of context are highly intertwined and cannot be considered independently when evaluating the influence of context on perception [12, 17]. To demonstrate the validity of this statement, consider the following example. Consider the following set of events:

1. a **monkey** that is *observing the surrounding environment* at a <u>beach on Phi Phi Island</u>
2. a **human** that is *observing the surrounding environment* at a <u>beach on Phi Phi Island</u>
3. a **human** that is *performing an experiment to assess an abnormality* at a <u>beach on Phi Phi Island</u>
4. a **human** that is *performing an experiment to assess an abnormality* on the <u>International Space Station</u>

The components of context that are included set of events just listed include observed subject (bold), situation/task (italic), and environment (underlined). Each event varies by one component of context from the event immediately proceeding it. Now consider potential perceptions for each of the events. In the first event, the monkey may seem intense and alert because it is monitoring for predators. In event 2, the human may seem lackadaisical and relaxed because they are watching the waves break on the beach. The action being performed by the human in event 3 might be perceived as directed or methodical as it may pertain to mundane work. The action in event 4 may be necessary to maintain the life of the astronauts and therefore might be interpreted as stressful. In each of these examples, it is demonstrated that perception is heavily dependent on multiple aspects of context. However, the individual effects of each component on the overall perception is unclear. In order to reduce the number of unknowns when investigating the compound effects of context on affect recognition, environment will be the only element of context that is examined in the pilot studies presented in this thesis.

Chapter 2

Frameworks for Quantifying Style and Affect

A major challenge in HRI studies is the ability to adequately assign quantitative metrics to represent qualitative traits. In this chapter, two existing framework for quantifying style and affect are presented. However, before discussing the frameworks, it is important to state the distinction between the terms style and affect. For the remainder of this paper, the term style will refer to physical variations in movement patterns that cause the same primary action (i.e. walk, throw, run, dive, etc.) to appear differently. The term affect will refer to an emotional response of a human that was in response to viewing some sort of stimulus. Section 2.1 presents a framework that is used in dance and movement analysis. Section 2.2 presents various methods used in modern psychology that are used to quantify affect.

2.1 Laban/Bartenieff Movement Studies (LBMS)

Different frameworks and models have been created across various disciplines in an attempt to quantify both style and affect. In dance and movement analysis, the Laban/Bartenieff Movement Studies (LBMS) is a framework that is used to classify and generate human movements. LBMS is divided into four categories: Body, Space, Shape, and Effort. Labanotation is a notation system that was developed to record movement using the LBMS framework. The movements are recorded on a staff where body parts are represented on the horizontal axis and the vertical axis represents time. In Labanotation, each LBMS category is represented by a different set of symbols. This section will provide a brief overview of each category and present robotics studies that have utilize aspects of each category.

The Body category is used to describe the organization of the human body during movement [18]. In [19], Huzaifa et al. proposed a strategy for creating bipedal walking that was initiated through core-actuation. Three of Bartenieff's Basic Six [20] principles of movement were used to analyze gait and the results of the analysis were used to develop a core-actuated bipedal walker simulation. In [21], LaViers et al. used high-level body movement abstractions, called Motif, to classify robotic actions. The symbols that were used are called the Basic Body Actions and included *flexion*, *extension*, *rotation*, *traveling*, *scatter*, and *gather*. This notation method was used so that the framework would not need to be altered if the robotic platform was changed.

The Space category is used to describe spatial orientation of the movement and mover [18]. In [22], Jang Sher proposed a platform-invariant framework for robotic spatial commands that was developed using 26 spatial directions from LBMS. The framework was validated in a user study where participants first selected Labanotation symbols for Body and Space to create a movement phrase for a specified context and then evaluated how well their intended movement was portrayed on two different robotic platforms (Baxter and NAO). In [23], Rett and Dias utilize a concept that describes Laban's spatial directions as lines of motion instead of points in space to develop a Bayesian model for gesture recognition. The previously described concept was developed in [24] by Longstaff, where he named lines of motion *Vector Symbols*.

The Shape category is used to describe how the body changes shape and moves within the environment [18]. In [25], Khoshhal et al. use characteristics of the Shape category to develop models that estimates human shape change in the vertical, horizontal, and sagittal planes. The resulting models were then used as features in a Bayesian Network for human motion recognition. In [26], Chi et al. present a system that uses features of the Shape and Effort categories to help create natural synthetic gestures in 3D character animations.

The Effort category is used to describe the qualitative aspect of movement [18]. The Effort category is the most widely used LBMS category in robotics studies. In [27], LaViers et al.

created a mapping between the Effort motion factors and the weights in a cost function that were used to create stylized robotic movements. In [28], Masuda et al. create equations (termed 'Laban's feature values') to represent the characteristics of the Effort category. The set of 'Laban's feature values' were implemented on a robotic platform in a user study where participants rated the level of emotion expressed through the robotic motion. The results were used to correlate emotion estimations with parameters of the Laban's feature value set. Masuda and Kato used [28] as a pilot study for [29], where they developed a method for adding target emotions to movements of a human form robot. In the RAD Lab at UIUC, quantified mappings of the the Effort motion factors weight, time, and space were created and are being used to create task-constrained motion generation of aerial robots. Effort is the LBMS category whose framework is most applicable to the topics discussed in this thesis. This section will conclude with a high-level overview of the Effort category framework.

2.1.1 LBMS Effort Category

Effort category is a framework that describes the inner intent of the mover as manifested in motion quality. The Effort category is composed of the Motion Factors (*Flow, Weight, Space,* and *Time*) [18]. Each motion factor is described by a continuum between an *Indulging* quality and a *Fighting* quality. The qualities associated with each motion factor are shown in Table 2.1.

Table 2.1: LBMS Effort Motion Factor Qualities

Motion Factor	***Indulging***	***Fighting***
Flow	Free	Bound
Weight	Light	Strong
Space	Indirect	Direct
Time	Sustained	Sudden

When the extreme notions of the Motion Factors *Weight*, *Space*, and *Time* are combined

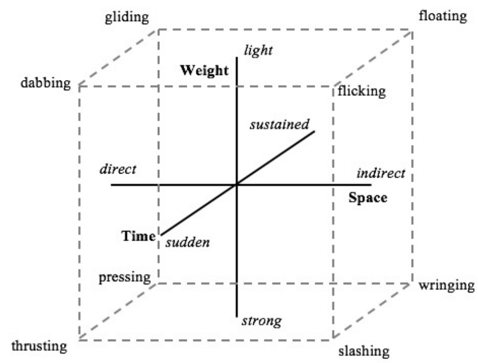

Figure 2.1: This diagram shows the Motion Factor qualities that each Basic Effort Action is comprised of. The Motion Factor qualities of a Basic Effort Action are the extremes closest to the corner where the Basic Effort Action is mapped. For instance, the motion factor qualities that describe 'dabbing' are *direct*, *sudden*, and *light*.

they create a set of descriptions known as the Basic Effort Actions. The Basic Effort Actions include dabbing, gliding, floating, flicking, thrusting, pressing, wringing, and slashing [27]. The Basic Effort Actions provide a set of style labels, rather than labels that are inherently affective, to describe movement quality. The motion factor combinations that produce each of the Basic Effort Actions are depicted in Fig. 2.1. As previously mentioned, features of the Motion Factors and the Basic Effort Actions have been used to create functions that map quantitative changes in kinematic parameters to qualitative styles.

2.2 Affective Domain of Modern Psychology

Modern psychology is often decomposed into three domains: the cognitive domain, the affective domain, and the conative domain. The cognitive domain encompasses processes that lead to the acquisition of knowledge, the affective domain deals with feelings and emotions, and the conative domain refers to the impulses and actions associated with knowledge (cognition) and feelings (affect) [30]. All three domains are highly intertwined, but this section will primarily focus on frameworks that have been developed to model the affective domain, as it relates most closely to the research documented in this thesis.

2.2.1 Semantic Differential Scales

Before discussing the specific frameworks, it is necessary to discuss semantic differential scales because they are utilized by all of the frameworks that are presented in this section. Semantic differential scales are commonly used in research studies that examine the affective domain. Semantic differential scales are rating scales that associates bipolar adjective pairs with numeric scales and are used to evaluate stimuli. Factor analysis is performed on the results of the semantic differential scales to quantify the relationships between the stimuli and/or determine the number of independent features in the stimuli that account for the majority of the variance in the results. Semantic differential scales have been utilized to rate a large range of stimuli including, but not limited to, emotions [31, 32], frequently used English words [33], images [15, 16], facial expressions [34], and human gait animations (part of the user studies presented in Chapter 6).

2.2.2 Pleasure-Arousal-Dominance (PAD) Emotional-State Model

In 1974, Russell and Mehrabian and claimed that only three bipolar dimensions were necessary to adequately describe emotions: pleasure, arousal, and dominance [31]. Pleasure is the positive to negative evaluation of a stimulus [35], arousal is the level of activation of the nervous system or the motivation to move towards or away from a stimulus [35], and dominance is the level of control or influence experienced when interacting with a stimulus [32]. There were originally 18 semantic differential scales (6 representing each dimension) used in the Pleasure-Arousal-Dominance (PAD) Emotional-State Model. After multiple studies were conducted, the scales in the PAD model were modified and the final PAD model developed by Mehrabian consisted of 47 semantic differential scales (24 pleasure, 8 arousal, and 15 dominance) [36]. The PAD model has been utilized in studies where models were created to predict physical attractiveness, name desirability, and product preference [37].

2.2.3 Circumplex Model of Affect

In 1980, James A. Russell developed a tool named the Circumplex Model of Affect that is used to map affective states to a two-dimensional chart using ratings of valence and arousal [38]. The first circumplex model was used to evaluate the affective ratings of 28 words [38]. Although this model excludes a dimension that was previously described as necessary for adequate emotional classification in the PAD model, this model was found to describe affective states reasonably well with the added benefit of considering fewer factors in the analysis [37]. There have been many studies that use aspects of the circumplex model to evaluate the affective states of stimulus including a study on affective classification of blog posts [39] and studies evaluating the affective ratings of images [15, 16].

Databases were created from the results of the studies evaluating affective ratings of images. These databases contain images with affective ratings of valence and arousal and have been created to assist in emotional research studies. The International Affective Picture System (IAPS) is the most widely reference affective visual stimulus database. However, IAPS is only made available to researchers by request and all information pertaining to the contents of the database are subject to copyright restrictions to help ensure the integrity of the affective ratings. The Geneva Affective PicturE Database (GAPED) [16] and the Open Affective Standardized Image Set (OASIS) [15] are open-access databases that contain affective ratings for 730 and 900 color images, respectively. These open-access databases are free to the public and do not have any copyright restrictions. Images from OASIS and GAPED were utilized as stimuli in the studies presented in this paper. Details about the images that were utilized are included in Section 4.1.

2.2.4 Towards Understanding How Environmental Context Influences Affect

There are studies that have subjectively evaluated affective ratings of various features of an environmental context including population density [37], auditory stimuli [40], and visual stimuli [15, 16]. However, the subjective nature of these evaluations can cause similar stimuli to have significantly different affective ratings. Replacing the subjective evaluation of affective ratings with an objective evaluation could help to reduce variation correlated to priming.

To the knowledge of this author, [41] is the only published research that attempts to include objectively quantifiable environmental parameters as variables in studies evaluating affective states. In [41], Kanjo et al. propose an approach for assessing the impact of environment and physiological changes on the affective state of a human using direct and continuous sensor data. The data was collected by having multiple participants walk around a city center with a Microsoft Band 2 to collect physiological data and an Android phone to collect environmental information and to self-report emotional state.

The only self reported variable in this study was the emotional response, which was defined using a valence scale. The physiological and environmental features were quantified using sensors that measured environmental noise, ambient light (UV), GPS location, air pressure, electro dermal activities (EDA), heart rate (HR), hand acceleration, and body temperature. The model that was developed using the results of the study was 86% accurate in determining a valence rating based on physiological and environmental feature values. It is important to note that the scope of this tool is limited to one specified environmental context and only considers one dimension of the affective domain. However, even with these limitations, Kanjo et al. demonstrate that it is possible to use quantifiable environmental parameters to assist in estimating subjective emotional responses in humans.

Chapter 3

Methods and Tools for Generating Data-Driven Stimuli

This chapter presents the methods that were used to create the video stimuli that was used in the studies presented in this paper and describes two MATLAB tools that were developed to aid in multimedia stimuli creation for future studies. Section 3.1 describes the process that was used to create the baseline stylized walking sequences utilized in the stimuli for the user studies presented in this thesis. Section 3.2 describes the process that was used to create the video stimuli for the user studies presented in this thesis. Section 3.3 describes the Viewing Simulations of Data Driven Trajectories in an Environment (V-SDDTE) MATLAB tool that was developed to display existing motion capture files in environmental contexts. Section 3.4 describes the Creating Simulations of Data Driven Trajectories in an Environment (C-SDDTE) MATLAB tool that was developed to assist in creating simulations of aerial avatar trajectories using manipulated motion capture data and then viewing the trajectories in various environmental contexts.

3.1 Generating Baseline Stylized Walking Sequences

The primary source of inspiration for the work presented in this thesis was developed during the a research seminar (ME 598 - High Level Robotic Control and Movement Representation) at the University of Illinois Urbana-Champaign. During the course of the semester long seminar, the relationship between movement profiles and context was probed with the intent of understanding how the two elements contribute to an outsiders perception of the affective intent of the movement. In generating the stimuli used in the user studies presented in

Chapters 4, 5, and 6, the authors sought to find existing animated movements that were previously verified as "stylized" in previous studies.

As previously mentioned, stylized walking sequences were created by expert animators and evaluated by participants that were uneducated in animation techniques in user study documented in [7]. The input and output walking sequences produced in [7] were used to generate the stimuli for the user studies presented in this paper (presented in Section 3.2). This section will summarize the process for creating and evaluating the stylized walking sequences that was presented in [7].

Etemad and Arya use a linear motion model to describe the various movement parameters involved in a 54 degree-of-freedom (DOF) skeleton walking. The linear model Etemad and Arya derive describes a complete motion sequence as a linear combination of primary features and secondary features. The primary features are associated with the base action of the movement sequence, such as throwing or walking. The set of primary features associated with the base action is referred to as the primary theme of the motion sequence. The secondary features are the components of movement associated with the style in which the action is performed. The weighted set of secondary features associated with one style is referred to as a secondary theme of the motion sequence. The model developed by Etemad and Arya permits multiple secondary themes to be simultaneously applied to a single primary theme.

The primary features used in [7] were based on neutral walking sequences that were captured using a retro-reflective marker-based Vicon motion capture system. The secondary features were created by 11 experienced animators. The animators were tasked with creating secondary features using radial basis functions to transform a neutral walking sequence into various stylized walking sequences. The animators were allowed to used up to three radial basis functions (RBFs) per DOF in the model to create the secondary feature sets for the secondary themes. The feature sets created by all of the animators for a particular secondary theme were averaged together to create one secondary feature set. The secondary themes that the animators were tasked with creating were labeled as *energetic*, *feminine*, *happy*, *masculine*, *sad* and *tired*.

For the remainder of this thesis the label pairs *feminine/masculine*, *energetic/tired*, and *happy/sad* will be considered bipolar pairs. Bipolar implies that if one style of a bipolar pair is embodied by a state at time t, then the other style in the bipolar pair cannot be embodied by the same state at time t. For example, a state could be viewed as both *happy* and *tired* but not as both *tired* and *energetic* at any given time.

The styles of the secondary themes were evaluated by 16 new participants. These participants used forced-choice scales to rate 12 stylized walking sequences for each secondary theme. The participants were told what the intended secondary theme was prior to rating. The walking sequences varied by number of secondary features (1, 4, 7, and 10) as well as weight w (0.4, 0.7 ,1.0, and 1.3) of the secondary features used to embody the intended affect of the animators. Please refer to [7] for further details on the generation and labeling of the context invariant stimuli leveraged in user studies presented in this paper.

3.2 Stimuli Creation Process for Thesis User Studies

The user studies that are presented in this thesis use still images of environments and video animations of stick figures walking in various environments as stimuli. This section details the methods that were used to create the video animations. The video stimuli consisted of the stylized walking sequences with 10 secondary features and weight $w = 1.0$ from [7] superimposed onto seven different background videos from YouTube (49 new animation videos) in an attempt to visually simulate an environmental context. The background videos were chosen because they contained similar characteristics to the 7 OASIS reference images. The selection process for the 7 OASIS references will be further described in Section 4.1 . The reference OASIS images and snapshots of the video stimuli that were created are shown in Fig. 3.1.

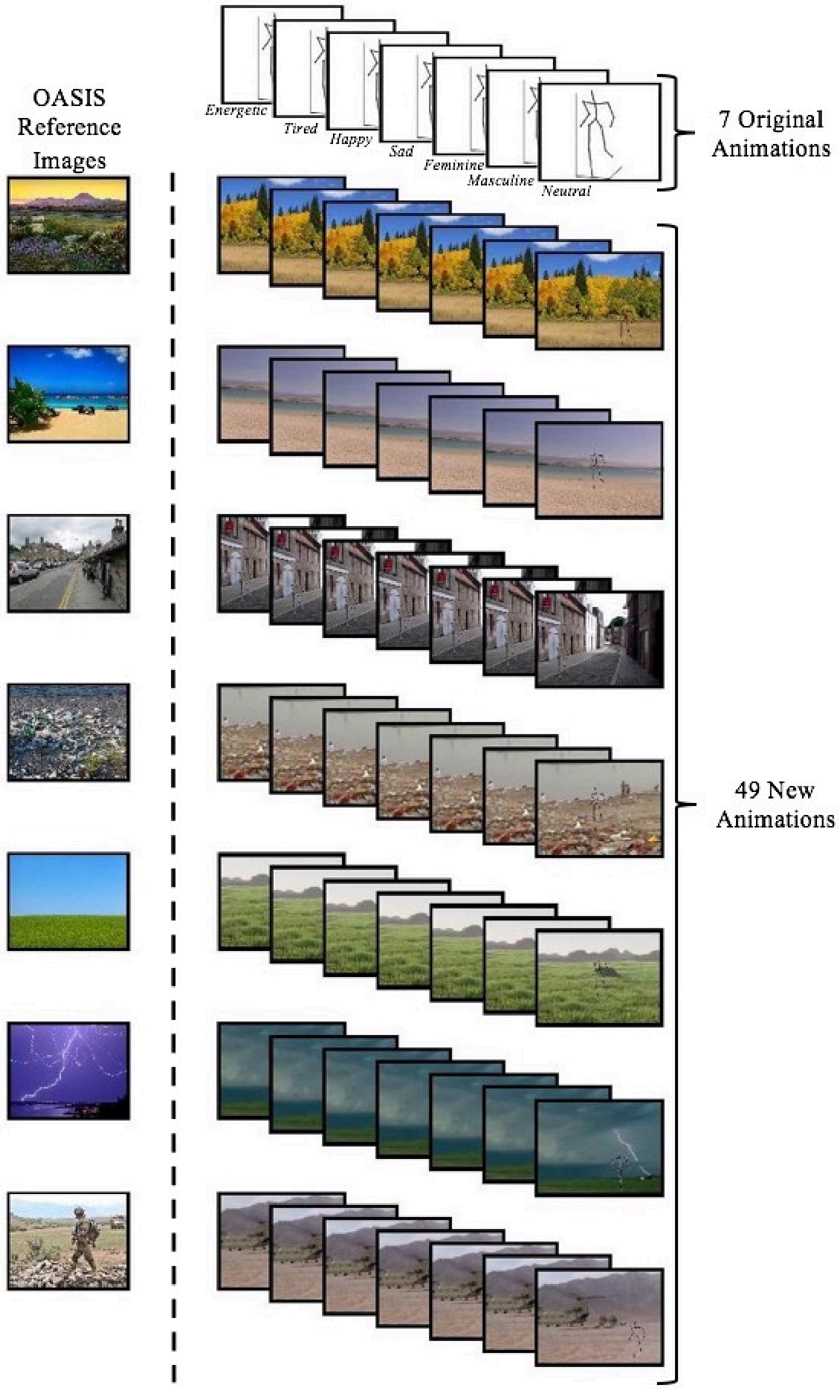

Figure 3.1: The first column contains the OASIS images that were used as reference. The labels of the OASIS images, from top to bottom, are *Nature 1*, *Beach 1*, *Street 1*, *Garbage Dump 2*, *Grass 6*, *Lightning 3*, and *Soldier 9* [15]. The second column contains snapshots of the 7 original animations [7] and the 49 new animations created for this user study. The environmental contexts in column 1 and column 2 that are in the same row have similar features. The labels of the snapshots, from top to bottom, are *white*, *aspen*, *beach*, *city*, *garbage*, *grass*, *lightning*, and *war*.

All of the user study video stimuli was created in iMovie using environmental contexts found in YouTube videos and the walking sequences created by animators in [7]. The original walking sequences from [7] were approximately 2 seconds long and consisted of a blue stick figure model completing one full gait cycle on a white background. The duration of each original walking sequence was extended to approximately 10 seconds by looping each animation 5 times. Color correction was then performed to change the color of the stick figure model to black so that it would be visible in all of the simulated environmental contexts. The 10 second videos with black stick figure models were used as the *white* background walking animations.

The white backgrounds had to be removed prior to superimposing the stick figure models onto the various background videos. This was achieved by adjusting the levels of color saturation and color contrast so that the white backgrounds became blue. The background was changed to blue because iMovie has a green screen feature that removes blue and green backgrounds of superimposed videos. The 10 second walking sequences with blue backgrounds were saved and used as the baseline walking animations that were superimposed onto the various background videos.

In some cases, the timing or orientation of the environmental context video was edited to match the gait cycle timing of the walking sequence. The timings of the walking sequences were never altered to ensure that the duration of each gait cycle remained constant in all of the video stimuli. It should be noted that the size and placement of the stick figure models was not constant in all simulated environmental contexts. However, the sizes and locations of all stick figure models within each environmental context were approximately the same.

3.2.1 Limitations of Method

The method for creating video stimuli just described was time-consuming and resulted in low-fidelity simulations. Some of the primary limitations that make this method non-ideal include:

- Inability to exactly sync timing of gait cycle with environmental context simulation
- Lack of appropriate, readily available environmental context videos
- Inability to simulate a believable 3-dimensional space
- Fixed orientations of the stylized walking sequences

These limitations inspired the development of two MATLAB tools that can be used to quickly and efficiently produce multi-dimensional (movement and context) animations for future user studies.

3.3 MATLAB Tool for Viewing Simulations of Data Driven Trajectories in Environments

One method of creating simulated trajectories is to use a data-driven approach. The MATLAB tools presented in this thesis utilize pre-existing motion capture data as the the baseline data. Motion capture data is commonly used in data driven approaches for creating realistic video game characters and animations [42]. Data-driven approaches are sometimes desirable because they enable the creation of complex trajectories without having to explicitly derive the equations of motion.

The MATLAB tool V-SDDTE, displayed in Fig. 3.2, was developed with the goal of becoming a platform where humans could quickly and easily investigate how varying the dimensionality of human movement and altering the environmental context influenced affect recognition. It should be noted that this tool does not account for any physical limitations (gravity, mass, hardware restrictions, etc.) that would be imposed on a physical system.

For the remainder of this chapter, the user input parameters for the MATLAB tools will be capitalized and italicized to emphasize the notion that the words represent parameter labels and not selection options. In this section, the high-level explanations of the capabilities and functionalities of each parameter will be followed by a more detailed discussion about design choices and code implementation. All of the user input parameters for this tool are shown in Fig. 3.2. For further formatting information that is not documented in this section, please refer to the `read_me` file that is outlined in Appendix A.

3.3.1 Motion Capture Data Library Creation

The current baseline motion capture library consists of 183 walking data files that were collected by 23 participants [43]. The library includes clockwise and counter-clockwise walking sequences that were performed in the various styles, including happy, sad, angry and neutral. The motion capture data files that are currently in the library were chosen based on accessibility. The tool is built to accommodate additional motion capture data files as long as they are formatted correctly and exist in the correct file locations.

The baseline motion capture data library that is currently being used was obtained as PTD files from [43]. MATLAB code was written to convert the PTD files into CSV files that were formatted and later used as the motion capture library for the MATLAB tools. The names of the CSV files are encoded with information about the action performed, the person who performed the action, the orientation of the action performed, and the style of the performed action.

3.3.2 Motion Capture File Selection for Simulation

The Data Selection parameters are used for defining the working file where the motion capture data will be extracted from. The *Action* and *Person* selection options are automatically generated each time the GUI code is run and determined by the CSV file names in the working directory. These parameters automatically update because the selected options are directly saved as the MATLAB handles storing the states of each parameter. The *Style* and *Walking Direction* selection options were manually defined during the GUI creation. These were manually defined because the handles storing the states of each parameter are

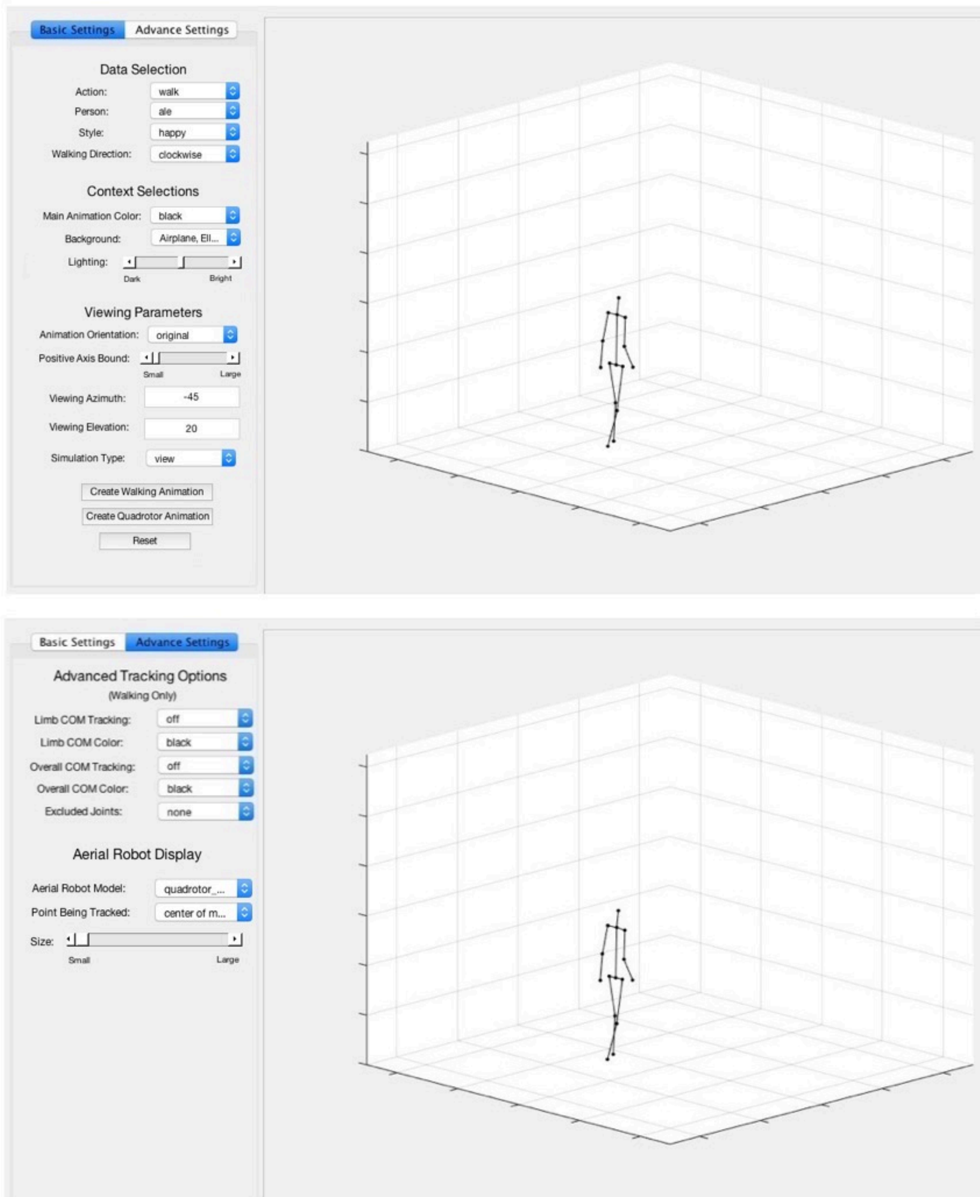

Figure 3.2: Screen captures of the MATLAB tool V-SDDTE

not set to the selected string. When a *Style* is selected, the MATLAB handle storing style information is set to a 2 letter code corresponding to the selected style. When the *Walking Direction* is changed, the MATLAB handle storing directional information is toggled to either 1 (counter-clockwise) or 2 (clockwise). The strings stored in each of the parameter handles are concatenated and used to define the name of the working file that was selected. In the future, the code may be updated to make all selection options automatically update. However, as of now, this format allows new motion capture data to be easily integrated into the existing library without requiring significant updates to the code (see Appendix A for details).

3.3.3 Simulated Environmental Contexts

The *Background* parameter contains selection options of images that can be used to simulate environmental contexts in the simulations. The *Lighting* parameter was also added to increase the believability of the simulations by adding ambient lighting in the environment and creating shadows on the aerial avatar. The *Lighting* parameter is defined using a slider to ensure that the selected value is valid.

Originally, the tool required three images for a background to be rendered during the simulation. The tool originally placed one image in each of the planes. This definition caused a very noticeable seam between the xz-plane and the yz-plane. To eliminate the seam, a mesh-grid was created so that one image would be projected continuously on both the xz and yz-planes. The mesh-grid only requires two images for a background to be rendered and creates a more convincing environment. Fig. 3.3 shows the same environment (constant *Lighting* value) displayed using both methods to show the benefit of using a mesh-grid.

The current *Background* selection options are primarily 2D in the sense that most of the selection options consist of 2 images projected onto 2D planes. In many cases, these backgrounds are not good representations of environments that are typically associated with HRI applications. Therefore, there is an ongoing effort to increase the number of HRI-relevant

Figure 3.3: Screen captures of the MATLAB tool V-SDDTE displaying the same environmental context using two different methods. The image on the left displays one image on each plane and the image on the right displays one image on the xy-plane and one image on a mesh-grid between the other planes.

3D environmental contexts that are available. The 3D environments include projections on each plane as well as objects that exist in multiple planes. The purpose of creating 3D environmental contexts in the MATLAB tool is not to try to replace high-fidelity platforms such as virtual reality (VR), but rather create mid-fidelity 3D environments where trajectories can be easily manipulated and viewed before they are imported to the high-fidelity 3D environments where trajectory manipulation is difficult. The MATLAB tool also enables the video simulations to be saved and used as stimuli in online user studies, which is something that is not possible in VR user studies. The 3D environment that is currently available in the MATLAB tools is displayed in Fig. 3.4

The *Background* selection options are automatically updated based on the names of the subfolders that are contained within the `background_images` folder that exists in the same directory location as the primary MATLAB file that initiates the GUI. The subfolder names are directly read into the MATLAB tools and set as the selection option string for the *Background* pop-up menu.

Figure 3.4: The 3D environment displayed in this screen capture of a aerial avatar trajectory simulation is a bedroom.

3.3.4 General Viewing Parameters

The tool includes a feature labeled *Simulation Type*, that allows the user to choose if the simulation is viewed or if it is viewed and saved. If the *Simulation Type* is set to 'save', the animation will be saved as an Uncompressed AVI. The *Positive Axis Bound/Plot Limits*, *Viewing Azimuth*, and *Viewing Elevation* all define basic viewing parameters of the animation. *Positive Axis Bound/Plot Limits* defines the boundaries of each axis. This parameter can be increased or decreased if the viewer wishes to see the animation from a more zoomed-out or zoomed-in perspective, respectively. This parameter is also used in the definition of the mesh-grid used for the background context display. *Viewing Azimuth* and *Viewing Elevation* define the viewing azimuth and viewing elevation of the 3D plot using the MATLAB function `view`.

The *Animation Type* parameter adjust the orientation of the avatar When the animation is initiated in the 'original' setting, the avatar transverses through space, following the same trajectory that was originally recorded. When the 'stationary' option is selected, the trajectory is translated and rotated so that the avatar appears to be moving but no locomotion occurs.

Before explaining the translation and rotation processes, it is necessary to establish the definitions that are used for the local coordinate system and the global coordinate system. The origin of the local coordinate system is defined as the pelvis in the motion capture data and the global coordinate system origin is at the (0, 0, 0) coordinate of the 3D plot. The translation and rotation are performed to ensure that the xz-planes and yz-planes of the local and global coordinate systems remain parallel at all time-steps. The translation and rotation only alter the x and y coordinates of the data points because if the z coordinates were altered, there would be no guarantee that the joint trajectories would remain above the global xy-plane in the simulation.

The translation is first performed by subtracting the x and y coordinates of the pelvis from the x and y coordinates of all joints at each given time stamp. This translation ensures that the x and y coordinates of the local and global coordinate system are equal. The rotation calculation that is performed at each time stamp to ensure that the xz-planes and yz-planes of the local and global coordinate systems are aligned is shown in (3.1), where the rotation angle θ at each time stamp is defined by (3.2).

$$\begin{bmatrix} x_{stat,i} \\ y_{stat,i} \end{bmatrix} = \begin{bmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{bmatrix} \begin{bmatrix} x_{trans,i} \\ y_{trans,i} \end{bmatrix} \tag{3.1}$$

$$\theta = -tan^{-1}\left(\frac{y_{LeftHip}}{x_{LeftHip}}\right) \tag{3.2}$$

3.3.5 Walking Figure Display Options

The *Excluded Joints* parameter is used to view a reduced dimensionality walking sequence. This parameter was included to investigate how perception of stylized movements changes as the number of DOF decreases. This could be useful when trying to use a data-driven approach to create a stylized trajectory for robots such as an Aldebaran NAO or a Rethink Robotics Baxter. The selection options that exist for this parameter include removing the 'elbows', 'knees', 'elbows and knees', 'upper body', or 'upper body and knees'.

When viewing a walking animation, the *Limb COM Tracking* and *Overall COM Tracking* can be enables so that the center of mass of the the limbs and/or entire body appear in the animation. The *Limb COM Tracking* and *Overall COM Tracking* are determined for the markers existing after the *Excluded Joints* option is selected. The COM locations were calculated using values in [44]. The parameters described as Advanced Tracking Options are only applied when the *Create Walking Animation* button is selected.

3.3.6 Aerial Avatar Display Options

The *Aerial Robot Model* parameter contains a list of STL files that are available to represent the aerial avatar in the simulations. The size of the aerial avatar can be altered by changing the position of the *Size* parameter slider. The *Point Being Tracked* parameter is used to select the joint trajectory that defines the aerial avatar trajectory. The aerial avatar simulations includes a shadow of the aerial avatar on the xy-plane to help orient the position of the aerial avatar in the 2-dimensional representation of a 3-dimensional volume.

The selection options for *Aerial Robot Model* are automatically populated (using the method previously described) and new STL files can be added to the simulations by including the STL file in a folder named `stl_files`. The chosen STL file is rendered using a toolbox of downloaded MATLAB functions [45].

3.4 MATLAB Tool for Creating Simulations of Data Driven Trajectories in Environments

When the fidelity of a movement model is diminished (reduced DOF, feature simplification, etc.), the style encoded in the movement model is altered. If motion capture data of human

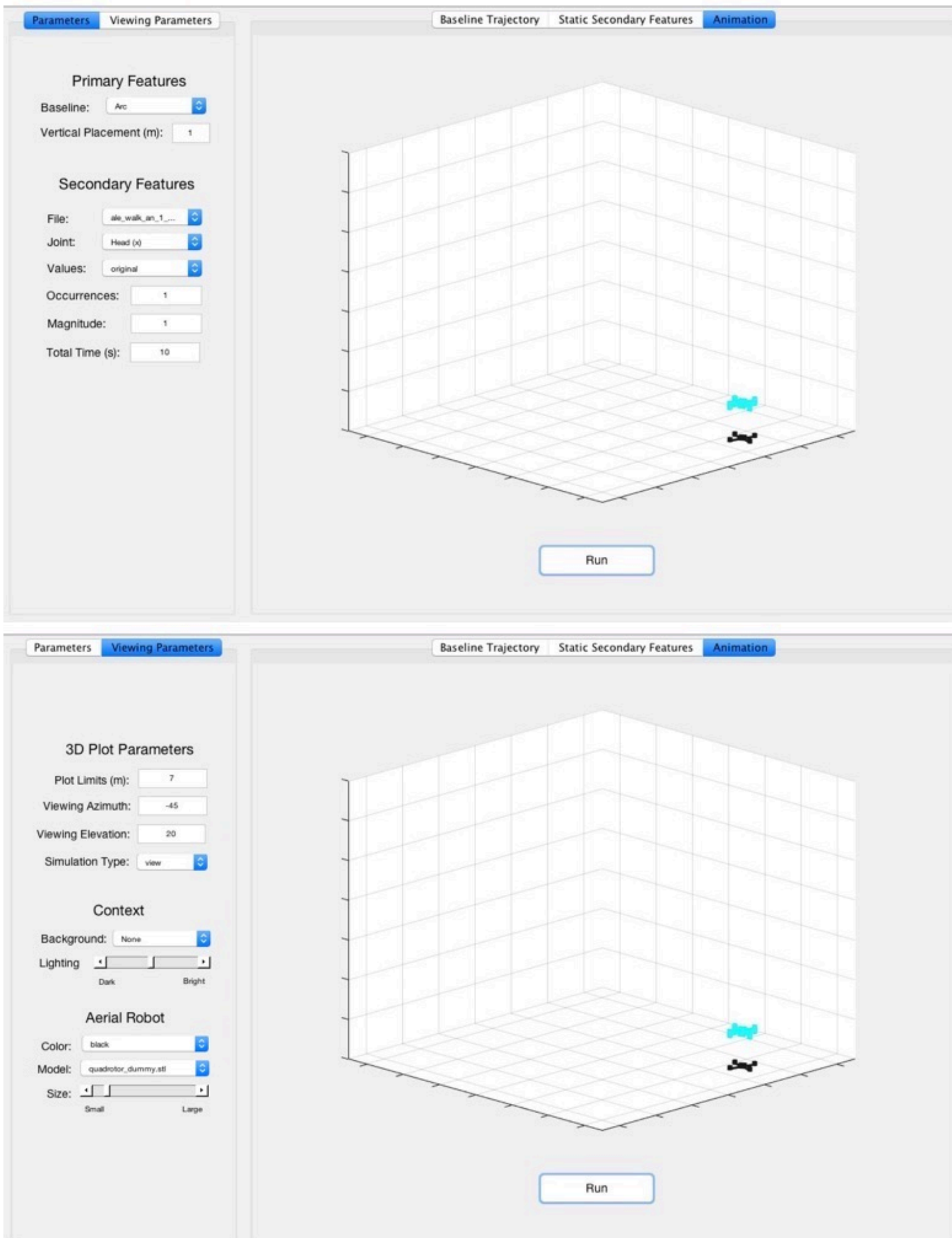

Figure 3.5: Screen captures of the MATLAB tool C-SDDTE.

movement were reduced to 6 DOF, it is likely that the intended style of the movement would be indistinguishable. This raises the question, is it possible to re-encode the style into a reduced-dimensionality trajectory by manipulating parameters of the motion capture data, such as frequency and magnitude? The MATLAB tool C-SDDTE was designed to see if motion capture data of a single human joint position can be manipulated to create stylized movement trajectories for simulated aerial avatars in different simulated environmental contexts. This approach for creating data-driven trajectories could be extremely beneficial for researchers searching to find affective aerial robotic trajectories that are suitable for various HRI settings. However, it should be noted that this tool does not account for any physical limitations (gravity, mass, hardware restrictions, etc.) that would be imposed on a physical system. The MATLAB tool C-SDDTE is shown in Fig. 3.5. For further formatting information that is not documented in this section, please refer to the `read_me` file that is outlined in Appendix B.

In [7], Etemad and Arya defined styles movements as a combination of two components: primary themes and secondary themes. The primary themes contained primary features that were critical to the action of the sequence (i.e. walking, throwing, etc) and the secondary themes contained secondary features that stylized the movements (i.e. happy, sad, etc.). The entire movement sequence was defined as a linear combination of the primary features and secondary features. In C-SDDTE, the primary features are defined as baseline trajectories of a aerial avatar that transverse through space. The primary features are created from very simple equations and consist of straight lines and curves. The primary features are meant to function as the necessary path the aerial avatar must follow in order to get from its starting position to its ending position. The secondary features are defined using modified joint positions corresponding to a single gait cycle of a human. Using C-SDDTE, a participant can change the selected joint, the magnitude of the selected joint positions, and the number of gait cycles used to easily define a new secondary feature for the aerial avatar. The trajectories can be simulated in 2D and 3D environmental context.

The remainder of this section will describe the motion capture library used in this tool as well as the available features. All of the viewing parameters in this tool are identical to the viewing parameters in V-SDDTE and will not be discussed in this section.

3.4.1 Motion Capture Data Library Creation

The baseline motion capture data library that was described in Section 3.3.1 was manipulated to create the input data for this tool. The input data for C-SDDTE is the average of the cyclical joint trajectories corresponding to a gait cycle. The preprocessing that takes place to create the baseline motion capture data library for this tool includes translation, rotation, gait segmentation, time-warping, and cycle averaging. This section will conclude with a detailed description of the gait segmentation, time-warping, and cycle averaging processes. The translation and rotation processes were previously described in Section 3.3.6.

Bipedal walking is characterized by periodic movement that composed of two phases: stance phase and swing phase. These phases can be further decomposed into the eight phases shown in Fig. 3.6. The *initial swing* phase occurs immediately before toe-off. Since no toe information was included in the motion capture data, the *initial swing* phase for the motion capture data was defined as the moment in time when the position of the right ankle reached a local maxima in the y-direction.

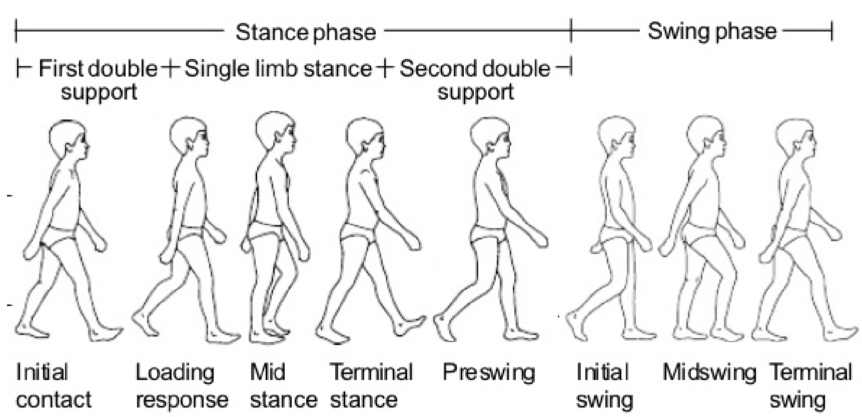

Figure 3.6: Diagram showing the phases of gait [46].

MATLAB code was written to determine the set of time-stamps $\{m_1, m_2, ..., m_n\}$ of all local maxima in the y-direction right ankle data. The data for joint position j_i was then segmented at each of these time stamps to create n different sets of joint position data, one for each gait cycle existing in the motion capture data file. The set of joint position data cycles can be represented by $\{j_{i,1}, j_{i,2}, ..., j_{i,n}\}$. It is important to note that the number of data points in each of the gait cycles was not necessarily equal. Therefore, in order to average the joint position j_i over all cycles, it was necessary to first perform time-warping to normalize the lengths of the gait cycles and match similar joint characteristics.

Since bipedal gait is considered a periodic movement, it was assumed that similar joint characteristics would occur at similar phases of each gait cycle. The first step of the time-warping process was to perform upsampling and downsampling so that each cycle contained the same number of data points. Currently, the desired number of data points per cycle is set to $p = 100$. However, this number is arbitrary and could be set to any reasonable number. The upsampling and downsampling multipliers were determined automatically based on number of data points in each cycle. When the time-warping was complete, the average of all of the cycles at each post-time-warped time stamp was calculated using (3.3) The averages of all of the cycles for the joints for each walking sequence were saved in new excel documents and used as the baseline motion capture data library for this tool.

$$j_{i,ave}(t) = \left\{ \sum_{a=1}^{n} j_{i,a}(t) \middle| 1 \leq t \leq p \right\} \tag{3.3}$$

3.4.2 Creating and Simulating Aerial Avatar Trajectories

The primary features currently available in this tool include *Baseline* and *Vertical Placement (m)*. The *Baseline* selection options are the primary trajectories of the aerial avatar. The value of *Vertical Placement (m)* determines the offset of the primary trajectory from the

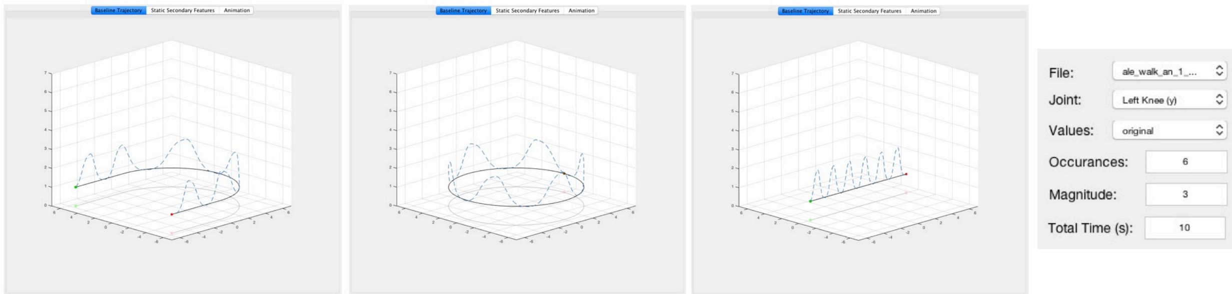

Figure 3.7: The primary trajectories that are currently available as selection options for *Baseline* in C-SDDTE. From left to right, the primary trajectories are ‘Arc’, ‘Circle’, and ‘Line’. In all of the trajectories the black line is the primary trajectory, the blue dotted line is the aerial avatar trajectory, the green dot is the start, and the red dot is the end. The trajectory, start, and end points are also reflected to the xy-plane in slightly lighter colors to enhance the representation of a 3D volume in 2D. The secondary feature parameters for the display trajectories are also included in the figure.

global xy-plane. The trajectories that are currently available are shown in Fig. 3.7.

The selection options for *Baseline* automatically update based on the field names of the structures defined in `baseline_trajectories.m`. The MATLAB code `baseline_trajectories.m` is used to define the data points for the primary trajectories. If a new trajectory is desired, the user can define the new data points in the MATLAB code. Before defining a new trajectory, the user should first read the `read me` file included in the directory containing the MATLAB code.

The panel in the C-SDDTE GUI that displays the secondary feature selection options is shown in Fig. 3.8. The *File* parameter is used to select the motion capture data file that will be used to create secondary features. The *Joint* option that is selected is used as the original secondary trajectory. The *Joint* parameter contains x, y and z positions the following data points: head, neck, (right/left) shoulder, (right/left) elbow, (right/left) wrist, pelvis, (right/left) hip, (right/left) knee, (right/left) ankle, COM, upper body COM, and lower body COM.

The *Occurrences* parameter determines the number of times the original secondary feature will occur along the primary trajectory. The *Magnitude* parameter defines the multiplier that is applied to the original position data to obtain the modified position data. The *Total Time*

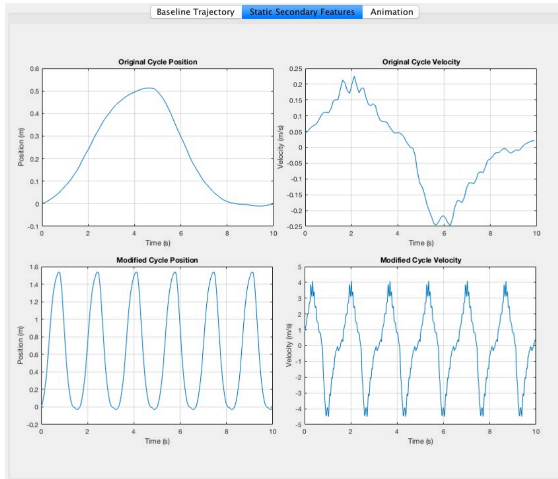

Figure 3.8: The secondary feature panel in C-SDDTE contains position and velocity graphs for the original and modified trajectories. The secondary feature displayed in this figure are defined in Fig. 3.7

(s) parameter determines the duration of the simulation. It also alters the velocity of the aerial avatar.

The *Values* parameter acts as an operator on the secondary feature. The selection options include 'original' (unmodified joint data), 'inverse' (joint data is multiplied by -1), 'positive' (absolute value of joint data is multiplied by -1), and 'negative' (absolute value of joint data is multiplied by -1).

When the 'Run' button is selected, the orientation of the local coordinate system of the aerial avatar is rotated so that the x-axis aligns with the vector tangent to the generated trajectory at each time stamp. The rotation matrix that was used to perform the calculations at each time stamp is shown in (3.4) and θ and ϕ are defined in (3.5) and (3.6), respectively. The subscript *traj* corresponds to the trajectory generated using the primary and secondary features and the subscript *orient* corresponds to the aerial avatar trajectory displayed in the animation.

$$\begin{bmatrix} x_{orient} \\ y_{orient} \\ z_{orient} \end{bmatrix} = \begin{bmatrix} cos(\theta)cos(\phi) & -sin(\theta) & cos(\theta)sin(\phi) \\ sin(\theta)cos(\phi) & cos(\theta) & sin(\theta)sin(\phi) \\ -sin(\phi) & 0 & cos(\phi) \end{bmatrix} \begin{bmatrix} x_{traj} \\ y_{traj} \\ z_{traj} \end{bmatrix} \tag{3.4}$$

$$\theta(t_i) = tan^{-1}\left(\frac{y_{traj}(t_i + t_{i+1}) - y_{traj}(t_i)}{x_{traj}(t_i + t_{i+1}) - x_{traj}(t_i)}\right) \tag{3.5}$$

$$\phi(t_i) = tan^{-1}\left(\frac{z_{traj}(t_i + t_{i+1}) - z_{traj}(t_i)}{t_{i+1} - (t_i)}\right) \tag{3.6}$$

Chapter 4

Study 1: Exploring Context-Dependency of Previous Labeling of Affective Gait

The primary goal of this study was to investigate a hypothesis that priming participants with an affective label would influence the environment where they envisioned a stylized walking sequence occurring. If this hypothesis were supported by the data, we would see trends that indicate that the images depicting environmental contexts were not selected randomly. However, if the image selection appears to be random, it does not necessarily mean that the hypothesis should be rejected. Random selection may also indicate that the set of images did not adequately represent a variety of different environments. In this user study, affective image databases were utilized to ensure that a variety of different environments were included in the image sets presented to participants.

The process that was followed when selecting the images is presented in Section 4.1. Section 4.2 presents the questionnaire format and is followed by a discussion of the Study 1 results in Section 4.3.

4.1 Images Selected to Simulate Environmental Context

Before any images were selected, the affective ratings of all the applicable images in the OASIS database [15] were mapped to a modified circumplex model that was created in [47]. Initially, the mapping was used to select five images that were mapped to regions of the circumplex model that are associated with the affective labels *energetic*, *happy*, *neutral*, *sad*, and *tired*. The mapping of the affective ratings were not initially considered when select-

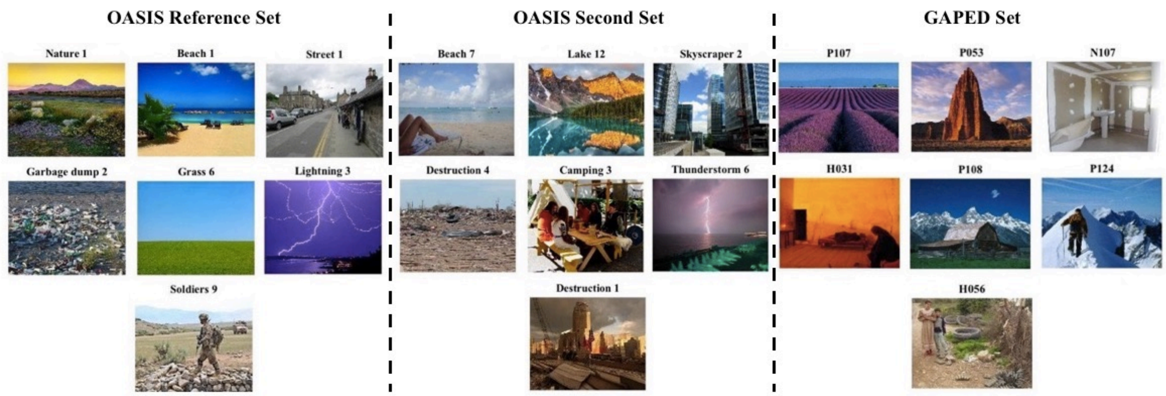

Figure 4.1: Images in the OASIS reference image set [15], the OASIS second image set [15], and the GAPED image set [16].

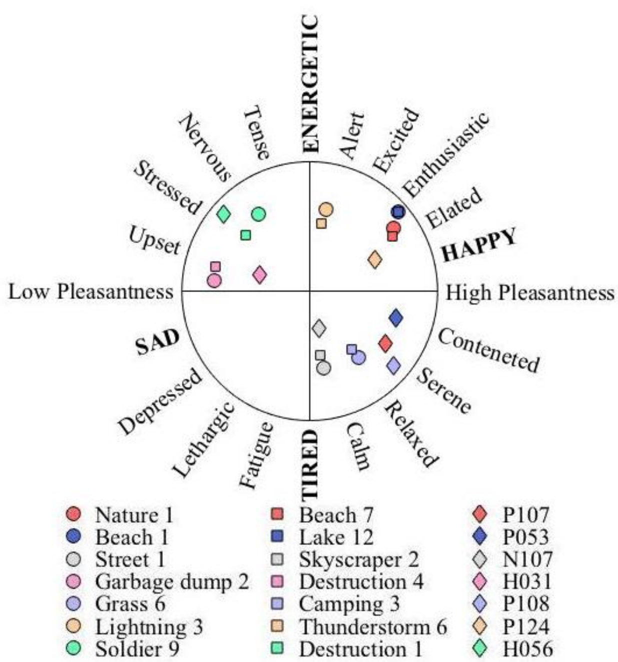

Figure 4.2: A mapping of the affective ratings of the images used in user studies to the circumplex model of affect. The OASIS reference set is represented by circular markers [15], the second OASIS image set is represented by square markers [15], and the GAPED image set is represented by diamond markers [16]. The similar colored markers were deemed similar during the selection process.

ing the images that were associated with the labels *feminine* and *masculine* because these labels describe physical states rather than affective states. The image that selected to correspond with the *masculine* label was selected because it depicts a scenario that is historically stereotyped as masculine [48]. Additionally, in [49] it was shown feminine traits generally had higher valence ratings and a lower arousal ratings than masculine traits. Therefore, the image associated with the *feminine* style was chosen to have a higher valence rating and a lower arousal rating than the image associated with the *masculine* style. The first set of seven images that were selected from the OASIS database [15] will be referred to as the OASIS reference image set.

The affective ratings of the images in OASIS reference image set were used select the images for two additional image sets: the OASIS second image set and the GAPED image set [16]. The images in each image set are shown in Fig. 4.1 and the circumplex mapping of the affective ratings of each image is presented in Fig. 4.2.

4.2 Questionnaire Format

The questionnaire given to the participants consisted of 36 questions. The first eight questions were standard demographics questions. The remaining 28 questions consisted of 3 types of questions:

1. Selecting the affective label that best represented the *white background neutral* walking animation (1 question)

2. Rating how "affective" the stylized *white background* walking sequences appeared on a scale of 1 (not affective) to 10 (extremely affective). (6 questions)

3. Selecting an image from each image set that depicted the environmental context that they believed best fit the prescribed affect of each stylized walking animation (21 questions)

These second type of question was similar to one of the types of questions that was asked in [7]. The results from these questions were used to see if the ratings obtained in Study 1 were comparable to the ratings found in [7]. Screen captures of the types of questions in Study 1 are included in Appendix C.

All questions pertaining to one walking animation were contained on the same page and the order of the pages was randomized. Additionally, the order of questions on each page and the order of the selection options for each question were also randomized. The participants were allowed to rewatch each stylized walking sequence as many times as they deemed necessary but were not allowed to return to questions on pages that they had previously submitted.

A total of 10 participants, 5 male and 5 female, were recruited at the University of Illinois Urbana-Champaign and completed Study 1. The participants were between the ages of 18 and 23 years old, with an average age of 19.4 years old and a standard deviation of 1.7 years. Although English was not the native language of every participant, 9 of the participants consider themselves fluent in English. Participants were compensated for their time with a $15 Starbucks gift card.

4.3 Results and Discussion

Immediately following the demographics questions, the participants were asked to select the affect label that they believed best corresponded to the *white background, neutral* walking sequence. It is important to note that this question did not subliminally prime the participants by informing them that the sequence was pre-labeled in as *neutral* [7]. The results from this question are shown in Table 4.1. A similar question was asked by Etemad and Arya [7] and in Study 2. The results of these will be discussed in Section 5.2, after the Study 2 results are presented.

Both Study 1 and [7] included questions asking the participants to rate the magnitude of the affect displayed in walking sequences created by the animators on a fixed scale. In

Table 4.1: Affect Selection for Neutral Walking Sequence in Study 1

Energetic	*Feminine*	*Happy*	*Masculine*	*Neutral*	*Sad*	*Tired*
10%	0%	0%	30%	0%	30%	30%

Study 1, these questions subliminally primed the participants by informing them of the style that was intended by the animators and asked them to rate how well the affective label was embodied by the stylized walking sequence on a scale of 1 (poor representation) to 10 (perfect representation). The average ratings of the walking sequences from Study 1 are presented in Table 4.2. The order of the affective rankings (most affective to least affective) for the *white background* walking sequences in both Study 1 and [7] were almost identical. The only difference was that in Study 1 *energetic* had the lowest ranking and in [7] *energetic* was ranked fourth.

Table 4.2: Affect Ratings of White Background Sequences

Energetic	*Feminine*	*Happy*	*Masculine*	*Sad*	*Tired*
4.8	7.4	5.0	8.1	7.3	6.6

The remainder of the questions in Study 1 asked the participants to select one image from each image set where they envisioned each pre-labeled walking sequence. The results of these questions were grouped according to similarities in affective ratings and are presented in Fig. 4.3.

The lengths l of the error bars in the plots in Figs. 4.3, 5.2, and 5.3 were calculated using

$$l = \pm 1.96 * \sqrt{\frac{\hat{p}(1-\hat{p})}{N}}$$

where the variable $\hat{p}$ is the sample proportion, N is the sample size, and 1.96 is the multiplication factor associated with 95% confidence [50]. In these studies, the sample proportion is the selection percentage. Each plot also includes the affect selection rate and error that is expected for random selection for our sample size. This is plotted as a dotted line with a

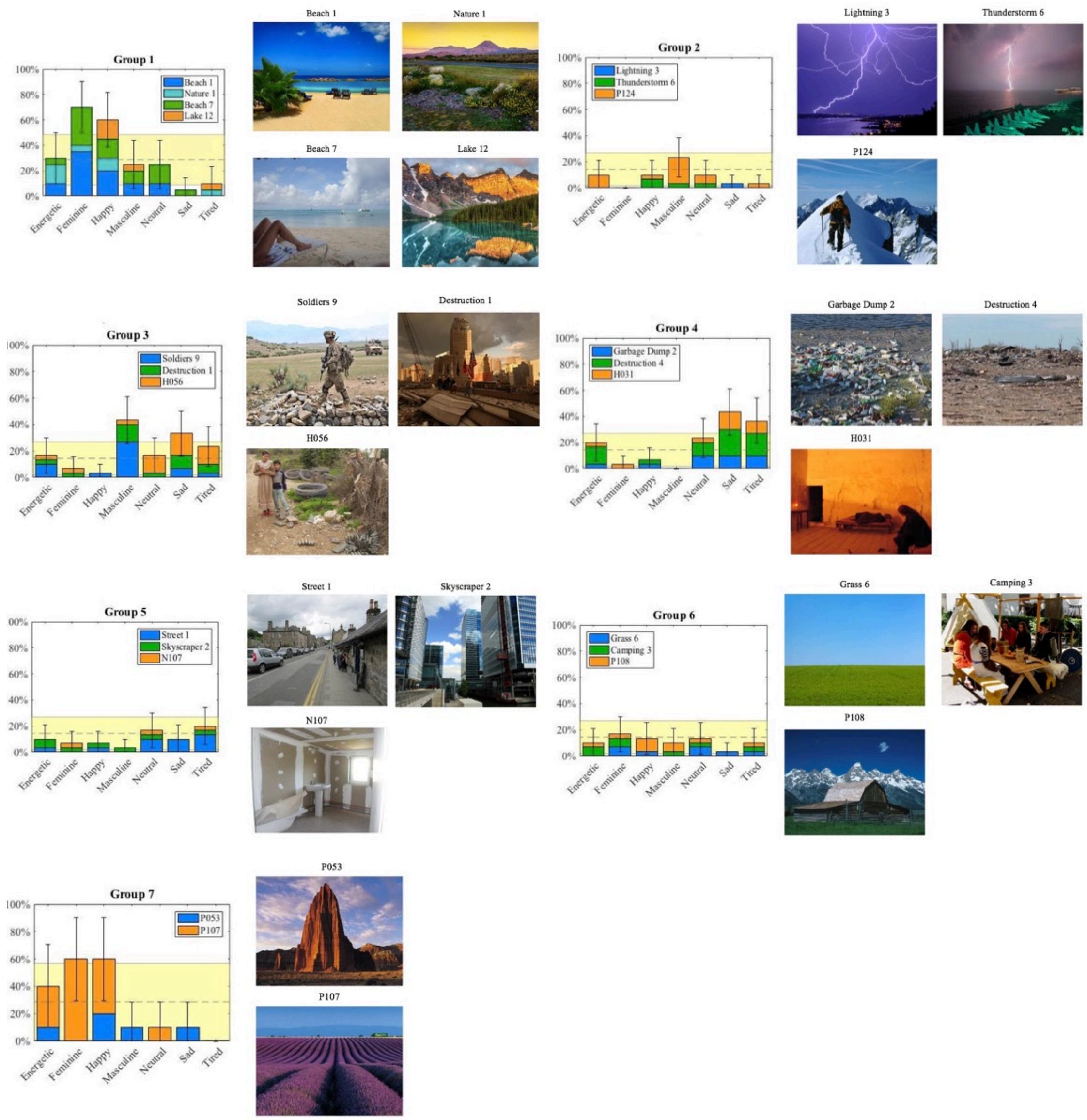

Figure 4.3: Plots of background selection after walking sequence priming in Study 1. Background images have been grouped according to their position in the circumplex model. The labels on the x-axis represent the walking sequence in the priming question and the values correspond to the percent of the time an image in the group was selected for said question. This figure shows that 1) some images were selected more frequently in general and 2) for certain primed walking styles some images were selected more frequently than random.

manila yellow margin around it for reference.

The results presented in Fig. 4.3 indicate that the hypothesis posed at the beginning of the chapter holds true. These results were somewhat expected in the sense that the images selected to accompany the pre-labeled walking sequences often depicted environments that are commonly associated with labels that have affective ratings similar to the walking sequence label. For example, environments such as the beach or a scenic nature view that are often depicted as fun or relaxing were often associated with high valence labels while scenes of war and destruction that often depict pain and suffering were typically associated with low valence labels. In environmental psychology, these associations are described as *place attachment*. Place attachment is a term that is used to describe the subjective, emotional, and symbolic meaning associated with a place [51] and is a component of inherent priming. In summary, these results indicate at a high-level that subliminally priming participants with an affective label influences the environmental contexts where a movement is envisioned.

Chapter 5

Study 2: Forced Choice Labeling of Stylized Gait

The primary objective of Study 2 was to investigate the effects of environmental context on human perception of the affect displayed through stylized movement when the participants were not subliminally primed with an affective label. This chapter will present the format of the questionnaire in Section 5.1 as well as discussion of the results of Study 2 in Section 5.2. The work presented in this chapter was published in [52].

5.1 Questionnaire Format

The questionnaire given to the participants consisted of 64 questions. The first eight questions were standard demographics questions. The remaining 56 questions used the 56 stylized walking animation videos described in Section 3.2 as stimuli. The questions were of the same format as the second question type in Study 1 (refer to Appendix C for a screen capture of question format). Each question was contained on a separate page and the order that the questions were presented to each participant was randomized. The questions asked the participants to select the affective label that they felt best described the stimulus that appeared at the top of the page. The participants were allowed to rewatch each stylized walking sequence as many times as they deemed necessary but were not allowed to return to questions that they had previously answered and submitted. The order that the affect label choices were displayed for each question was randomized.

A total of 20 participants, 8 males and 12 females, were recruited and completed the questionnaire at the University of Illinois Urbana-Champaign. The participants were between the ages of 19 and 40 years old, with an average age of 21.9 years old and a standard deviation of 4.5 years. Although English was not the native language of every participant, all of the participants consider themselves fluent in English. Participants were compensated for their time with a $15 Starbucks gift card.

5.2 Results and Discussion

The study conducted by Etemad and Arya in [7], Study 1, and Study 2 all asked participants to select the affective label that best described the *white background, neutral* walking animation. The results from Study 2 are presented in Table 5.1. The results presented by Etemad and Arya show that *neutral* was selected 75% of the time and the other selections were spread relatively evenly across the remaining affect labels. The results from Study 1 and Study 2 were significantly different than the results presented by Etemad and Arya.

In Study 1, *neutral* was never selected as the affective label that best described the *neutral* walking sequence. In Study 2, the *neutral* and *masculine* labels had the highest selection rates (25% each) and the other selections were relatively evenly dispersed among the remaining affect labels. The only difference in subliminal priming between the questions from Study 1 and Study 2 was the question ordering. In Study 1, this question immediately followed the demographics section and in Study 2 the order this question appeared was randomized for each participant. Since all other parameters were held constant, the variations in selection rates between Study 1 and Study 2 indicate that the question ordering impacted the selection rates for the *white background, neutral* walking animation. It is worth noting that the order that the stimuli were presented to the participants in Study 2 was randomized in an attempt to mitigate the error caused by question ordering. The details about the subliminal priming of this question in [7] are unknown and may account for some of the variation between the results documented in [7] and the results presented in this thesis.

The affective labels that participants could select for each animation stimulus included

Table 5.1: Affect Selection for Neutral Walking Sequence in Study 2

Energetic	*Feminine*	*Happy*	*Masculine*	*Neutral*	*Sad*	*Tired*
10%	15%	5%	25%	25%	10%	10%

energetic, *feminine*, *happy*, *masculine*, *neutral*, *sad*, and *tired*. The total number of times each affective label was selected in Study 2 is presented in Fig. 5.1. The blue bars in Figs. 5.1, 5.2, and 5.3 indicate the instances when the affect label selection results from Study 2 matched the prescribed styles from [7]. The large variation between affective label selection rates may indicate a general preference towards specific labels that can be attributed to inherent priming.

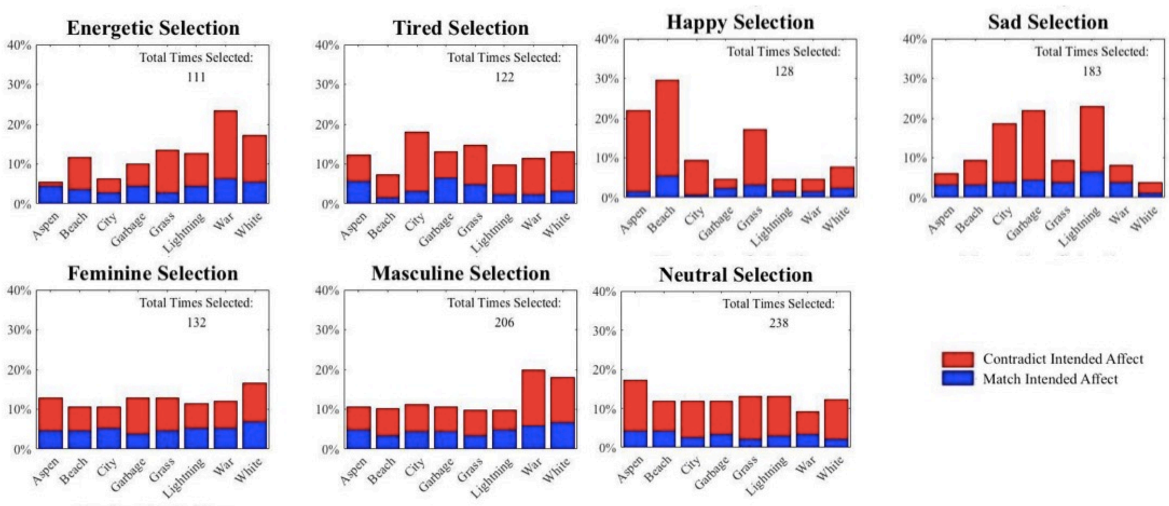

Figure 5.1: These plots show the results of study two. Each plot shows user labeling in a forced choice task to give previously validated walking sequences an affective label. These results show that people typically selected the 'incorrect' label for each walking style. The total number of times each affect label was selected is displayed in each plot.

At least two observations are immediately evident when interpreting the results in Fig. 5.1: (1) people seldom interpreted the affect displayed through the animations as the affect originally intended by the animators and (2) the affect label selection rate was influenced by the environmental context. Both observations are exemplified by the plot labeled "Happy Selection". Approximately 30% of all the *happy* label selections correspond to the *beach* environment while the *garbage*, *lightning*, and *war* environments each account for less than

5% of all the *happy* label selections. As with Study 1, these results were somewhat expected in the sense that the labels selected for an environmental context generally had comparable affective ratings to labels that are commonly associated with the environmental context. Another interesting observation is that the average variance associated with the label selection rates of the bipolar label pair *happy* and *sad* was over 3 times larger than the average variance associated with the bipolar label pair *energetic* and *tired* and over 7 times larger than the average variance associated with the bipolar label pair *feminine* and *masculine*.

The affective label selection results were further analyzed with respect to each stylized walking sequence (Fig. 5.2) and then with respect to each environmental context (Fig. 5.3). The data presented in Fig. 5.2 represents the perceived affects of each walking sequence across all environmental contexts. Overall, the *masculine* sequence had the highest correct identification rate (approximately 50%) and the *happy* sequence had the lowest correct identification rate (approximately 15%). The results for the *happy* and *tired* sequences are interesting in the sense that the affective labels that were selected the most frequently did not match the affective intent of the animators from [7]. The *happy* sequence was viewed as *feminine*, *masculine*, and *neutral* more often than it was seen as *happy* and the *tired* sequence was viewed as *sad* more often than it was seen as *tired*.

The data presented in Fig. 5.3 organizes the Study 2 results so that the affective label selection rates within each environmental context can be compared. The purpose of Fig. 5.3 is to enforce the idea that the affective label selection rates were influenced by the environmental context. The conclusions that can be drawn from Fig. 5.3 are similar to the conclusions formulated from the results presented in Fig. 5.1. The primary benefit of Fig. 5.3 is that significant variability in label selection rates in each environmental context can be determine simply by looking at each subplots.

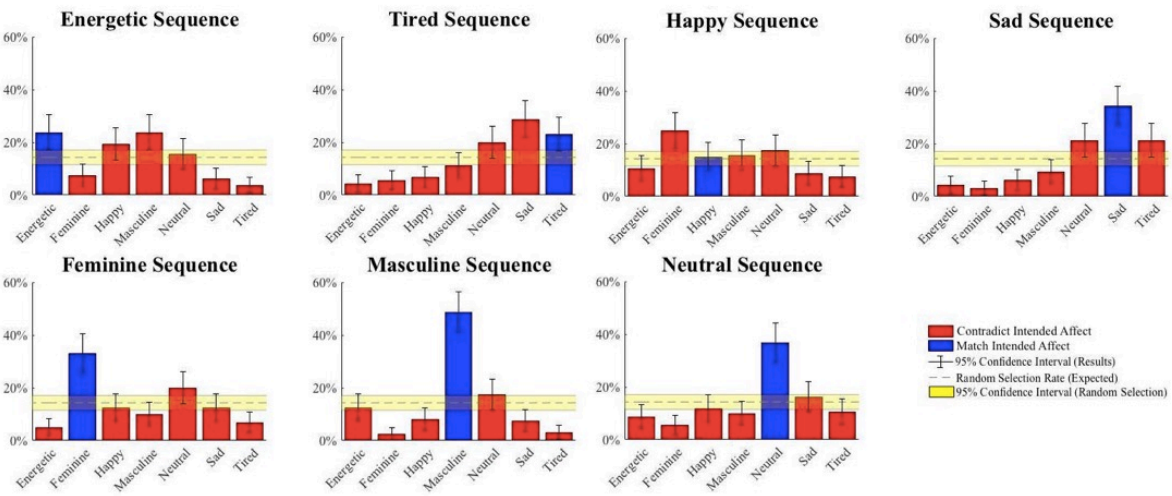

Figure 5.2: These bar graphs present the results of Study 2 with respect to each walking sequence. These plots can be used to examine the affect label selection rates associated with each walking sequence and quickly evaluate the level of agreement between animators and participants with regards to the style portrayed in the walking sequence.

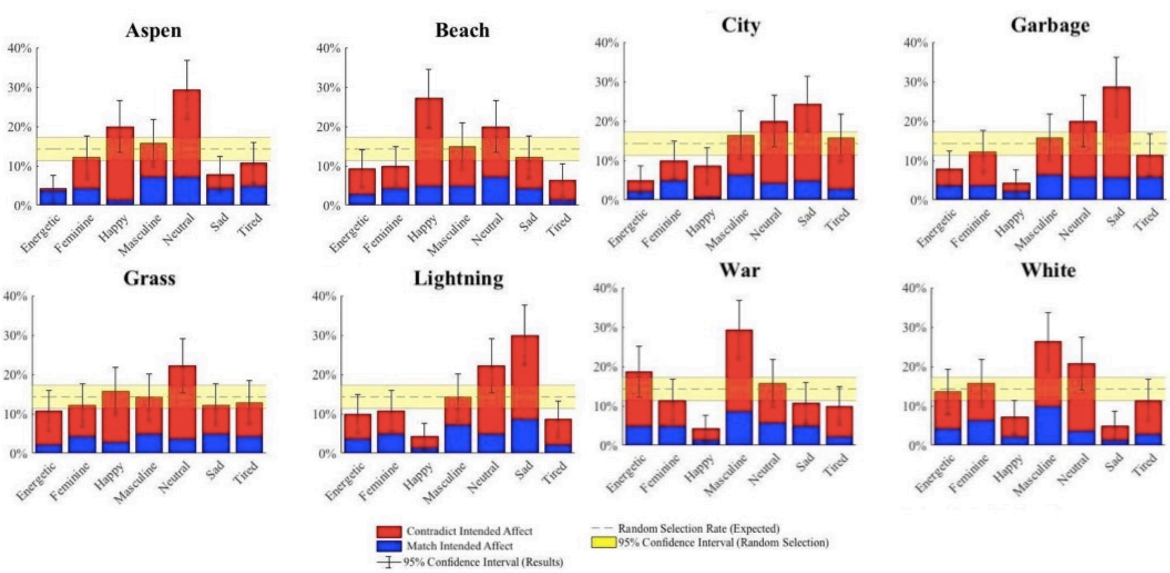

Figure 5.3: These bar graphs present the results of Study 2 with respect to each environmental context. These plots can be used to examine the affect label selection rates associated with each environmental context.

Chapter 6

Study 3: Modeling the Compound Effects of Environment and Gait on Perceived Affect

The primary goal of Study 3 was to utilize the circumplex model of affect to gain greater insight into the compound effects of environmental context and stylized walking sequences on the affect recognition of animations. The study utilized semantic differential rating scales from [53] to collect affective ratings of the environmental contexts, the *white background* walking sequences, and the stylized walking sequences in environmental contexts that were used in Study 2. The format of each questionnaire is discussed in Section 6.1, the method for calculating the affective ratings of each stimuli is presented in Section 6.2, and a discussion of the preliminary findings is included in Section 6.3.

6.1 Questionnaire Format

This user study was comprised of 9 study trials, each containing a questionnaire with the same 100 questions but with respect to different stimuli. The first question in all of the questionnaires asked the participants to confirm that they had read and agreed to the conditions in the consent form. If they accepted the terms, they were taken to the next page where they were presented with 8 demographics questions. Following the demographics questions, the participants were asked to rate 7 stimuli using the 12 semantic differential scales in Table 6.1.

One of the study trials utilized the stylized walking sequences as stimuli, one study trial utilized the environmental context videos as stimuli, and the remaining seven study trials utilized groups of the animations created for Study 2 as stimuli. The groupings for the

Table 6.1: Semantic Differential Rating Scales and Factor Loadings from [53] that were used in Study 3

Low (-4)	**High (4)**	**Valence Factor (f_{val})**	**Arousal Factor (f_{aro})**
Unhappy	Happy	0.914	0.063
Annoyed	Pleased	0.833	0.068
Unsatisfied	Satisfied	0.868	0.144
Melancholic	Contented	0.725	0.095
Despairing	Hopeful	0.858	0.063
Bored	Relaxed	0.580	0.372
Relaxed	Stimulated	-0.211	0.774
Calm	Excited	-0.181	0.793
Sluggish	Frenzied	0.268	0.771
Dull	Jittery	-0.211	0.793
Sleepy	Wide Awake	-0.046	0.810
Unaroused	Aroused	0.051	0.827

Table 6.2: Groupings of Stimuli with both Walking Sequences and Environmental Contexts used in Study 3

		Environmental Context						
		Aspen	**Beach**	**City**	**Garbage**	**Grass**	**Lightning**	**War**
Grouping	**1**	Feminine	Happy	Masculine	Neutral	Sad	Tired	Energetic
	2	Happy	Masculine	Neutral	Sad	Tired	Energetic	Feminine
	3	Masculine	Neutral	Sad	Tired	Energetic	Feminine	Happy
	4	Neutral	Sad	Tired	Energetic	Feminine	Happy	Masculine
	5	Sad	Tired	Energetic	Feminine	Happy	Masculine	Neutral
	6	Tired	Energetic	Feminine	Happy	Masculine	Neutral	Sad
	7	Energetic	Feminine	Happy	Masculine	Neutral	Sad	Tired

stimuli that were created for Study 2 were used in seven of the study trials are presented in Table 6.2.

In all of the study trials, each stimulus was contained on a separate page in the questionnaire and the order of the semantic differential scales was randomized. In addition to the stimulus and the semantic differential scales, each page contained a simple human verification question that was used to help ensure that the participants were not randomly selecting. The human verification questions were only used to validate or invalidate each questionnaire and did not effect the ratings of the stimuli. Screen captures of the types of questions asked in Study 3 are included in Appendix C.

Amazon Mechanical Turk and SurveyMonkey were used to advertise and administer the questionnaires, respectively. To maximize the probability that participants were actually considering the questions and not randomly selecting, only surveys that were completed in a time over 4.5 minutes were kept and used in the analysis. Additionally, if a participant answered more than one of the human verification questions incorrectly, the questionnaire they completed was invalidated and discarded prior to analysis. Since some questionnaires were discarded, the size of the datasets pertaining to each questionnaire varied. The numbers of questionnaires kept and analyzed for each of the stimuli groupings are presented along with other statistical data in Table 6.3. Although English was not the native language of

Table 6.3: Demographic Statistics for each Grouping

		Age				***Gender***		
		Min	**Max**	**Mean**	**SD**	**Female**	**Male**	**Total**
Grouping	**1**	22	57	33.6	8.8	6	14	20
	2	23	57	33.6	9.3	12	13	25
	3	22	68	38.5	11.6	13	14	27
	4	21	64	31.8	10.9	7	18	25
	5	22	63	32.3	11.0	10	16	26
	6	20	60	32.5	11.3	11	15	26
	7	24	66	36.3	11.8	11	17	28
	Walking	25	43	32.9	5.3	7	13	20
	Environment	23	63	34.0	10.2	7	13	20
Overall		20	68	34.0	10.4	84	133	217

all participants, all participants considered themselves fluent in English. Participants were compensated for their time with approximately \$0.10 per minute. Participants were not allowed to take the same survey more than one time, but they were allowed to complete multiple different surveys. Since participation was anonymous, we treated each survey submission as though it was completed by a new participant when this may not actually not the case. The possibility of a single participant completing more than one survey was not initially considered, but if future studies are conducted, we may only allow participants to complete one study.

6.2 Calculating Affective Ratings

The affective ratings $x = [x_{val} \;\; x_{aro}]^T$ for each of the stimuli were calculated using the ratings r collected in the questionnaires and the scaling factors $f = [f_{val} \;\; f_{aro}]^T$ from [53] that are presented in Table 6.1. The valence and arousal ratings of the stimuli were calculating using

$$x_{val} = \frac{\sum_{j=1}^{m} \sum_{n=1}^{N} r_j(n) f_{val,j}}{4N \sum_{j=1}^{m} f_{val,j}}, \tag{6.1}$$

and

$$x_{aro} = \frac{\sum_{j=1}^{m} \sum_{n=1}^{N} r_j(n) f_{aro,j}}{4N \sum_{j=1}^{m} f_{aro,j}}, \tag{6.2}$$

where m is the number of semantic differential scales, N is the number of ratings that exist for each animation, and 4 is used to scale the ratings so that they exist in the interval $[-1, 1]$. For the remainder of this chapter, the affective ratings x will be separated into three sets: environmental contexts ($c = [c_{val} \;\; c_{aro}]^T$), stylized waking sequences ($w = [w_{val} \;\; w_{aro}]^T$), and the animations created for Study 2 ($a = [a_{val} \;\; a_{aro}]^T$).

6.3 Results and Discussion

This section will be decomposed into subsections that discuss the results of the environmental context rating questionnaire, the results of the *white background* stylized walking sequence rating questionnaire, and the least squares approximations that model the compound effects of environmental context and stylized walking sequences on the affective ratings of stimuli in Study 3.

6.3.1 Rating Environmental Contexts

The purposes of this study trial were (1) to compare the affective ratings of the environmental contexts used in Study 2 with similar OASIS images and (2) to obtain baseline affective ratings for the environmental contexts that will be used to develop a model for estimating affective ratings of animations using least squares approximations.

Fig. 6.1 contains a circumplex mapping of the affective ratings for the environmental contexts used in Study 2 and the reference OASIS image set. The ratings for the environmental contexts used in Study 2 were calculated using Equations 6.1 and 6.2. Before the affective ratings of the OASIS reference images were mapped to the circumplex model, they were linearly shifted and scaled using

$$\tilde{o} = \frac{o}{3.5} - 1, \tag{6.3}$$

where $\tilde{o} = [\tilde{o}_{val} \quad \tilde{o}_{aro}]^T$ is the affective rating that is mapped to the circumplex model, $o = [o_{val} \quad o_{aro}]^T$ is the original rating obtained from the OASIS database [15], and 3.5 is the scaling factor needed based on the work in [15].

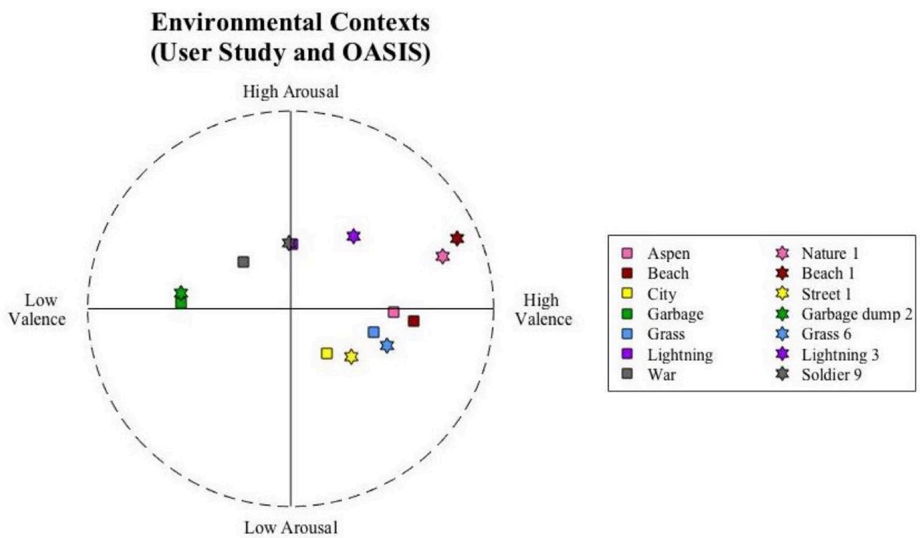

Figure 6.1: The ratings of the environmental contexts that were used as stimuli in Study 2 and Study 3 are displayed as squares and the OASIS reference images [15] are represented by six point stars. The similar colored markers were deemed similar during the image selection process. Each data set was scaled so that the range of each affective dimension was [-1,1].

The results in Fig. 6.1 show that for each corresponding image pair, the OASIS images had affective ratings with larger magnitudes than the environmental contexts used in Study 2 and Study 3. The images in the pairs *city*/Street 1, *garbage*/Garbage Dump 2, and *grass*/Grass

6 had similar valence and arousal ratings to the other image in the pair. The images in each of these pairs was contained within the same quadrant of the circumplex model. The pairs *lightning*/Lightning 3 and *war*/Soldier 9 each contained one image that had a valence rating of approximately zero. The images within these pairs contained comparable arousal ratings but significantly different valence ratings. An interesting observation is that the affective rating of *lightning* and Soldier 9 were approximately the same. The images within the pairs *aspen*/Nature 1 and *beach*/Beach 1 had significantly different valence and arousal ratings. The OASIS images in these pairs had positive arousal ratings while the corresponding environmental contexts from Study 2 and Study 3 had negative arousal ratings.

However, it can be seen in Fig. 3.1 that the *aspen* and *beach* contain similar settings as in Nature 1 and Beach 1, respectively, but vary in manners such as vibrance of the picture, specific objects in the environment, and lighting. When comparing the ratings of *aspen* and *beach* to ratings of other similar images in OASIS like Beach 4 and Lake 3 , the affective ratings of the pairs are closer. The affective ratings of an alternate set of OASIS images that more closely resembles the environmental contexts in Study 2 and Study 3 are mapped to the circumplex model in Fig. 6.2. The results shown in Fig. 6.1 and Fig. 6.2, helps to conclude that the affective ratings of the environmental contexts used in Study 3 were comparable to affective ratings of similar images from the OASIS database.

The affective ratings of the environmental contexts were normalized prior to the creation of the least squares approximation model using

$$\tilde{c_i} = \frac{c_i}{\|c_{max}\|_2}, \tag{6.4}$$

where $\tilde{c_i} = [\tilde{c}_{val,i} \quad \tilde{c}_{aro,i}]^T$ is the normalized affective rating of environmental context i, $c_i = [c_{val,i} \ \ c_{aro,i}]^T$ is the affective rating of environmental context i that was calculated using Equations 6.1 and 6.2, and c_{max} is the environmental context ratings with the largest vector magnitude. The normalizations were performed so that affective ratings corresponding to

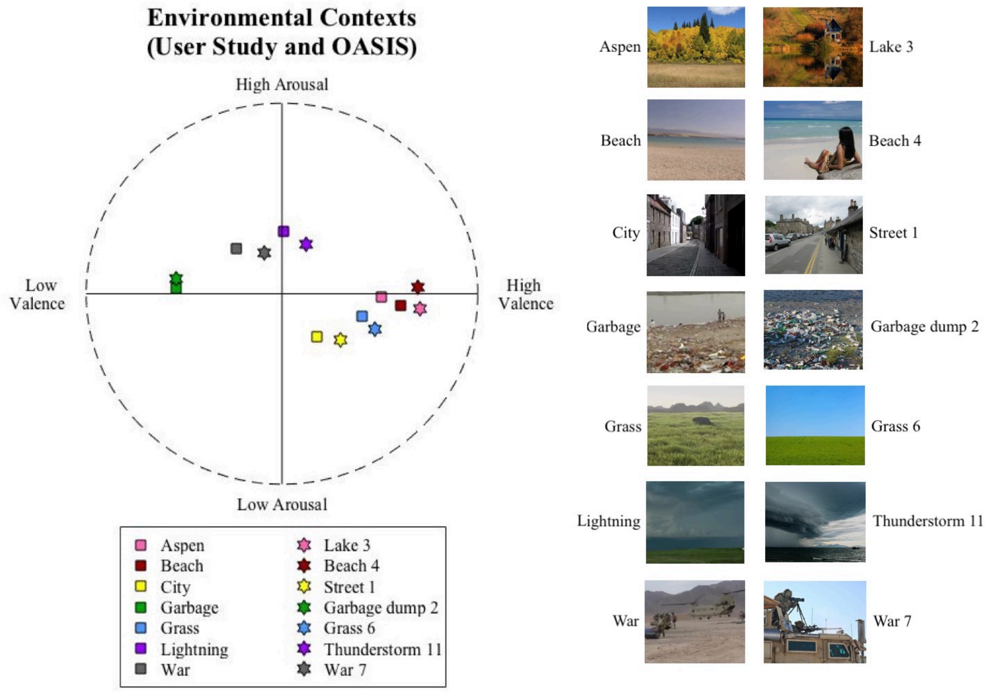

Figure 6.2: The ratings of the environmental contexts that were used as stimuli in Study 2 and Study 3 are displayed as squares and the OASIS images [15] are represented by six point stars. The images that differ from the previous figure include Lake 3, Beach 4, Thunderstorm 11, and War 7. The first column of images depict stimuli in Study 2 and Study 3 and the second column of images alternate OASIS images. Each data set was scaled so that the range of each affective dimension was [-1,1].

the set of environmental contexts would span a larger range of the circumplex model. The differences between the original ratings and the normalized ratings are depicted in Fig. 6.3.

6.3.2 Rating Stylized Walking Sequences

A second study trial contained the stylized walking sequences as stimuli. The purpose of this study trial was to collect baseline affective ratings for the stylized walking sequences that could be later used to develop a least squares approximation model. Similarly to the environmental context ratings, the affective ratings of the stylized walking sequences were normalized prior to the creation of the least squares approximation model using

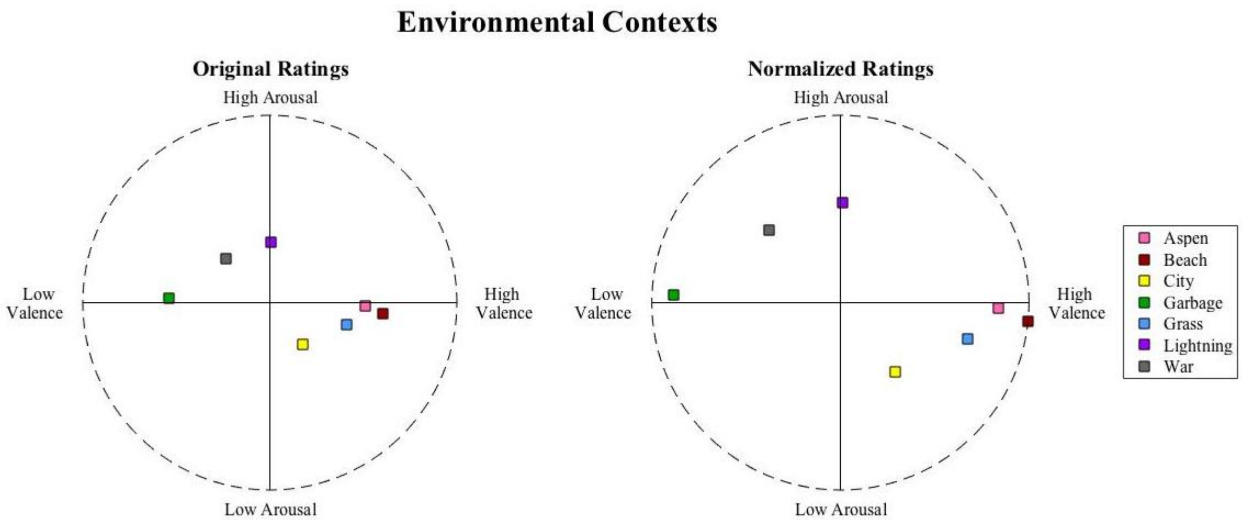

Figure 6.3: The mappings show the original ratings (left) and the normalized ratings (right) of the environmental contexts that were used as stimuli in Study 2 and Study 3. The ratings were normalized using Equation 6.4.

$$\tilde{w_j} = \frac{w_j - w_{neutral}}{\|w_{max} - w_{neutral}\|_2}, \tag{6.5}$$

where $\tilde{w_j} = [\tilde{w}_{val,j} \;\; \tilde{w}_{aro,j}]^T$ is the normalized affective rating of stylized walking sequence j, $w_j = [w_{val,j} \;\; w_{aro,j}]^T$ is the affective rating of stylized walking sequence j that was calculated using Equations 6.1 and 6.2, $w_{neutral} = [w_{val,neutral} \;\; w_{aro,neurtral}]^T$ is the affective rating of the *neutral* walking sequence that was calculated using Equations 6.1 and 6.2, and w_{max} is the stylized walking sequence rating with the largest vector magnitude. The normalizations were performed so that affective ratings corresponding to the set of stylized walking sequences would span a larger range of the circumplex model and so that the *neutral* walking sequence ratings would coincide with the origin of the circumplex mapping. Circumplex mappings of the original and normalized affective ratings associated with each *white background* stylized walking sequence are included in Fig. 6.4.

The normalized affective ratings mapped the stylized walking sequences to the quadrants of the circumplex model that were associated with the style label. The *sad* and *tired* walk-

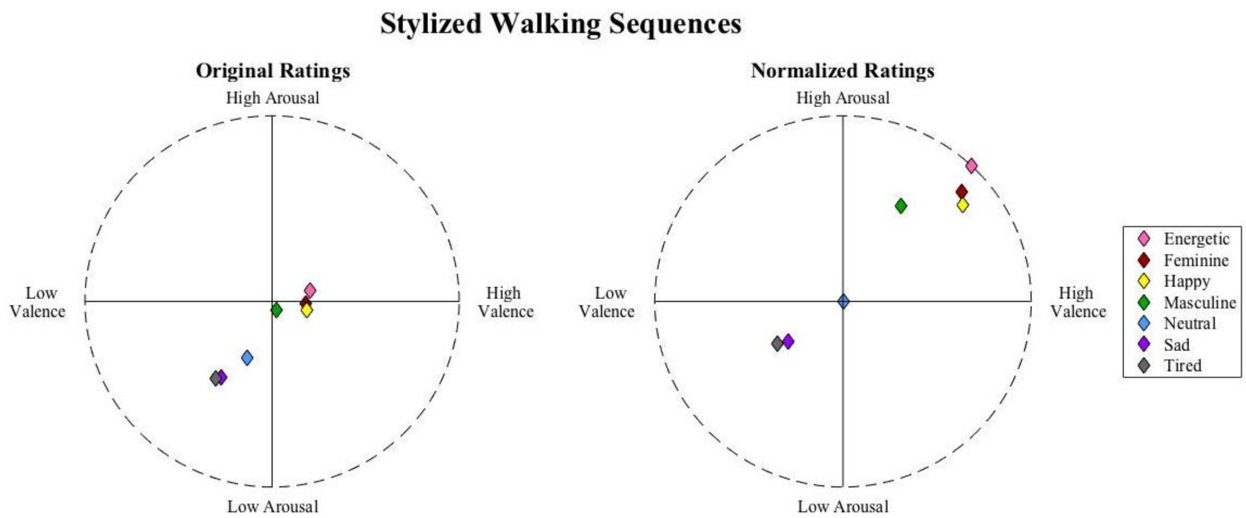

Figure 6.4: The mappings show the original ratings (left) and the normalized ratings (right) of the stylized walking sequences that were used as stimuli in Study 1, Study 2, and Study 3. The ratings were normalized using Equation 6.5.

ing sequences were in the quadrant associated with negative valence and arousal while the *energetic* and *happy* walking sequences were mapped to the quadrant with positive valence and arousal. It is worth noting that the normalized affective ratings only mapped the stylized walking sequences to two of the four quadrants. In fact, each walking sequence has a normalized valence rating that is approximately the same the normalized arousal rating.

6.3.3 Rating Animations

The remaining seven study trials contained the new animations that were created for Study 2 as stimuli. The purpose of these study trials was to collect affective ratings for each of the 49 new animations that could be used to develop least squares approximation models. The affective ratings of the animations were normalized using

$$\tilde{a_k} = \frac{a_k}{\|a_{max}\|_2}, \tag{6.6}$$

where $\tilde{a_k} = [\tilde{a}_{val,k} \;\; \tilde{a}_{aro,k}]^T$ is the normalized affective rating of animation k, $a_k = [a_{val,k} \;\; a_{aro,k}]^T$

is the affective rating of animation k that was calculated using Equations 6.1 and 6.2, and a_{max} is the animation rating with the largest vector magnitude. The normalized affective ratings for the animations are mapped to circumplex models that are separated by environmental context in Fig. 6.5. The mappings in Fig. 6.5 indicate that some of the environmental contexts influenced the affective ratings of the animations significantly more than the stylized walking sequences. This influence manifests as clusters of affective ratings and is evident in at least two environmental contexts: *city* and *war*. It is worth noting that there seems be loose groupings and/or trends in each of the circumplex mappings in Fig. 6.5. For example, the affective ratings in *aspen* and *garbage* are generally close to imaginary lines with 30° and 45° inclinations, respectively, that runs through the origin of the circumplex model. Additional circumplex mappings for the animations that contain the normalized affective ratings separated by walking sequence and the original affective ratings separated by both environmental context and walking sequence are included in Appendix D.

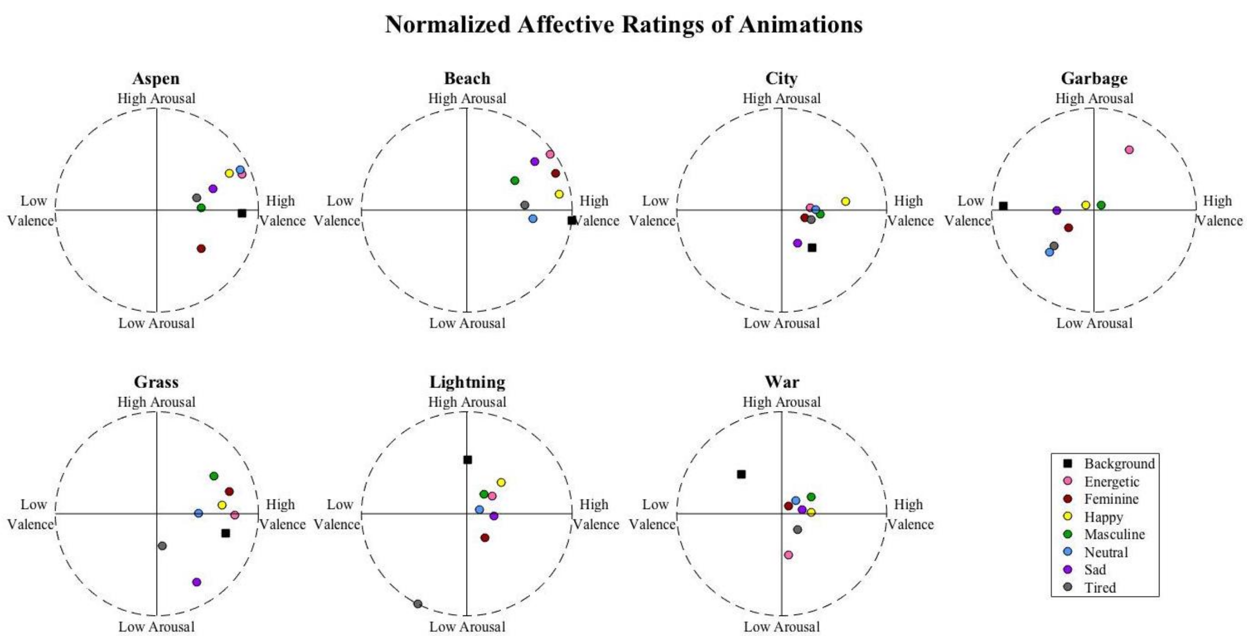

Figure 6.5: The mappings show the normalized ratings of the animations that were created for Study 2. The ratings of the animations are separated by environmental context and are displayed as circles. The normalized ratings of the environmental contexts are also included and are represented by black squares. The animation ratings were normalized using Equation 6.6.

6.3.4 Preliminary Least Squares Approximation Model for Predicting Affective Ratings of Stimuli

The primary goal of this user study was to model the compound effects of stylized walking sequences and environmental contexts on the affective ratings of the animations leveraged in Study 2. To accomplish this, two models were explored. Model 1 was developed with the assumption that the affective dimensions are orthogonal and Model 2 was developed with the assumption that the affective dimensions of the circumplex model may perfectly describe the collected data. The goal for Model 1 was to gain intuition about the compound effects of stylized walking sequences and environmental contexts on the affective ratings of the animations and the goal for Model 2 was to minimize the affective rating approximation error. The equations that were used to create the models are

$$\min_{W,C} \left(\sum_{k=1}^{M} \|\tilde{a}_k - \hat{a}_k\|_2^2 \right), \tag{6.7}$$

$$\hat{a}_k = \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix} \tilde{w}_k + \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \tilde{c}_k, \tag{6.8}$$

where $\hat{a}_k = [\hat{a}_{val,k} \quad \hat{a}_{aro,k}]^T$ is the least squares approximation of an affective rating for animation k, $\tilde{w}_k$ and $\tilde{c}_k$ are the normalized ratings for the walking sequence and environmental context (respectively) that correspond to animation k, $\mathbf{W}$ and $\mathbf{C}$ are the coefficient matrices that solve the least squares problem, $\tilde{a}_k$ is the normalized rating of animation k, and M is the number of animations evaluated in Study 3.

Model 1 treated the valence and arousal ratings as orthogonal components, meaning arousal ratings don't influence valence ratings, and vice versa. This was done so that the valence and arousal dimensions that are calculated stay consistent with the definitions in existing literature [31, 35]. For this model, the off diagonal values in the coefficient matrices in Equation 6.8 were set to zero (i.e. $W_{12} = W_{21} = C_{12} = C_{21} = 0$) prior to the least squares calculations. The results from the least squares calculations were used to construct the following model:

$$\begin{bmatrix} \hat{a}_{val,k} \\ \hat{a}_{aro,k} \end{bmatrix} = \begin{bmatrix} 0.415 & 0 \\ 0 & 0.224 \end{bmatrix} \begin{bmatrix} \tilde{w}_{val,k} \\ \tilde{w}_{aro,k} \end{bmatrix} + \begin{bmatrix} 0.489 & 0 \\ 0 & -0.093 \end{bmatrix} \begin{bmatrix} \tilde{c}_{val,k} \\ \tilde{c}_{aro,k} \end{bmatrix}. \tag{6.9}$$

The average least squares error for the Model 1 approximations was 0.275 (0.178 (*valence*), 0.209 (*arousal*)). The coefficients of the least squares fit indicate that the environmental contexts influenced the valence ratings of the animations more than the stylized walking sequences and the stylized walking sequences influenced the arousal ratings of the animations more than the environmental contexts. In [54], Robinson proposes that when humans are generating an emotional response to a stimulus, preattentive processing is initially used to assess the valence and urgency (arousal) of the stimulus. The preattentive valence detection mechanism serves to initially determine the potential significance of a stimulus and and preattentive urgency (arousal) detection mechanism assesses and responds to a stimulus based on the perceived significance. Applying this theory to the results of Model 1, it makes sense that valence is influenced more by the relatively static environmental contexts and arousal is primarily influenced by the dynamics stick figure animation. In the environmental context stimuli it is likely that the valence detection is the primary focus during the initial affective evaluation due to a lack of dynamic components. In the walking sequences the only stimulus is a dynamic walking stick figure, making it likely that urgency detection will be the primary focus during the initial affective evaluation. Using this logic, when the walking sequences and environmental contexts are combined to create an animation, it makes sense that the environment would be the focal point in valence detection and the dynamic stick figure would be the focal point in arousal detection.

Model 2 was created with the goal of minimizing the least squares error for the predictions. This was achieved by treating all of the indices of the coefficient matrices as free variables during the least squares calculations. The results from the least squares calculation were used to create the following model:

$$\begin{bmatrix} \hat{a}_{val,k} \\ \hat{a}_{aro,k} \end{bmatrix} = \begin{bmatrix} -0.175 & 0.615 \\ 0.211 & -0.017 \end{bmatrix} \begin{bmatrix} \tilde{w}_{val,k} \\ \tilde{w}_{aro,k} \end{bmatrix} + \begin{bmatrix} 0.488 & 0.081 \\ 0.096 & -0.002 \end{bmatrix} \begin{bmatrix} \tilde{c}_{val,k} \\ \tilde{c}_{aro,k} \end{bmatrix}. \tag{6.10}$$

The average least squares error for the Model 2 approximations was 0.265 (0.177 (*valence*), 0.197 (*arousal*)), which is only marginally better than Model 1. Additionally, there aren't any distinct patterns in Model 2 that provide further insight into the compound effects of the stylized walking sequences and the environmental contexts on the affective ratings of the animations. These findings suggest that the assumption made during the development of Model 1 were accurate enough to produce an adequate model. These findings also seem to support the claim that the dimensions of valence and arousal are orthogonal.

Chapter 7

Towards Understanding the Relationship Between Movement, Context, and Affect Recognition

The contributions that are presented in this thesis can be broken down into two components: development of MATLAB tools to assist in stimulus creation and user studies that investigate the role of environmental context in affect recognition of stylized movements. The MATLAB tools provide a platform where data-driven (motion capture data) trajectories can be created and simulated in various environmental contexts. The trajectories that are simulated can then be saved and used as stimuli for user studies. Additionally, the tools were designed with the goal of making the tools highly customizable in the sense that it is simple to integrate new data, new backgrounds, and add additional dimensions (i.e. audio, etc.). The MATLAB tools, as well as the user studies, can be used as the foundation of further research that seeks to understand the relationship between movement, context, and affect.

The three user studies were conducted in order to gain additional insight into:

1. environmental context selection when subliminally primed with an affective label,
2. affective label selection rates of various combinations of stylized walking sequences and environmental contexts,
3. the compound effects of an environmental context and a stylized walking sequence on affective ratings of multi-dimensional stimuli (i.e. a stylized walking sequence in an environmental context).

There were two high-level take aways from the first two studies that aided in the development of the third study. The first take-away was that environmental context does influence a human's perception of the affect displayed through a stylized walking sequence. The second take-away was that it is not possible to adequately characterize a multi-dimensional stimulus using a single affective label. This was made evident by the large number of 'misclassifications' that were found in Study 2.

With the high-level take-aways from the first two studies in mind, a third user study was conducted that assigned two-dimensional, semi-quantitative affective ratings to stimuli based on semantic differential scale results collected in a user study. The goals were (1) to gain intuition about the compound effects of stylized walking sequences and environmental contexts on the affective ratings of the animations and (2) to develop a model that accurately approximated the affective ratings of animations using the affective ratings of stylized walking sequences and environmental contexts as inputs. Two models were originally developed, one to solve each of the goals, but it was discovered that Model 1 performed as well, or better, than Model 2 with respect to both goals. Model 1 indicated that the environmental context contributed more to the valence rating of the animation and the stylized walking sequence contributed more to the arousal rating of the animation. Additionally, the average least squares error associated with the Model 1 approximations was 0.275.

If time would have permitted, follow-up studies with the following goals may have been conducted:

- Develop a model to predict affective ratings of environmental contexts based on quantifiable metrics such as light intensity, temperature, noise, entropy, etc.
- Develop a model to predict affective ratings of movement trajectories based on quantifiable metrics such as position, velocity, acceleration, heading, etc.
- Investigate the influence of additional contextual parameters (i.e. situational context, multiple agents present, audial dimension of environmental context, etc.) on perceived affect.

In closing, the work presented in this thesis creates a foundation for future studies that investigate the relationship between movement, context, and affect recognition. The MATLAB tools presented in Chapter 3 can be used to quickly and efficiently create multi-dimensional stimuli to be used in future research studies. Similarly to the affective image databases, the results from the user studies presented in Chapters 4, 5, and 6 can be used as reference sets of affectively rated stimuli. In addition the the reference ratings, the model proposed in Chapter 6 can be used as a basis for creating a robust model for predicting the affective ratings of a multi-dimensional stimulus using the affective ratings of one-dimensional stimuli.

7.1 Publications

The work done in this thesis produced the following publication:

- M. Heimerdinger and A. LaViers, "Influence of Environmental Context on Recognition Rates of Stylized Walking Sequences," in *International Conference on Social Robotics*. Springer International Publishing, 2017, pp. 272-282.

The work done in this thesis also contributed to the following papers that are under preparation:

- A. LaViers *et al.*, "Choreographic and Somatic Practices Toward the Development of Expressive Robotic Systems," *Arts, The Machine as Artist (for the 21st Century)*, under preparation.

- M. Heimerdinger and A. LaViers, "Modeling the Compound Effects of Environment and Gait on Perceived Affect," in *IEEE Transactions on Human-Machine Systems*, under preparation.

Appendix A

Read Me for V-SDDTE

This file will discuss 6 topics:

1. The MATLAB files that you need to edit
2. List of `.m` files and folder structure required to run `Viewing_Mocap _Trajectories.m`
3. Required formatting and naming for CSV files for code AS IS
4. Changes required if you want to change the CSV file format
5. How to add additional background images
6. How to add additional quadrotor models

1. The MATLAB files that you need to edit

The only files that you need to open and edit are the files called `Viewing_Mocap _Trajectories.m` and `Viewing_Mocap_Trajectories.fig`.

To properly open and edit `Viewing_Mocap_Trajectories.fig`, open MATLAB, right click `Viewing_Mocap_Trajectories.fig`, and click "Open in `GUIDE`". The `Viewing_Mocap _Trajectories.fig`, file is the graphical interface that is used to change the physical appearance of the GUI. There are ways to add components to the GUI by adding additional commands to `Viewing_Mocap_Trajectories.m`, but this GUI was constructed primarily in GUIDE. If you edit the GUI, it is suggested that you also use `GUIDE`.

The file `Viewing_Mocap_Trajectories.m` contains the code that is associated with the GUI defined in `Viewing_Mocap_Trajectories.fig`. The code in `Viewing_Mocap _Trajectories.m` corresponds to each of the dynamic components (i.e. edit text, push buttons, pop up menus, etc.) that were added. The static components (i.e. static text, panels, etc) do not produce any code that is visible in `Viewing_Mocap_Trajectories.m`.

The next couple of paragraphs will describe key functions of the code in `Viewing_Mocap _Trajectories.m`.

The first function that you want to concern yourself with is as follows:

```
function Viewing_Mocap _Trajectories_OpeningFcn (hObject, eventdata, handles,
varargin)
```

This function is used to define variables prior to the GUI becoming visible. If you want to define any variables that are defined after the GUI is built but prior to the GUI being visible to the user, this is the section where you define them (see code for examples).

There are two additional types of functions that are currently (as of June 7, 2017) being used in the GUI. They are `Callback` and `CreateFcn` and an example for a component with the tag, `COMPONENT_TAG` are shown below.

```
function COMPONENT_TAG_Callback(hObject, eventdata, handles)
function COMPONENT_TAG_CreateFcn(hObject, eventdata, handles)
```

The `CreateFcn` is used to create the physical component that is seen in the GUI. This section can be used to define additional characteristics to a component that weren't previously defined in `Viewing_Mocap_Trajectories.fig`. For example, in function `person _popupmenu_CreateFcn(hObject, eventdata, handles)` the options that the user is able to select are defined dynamically. This is done because the choices may change over time. Therefore, the choices will update automatically without the user having to manually change the possible selection choices.

The `Callback` is used to update variables when the state of a component has been changed (i.e. someone changes a value in an edit text box, someone selects a new option from a popup menu, etc.). This type of function is pretty straight forward. See `Viewing_Mocap _Trajectories.m` for examples.

Adding Styles Note:

If you want to add/change styles to the GUI, add them to the `String` parameter defined in `Viewing_Mocap_Trajectories.fig`. Once the style is added to the GUI interface, add a case to the function function `style_popupmenu_Callback(hObject, eventdata, handles)` to reflect the added style so it can be incorporated

IMPORTANT NOTE: After you define or update any handles variable, you must include the following text at the end of the function: `guidata(hObject,handles)`. This text saves the variable so that it can be used in other functions. If you do not include this text, the variable will not be stored after the completion of the function.

2. List of files and folder structure required to run `Data_Driven_Quad.m`

Required files

`get_cntr_of_mass.m`

`get_data_pts.m`

`get_rot_stat_pts.m`

`get_stat_pts.m`

`get_stick_fig_pts.m`

`plot_walking_animation.m`

`quadrotor_animation_all_affect.m`

`quadrotor_animation_one_affect.m`

`stlRead.m`

`stl.GetFormat.m`

`stlReadAscii.m`

`stlReadBinary.m`

`TabManager.m`

`Viewing_Mocap _Trajectories.fig`

`Viewing_Mocap _Trajectories.m`

`walking_animation_all_affect.m`

`walking_animation_one_affect.m`

`quadrotor_dummy.stl`

Folders

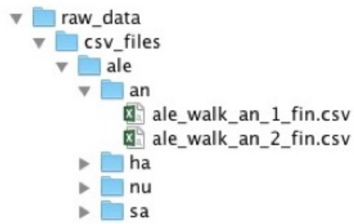

Figure A.1: Subfolder structure for V-SDDTE

In the folder containing all of the .m files, there must also be three folders. They MUST be named `background_images`, `stl_files`, and `raw_data`.

The `background_images` folder will contain subfolders that describe various backgrounds. IT IS VERY IMPORTANT THAT YOU NEVER DELETE THE SUBFOLDER LABELED "`None`". The names of these subfolders are used to generate background selection options in the GUI. Within each of these subfolders, there are 2 pictures corresponding to images that will be projected onto the xy-plane and the xz/yz-planes.

The `stl_files` folder contains the models that are used to create the quadrotor simulation. Theoretically, any `.stl` model will work, but if it is really large, it may cause the system to be incredibly slow or crash. Try to keep the features to a minimum to reduce the number of vertices and faces in the STL file.

The `raw_data` folder must contain a subfolder named `csv_files` for the GUI to run. Within the `csv_files` folder, there are subfolders that correspond to the person (3 character code) that was filmed in the motion capture. Within the `person` subfolder, there are additional subfolders that correspond to the style of the walking sequences (2 character code). The CSV files with the walking data are corresponding to the person and style are contained within the `style` subfolders. There is a screen capture of the subfolder structure shown below in Fig. A.1. The naming convention for the CSV files is discussed in the next section.

3. Required formatting and naming for CSV files for code AS IS

The CSV files must be names in the following format:

`AAA_walk_BB_C_One_Cycle.csv`

Where `AAA` corresponds to a 3 character code associated with the person who was recorded, `BB` is a 2 character code associated with the style or affect of the movement, and `C` corresponds to the direction the person walks during the movement capture recording and is either 1 (counter-clockwise), 2 (clockwise), or 3 (straight).

If you do not name the CSV files correctly, or put them in the incorrect folder, the GUI will not run.

The columns of the CSV file must be as follows:

columns 1-3: head (x,y,z)

columns 4-6: neck (x,y,z)

columns 7-9: left shoulder (x,y,z)

columns 10-12: left elbow (x,y,z)

columns 13-15: left wrist (x,y,z)

columns 16-18: right shoulder (x,y,z)

columns 19-21: right elbow (x,y,z)

columns 22-24: right wrist (x,y,z)

columns 25-27: pelvis (x,y,z)

columns 28-30: left hip (x,y,z)

columns 31-33: left knee (x,y,z)

columns 34-36: left ankle (x,y,z)

columns 37-39: right hip (x,y,z)

columns 40-42: right knee (x,y,z)

columns 43-45: right ankle (x,y,z)

columns 46-48: COM (x,y,z)

columns 49-51: lower body COM (x,y,z)

columns 52-54: upper body COM (x,y,z)

4. How to get CSV files with individual cycle data if the folder One_Cycle does not exist

If you just want to change the naming format but you want to keep the contents of the file the same, you can but you will have to edit code in `Viewing_Mocap _Trajectories.m`, `walking_animation_one_affect.m`, and `walking _animation_all_affect.m`

If you want to change the contents of the file (i.e. add additional joints or remove existing joints) you need to reformat `get_cntr_of_mass.m`, `get_data_pts.m`, `get_rot_stat_pts.m`, `get_stat_pts.m`, `get_stick_fig_pts.m`, `walking_animation_one_affect.m`, and `walking _animation_all_affect.m`

5. How to add additional background images

The first paragraph is directly copied from a previous section. This was done in case you did not read said section or needed a refresher. The `background_images` folder will contain subfolders that describe various backgrounds. IT IS VERY IMPORTANT THAT YOU NEVER DELETE THE SUBFOLDER LABELED "`None`". The names of these subfolders are used to generate background selection options in the GUI. Within each of these subfolders, there are 2 pictures corresponding to images that will be projected onto the xy-plane and the xz/yz-planes.

To create a new background, the first thing you need to do is add a subfolder to the `background_images`. The name of the folder should be the background description/name. The folder name will be used to create the background selection option in background context popup menu in the GUI.

Within the folder you just created, you need to add 2 pictures corresponding to images that will be projected onto the xy-plane and the xz/yz-planes (sides). They need to be named `xy.jpg` for the xy-plane image, `side.jpg` for the xz/yz-planes image. If you do not name these correctly, or put them in the incorrect folder, the GUI may not run.

6. How to add additional aerial models

To create a new aerial model, you just need to add the STL file for the model to the folder labeled `stl_files`. The name of the STL file will automatically appear in the GUI the next time you run it. The naming format does not matter as long as it ends in `.stl`.

Appendix B

Read Me for C-SDDTE

This file will discuss 6 topics:

1. List of .m files and folder structure required to run Data_Driven_Quad.m

2. Required formatting and naming for CSV files for code AS IS

3. How to get CSV files with individual cycle data if the folder One_Cycle does not exist

4. Defining new primary trajectories

5. How to add additional background images

6. How to add additional quadrotor models

1. List of files and folder structure required to run Data_Driven_Quad.m

Required files

Data_Driven_Quad.m

Data_Driven_Quad.fig

baseline_trajectories.m

populate_GUI.m

run_GUI_animation.m

TabManager.m

working_cycle.m

quadrotor_dummy.stl

stlRead.m

stlGetFormat.m

stlReadAscii.m

`stlReadBinary.m`

`stlSlimVerts.m`

Folders

In the folder containing all of the .m files, there must also be three folders. They MUST be named `background_images`, `stl_files`, and `One_Cycle`.

The `background_images` folder will contain subfolders that describe various backgrounds. IT IS VERY IMPORTANT THAT YOU NEVER DELETE THE SUBFOLDER LABELED "`None`". The names of these subfolders are used to generate background selection options in the GUI. Within each of these subfolders, there are 2 pictures corresponding to images that will be projected onto the xy-plane and the xz/yz-planes.

The `stl_files` folder contains the models that are used to create the quadrotor simulation. Theoretically, any `.stl` model will work, but if it is really large, it may cause the system to be incredibly slow or crash. Try to keep the features to a minimum to reduce the number of vertices and faces in the STL file.

The `One_Cycle` folder must contain all of the CSV files directly in it. The CSV files can not be contained in subfolders of the program will not be able to find them. The naming convention for the CSV files is discussed in the next section.

2. Required formatting and naming for CSV files for code AS IS

The CSV files must be names in the following format:

`AAA_walk_BB_C_One_Cycle.csv`

Where `AAA` corresponds to a 3 character code associated with the person who was recorded, `BB` is a 2 character code associated with the style or affect of the movement, and `C` corresponds to the direction the person walks during the movement capture recording and is either 1 (counter-clockwise), 2 (clockwise), or 3 (straight).

If you do not name the CSV files correctly, or put them in the incorrect folder, the GUI will not run.

The columns of the CSV file must be as follows:

columns 1-3: head (x,y,z)

columns 4-6: neck (x,y,z)

columns 7-9: left shoulder (x,y,z)

columns 10-12: left elbow (x,y,z)

columns 13-15: left wrist (x,y,z)

columns 16-18: right shoulder (x,y,z)

columns 19-21: right elbow (x,y,z)

columns 22-24: right wrist (x,y,z)

columns 25-27: pelvis (x,y,z)

columns 28-30: left hip (x,y,z)

columns 31-33: left knee (x,y,z)

columns 34-36: left ankle (x,y,z)

columns 37-39: right hip (x,y,z)

columns 40-42: right knee (x,y,z)

columns 43-45: right ankle (x,y,z)

columns 46-48: COM (x,y,z)

columns 49-51: lower body COM (x,y,z)

columns 52-54: upper body COM (x,y,z)

3. How to get CSV files with individual cycle data if the folder `One_Cycle` does not exist

In the folder containing `Data_Driven_Quad.m` there is a subfolder named `Format Raw Data`. If `One_Cycle` does not exist, open `create_individual_cycle.m` and run it. This will create the `One_Cycle` folder and format all the data correctly. The list of files below are required to run the `create_individual_cycle.m` code.

`create_individual_cycle.m`

`get_cntr_of_mass.m`

`get_data_pts.m`

`get_rot_stat_pts.m`

`get_stat_pts.m`

In addition to the above files, there needs to be a folder named `raw_data` that contains a subfolder named `csv_files`. There should be no subfolders in `csv_files`, only CSV files. If the subfolder `csv_files` doesn't exist, open `convert_ptd.m` and run it. When it is finished running, make sure that the the contents of the `csv_files` folder are formatted correctly.

4. Defining new primary trajectories

To define new baseline trajectories, open the `.m` file named `baseline_trajectories.m`. As of 08/25/2017, there are 3 baseline trajectories already defined (Arc, Circle, Line).

When creating a new trajectory, it is important to make sure that the length of the vectors defining the trajectory are equal to the variable `mod_len`. Additionally, the vectors for the new trajectory must be included in a structure. See the example below. In example below, NNNNN is the name that you want to appear in the MATLAB GUI.

```
NNNNN_x= desired x trajectory; %EDIT - x coordinates for baseline trajectory
NNNNN_y=desired y trajectory; %EDIT - y coordinates for baseline trajectory
NNNNN_z=ones(mod_len,1)*hand.z_height; %DO NOT EDIT - z coordinated for baseline trajectory
data.NNNNN=struct('x',NNNNN_x,'y',NNNNN_y,'z',NNNNN_z);%DO NOT EDIT - defines the baseline trajectory and outputs it from the function
```

5. How to add additional background images

The first paragraph is directly copied from a previous section. This was done in case you did not read said section or needed a refresher. The `background_images` folder will contain subfolders that describe various backgrounds. IT IS VERY IMPORTANT THAT YOU NEVER DELETE THE SUBFOLDER LABELED "`None`". The names of these subfolders are used to generate background selection options in the GUI. Within each of these subfolders, there are 2 pictures corresponding to images that will be projected onto the xy-plane and the xz/yz-planes.

To create a new background, the first thing you need to do is add a subfolder to the `background_images`. The name of the folder should be the background description/name. The folder name will be used to create the background selection option in background context popup menu in the GUI.

Within the folder you just created, you need to add 2 pictures corresponding to images that will be projected onto the xy-plane and the xz/yz-planes (sides). They need to be named `xy.jpg` for the xy-plane image, `side.jpg` for the xz/yz-planes image. If you do not name these correctly, or put them in the incorrect folder, the GUI may not run.

6. How to add additional aerial models

To create a new aerial model, you just need to add the STL file for the model to the folder labeled `stl_files`. The name of the STL file will automatically appear in the GUI the next time you run it. The naming format does not matter as long as it ends in `.stl`.

Appendix C

Screen Captures of User Study Questions

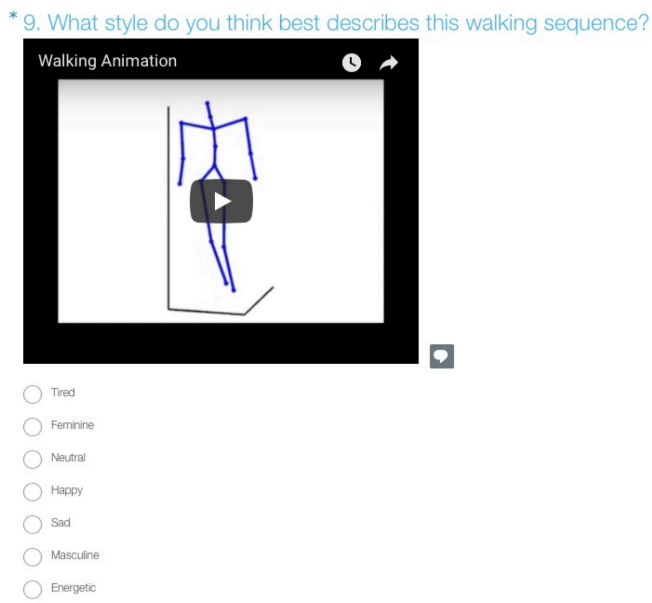

Figure C.1: This figure shows a screen capture of the first type of question in Study 1 as the type of question asked in Study 2

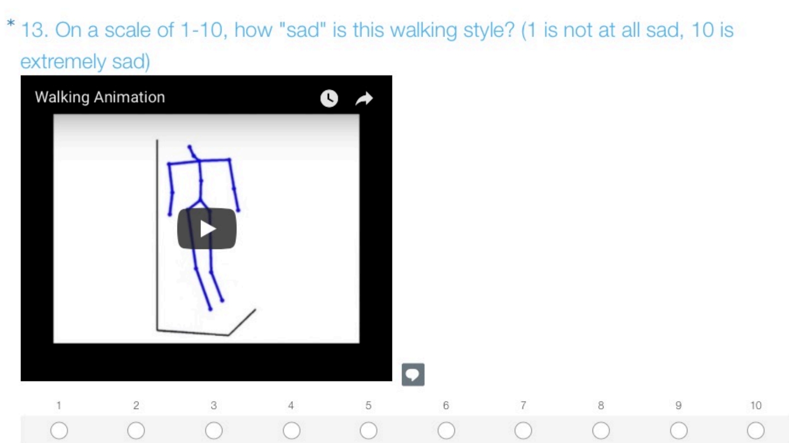

Figure C.2: This figure shows a screen capture of the second type of question in Study 1.

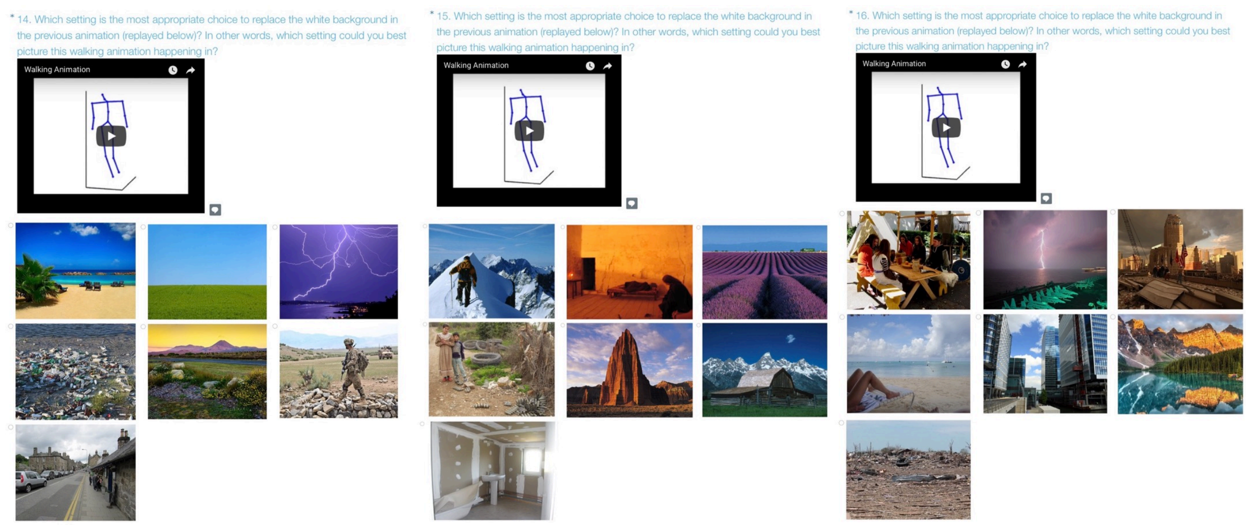

Figure C.3: This figure shows a screen captures of the third type of question in Study 1. This example contains the OASIS reference image set [15] (left), the GAPED image set [16] (center), and the the OASIS second image set [15] (right).

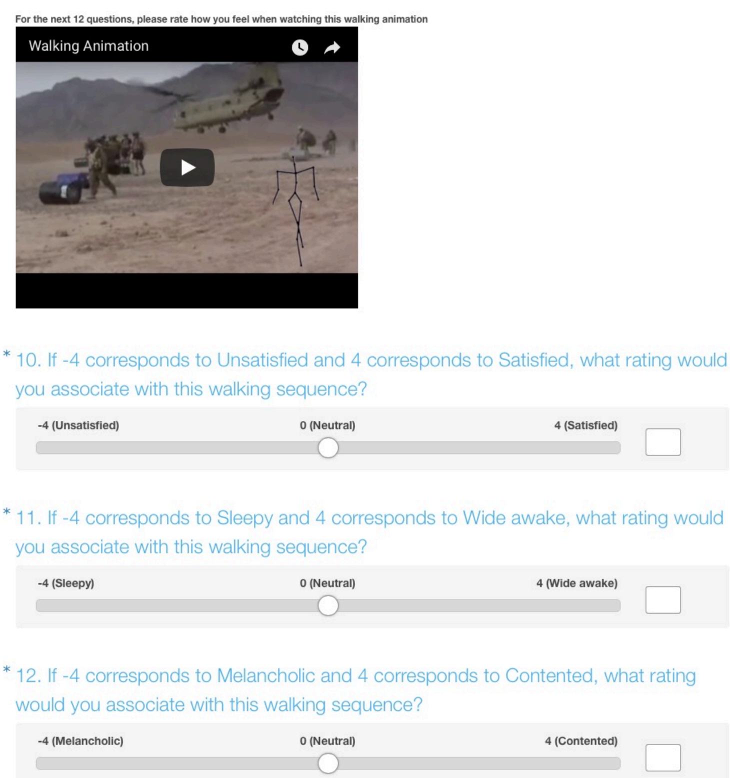

Figure C.4: This figure shows a screen captures of some of the semantic differential scale questions asked in Study 3. The screen capture also includes the video being evaluated at the top of the figure.

* 22. How many people are in this photo?

- ◯ 5
- ◯ 6
- ◯ 2
- ◯ 1
- ◯ 3
- ◯ 4

Figure C.5: This figure shows a screen captures of one of the questions that was asked in Study 3 to help reduce the likelihood that the participants were not random selecting.

Appendix D

Additional Results from User Studies

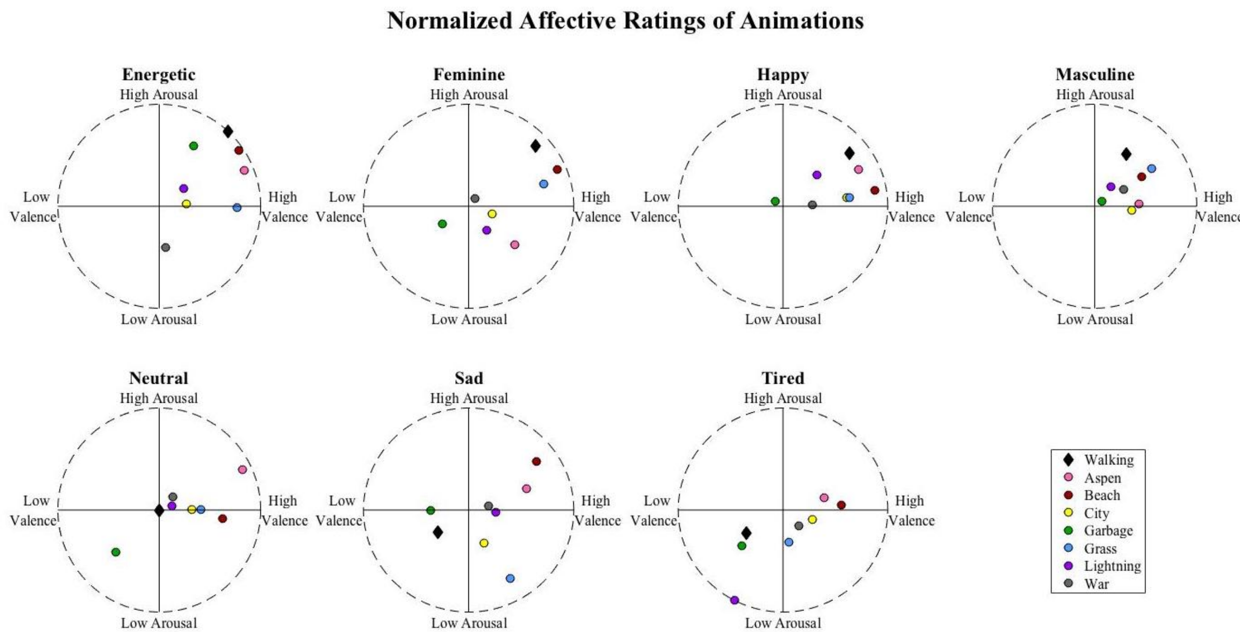

Figure D.1: The mappings show the normalized ratings of the animations that were created for Study 2. The ratings of the animations are separated by stylized walking sequence and are displayed as circles. The normalized ratings of the stylized walking sequences are also included and are represented by black diamonds. The animation ratings were normalized using Equation 6.6.

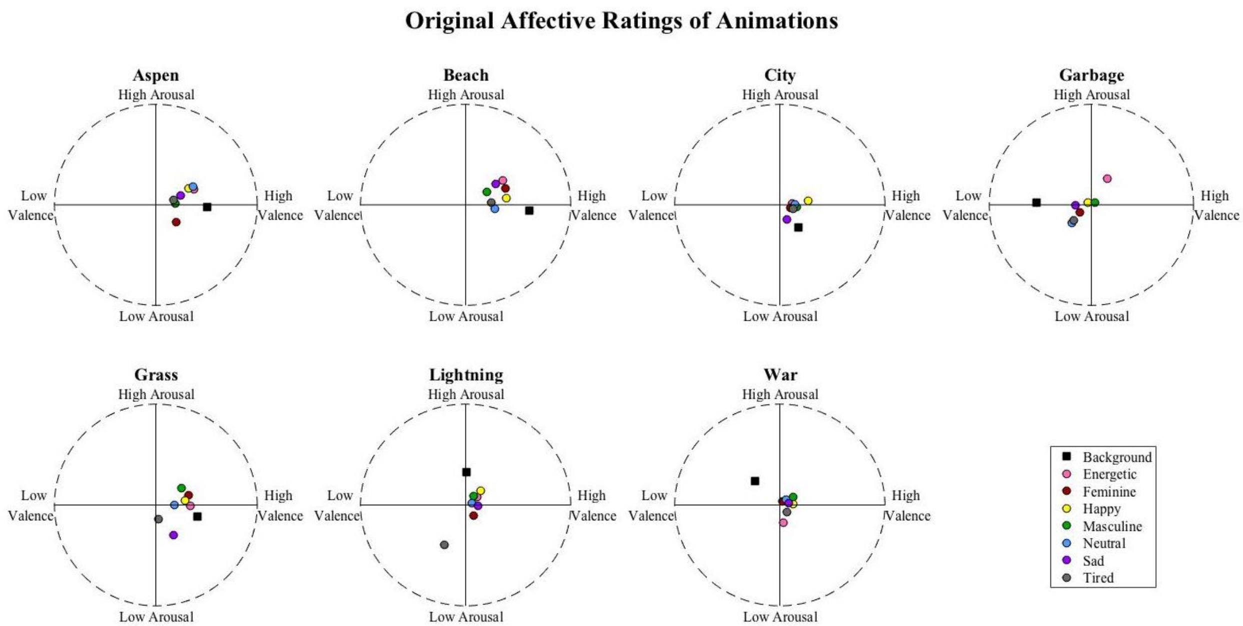

Figure D.2: The mappings show the original ratings of the animations that were created for Study 2. The ratings of the animations are separated by environmental context and are displayed as circles. The original ratings of the environmental contexts are also included and are represented by black squares.

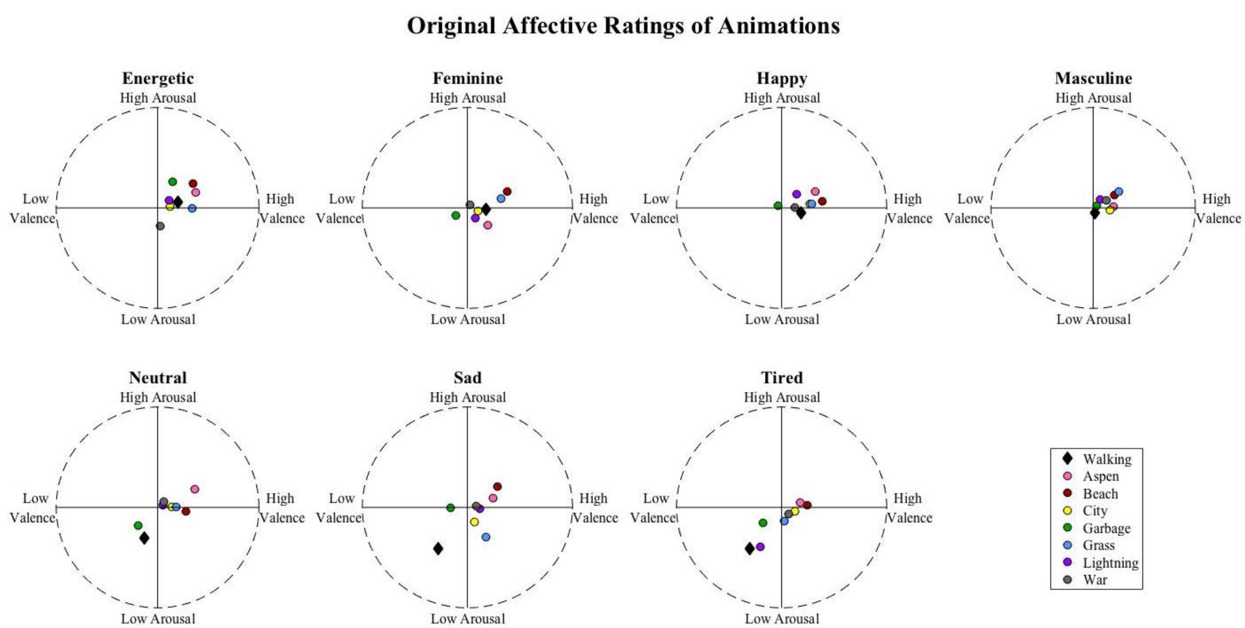

Figure D.3: The mappings show the original ratings of the animations that were created for Study 2. The ratings of the animations are separated by stylized walking sequence and are displayed as circles. The original ratings of the stylized walking sequences are also included and are represented by black diamonds.

References

[1] D. J. Hand, H. Mannila, and P. Smyth, *Principles of data mining*. MIT press, 2001.

[2] A. Y. Ng and M. I. Jordan, "On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes," in *Advances in neural information processing systems*, 2002, pp. 841–848.

[3] D. Bernhardt and P. Robinson, "Detecting affect from non-stylised body motions," in *International conference on affective computing and intelligent interaction*. Springer, 2007, pp. 59–70.

[4] D. Bernhardt and P. Robinson, "Detecting emotions from connected action sequences," *Visual Informatics: Bridging Research and Practice*, pp. 1–11, 2009.

[5] S. A. Etemad and A. Arya, "Classification and translation of style and affect in human motion using rbf neural networks," *Neurocomputing*, vol. 129, pp. 585–595, 2014.

[6] Z. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang, "A survey of affect recognition methods: Audio, visual, and spontaneous expressions," *IEEE transactions on pattern analysis and machine intelligence*, vol. 31, no. 1, pp. 39–58, 2009.

[7] S. A. Etemad and A. Arya, "Expert-driven perceptual features for modeling style and affect in human motion," *IEEE Transactions on Human-Machine Systems*, vol. 46, no. 4, pp. 534–545, 2016.

[8] S. A. Etemad, A. Arya, A. Parush, and S. DiPaola, "Perceptual validity in animation of human motion," *Computer Animation and Virtual Worlds*, vol. 27, no. 1, pp. 58–71, 2016.

[9] T. Ribeiro and A. Paiva, "The illusion of robotic life: principles and practices of animation for robots," in *Human-Robot Interaction (HRI), 2012 7th ACM/IEEE International Conference on*. IEEE, 2012, pp. 383–390.

[10] A. Van Breemen, "Bringing robots to life: Applying principles of animation to robots," in *Proceedings of Shapping Human-Robot Interaction workshop held at CHI 2004*, 2004, pp. 143–144.

[11] S. A. Etemad and A. Arya, "Modeling and transformation of 3d human motion." in *GRAPP*, 2010, pp. 307–315.

[12] C. Rollero and N. De Piccoli, "Place attachment, identification and environment perception: An empirical study," *Journal of Environmental Psychology*, vol. 30, no. 2, pp. 198–205, 2010.

[13] J. P. Redden, "Reducing satiation: The role of categorization level," *Journal of Consumer Research*, vol. 34, no. 5, pp. 624–634, 2007.

[14] F. K. McSweeney and S. Swindell, "General-process theories of motivation revisited: the role of habituation." *Psychological Bulletin*, vol. 125, no. 4, p. 437, 1999.

[15] B. Kurdi, S. Lozano, and M. R. Banaji, "Introducing the open affective standardized image set (oasis)," *Behavior research methods*, vol. 49, no. 2, pp. 457–470, 2017.

[16] E. S. Dan-Glauser and K. R. Scherer, "The geneva affective picture database (gaped): a new 730-picture database focusing on valence and normative significance," *Behavior research methods*, vol. 43, no. 2, p. 468, 2011.

[17] M.-L. Félonneau, "Love and loathing of the city: Urbanophilia and urbanophobia, topological identity and perceived incivilities," *Journal of environmental psychology*, vol. 24, no. 1, pp. 43–52, 2004.

[18] K. Studd and L. L. Cox, *Everybody is a Body*. Dog Ear Publishing, 2013.

[19] U. Huzaifa, C. Bernier, Z. Calhoun, G. Heddy, C. Kohout, B. Libowitz, A. Moenning, J. Ye, C. Maguire, and A. LaViers, "Embodied movement strategies for development of a core-located actuation walker," in *Biomedical Robotics and Biomechatronics (BioRob), 2016 6th IEEE International Conference on*. IEEE, 2016, pp. 176–181.

[20] I. Bartenieff and D. Lewis, *Body movement: Coping with the environment*. Psychology Press, 1980.

[21] A. LaViers, L. Bai, M. Bashiri, G. Heddy, and Y. Sheng, "Abstractions for design-by-humans of heterogeneous behaviors," in *Dance Notations and Robot Motion*. Springer, 2016, pp. 237–262.

[22] J. Sher and A. Jang, "An embodied, platform-invariant architecture for robotic spatial commands," Ph.D. dissertation, 2017.

[23] J. Rett and J. Dias, "Human-robot interface with anticipatory characteristics based on laban movement analysis and bayesian models," in *Rehabilitation Robotics, 2007. ICORR 2007. IEEE 10th International Conference on*. IEEE, 2007, pp. 257–268.

[24] J. S. Longstaff, "Translating "vector symbols" from labans (1926) choreography," in *Twenty-Second Biennial Conference of the International Council of Kinetography Laban, ICKL, 26 July-2 August, Ohio State University, Columbus, Ohio. USA*, 2001, pp. 70–86.

[25] K. Khoshhal, H. Aliakbarpour, J. Quintas, M. Hofmann, and J. Dias, "Probabilistic lma-based human motion analysis by conjugating frequency and spatial based features," in *WIAMIS 2011: 12th International Workshop on Image Analysis for Multimedia Interactive Services, Delft, The Netherlands, April 13-15, 2011.* TU Delft; EWI; MM; PRB, 2011.

[26] D. Chi, M. Costa, L. Zhao, and N. Badler, "The emote model for effort and shape," in *Proceedings of the 27th annual conference on Computer graphics and interactive techniques.* ACM Press/Addison-Wesley Publishing Co., 2000, pp. 173–182.

[27] A. LaViers and M. Egerstedt, *Controls and Art: Inquiries at the Intersection of the Subjective and the Objective.* Springer Science & Business Media, 2014.

[28] M. Masuda, S. Kato, and H. Itoh, "Emotion detection from body motion of human form robot based on laban movement analysis." in *PRIMA.* Springer, 2009, pp. 322–334.

[29] M. Masuda and S. Kato, "Motion rendering system for emotion expression of human form robots based on laban movement analysis," in *RO-MAN, 2010 IEEE.* IEEE, 2010, pp. 324–329.

[30] W. Huitt and S. Cain, "An overview of the conative domain," *Educational psychology interactive*, pp. 1–20, 2005.

[31] A. Mehrabian and J. A. Russell, *An approach to environmental psychology.* the MIT Press, 1974.

[32] J. A. Russell and A. Mehrabian, "Evidence for a three-factor theory of emotions," *Journal of research in Personality*, vol. 11, no. 3, pp. 273–294, 1977.

[33] D. R. Heise, "Semantic differential profiles for 1,000 most frequent english words." *Psychological Monographs: General and Applied*, vol. 79, no. 8, p. 1, 1965.

[34] C. E. Osgood, "Dimensionality of the semantic space for communication via facial expressions," *Scandinavian journal of Psychology*, vol. 7, no. 1, pp. 1–30, 1966.

[35] E. Harmon-Jones, P. A. Gable, and T. F. Price, "Does negative affect always narrow and positive affect always broaden the mind? considering the influence of motivational intensity on cognitive scope," *Current Directions in Psychological Science*, vol. 22, no. 4, pp. 301–307, 2013.

[36] A. Mehrabian, "Measures of individual differences in temperament," *Educational and Psychological Measurement*, vol. 38, no. 4, pp. 1105–1117, 1978.

[37] A. Mehrabian, "Framework for a comprehensive description and measurement of emotional states." *Genetic, social, and general psychology monographs*, 1995.

[38] J. A. Russell, "A circumplex model of affect," *Journal of Personality and Social Psychology*, vol. 39, no. 6, pp. 1161–1178, 1980.

[39] G. Paltoglou and M. Thelwall, "Seeing stars of valence and arousal in blog posts," *IEEE Transactions on Affective Computing*, vol. 4, no. 1, pp. 116–123, 2013.

[40] M. Bradley and P. J. Lang, *The International affective digitized sounds (IADS)[: stimuli, instruction manual and affective ratings.* NIMH Center for the Study of Emotion and Attention, 1999.

[41] E. Kanjo, E. M. Younis, and N. Sherkat, "Towards unravelling the relationship between on-body, environmental and emotion data using sensor information fusion approach," *Information Fusion*, vol. 40, pp. 18–31, 2018.

[42] V. B. Zordan and J. K. Hodgins, "Motion capture-driven simulations that hit and react," in *Proceedings of the 2002 ACM SIGGRAPH/Eurographics symposium on Computer animation*. ACM, 2002, pp. 89–96.

[43] Y. Ma, H. M. Paterson, and F. E. Pollick, "A motion capture library for the study of identity, gender, and emotion perception from biological motion," *Behavior Research Methods*, vol. 38, pp. 134–141, February 2006.

[44] "Anthropometric data," Jan 2017. [Online]. Available: http://www.ele.uri.edu/faculty/vetter/BME207/anthropometric-data.pdf

[45] P. Micó, "stltools," June 2015, mATLAB File Exchange. [Online]. Available: https://www.mathworks.com/matlabcentral/fileexchange/51200-stltools

[46] C. L. Vaughan, B. L. Davis, C. Jeremy et al., *Dynamics of human gait.* Kiboho publishers, 1999.

[47] S. G. Barsade and D. E. Gibson, "Why does affect matter in organizations?" *The Academy of Management Perspectives*, vol. 21, no. 1, pp. 36–59, 2007.

[48] S. Macdonald, "Drawing the lines—gender, peace and war: An introduction," in *Images of Women in Peace and War.* Springer, 1987, pp. 1–26.

[49] E. J. Pedhazur and T. J. Tetenbaum, "Bem sex role inventory: A theoretical and methodological critique." *Journal of Personality and Social psychology*, vol. 37, no. 6, p. 996, 1979.

[50] A. Agresti and M. Kateri, "Categorical data analysis," in *International encyclopedia of statistical science.* Springer, 2011, pp. 206–208.

[51] D. R. Williams and J. J. Vaske, "The measurement of place attachment: Validity and generalizability of a psychometric approach," *Forest science*, vol. 49, no. 6, pp. 830–840, 2003.

[52] M. Heimerdinger and A. LaViers, "Influence of environmental context on recognition rates of stylized walking sequences," in *International Conference on Social Robotics.* Springer, 2017, pp. 272–282.

[53] M. M. Bradley and P. J. Lang, “Measuring emotion: the self-assessment manikin and the semantic differential,” *Journal of behavior therapy and experimental psychiatry*, vol. 25, no. 1, pp. 49–59, 1994.

[54] M. D. Robinson, “Running from william james’ bear: A review of preattentive mechanisms and their contributions to emotional experience,” *Cognition & Emotion*, vol. 12, no. 5, pp. 667–696, 1998.