CHOREOGRAPHIC ABSTRACTIONS FOR STYLE-BASED ROBOTIC MOTION

A Dissertation
Presented to
The Academic Faculty

By

Amy LaViers

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
in
Electrical and Computer Engineering

School of Electrical and Computer Engineering
Georgia Institute of Technology
May 2013

CHOREOGRAPHIC ABSTRACTIONS FOR STYLE-BASED ROBOTIC MOTION

Approved by:

Dr. Magnus Egerstedt, Advisor
Professor, School of ECE
Georgia Institute of Technology

Dr. Ayanna Howard
Professor, School of ECE
Georgia Institute of Technology

Dr. Fumin Zhang
Assistant Professor, School of ECE
Georgia Institute of Technology

Date Approved: April 15, 2013

PREFACE

This thesis adds a drop to the ocean of literature that grapples with how to automate robotic systems such that these systems have the capability to move in a certain way. The focus of this thesis is outlining, quantitatively, an interpretation of the enigmatic phrase "in a certain way" – particularly for the case of human shaped robots. Typically, robotic algorithms deal with task dependent quantities like achieving a spatial goal and moving with minimum energy. However, some applications, like the one pictured in Fig. 1, may require parameters which offer control over the sequence and expression of actions that are sensitive to how these motions are perceived by a viewer. That is, we would like to control the *style* of motion.

Figure 1. In the image above, a human-shaped robot interacts with a human by moving *in a certain way*. Photo credit: LucasArts.

This thesis proposes an interpretation of stylistic motion, a two-part control scheme for *generation*, and an inverse optimal control problem for *interpretation*. The two methods proposed, one for *generation* (Ch. 2) and one for *interpretation* (Ch. 3), rely on a separation between movement sequence and execution. Namely, so-called primary motions are sequenced and then scaled according to *stylistic* parameters, and a method for extracting

this scaling and basic movements from real movement is outlined. This stance and, thus, these parameters imitates the methods employed by dance professionals – people who make their living by producing stylistic motion, giving particular attention to the effect on their audience. Thus, the interpretations extracted from real data at the end of Ch. 3 are rooted in abstractions from dance theory in order that the framework might be aligned with the kinesthetic experience contained in the embodied explorations that drive the way dancers think about movement.

Chapter 4 discusses the application of the quantitative notions of style presented in this dissertation to a dance performance that incorporates a humanoid robot. Then, a dynamic, self-tuning clustering algorithm is presented in Ch. 5 that segments – or interprets – data in a very particular way, grouping data points into aggregate representations that may hold over or disperse over time. This algorithm utilizes the same machinery used to understand style in this thesis and gives a notion of temporal data organization precise according to a graph-theoretic notion of connectedness (answering questions like: how many clusters is enough? and how long should temporal interactions be taken into account?).

Understanding how humans move and create movement is a fundamental quest in both artistic and technical fields. The tool presented here is a step towards uniting these fields in this search. It is well known that body language (as evidenced by public orators in political speeches) and even choreographed movements (as evidenced by dancers in performing arts) are key aspects of human-to-human interaction – and will in be in human-to-humanoid interaction. When robots can tap into channels of nonverbal interaction, automated systems will become more than button-based, frontal terminals, such as ice makers and ATMs. Tools like the one presented in this thesis, which take into consideration the great body of knowledge in the arts and humanities, will take on greater importance as we approach a society with greater interface with robotic systems – as imaginatively pictured in Fig. 1.

ACKNOWLEDGMENT

First and foremost, I would like to thank my adviser, collaborator, and mentor, Magnus Egerstedt, for creating this body of research with me. In our first meeting ever, he showed me an automaton (I immediately went to the library to check out [1]), and four years later, I think I finally understand what a derivative is. I would also like to recognize Yushan Chen and Calin Belta, our co-authors on the content of Secs. 2.1.1 and 2.2.2, for their input into the material of this thesis. Thirdly, Lori Teague, a dance professor at Emory University, had a few conversations with us that have been very valuable to the direction of this work.

My time at Georgia Tech has been sponsored by the National Science Foundation (grant number 0757317), and I would like to thank them for their monetary support and faith in out-of-the-box research. The performance of *Automaton* has been funded by the Electrical and Computer Engineering department. This award has not only produced a show, but funded rehearsals where I have re-thought every component of this thesis from a new angle. I cannot say enough about what their sponsorship has done for the work (in part, because it's hard to quantify and in part, because it is such a big contribution). I would also like to thank Ann Terka, the Ferst Center for the Arts, Clough Undergraduate Learning Commons, and Jack Rogers for their administrative and production support of *Automaton*.

My collaborators in movement, Alex Abarca, Helen Hale, Camille Jackson, and Erik Thurmond: I thank each of you for bringing your energy, thought, and bodies to this research. I also thank my colleagues in the Georgia Robotics and Intelligent Systems (GRITS) Lab for listening to each phase of this research from the beginning and providing invaluable input on each phase. I would also like to thank Chris Malbrue in the ECE office for always having a solution to any administrative snag and a word of encouragement.

Thanks most of all to my family and friends – Mom, Dad, Lisa, and Eric, in particular – for believing in me.

TABLE OF CONTENTS

PREFACE iii

ACKNOWLEDGMENT v

LIST OF FIGURES vii

SUMMARY ix

CHAPTER 1 BACKGROUND 1
1.1 From Data to Doing: Inverse Problems for Robotics 1
1.2 Abstractions for Automated Movement 3
1.3 Choreographing Movement: Dance Theory 7

CHAPTER 2 THE FORWARD PROBLEM: A GENERATIVE DEFINITION OF STYLE 13
2.1 A Definition of Movement for Stylistic Specification 13
2.2 Motion Sequencing 19
2.3 Motion Modulation 30
2.4 An Illustrative Example for Robotics 34

CHAPTER 3 THE INVERSE PROBLEM: A STYLISTIC INTERPRETATION OF HUMAN MOVEMENT 44
3.1 Style-based Abstractions for Human Motion Analysis 44
3.2 Quality Matching: The Inverse Problem 46
3.3 A Style-based Segmentation Problem 50
3.4 Application to Human Movement 58

CHAPTER 4 *AUTOMATON*: AN APPLICATION IN ROBOTICS AND DANCE PERFORMANCE 68
4.1 Expectations and Reactions: Programmed Behavior 69
4.2 Rule-based Movement Generation 72
4.3 What did the audience see?: A Human Study 75

CHAPTER 5 EXTENSIONS: BEYOND STYLE 80
5.1 Dynamic Clustering 80
5.2 Automatic Time-scale Extraction 87

APPENDIX A THE PROGRAM FOR *AUTOMATON* 89

APPENDIX B HUMAN STUDY IRB FORMS 91

REFERENCES 97

LIST OF FIGURES

Figure 1 Luke Skywalker and C3PO iii

Figure 2 Ballet Barre Exercise 8

Figure 3 Tai Chi Form 8

Figure 4 The Dynamosphere 10

Figure 5 The Unusual Movement of Merce Cunningham 12

Figure 6 The Barre Automaton 14

Figure 7 The Augmented Barre Automaton 18

Figure 8 Linear Temporal Logic Specification Scheme 25

Figure 9 Five Sequences Showing the Effect of Specifications 28

Figure 10 Occurrences of Desired Transitions 29

Figure 11 Occurrences of Desired Transition Series 30

Figure 12 Representation of Robot Arm Pose 35

Figure 13 Discretization of Robot Arm State Space 36

Figure 14 Disco Automaton 37

Figure 15 Cheerleading Automaton 38

Figure 16 Disco and Cheerleading Behavior 42

Figure 17 Movement Generation Scheme 43

Figure 18 A Segmentation Problem 45

Figure 19 Human Motion Capture Data 60

Figure 20 Converging Implementation of the Inverse Problem 64

Figure 21 One-dimensional Example of Classification 64

Figure 22 Nine-dimensional Example of Classification 66

Figure 23 Animation of Simulated Leg Motion 66

Figure 24 Movement Generation and Interpretation Scheme 67

Figure 25 The Cast of *Automaton* 70

Figure 26 The Robot as an Onstage Character 71
Figure 27 Poses for Human and Robot Movement 73
Figure 28 Style I 74
Figure 29 Style II 74
Figure 30 Style III 74
Figure 31 Style IV 74
Figure 32 Five Styles, Performed Live 75
Figure 33 Histogram of Human Study Data, Question 1 76
Figure 34 Histogram of Human Study Data, Question 2 76
Figure 35 Histogram of Human Study Data, Question 3 77
Figure 36 Histogram of Human Study Data, Question 4 77
Figure 37 Histogram of Human Study Data, Question 5 78
Figure 38 Static vs. Dynamic Clustering 87
Figure 39 The Front of the Program 89
Figure 40 The Back of the Program 90
Figure 41 Program Insert 90
Figure 42 *Automaton* Human Study Approval 92
Figure 43 *Automaton* Human Study Consent Form 93
Figure 44 *Automaton* Questionnaire 94
Figure 45 *Automaton* Questionnaire cont. 95
Figure 46 Human Study Data 96

SUMMARY

What does it mean to do the disco? Or perform a cheerleading routine? Or move in a *style* appropriate for a given mode of human interaction? Answering these questions requires an interpretation of what differentiates two distinct movement styles and a method for parsing this difference into quantitative parameters. Furthermore, such an understanding of principles of style has applications in control, robotics, and dance theory. This thesis present a definition for style of motion that is rooted in dance theory, a framework for stylistic motion generation that separates basic movement ordering from its precise trajectory, and an inverse optimal control method for extracting these stylistic parameters from real data. On the part of generation, the processes of sequencing and scaling are modulated by the stylistic parameters enumerated: an automation that lists basic primary movements, sets which determine the final structure of the state machine that encodes allowable sequences, and weights in an optimal control problem that generates motions of the desired quality. This generation framework is demonstrated on a humanoid robotic platform for two distinct case studies disco dancing and cheerleading. In order to extract the parameters that comprise the stylistic definition put forth, two inverse optimal control problems are posed and solved – one to classify individual movements and one to segment longer movement sequences into smaller motion primitives. The motion of a real human leg (recorded via motion capture) is classified in an example. Thus, the contents of the thesis comprise a tool to produce and understand stylistic motion.

CHAPTER 1

BACKGROUND

This review discusses various methods for creating movement (both for robots and humans). Even when the platform is an alien one, whose movements will look nothing like human ones due to limited and alternate actuation, the success of such methods is easily verifiable by a human viewer. Michalowski et. al. [2] show that an unusual robotic platform, the squishy yellow Keepon, has modes of operation which are preferable to human viewers. Such movements thus foster successful human-robot interaction: just watching the Keepon move is mesmerizing.

Section 1.1 describes methods for incorporating real human data into robotic movement generation. Then, research at different abstraction layers (low-level dynamics-based controllers vs. high-level planning methods) is reviewed. Finally, methods from dance theory are discussed. Methods from each of these categories have directly influenced the research presented in the remainder of this thesis. The ultimate thrust of the thesis is to provide a "closed-loop" where the same parameters than generate movement can be extracted from human movers; thus, the work in the first section is relevant. Sections 1.2 and 1.3 work together to inspire the method for movement generation in this thesis, which aims to have abstractions based in dance theory.

1.1 From Data to Doing: Inverse Problems for Robotics

The works presented in this section have the basic aim to take data of human subjects, i.e., from motion capture, and automatically translate it for robotic motion. First, methods which spend much of their effort on understanding the guiding principles – and formulating simplifications – of human movement are discussed. Several of these methods might have easily been put in the next section as their work can be seen as elucidating *motion primitives*, which will be discussed in Sec. 1.2.1. Finally, methods in inverse optimal control are

reviewed as they are mathematically similar to the work presented in Ch. 3.

Although the discussions in this thesis will largely take place at the level of joint angles, it is important to note that reduced representations of human movement can play a key role in robotic imitation of that movement. A low dimensional representation of human movement could supplement the work discussed here and keep the computational tractability of complex instantiations of this work feasible. For example in [3], LaViers demsontrates differences between styles of movement via lower dimensional representations using Gaussian process basis functions to learn manifolds alternate to joint angle representation.

Learning to imitate human task completion via training by human subjects is a goal of Thomaz et. al.'s work [4, 5, 6]. Their concept of complex tasks is defined via *keyframes*. Robot poses are extracted from input by a human user over repeated trials and then clustered; cluster representatives then become these keyframes. These poses then form the basis of their task representation allowing them to repeat the motion in future similar tasks.

Another data-driven approach to generating human-like behavior is the work of Ames et. al. [7, 8, 9, 10]. Ames seeks "canonical walking functions" to characterize human walking in order that robots may better imitate it. In this framework, these functions, which are formulated as desired outputs for the robotic system to match, satisfy three criteria; they (1) are functions of the joint angles, (2) have simple time-based representation, and (3) are mutually exclusive (the decoupling matrix associated with the outputs is non-singular). Four of these canonical walking functions are then tracked by an optimal control problem such that the robotic system mimics their behavior.

The work of Gillies [11] is an example of how pattern recognition can be key to sorting through user-generated data in order to find the sparse representations necessary to encode desired behavior – this time at a higher level of abstraction. Here motion capture data is clustered in order to find key, repeated poses that will serve as states in their transition system model of movement. This system is akin to the one in [12] but is populated with real action behavior of human actors. These problems can be seen as applications of the general

theory of constructing automata from finite samples of behavior [13, 14, 15]. Progress has been made toward a general automatic implementation of this theory as in [16].

In general, a problem that attempts to determine a cost function that produces an output that recreates real, observed data can be considered an inverse optimal control problem [17]. Many of these problems have even sought to model human movement [18, 19, 20]. Mombaur et al. take human locomotion paths through space and animate them on a humanoid [21].

The work of Nakazawa et al. [22, 23] aims to mimic human dancing in a method that mirrors that in this thesis, at a high-level, most closely. This research extracts motion primitives from motion capture data and recombines these primitives into new sequences for robots. The methods in [22] search the data for moments where the person is completely still in order to segment movements. Conversely, the stance in this thesis is that a person may never be completely still and yet may consider, from a kinesthetic standpoint, their movement comprised of many smaller movements.

1.2 Abstractions for Automated Movement

Robotic movement generation often relies on abstraction levels to keep problems tractable; the same framework typically does not compute commands to motors as well as an overall motion path or plan to achieve a given task. Abstraction layers allow a low-level controller to be designed for certain basic outcomes for a system; then, a high-level planner may stitch those outcomes together in order to navigate some task. For example, for a wheeled robot equipped with proximity detectors that is tasked to navigate an obstacle course, low level controllers may be designed to drive to a certain point or follow the wall of an obstacle.

Research in robotic motion typically takes place at one level or another or deals with how these abstractions should be envisioned. Seminal works in these areas include the optimal control problem designed in [24] that produces human-like motion by minimizing jerk in order to match data of human movement, the geometric methods in [25] which help

envision the capabilities of a multi-degree of freedom manipulator relative to its task, and architecture design as in [26, 27] that formulates *behaviors* as an abstraction to separate high-level and low-level tasks.

1.2.1 Motion Primitives

The ideology in this thesis is a natural match to some concepts from the body of work that has attempted to enumerate dynamic motion primitives which may be recombined to create full-fledged movement sequences. The goal is often to develop an appropriate mathematical definition and then segment these primitives from real data – in either two (i.e., video) or three (i.e., motion capture) dimensions. In [28] Bregler coins the term *movemes* – in analogy to the linguistic concept of a phoneme – which is cited in many following publications, often offering an alternate technical definition to this same general concept. The concept pops up in computer vision [29, 30, 31, 32, 33] as well as research in robotics [34, 35, 36, 37, 38] and gait analysis [39, 40]. The basic search that pervades even between these academic arenas is for atomic nuggets of movements which can recombine, in the same way phonemes are the sounds that create words, to successfully recreate and form novel full-fledged movement sequences.

The concept of a moveme as used by Del Vecchio in [29, 41, 42, 43] is one that strives to interpret real data using a control theoretic description for the primitives. Namely, parameters for LTI dynamical systems are learned via system identification techniques for three actions in a drawing application controlled by a computer mouse. The learned description of these three actions, reach (when the user simply moved to another area of the screen), draw (when the user drew a straight line), and circle (when the user drew a circle) were found to be fairly invariant between users – indicating that their formulation was capturing something more universal than patterns of individual movers. Such an automatic classification and segmentation would help provide better understanding of how users interact with various software and websites.

1.2.2 Formal Methods for High-level Task Completion

At a higher level of abstraction lies work which attempts to order these motion primitives in a way such as to achieve loftier goals. In particular, temporal logics, such as Linear Temporal Logic (LTL) have been suggested as motion specification languages [44, 45, 46, 47, 48]. The use of such logics allows for a large spectrum of high-level specifications which include: choice of a goal: "go to either A or B"; convergence to a region: "reach A eventually and stay there"; visiting targets sequentially: "reach A, and then B, and then C"; surveillance: "go back and forth between A and B forever"; and the satisfaction of more complicated temporal and logic conditions about the reachability of regions of interest: "Never go to A. Don't go to B unless C or D were visited first." Such robot motion planning and control objectives are achieved based on algorithms inspired from model checking [49] and temporal logic games [50].

LTL formulas are built from a set of atomic propositions Π, boolean operators $\neg$ (negation), $\vee$ (disjunction), $\wedge$ (conjunction), $\rightarrow$ (implication) and temporal operators $\mathbf{X}$ (next), $\mathcal{U}$ (until), $\mathbf{F}$ (eventually), $\mathbf{G}$ (always). The semantics of LTL formulas are given over infinite words generated by a labeled transition system (T); states of the transition system are labeled with atomic propositions. We can then watch T evolving according to these propositions: that is, a word of T corresponds to a (infinite) sequence of sets of the propositions in Π.

A word satisfies an LTL formula ϕ if ϕ is true at the first position of the word; $\mathbf{X}\phi$ states that at the next state, an LTL formula ϕ is true; $\mathbf{F}\phi$ means that ϕ eventually becomes true in the word; $\mathbf{G}\phi$ means that ϕ is true at all positions of the word; $\phi_1 \,\mathcal{U}\phi_2$ means ϕ_2 eventually becomes true and ϕ_1 is true until this happens. More expressiveness can be achieved by combining the above with Boolean operators.

Runs satisfying the original transition system and the specification can be found using techniques inspired by LTL model checking [49], which checks whether all the words of a transition system satisfy an LTL formula ϕ over its set of propositions. Central to the

LTL model checking problem is the construction of a Büchi automaton that accepts all and only words satisfying ϕ. An off-the-shelf software tool, such as LTL2BA [51], can be used to generate such a Büchi automaton. The product automaton, $\mathcal{A}$, between the transition system and the Büchi automaton accepts all and only the runs of the transition system whose words satisfy ϕ. In other words, $\mathcal{A}$ encodes the system dynamics of T and the specification contained in ϕ; another way of thinking of it is that the Büchi automaton generated by ϕ is a kind of supervisor for the system T.

Acceptable runs of $\mathcal{A}$ are thought of as output behavior. This output may be flavored in the following sense: instead of producing sequences that are simply a random walk over the resulting "supervised" transition system, we can assign rewards to desired states and, through the use of a receding-horizon optimal controller, visit these desired states more frequently in our output sequences. This capability is developed and detailed in [48].

1.2.3 Style for Movement Animation

A large body of research has aimed at understanding motion patterns for application to vision and animation with the hopes of implementing character to on screen avatars. Some work [52, 23, 53] has a more statistical bent while others [54, 12, 55, 11, 56, 57] aim more directly at generation of novel motion sequences according to various constraints and character-driven rules. Similarly, attempts at "stylizing" robotic motion such as [58], aim to facilitate better human-robot interaction through a similar, character-driven philosophy (i.e., robots that move unpredictably or in a "strange" way are hard to interpret and suspicious to users).

In particular, the formulation of Brand and Hertzman [52] is particularly pleasing in that their *style machines* define a space of movement styles (analogous to the vision of Laban in Fig. 4). These so-called stylistic hidden Markov models (SHMMs) learn a generic model and a set of style-specific models that are present in the training data. This generic model defines a space of HMMs in which each of the style-specific models live governed by an additional parameter. Choosing this parameter produces a unique HMM thereby

fixing the style of motion governed by it. Thus, new styles can be generated by selecting recombinations of known movement styles in this space. New motion styles are animated via a random walk through the desired HMM state machine. This allows for a learned model which assumes very little about the way humans move and modify the style of their movement.

1.3 Choreographing Movement: Dance Theory

When dancers learn a piece, they first learn the order of the various steps that a choreographer has arranged. Then, they refine the dynamic qualities which they impart to these sequences of steps according to the choreographer's desire. In this sense, they rely on a specific, internal notion of basic steps to be strung together that is separate from each movement's timing and execution. It is through this sequencing and execution that a choreographer is able to construct the movement that is a key component of their artistic expression and resultant effect on their audience[1]. This process inspires two notions of dance which we review further here: how dancers think of individual movement *sequencing* and *execution*.

1.3.1 Canonical Exercises

One of the canonical warm up routines employed by dancers of many styles and – even figure skaters and professional athletes – is the classical ballet barre during which dancers practice small snippets of movement over and over. These basic movement snippets (one is demonstrated in Fig. 2) are then recombined later in class and during performances to create more complex, full-fledged movements and movement phrases – thus, we think of them as *primary*. Indeed warm-up routines and training exercises seed the resulting patterns of movement in many highly stylized movement genres. Martha Graham's floor work series, the karate kata, or tai chi's forms (Fig. 3) are other examples of canonical

[1] Along with a plethora of other theatrical devices like but not limited to: lighting, costumes, dancer virtuosity, staging, program notes, and music.

exercises.

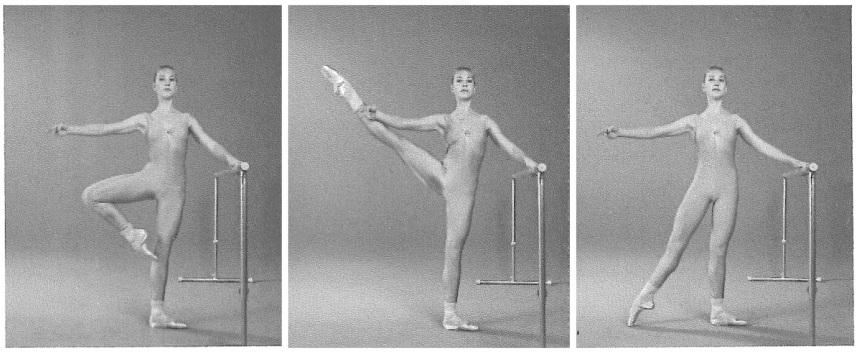


Figure 2. The dancer in this figure demonstrates proper technique for a *developpé*, an exercise that is part of the cannon of exercises called the ballet barre. [59]

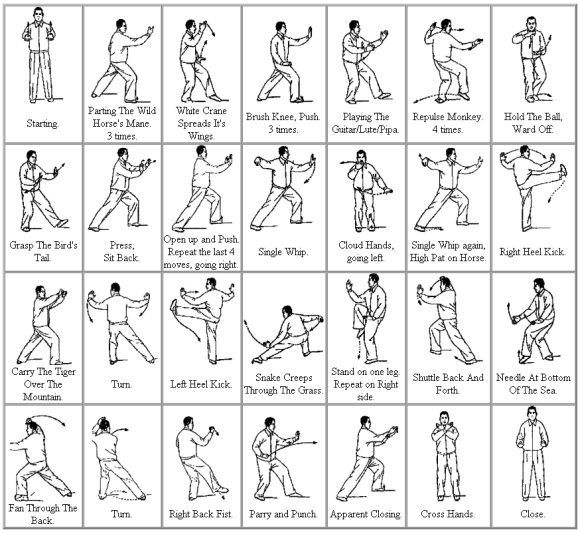


Figure 3. This image depicts some of the discrete building movements that make up this canonical form for training in tai chi. [60]

For example, in ballet, the skill required to execute a grand jeté, a majestic split-legged leap common in ballet choreography, is honed through grand battements, straight-legged high kicks, at the barre. Performing repeated kicks forward, backward, and to the side with one leg, while the other, standing leg provides static support allows the dancer to develop these muscles. Then, away from the barre, the dancer kicks one leg forward and one leg backward simultaneously to perform the grand jeté, a combination of these primary motions that invokes both practiced muscle groups at the same time.

1.3.2 Movement *Efforts*

On the part of generation, we look to dance theory. Academics and theoreticians in the field of dance typically have one of two foci: either they study the effect of a choreographer's work, its place in history and culture, and the novel mechanisms for composition employed by the artist; or they study the mechanics and psychology of human movement, producing effective training programs and exercises that allow dancers to exhibit a greater control over a greater range of motions and motion qualities. It is the latter focus, where one of the most notable scholars is Rudolf Laban, that is of interest here. Laban is perhaps most famous for his dance notation system which is built upon his principles of the body's geometry and movement's dynamic quality[2]. This thesis will draw direct inspiration from his work in movement quality; and, thus his system of movement *efforts* is reviewed here. [61, 62, 63]

The effort system laid out by Rudolf Laban in the early 20th century codified terms and concepts and continues to influence how dancers describe their movement and how they train themselves to perform a greater range of movements[3]. Namely, this vocabulary entitles dancers to a detailed description of movement that measures dynamic quality thus helping to determine an execution that produces the desired effect on the viewer.

Laban names four categories of effort or *motion factors*: *space*, *weight*, *time*, and *flow*. Space, weight, and time deal with individual movements, interrelated via the structure in Fig. 4, while flow describes the connection among a succession of movements. In a given instance of a movement, each factor may take on one of two qualities. These qualities represent the extreme notions of each motion factor.

The extremes of the first three motion factors combine pairwise to form the eight basic efforts: dabbing, gliding, floating, flicking, thrusting, pressing, wringing, and slashing. Each of these terms corresponds to a familiar pedestrian action, highlighting, even to a lay audience, the nature of the dynamosphere arrangement: changing the quality of one motion

[2]These have some analog with the notions of kinematics and dynamics in engineering, respectively.

[3]Today, Laban's direct influence is seen through Certified Laban Movement Analyst (CMA) training programs where Laban and his most influential students' - such as Irmgard Bartenieff [61]- technique and theory are taught. CMAs often go on to be university dance program professors and directors.

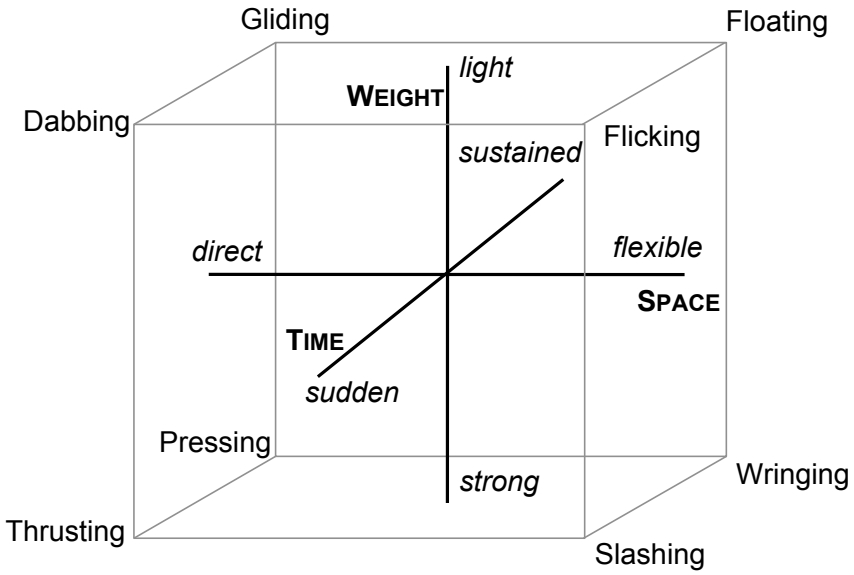


Figure 4. The dynamosphere is Laban's arrangement of eight basic efforts according to the axes of space, weight, and time. In bold font are the three Laban *motion factors* which deal with single movements; in italics are the two qualities Laban associates with each factor; and in plain font are the eight basic efforts which result from the pairwise combination of each quality.

factor moves around the cube to a different basic effort. These three motion factors, and the fourth factor, flow, which describes the quality of the connection between movements, are described – along with some intuition behind our mathematical interpretation – in the next four paragraphs; we continue to base our discussion of Laban in [63].

The space axis describes how the dancer's attitude toward space is perceived. *Flexible* movements seem more carefree, meandering, and indirect; *direct* motions appear more matter of fact and judicious with their use of space. The axis of weight deals with the emanated sense of weight in the dancer's body during the movement. *Light* movements look as though they are less influenced by gravity – perhaps they are effortless – whereas *strong* movements are muscular and seem taxing on the body to perform. Thirdly, Laban prescribes a time axis on which the movement may either be *sudden* or *sustained*. This describes a quality which, while it adheres closely to the colloquial notion of these terms, is more subtle than just the duration of the movement; that is, a movement which lasts five seconds may be executed with a sudden or sustained quality. Finally, Laban describes transitions between movements with flow; these may either be *free* or *bound*. In free flow

a dancer seems to move through movements confidently with less care for precise execution; while in bound flow, the dancer appears more careful to execute the succession of movements precisely.

1.3.3 Algorithms for Art

Algorithmic creation of art has perhaps been most explored in music technology [64, 65], but there are examples of movement creation as in Bradley et al. [54]. Such attempts may be viewed as being in line with methods in dance choreography that allow a choreographer to distance their own body from the creation of movement – in order to produce more novel sequences that are outside the natural patterns of their own body's movement. Two choreographers with notable, technological efforts in this realm are Merce Cunningham and William Forsythe. Cunningham [66] famously used dice to introduce chance and randomness into his work; a roll of the dice would determine the next movement in a sequence or the order of a show on a given night. These techniques produced new and exciting movements as shown in Fig. 5. Later in his career, he also developed technology, LifeForms which allowed for movement generation on a computer, to facilitate his choreographic process. Forsythe has worked to produce technologically motivated visualizations of his improvisation techniques[67] and his choreography [68].

Figure 5. Merce Cunningham, "Scenario" 1997. These dancers are bending in the coronal plane of the body, a movement pattern previously not found in Western culture that was developed through Cunningham's chance-based process. Photo by Jacques Moatti.

CHAPTER 2

THE FORWARD PROBLEM: A GENERATIVE DEFINITION OF STYLE

Several previous publications [69, 70, 71, 72] introduce a method for producing stylistic movement phrases in a principled way. The method uses a notion of human movement, outlined in this section, that separates movement ordering from exact execution; this implicit segmentation of movement into distinct "moves" or *primary motions* is tailored for handling our notion of *style* in a quantitative way. That is, the order of moves and the way each move is executed defines a style of movement and, thus, these quantities should be allowed to vary independently. As such, the stylistic parameters fall into two categories: those which govern sequencing rules, Sec. 2.2, and those which specify each movement trajectory, Sec. 2.3. The details of these parameters are described in the context of their motion generation platforms in the following two sections. First, rhe overarching movement model that each of these schemes utilizes is described.

2.1 A Definition of Movement for Stylistic Specification

The most basic movements are defined in the principal element of the movement model: a discrete event description of primary motions. These pre-defined motions are initially abstracted as events in a transition system – defined only by their starting and ending poses. Structured combinations of these transition systems will produce more complication phrases of movement, which are initially only sequences of poses. Later, interpolations between these poses are used as reference signals for an optimal control problem that can modulate exactly how each motion is executed. Thus, from the simple system first enumerated, much more complex movement phrases that are structured combinations (see Sec. 2.2) and modulations (see Sec. 2.3) of these primary movements can be produced.

This stance is directly inspired by the method of training used in classical ballet; thus,

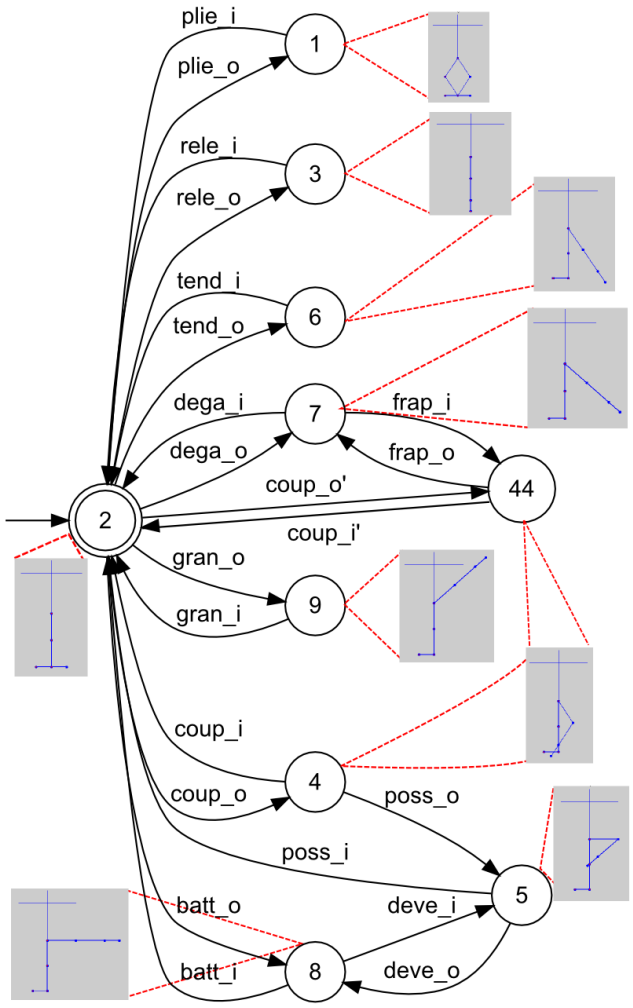


Figure 6. A transition system which models the working leg (right leg) of a dancer during a ballet barre exercise. This figure illustrates the basic element of our movement model: a discrete event description of primary motions.

a rational choice for the primary motions, and thus state transitions, in a model describing free flowing motion in the style of classical ballet are the movements from the barre exercises. These movements [59], listed in Table 1[1], connect ten basic body positions which are thus natural choices for states in this example as in Fig. 6 and [69]. Additionally, we distinguish two transitions for each movement listed in the table using a subscript to indicate an *in* and *out* variant. The variants stem from the fact that each primary movement should be unique. Thus, a movement is either going towards or away from each of the two states (body positions) it connects, and the next state depends on which of these is the case.

Table 1. A list of barre exercises and their corresponding transition labels.

Movement	Transition Labels
plié	$plie_i$, $plie_o$
relevé	$rele_i$, $rele_o$
battement tendu	$tend_i$, $tend_o$
degajé	$dega_i$, $dega_o$
coupé	$coup_i$, $coup_o$
frappé	$frap_i$, $frap_o$
passé	$pass_i$, $pass_o$
développé	$deve_i$, $deve_o$
battement	$batt_i$, $batt_o$
grand battement	$gran_i$, $gran_o$

In order to demonstrate how a sample path through the system works, consider, for example, a développé; this movement is found both in barre exercises and more complex ballet movement phrases. A développé is the action when the working leg's foot is moved from the knee and then extends from the body. At the barre this motion is practiced by moving the working foot from standing to the ankle, then the knee, then extends from the body so that the leg is parallel to the floor. Lifting the foot to the ankle or knee (without, for example, any extension to follow) are allowable movements called coupé and possé, respectively. Thus, our uniquely defined transitions (and trajectories) are three separate

[1]Passé is the movement referred to more formally as sur le coup de pied in [59].

events for the working leg: $coup_o$, $poss_o$, $deve_o$. Next, the dancer performs a closing movement where the foot remains extended from the body and the leg is lowered till the foot is returned to the starting stance. This is modeled as the event $batt_i$ – the transition from pose 8 directly to pose 2 with a label that corresponds to the *in*-trajectory of a battement (a simpler movement that looks like a high straight-legged kick). The events $coup_i$, $poss_i$, and $deve_i$ are also defined, that is, the reverse pose transitions are allowed and used for more complex movements. And, a special event is enumerated to describe the inaction of the standing leg.

Next, two similar, but distinct, formal definitions for this discrete event structure are described. Each definition will lead to two methods for producing more complex movement sequences from this primary element. The difference in these methods really lies in how each takes into account additional information for limiting the product of two such elements. The output function [69, 72] of the first produces a trajectory based elimination scheme while the labeling function [70, 71] in the second allows for specifications to be scripted over these labels in a formal language such as LTL.

2.1.1 Representation via a Finite State Machine

As in [69] this element maybe represented as a finite state machine

$$\mathcal{G} = (X, E, O, f, \Gamma, o, x_0, X_m, \epsilon, \omega), \tag{1}$$

where X is the finite state space, E is the event set, O is an output set, $f : X \times E \to X$ is the state transition function, $\Gamma : X \to 2^E$ is the set of feasible events (at a given state), $o : X \times E \to O$ is the output map, $x_0 \in X$ is the initial condition, and $X_m \subseteq X$ is a set of marked states. In order to allow for both synchronous and asynchronous transitions in the Cartesian composition of two such systems, "empty" transitions, which are defined by the symbol ϵ, are used. The interpretation is that for the finite state machine, we insist on $\epsilon \in E$, with the result that $\epsilon \in \Gamma(x)$ as well as $f(x, \epsilon) = x, \ \forall \ x \in X$. Moreover, $\omega \in O$ is associated with the outputs from "empty" events, i.e., $o(x, \epsilon) = \omega, \ \forall \ x \in X$.

We will use the output map as a means of preventing physically infeasible and stylistically inappropriate motions from happening when composing two automata in order to emulate more complex routines. For this, consider the continuous set $\mathcal{X}$ which spans the entire physical space of configurations - namely, the ones describing the primary motions that are abstracted away by the state machine. This additional state space may be defined as

$$\mathcal{X} = \{(\theta_1, ..., \theta_n) \mid \theta_i \in [\theta_{i_{min}}, \theta_{i_{max}}]\}. \tag{2}$$

For example, this set may describe body shapes, defined in terms of the joint angles which are limited by general kinematic constraints of humanoid geometry. We can now think of transitions between the discrete states of the automaton as tracing paths[2] through $\mathcal{X}$. And, as the formulation associates every event with a unique state, we can associate the output map with the path that the corresponding pose transition sweeps in $\mathcal{X}$.

Sometimes a physically infeasible (i.e. trajectories which intersect) or aesthetically undesirable trajectory pair will exist. These trajectories are kept track of in O as follows: For $e \in E, x \in X$, let

$$o(x, e) \in 2^X = O. \tag{3}$$

Thus, in the next section when we consider $\mathcal{G} \times \mathcal{G}$, the set O will correspond to trajectory pairs. These will be trimmed systematically according to the stylistic parameters.

2.1.2 Representation via a Labeled Transition System

On the other hand, as in [70] we may think of this element as a labeled transition system with the following formal properties:

$$T = (Q, q_0, \rightarrow, \Pi, h), \tag{4}$$

where $Q = \{q_1, \ldots, q_N\}$ is the set of N finite states, q_0 is the initial state representing the initial pose, $\rightarrow \subseteq Q \times Q$ is the reflexive transition relation (i.e., each state has a self

[2] A relaxed version of this may take into account that the angles corresponding to the states in the automaton are, in practice, approximate. In this case, think of these paths as "tubes."

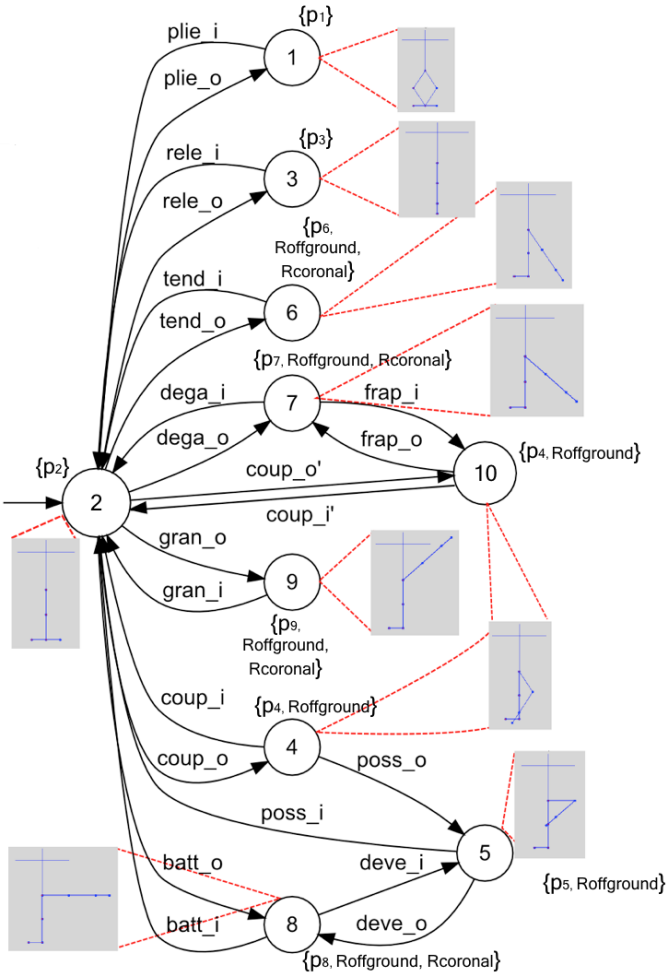


Figure 7. The set of primary motions augmented with a set of atomic propositions over which a stylistic specification may act. The propositions $\{p_1, \ldots, p_9\}$ correspond to poses defined by three joint angles: hip, knee, and ankle. While Roffground and Rcoronal indicate that the *right* leg is off the ground or extended in the coronal plane of the body, respectively.

transition), $\Pi = \{p_1, \ldots, p_M\}$ is a finite set of M atomic propositions, and $h : Q \mapsto 2^{\Pi}$ is a satisfaction (output) map, where state q_i satisfies the set $h(q_i)$ of propositions from Π as shown in Fig. 7 for the case of the ballet barre. These include a label for each pose and two additional labels, Roffground and Rcoronal, which indicate that the right (or left) leg is off the ground or extended in the coronal plane. The next section describes how LTL formula can eliminate unwanted behavior from $\mathcal{G} \times \mathcal{G}$.

2.2 Motion Sequencing

In this section, a general method for combining the elemental unit of the movement model presented in the previous section such as to impart rules – which may be style based – to sequences in the combined state space is described. In order to capture the fact that these primary motions may be performed at the same time and thus produce more complex movement sequences, we want to be able to compose two state machines, of the form in Eq. 1 or Eq. 4, while transitions that are deemed impossible or incorrect are removed from the composed system. For example, we may aspire to go from a one-legged automaton describing the primary movements of that single leg to a system involving two legs without violating the laws of physics or rules that ensure we stay faithful to a given style of movement. As such, we need to introduce some compositional operations. As before, we will first introduce the direct method of [69],[72] and second the more involved method in [70], [71].

2.2.1 Supervisory Control via Direct Automata Composition

Given two finite automata

$$\mathcal{G}_i = (X_i, E_i, O_i, f_i, \Gamma_i, o_i, x_{i,0}, X_{i,m}, \epsilon_i, \omega_i),\ i = 1, 2,$$

let the Cartesian composition of these two systems be given by

$$\mathcal{G} = \mathcal{G}_1 \times \mathcal{G}_2 = (X, E, O, f, \Gamma, o, x_0, X_m, \epsilon, \omega),$$

where

$$
\begin{aligned}
X &= X_1 \times X_2 \\
E &= E_1 \times E_2 \\
O &= O_1 \times O_2 \\
f((x_1, x_2), (e_1, e_2)) &= (f_1(x_1, e_1), f_2(x_2, e_2)) \\
\Gamma((x_1, x_2)) &= \Gamma_1(x_1) \times \Gamma_2(x_2) \\
o((x_1, x_2), (e_1, e_2)) &= (o_1(x_1, e_1), o_2(x_2, e_2)) \\
x_0 &= (x_{1,0}, x_{2,0}) \\
X_m &= X_{1,m} \times X_{2,m} \\
\epsilon &= (\epsilon_1, \epsilon_2) \\
\omega &= (\omega_1, \omega_2).
\end{aligned}
\tag{5}
$$

Note that this is a synchronous composition in the sense that events have to happen to both of the two systems in order for a transition to happen. However, through the introduction of the empty word, we can produce asynchronous transitions directly through the use of the events (e_1, ϵ_2) or (ϵ_1, e_2) in a straightforward manner.

If this operation is performed on two automata that describe the primary movements at the ballet barre, one for the right leg and one for the left, as $\mathcal{G} = \mathcal{G}_{barre_1} \times \mathcal{G}_{barre_2}$, we obtain a system that no longer is a one-legged warm-up routine, but rather a two-legged dance model. For example, *pas de chat*[3] is a jump in which the dancer picks up each leg in sequence causing him or her to move side to side. It is modeled for either leg in this automaton as the event sequence

$$coup_o, poss_o, poss_i, plie_o, plie_i \in E^{\star}_{barre},$$

where $\star$ denotes the Kleene closure. The corresponding event string for the composite system would be

$$(\epsilon_1, coup_o), (coup_o, poss_o), (poss_o, poss_i), (poss_i, plie_o), (plie_o, plie_i), (plie_i, \epsilon_2).$$

[3]Literal translation from French: "step of the cat."

Note that both of these joint event strings belong to the composite set $(E_{barre} \times E_{barre})^\star$ and that these strings correspond to sequences of the primary movements.

We now introduce two operations, *trans* and *pose*, which will whittle away undesirable sequences and narrow the defined style to something more specific than strings in $(E_{barre} \times E_{barre})^\star$. Let

$$\mathcal{G} = (X, E, O, f, \Gamma, o, x_0, X_m, \epsilon, \omega)$$

be a finite state machine as per Eq. 1 and let $O_{infeas} \subset O$ and $O_{unaesth} \subset O$ be subsets of the output set that do not contain ω. The *trans* operation will take $\mathcal{G}$ and $O_{infeas} \cup O_{unaesth}$ as arguments and return a new finite state machine,

$$trans(\mathcal{G}, O_{infeas} \cup O_{unaesth}) = (X, E, O, f, \hat{\Gamma}, o, x_0, X_m, \epsilon, \omega),$$

where we only have changed the definition of Γ i.e., the set of events that are allowed to happen at a given state. The new such set is given by

$$e \in \hat{\Gamma}(x) \;\Leftrightarrow\; e \in \Gamma(x) \text{ and } o(x, e) \notin O_{infeas} \cup O_{unaesth}. \tag{6}$$

Thus, *trans* removes transitions from the state machine that represent joint motions that are physically impossible or not of our defined aesthetic.

Secondly, certain poses when composed with one another may not be physically feasible or aesthetically desired. Thus, the system is limited from entering such poses through the use of a supervisory controller and two sets of states X_{infeas} and $X_{unaesth}$ which restrict each type of joint state, respectively. Formally speaking, given a finite state machine together with the sets $X_{infeas} \subset X$ and $X_{unaesth} \subset X$, with $x_0 \notin X_{infeas} \cup X_{unaesth}$, we let the operation

$$pose(\mathcal{G}, X_{infeas} \cup X_{unaesth})$$

be the supervised system $G \backslash S$, where S is the maximally permissive, non-blocking supervisor that ensures that the set $X_{infeas} \cup X_{unaesth}$ is never reached. Note that such supervisors can be automatically generated, e.g., [73].

The final system is then given by

$$pose(trans(\mathcal{G}_1 \times \mathcal{G}_2, O_{infeas} \cup O_{unaesth}), X_{infeas} \cup X_{unaesth}). \tag{7}$$

Returning to our example, the set O^{ballet}_{infeas}, defined for the ballet model, allows us to whittle away excess transitions in order to produce system behaviors consistent with the physical capabilities of a bipedal geometry. For example, lifting a flat-footed leg off the ground, in a manner which indicates a jump (or, equivalently, raising the leg when the other is already in the air), without a bend in the knees to provide spring for the jump, is physically impossible. This corresponds to state 2 (foot flat on the ground with a straightened knee), and the infeasible event is an extension of the leg from the ground to state 8 (extended away from the body, parallel to the floor). Similar such simple rules used to govern a leg providing critical support (a leg in state 1, 2, or 3 when the other is in any of the others, states 44 and 4-8) such as "no coronal extension of the supporting leg from states 1 and 2" can easily be translated into regions of $2^X \times 2^X$.

Thus, pairs of regions of continuous space define O^{ballet}_{infeas} and correspond to disallowed synchronous leg paths. Likewise, we may enumerate $O^{ballet}_{unaesth}$ to describe any undesired synchronous motions that, while possible to execute, are not in the desired style of movement. Since we began this example from such a narrow set of primary motions which are, by design, intended to be performed synchronously, we leave this set empty.

Each joint state in the two-legged barre automaton is physically feasible; thus, X^{ballet}_{infeas} is an empty set. The definition of $X^{ballet}_{unaesth}$ in the ballet model is a straightforward list of two-legged states which are perhaps considered ugly as judged by the metric of ballet; often, these are asymmetrical poses or poses which cannot be seen from the audience's distant perspective.

Finally, with both of these components in place, Eq. 7 reduces to

$$pose(trans(\mathcal{G}_{barre_1} \times \mathcal{G}_{barre_2}, O^{ballet}_{infeas}), X^{ballet}_{unaesth}).$$

This describes an automaton that accepts strings of body poses that follow our ballet-inspired rule set.

2.2.2 Supervisory Control via Formal Language Specification

It is worth asking the question of how easy this method is to implement – for example, is naming the sets O_{infeas}, $O_{unaesth}$, X_{infeas}, and $X_{unaesth}$ in situations where the desired motion style is less formalized a manageable task? The discrete event model for motion sequencing can also be assembled through more natural statements scripted in linear temporal logic (LTL). These statements operate on atomic truth propositions that are appended to a transition system describing the organization of primary movements. Thus, one single LTL statement may eliminate many transitions and states. The mechanics of this process eliminate joint transitions and states of the system through a supervising automaton, called a Bucchi automaton, that encodes the specification on the atomic propositions.

Instead of encoding disallowed states and transitions directly in sets like O_{infeas}, $O_{unaesth}$, X_{infeas}, and $X_{unaesth}$ our stylistic parameter now takes the form of an LTL formula, ϕ. In order to accommodate this new type of specification, our system has a different formal representation as well; namely, we augment the states in our transition system with atomic propositions over which ϕ may act.

The atomic propositions are statements which are either true or false about every state of our system. We use the power of temporal logic to watch the evolution of our system in terms of these statements of particular interest; hence, we often think of these propositions as observations. For our case study of the ballet barre (see Fig. 7), the pose propositions $\{p_1, \ldots, p_9\}$ for each leg are satisfied only when that leg is in the corresponding pose. State q_{10} is a slight variation of state q_4[4] and in this two dimensional framework we represent them with the same pose. Hence, proposition p_4 is true at state q_4 and q_{10}. The additional

[4]The pose we call coupé, more formally known as sur le cou-de-pied, has two variations: in one the foot of the working leg is wrapped around the ankle while in the other it is fully pointed and placed next to the ankle. While two states are necessary to prevent nonsense barre sequences from being accepted by the system, the resolution of our model is not intended to capture such subtle differences in pose.

propositions, Roffground ("the right leg is off the ground") and Rcoronal ("the right leg is in a coronal extension away from the body"), help script statements about the system more concisely as they apply to several states each and have corresponding meaning for the left leg system. When needed, a script "R" or "L" is added to the states and propositions corresponding to T_R and T_L, respectively; however, generally speaking, moves valid (or invalid) on one leg are likewise allowed (or disallowed) on the other so often these scripts will be neglected as our specifications are symmetrical.

The next task is to compose these two systems using a synchronous product; formally, the synchronous product[5] of the two transition systems T_L and T_R, denoted as $T_L \otimes T_R$, is a new transition system with $(Q_P, q_{0_P}, \rightarrow_P, \Pi_P, h_P)$. The states are Cartesian pairs of the single leg states, i.e. $Q_P \subseteq Q_L \times Q_R$, likewise $q_{0_P} = (q_{0_L}, q_{0_R})$. Transitions exist between these joint states if and only if a transition existed between both single states, i.e. $\rightarrow_P \subseteq Q_P \times Q_P$ is defined by $(q, q') \in \rightarrow_P$ if and only if $q \neq q'$, $(q_L, q'_L) \in \rightarrow_L$ and $(q_R, q'_R) \in \rightarrow_R$, where $q = (q_L, q_R)$ and $q' = (q'_L, q'_R)$. The set of propositions, Π_P, becomes $\Pi_R \cup \Pi_L$ with possible additional propositions, such as Spose which is used in the union of the single leg transitions for the left and right leg with an added proposition, Spose, which is satisfied when both legs are in the same positions. The labeling function associates any proposition that was true for either single leg state with the joint state. This scheme is summarized in Fig. 8.

This thesis presents two case studies that divide their specifications into two categories: hard and soft. Hard specifications are those encoded in a Bucchi automaton and must be met. Soft specifications are not requirements but simply encouragements that the final output have more of a certain type of sequence in it than usual. These are enforced during output selection by a receding horizon optimal controller.

The hard specifications, which are applied to both case studies and ensure that the basic style objectives are met, are given according to a collection of physically and aesthetically

[5] Asynchronous transitions of the original single leg systems are also allowed since the reflexive transitions defined in Eq. 4 (iii) establish self-loops at each state.

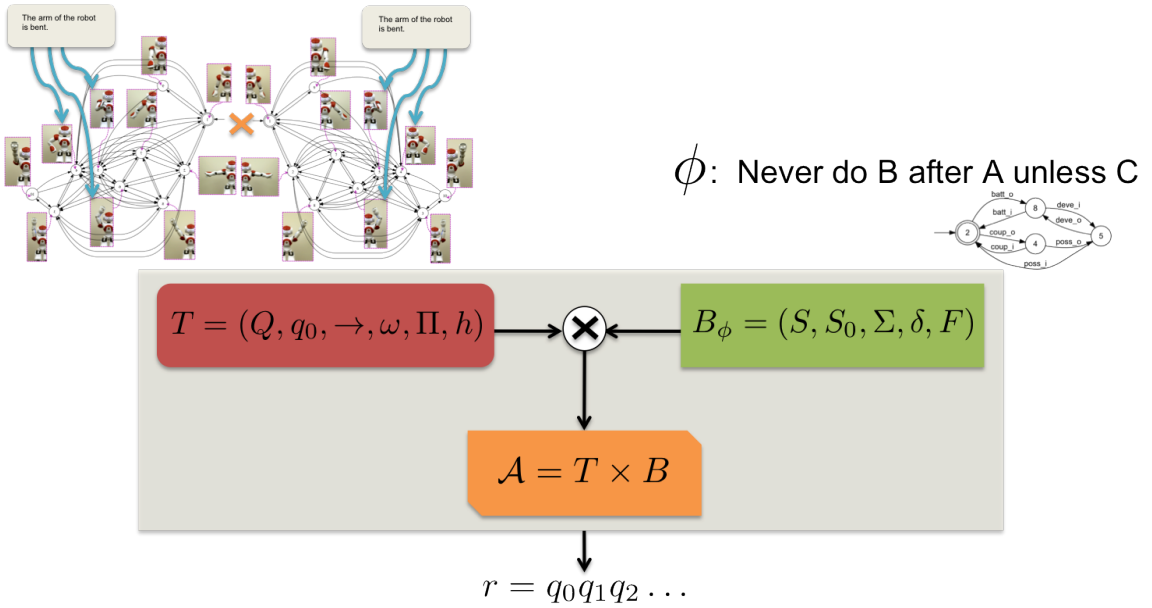


Figure 8. This graphic demonstrates how this scheme captures both the patterns in the system dynamics and those sequences of atomic propositions which are allowable by the LTL specification (and are encoded via a Bucchi automaton).

driven rules and translate to LTL formulas as following:

1. We first ensure that no physically infeasible joint-poses are entered by the system. Such poses include (i) those poses which imply impossible jumping positions, i.e., state (q_6, q_7), (ii) poses impossible to hold because they are neither jumping poses nor feasible standing poses, i.e., both legs attempting to touch the ankle of the other as in state (q_4, q_4), and (iii) poses that the geometry of the biped's hips will not easily allow and thus form awkward poses which are ugly in the eyes of traditional ballet choreography, i.e., state (q_1, q_3). Thus, a set of poses are disallowed:

   (i) *"always avoid both legs off ground except poses in which two legs are in the same position, and poses satisfying* $p_7 \wedge p_{10}$, $p_5 \wedge p_8$ *or* $p_4 \wedge p_5$*"*

   $$\mathbf{G}\ \neg(\text{Roffground} \wedge \text{Loffground} \wedge (\neg\text{Spose}) \wedge \neg(p_7 \wedge p_{10}) \wedge \neg(p_5 \wedge p_8)) \wedge \neg(p_4 \wedge p_5))$$

   (ii) *"always avoid both legs in poses* p_4, p_6, *and* p_{10}*"*

   $$\mathbf{G}\ \neg(\text{Spose} \wedge (p_4 \vee p_6 \vee p_{10}))$$

(iii) *"always avoid both legs in poses satisfying $p_1 \wedge p_2$, $p_1 \wedge p_3$, or $p_2 \wedge p_3$"*

$$\mathbf{G}\ \neg((p_1 \wedge p_2) \vee (p_1 \wedge p_3) \vee (p_2 \wedge p_3))$$

2. Next, we prevent the system from switching to a pose which, given a current pose, would cause a biped to fall over. We restrict transitions from states that have only one leg supporting the body; these include (i) jumping from a leg in relevé - this would require just the toe joint of one leg to lift the entire body weight into the air, i.e., (q_3, q_8) to (q_8, q_5), (ii) extending the supporting leg when it is in plié, i.e., (q_1, q_5) to (q_8, q_5), and (iii) extending the supporting leg when it is on flat, i.e., (q_2, q_5) to (q_8, q_5). Thus, a collection of sequences of poses are disallowed:

   (i) *"when one leg is in pose p_1 and the other leg is already off ground, always avoid lifting both legs off ground"*

   $$\mathbf{G}\ (\neg((p_1 \wedge (\text{Roffground} \vee \text{Loffground}))\ \rightarrow \mathbf{X}\ (\text{Roffground} \wedge \text{Loffground})))$$

   (ii) *"when the right leg is in pose p_2 and the left leg is currently off ground, always avoid performing coronal extensions using the right leg without putting down the left leg first"*

   $$\mathbf{G}\ ((p_2 \wedge \text{Roffground})\ \rightarrow \mathbf{X}\ \neg(\text{Roffground} \wedge \text{Lcoronal}))$$

   (iii) *"when the left leg is in pose p_2 and the right leg is currently off ground, always avoid performing coronal extensions using the left leg without putting down the right leg first"*

   $$\mathbf{G}\ ((p_2 \wedge \text{Loffground}) \rightarrow \mathbf{X}\ \neg(\text{Loffground} \wedge \text{Rcoronal}))$$

3. Finally, we ensure that the system puts both legs standing on the ground once in a while. This is phrased as a requirement to visit the standing pose, state (q_2, q_2), infinitely often:

   "visit the pose where both legs are in pose p_2 infinitely many times"

   $$\mathbf{G}\ \mathbf{F}\ (\text{Spose} \wedge p_2)$$

Now, we will use the soft specifications and an optimal controller to produce sequences which satisfy more subtle style objectives. These allow us to, in a sense, flavor the output with specific poses or sequences of poses – without breaking any of the hard specifications.

We first provide a case study that demonstrates how small changes in style specifications can lead to a sequence with a new emotive effect. Ultimately, it is this type of coupling between choreographic structure and emotive effect that this framework aims to crystallize. This case study demonstrates the effect of changing just one aspect of a specification. We encourage the movement called développé. This movement does not specify the configuration of the other leg (the supporting leg). In fact, the supporting leg may either be in pose 1 (in plié), 2 (on flat), or 3 (on relevé) - a choice left up to the choreographer.

The choice of the standing leg position determines whether a sharp depression or elevation of the dancer's body occurs. Changing levels is a well-known tool in choreography: if an entire phrase of movement occurs on the same level, it may be perceived as boring or monotone. Adding elevation can be thought of as adding high notes in a musical score - which typically brings images which are light and happy to the audience's mind. On the other hand, low notes on the bass clef in music imbue a more somber, mysterious tone to the score.

The ability of the machinery presented here to produce such sequences is demonstrated in Fig. 9 and further analyzed in Fig. 10. In these cases, rewards have been added directly to the composed transition system. These soft specifications encourage specific movements to occur and are not phrased in LTL statements like the hard specifications - instead we employ the receding-horizon controller to collect the rewards on transitions and thus produce a desired output sequence. Namely, in each case we encourage the transitions $coup_o$, $poss_o$, and $deve_o$ when the standing leg is in the desired pose. For Case 1 this corresponds to six transitions: $(q_2, q_2) \to (q_1, q_4)$, $(q_1, q_4) \to (q_1, q_5)$, $(q_1, q_5) \to (q_1, q_8)$, and the three transitions which correspond to doing the movement on the other leg. The soft specification is summarized in [48, 71].

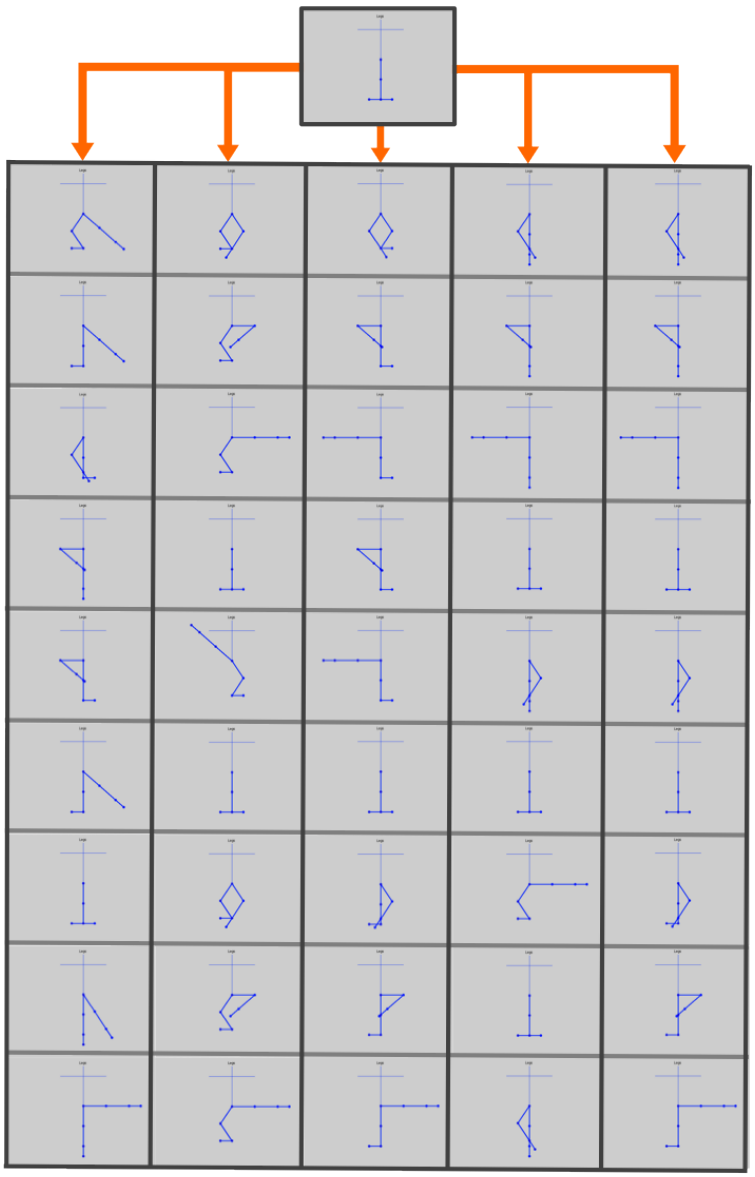


Figure 9. Five sequences demonstrate how a simple change in specification produces great variation in the output sequence. From left: (1) Case 0 (base case), where only the hard specifications are applied. (2) Case 1, encourages more développés in plié; (3) Case 2, encourages développé on flat; (4) Case 3, with elevated développés on relevé; (5) Case 4, the rightmost sequence, where the développé is encouraged without regard to the standing leg's position.

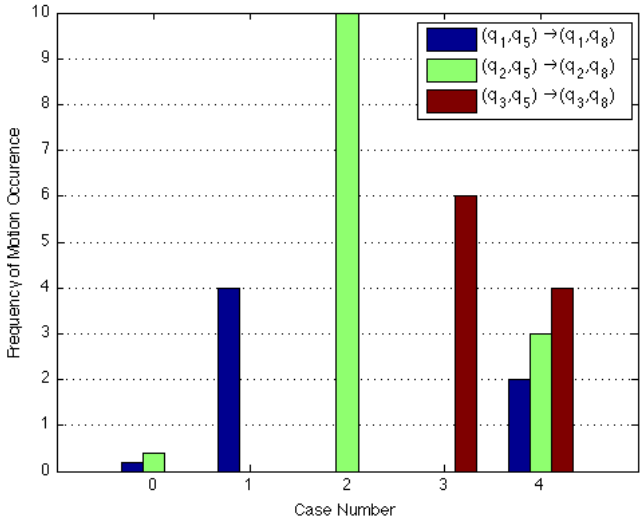


Figure 10. Occurrences of the transitions satisfying the soft specification for different sample paths show how this specification can encourage certain behaviors of the system. Each sample path contains 52 steps. Case 0 shows the average of 10 runs without any soft specification, showing the tendency of the system without any encouragement from the soft specifications. The occurrences of développé in plié, on flat, and on relevé are shown for the remaining four cases in blue, green, and red, respectively.

Secondly, we present a case study where, by modifying the hard specifications, we encourage full sequences of desired movements. Specifically, we add another rule which translates to an LTL formula as given below and encourage the pose (q_7, q_7).

(i) Any time a *glissade* happens, follow it with *plié grand jeté* or *plié pas de chat développé*

$$\mathbf{G}((p_7 \wedge \text{Spose}) \rightarrow \mathbf{X}((p_2 \wedge \text{Spose}) \wedge \mathbf{X}((p_1 \wedge \text{Spose}) \wedge ...$$

$$...\mathbf{X}((p_2 \wedge \text{Spose}) \wedge \mathbf{X}(((p_2 \wedge p_4) \wedge \mathbf{X}((p_4 \wedge p_5) \wedge \mathbf{X}(p_5 \wedge p_8))) \vee (p_8 \wedge \text{Spose}))))))$$

The encouraged sequences of movement, glissade plié grand jeté, corresponds to:

$$(q_7, q_7)(q_2, q_2)(q_1, q_1)(q_2, q_2)(q_8, q_8)$$

and the other, glissade plié pas de chat développé, corresponds to:

$$(q_7, q_7)(q_2, q_2)(q_1, q_1)(q_2, q_2)(q_2, q_4)(q_4, q_5)(q_5, q_8)(q_8, q_8).$$

Note that the specification is symmetric so a sequence with the left and right leg roles reversed is also encouraged.

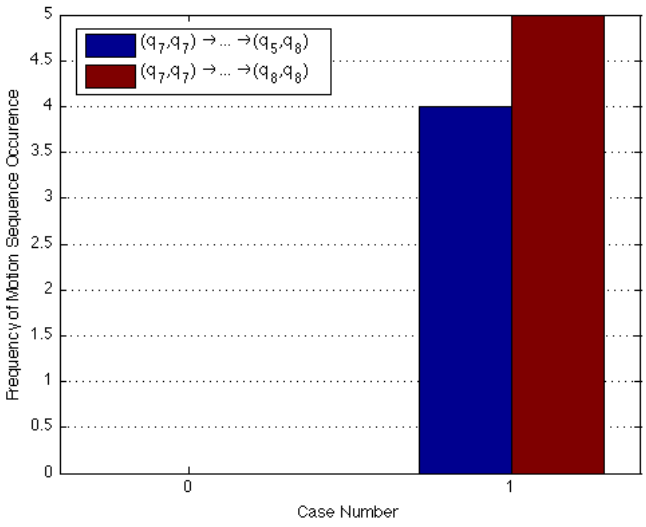


Figure 11. Occurrences of the transitions satisfying the soft specification; Each output sequence contained 100 steps. Without the specification neither sequence showed up in the output sequence. With the specification glissade plié pas de chat développé (blue) appeared 4 times and glissade plié grand jeté (red) appeared 5 times.

With this added hard specification, we are able to ensure that one of these movement sequences will occur every time both legs are in pose 7. Encouraging this pose in the soft specification will thus encourage the frequency of the desired sequences. The results from this case study are presented in Fig. 11. The resulting sequences exhibit frequent use of this series of jumps which may comprise the virtuousic displays that characterize certain sections of evening length ballets.

2.3 Motion Modulation

As outlined by dance scholar Rudolf Laban, motions may be performed with different dynamic quality: consider the difference between dabbing, flicking, and wringing. Each of those movements could have the same basic start and end pose yet be completely different in between. Thus, the basic primary motions – which are abstracted away in the discrete motion sequencing framework from the previous section – tell a limiting story of stylistic motion. In this section, we use a linear quadratic optimal control framework to find new time-varying trajectories between static poses through a mapping between Laban's effort

system and the weights in our cost function that was first proposed in [69][6].

We pair Laban's concept of the motion factor *space* with a system's notion of reference tracking; direct movements track their path more aggressively than flexible ones. Thus, we will make use of nominal trajectories away from which our solutions may deviate or adhere closely. We interpret the motion factor *weight* as a specification for how much control effort is used to perform the movement, or, in terms of a cost function, how "cheap" it is to increase the magnitude of the control signal. We interpret the factor *time* as a metric over how much the state of the system is allowed to change: during a sustained movement it should change less while in a sudden movement it may deviate wildly to produce a trajectory that appears to the audience as frantic. Finally, we interpret *flow* from a systems perspective as a lesser (or greater) desire for the dancer to hit poses between movements exactly. In the sequencing framework we will employ, this translates to varying the weights on a terminal pose.

To make these concepts mathematically rigorous, consider a linear system with an input $u \in \mathbb{R}^m$, a state $x \in \mathbb{R}^n$, and an output $y \in \mathbb{R}^l$ which tracks a reference signal $r \in \mathbb{R}^l$. We establish a quadratic cost function

$$J = \frac{1}{2}\int_0^{T_f} \left[(y-r)^T Q(y-r) + u^T R u + \dot{x}^T P \dot{x}\right] dt + \frac{1}{2}(y-r)^T S\,(y-r)\Big|_{T_f} \tag{8}$$

in order to find an input u principled on the weight matrices $Q \in \mathbb{R}^{l\times l}$, $R \in \mathbb{R}^{m\times m}$, $P \in \mathbb{R}^{n\times n}$, and $S \in \mathbb{R}^{l\times l}$. By construction, each of these matrices are positive definite and symmetric. Furthermore, their entries create a continuous, quantitative version of Laban's effort system and will determine which movement qualities are exhibited by the optimal trajectory, i.e. the trajectory may be bound, direct, sudden, and strong.

Based on the previous discussion, we associate the weight Q to the Laban's motion factor space, R with the factor weight, P with time, and S with flow[7]. These weights

[6] Which may be considered an alternate mapping to the ad-hoc one in [74].

[7] An alternate mapping to the one in [74].

correlate with the quality of each factor as follows:

$$Q \sim direct \tag{9}$$

$$R \sim light \tag{10}$$

$$P \sim sustained \tag{11}$$

$$S \sim bound \tag{12}$$

where the opposite of the qualities listed, flexible, strong, sudden, and free, are achieved when these weights are relatively small, respectively.

Using these weights as the style-based parameters for varying the resulting trajectory, we solve the optimal control problem

$$\begin{aligned} &\min_u J \\ &s.t. \\ &\dot{x} = Ax + Bu \\ &y = Cx \end{aligned} \tag{13}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, and $C \in \mathbb{R}^{l\times n}$.

Differentiating the Hamiltonian

$$H = \frac{1}{2}[(y-r)^T Q(y-r) + u^T Ru + \dot{x}^T P\dot{x}] + \lambda f$$

with respect to u and x gives the first order necessary condition for optimality and the dynamics of the costate $\lambda = [\lambda_1, \lambda_2, ...\lambda_n]$:

$$\frac{\partial H}{\partial u} = u^T(R + B^T PB) + x^T A^T PB + \lambda B = 0$$

from which it follows that

$$u = -(R + B^T PB)^{-1}(B^T PAx - B^T \lambda^T) \tag{14}$$

and

$$\begin{aligned} \frac{\partial H}{\partial x} = -\dot{\lambda} = x^T(C^T QC + A^T PA) + u^T B^T PA \\ +\lambda A - r^T QC \end{aligned}$$

from which it follows that

$$\dot{\xi} = \dot{\lambda}^T = (A^T PB(R + B^T PB)^{-1} B^T PA - C^T QC$$
$$-A^T PA)x + (A^T PB(R + B^T PB)^{-1} B^T - A^T)\xi$$
$$+C^T Qr. \tag{15}$$

For the costate's boundary condition we obtain:

$$\xi(T_f) = C^T S\, C x(T_f) - C^T S\, r(T_f). \tag{16}$$

To solve this system for an optimal $x(t)$ we need to find ξ_0. Thus, assemble a new state $z = [x, \xi]^T$. Now

$$\dot{z} = Mz + Nr \tag{17}$$

where the entries of M and N are determined from Eqs. 13 - 15 and are given below:

$$M_{11} = A - B(R + B^T PB)^{-1} B^T PA \tag{18}$$

$$M_{12} = -B(R + B^T PB)^{-1} B^T \tag{19}$$

$$M_{21} = A^T PB(R + B^T PB)^{-1} B^T PA$$
$$-C^T QC - A^T PA \tag{20}$$

$$M_{22} = A^T PB(R + B^T PB)^{-1} B^T - A^T \tag{21}$$

$$N_1 = [0]_{n \times l} \tag{22}$$

$$N_2 = C^T Q. \tag{23}$$

We know that in general

$$z(T_f) = e^{MT_f} z_0 + \int_0^{T_f} e^{M(T_f - t)} Nr(t) dt, \tag{24}$$

which we denote by

$$z(T_f) = \Phi z_0 + q, \tag{25}$$

where

$$\Phi = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \tag{26}$$

$$q = \begin{bmatrix} q_1 \\ q_2 \end{bmatrix}. \tag{27}$$

Combining Eq. 16 with Eqs. 25, 26, and 27 we get

$$\begin{aligned} \xi_0 = (C^T S C \Phi_{12} - \Phi_{22})^{-1} [(\Phi_{21} - C^T S C \Phi_{11}) x_0 \\ + C^T S C q_1 + q_2 + C^T S r(T_f)]. \end{aligned} \tag{28}$$

Thus, we have found our initial condition $z_0 = [x_0, \xi_0]^T$. When combined with Eq. 110, this gives the optimal $x(t)$ – and thus our output $y(t)$ – as nominated by the weights in Eq. 8.

2.4 An Illustrative Example for Robotics

Now we demonstrate the combined utility of this framework for specifying desired stylistic robotic behavior by constructing a disco dancing and a cheerleading behavior as in [72]. We produce pose sequences that are perceived by viewers as being different - although they are actually composed from the same underlying poses - and animate them on the Aldebaran NAO robotic platform. Further, we simulate dynamic trajectories between the poses in these sequences which, by varying weights in a cost function, emote very different attitudes of movement execution.

First, we construct two instantiations of one-arm automata $\mathcal{G}^{disco}_{arm_{1,2}}$ and $\mathcal{G}^{cheer}_{arm_{1,2}}$, a continuous feasibility set O^{NAO}_{infeas} [8], and two corresponding sets of unaesthetic states $X^{disco}_{unaesth}$ and $X^{cheer}_{unaesth}$. These objects assemble to create an automata which establishes allowable sequences of movements for both arms. Then, we enumerate two sets of weights $\{Q, R, P, S\}_{disco}$ and $\{Q, R, P, S\}_{cheer}$ which determine the effort quality of our trajectories between poses according to a cost function corresponding to the one in Eq. 8. Thus, we specify rules for

[8] Which, in practice, is implemented by removing transitions from the joint arm automaton.

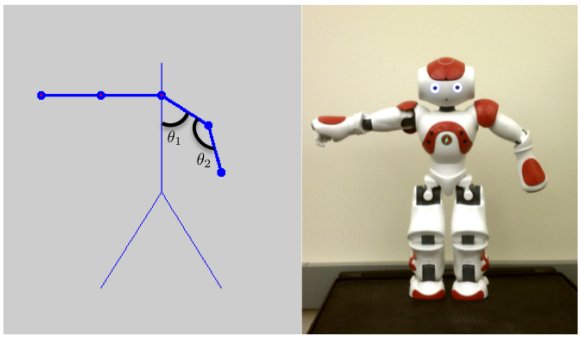


Figure 12. The discrete states are interpreted as poses corresponding to two joint angles: shoulder and elbow as restricted to the body's coronal plane. On the left is the simulated view of a pose, and on the right is a corresponding pose on an actual robotic platform.

motion sequencing and the dynamic timing each motion should exhibit.

Each state of the one-armed automata, which corresponds to a single arm pose, is constructed from a pair of joint angles, (θ_1, θ_2), as shown in Fig. 12. These poses were chosen such that the degrees of freedom of the arm (limited to the body's coronal plane) were discretized in a reasonable way and are shown in Fig. 13. Most of the poses have a fully extended elbow except some common examples of when the elbow bends in human behavior, i.e., "hands on hips" and "hands clasped at chest" (poses 4 and 5, respectively). More importantly, we represent shapes critical to the experience of behaviors such as cheerleading and disco-style dancing. From Fig. 12 and the general kinematic constraints of humanoid geometry, we see that the range of $\mathcal{X}$, for both behaviors, is limited, e.g.,

$$\mathcal{X} = \{(\theta_1, \theta_2) \mid \theta_1, \theta_2 \in [0, \pi]\}\,. \tag{29}$$

The state transitions (events) correspond to movements, i.e., "put hand on hip" and "extend arm straight out." That is, here, we enumerate primary motions that in this case are basically linear interpolations between their respective start and end poses. In particular, X^{disco}_{arm}, E^{disco}_{arm}, f^{disco}_{arm}, and Γ^{disco}_{arm} are defined in Fig. 14. Moreover, the rest position is given by the state labeled "1." As the system should always aim to return to that pose, the initial state is given by $x_0 = 1$, and the marked states are set to be $X_m = \{1\}$. The resulting marked language, i.e., the set of event strings that start at x_0 and end in X_m, produce feasible pose

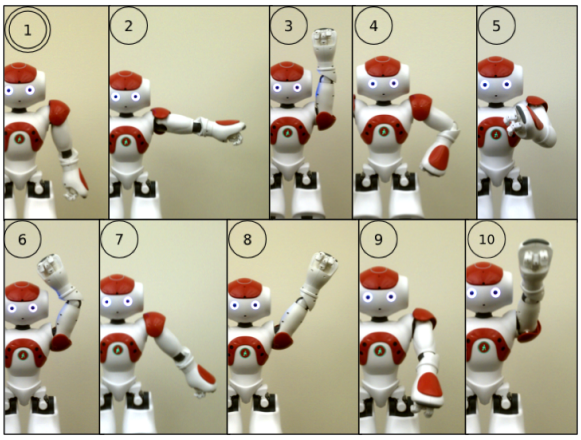


Figure 13. An illustration of the ten arm poses chosen as those of importance and their corresponding discrete states which are used throughout this section – namely, in Figs. 14 and 15.

sequences recognizable as being in a style of disco dance[9]. Likewise, X^{cheer}_{arm}, E^{cheer}_{arm}, f^{cheer}_{arm}, and Γ^{cheer}_{arm} are defined in Fig. 15. Note that $\mathcal{G}^{disco}_{arm_{1,2}}$ and $\mathcal{G}^{cheer}_{arm_{1,2}}$ have the same initial and marked states. Each of these one-armed systems are composed to an automaton describing potential two-armed motion sequences; however this composition will be greedy and allow for physically infeasible and stylistically insensitive sequences.

Further, we specify a region O^{NAO}_{infeas} which corresponds to infeasible trajectory pairs given our ten states and body geometry. This region is consistent with transitions whose trajectories will intersect, which for this set of states and our platform configuration, is any trajectories which end in both arms in pose 9 or both arms in pose 10. Finally, the two sets of unaesthetic, disallowed two-arm states are given in Eqs. 30 and 31.

[9] As disco is not as formalized as a genre like classical ballet (which also has many variants), this should be considered our own interpretation of disco dance as applied to this platform. Nevertheless, what we refer to as "disco" in this paper is a distinct style of movement from the style we call "cheer" and free-form unconstrained movement.

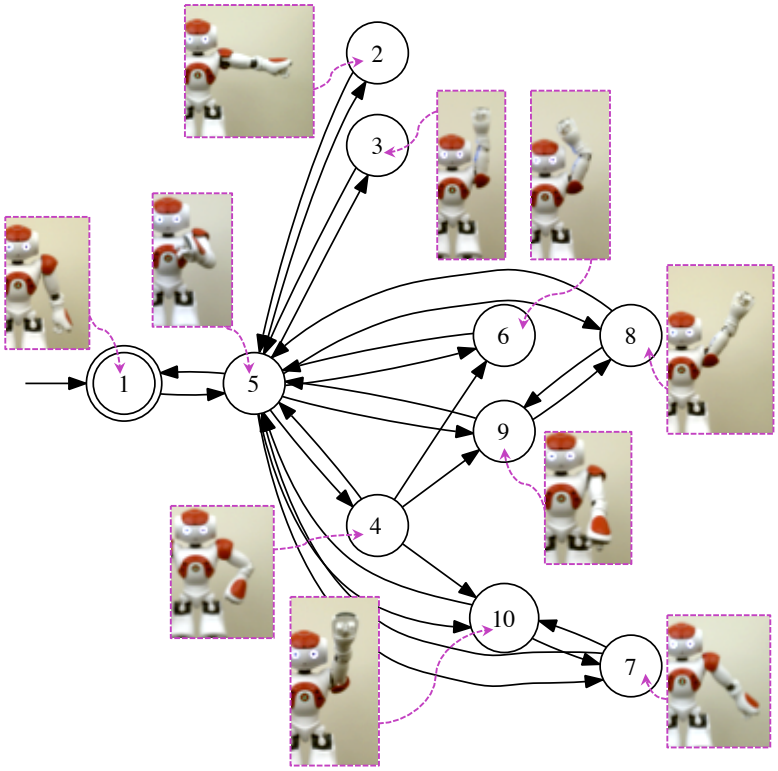


Figure 14. A discrete event model of one arm performing the "disco" behavior. States correspond to poses defined by two joint angles: shoulder (as limited to the coronal plane) and elbow. Events are given by primary movements plus the empty event (or hold), which corresponds to undrawn self-loops.

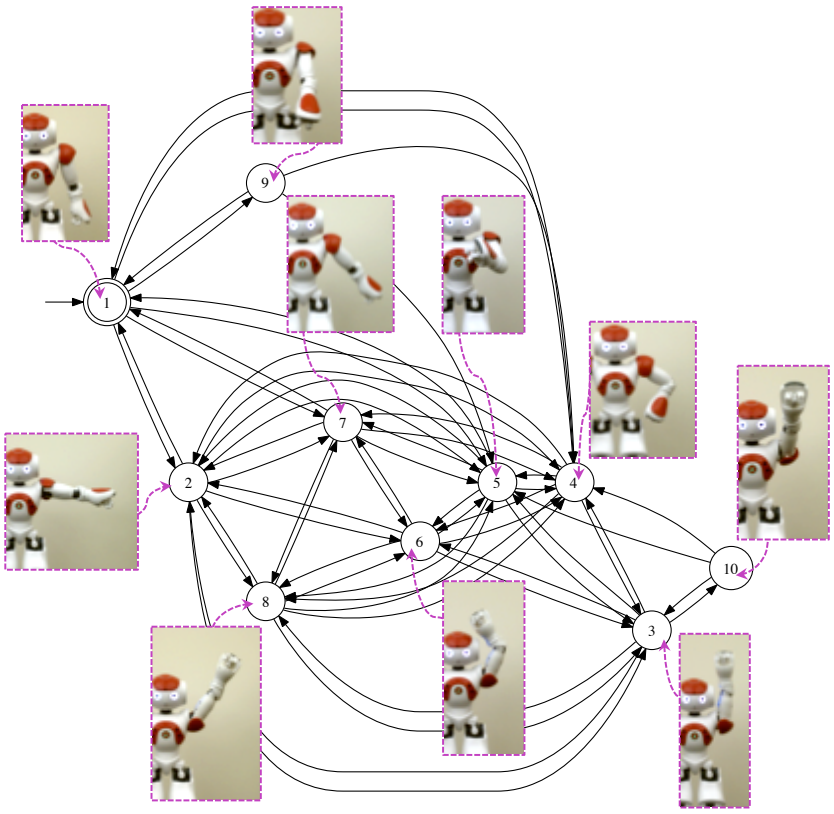


Figure 15. A discrete event model of one arm performing the "cheer" behavior. States correspond to poses defined by two joint angles: shoulder (as limited to the coronal plane) and elbow. Events are given by primary movements plus the empty event (or hold), which corresponds to undrawn self-loops.

$$X_{unaesth}^{disco} = \left\{ \begin{array}{l} (1,2)\ (1,3)\ (1,4)\ (1,5)\ (1,6)\ (1,7)\ (1,8)\ (1,9) \\ (2,1)\ (2,3)\ (2,4)\ (2,5)\ (2,6)\ (2,7)\ (2,8)\ (2,9)\ (2,10) \\ (3,1)\ (3,2)\ (3,3)\ (3,4)\ (3,5)\ (3,6)\ (3,7)\ (3,8)\ (3,9)\ (3,10) \\ (4,1)\ (4,2)\ (4,3)\ (4,5)\ (4,6) \\ (5,1)\ (5,2)\ (5,3)\ (5,4)\ (5,6)\ (5,7)\ (5,8)\ (5,9)\ (5,10) \\ (6,1)\ (6,2)\ (6,3)\ (6,4)\ (6,5)\ (6,7)\ (6,8)\ (6,9)\ (6,10) \\ (7,1)\ (7,2)\ (7,3)\ (7,5)\ (7,6)\ (7,7)\ (7,8)\ (7,9)\ (7,10) \\ (8,1)\ (8,2)\ (8,3)\ (8,5)\ (8,6)\ (8,7)\ (8,8)\ (8,9)\ (8,10) \\ (9,1)\ (9,2)\ (9,3)\ (9,5)\ (9,6)\ (9,7)\ (9,8)\ (9,10) \\ (10,1)\ (10,2)\ (10,3)\ (10,5)\ (10,6)\ (10,7)\ (10,8)\ (10,9) \end{array} \right\} \tag{30}$$

$$X_{unaesth}^{cheer} = \left\{ \begin{array}{l} (1,4)\ (1,5)\ (1,6)\ (1,7)\ (1,8)\ (1,9) \\ (2,4)\ (2,5)\ (2,6)\ (2,7)\ (2,8) \\ (3,4)\ (3,5)\ (3,6)\ (3,7)\ (3,8)\ (3,10) \\ (4,1)\ (4,2)\ (4,3)\ (4,5)\ (4,6) \\ (5,1)\ (5,2)\ (5,3)\ (5,4)\ (5,6)\ (5,7)\ (5,8)\ (5,9)\ (5,10) \\ (6,1)\ (6,2)\ (6,3)\ (6,4)\ (6,5)\ (6,7)\ (6,8)\ (6,9)\ (6,10) \\ (7,1)\ (7,2)\ (7,3)\ (7,5)\ (7,6)\ (7,10) \\ (8,1)\ (8,2)\ (8,3)\ (8,5)\ (8,6)\ (8,9) \\ (9,1)\ (9,5)\ (9,6)\ (9,8)\ (9,10) \\ (10,3)\ (10,5)\ (10,6)\ (10,7)\ (10,9) \end{array} \right\} \tag{31}$$

Pulling these components together, the final motion sequencing systems which evolve according to the styles of disco and cheerleading on the NAO robotic platform $\mathcal{G}_{disco}$ and $\mathcal{G}_{cheer}$ are given respectively by

$$pose(trans(\mathcal{G}_{arm_1}^{disco} \times \mathcal{G}_{arm_2}^{disco}, O_{infeas}^{NAO}), X_{unaesth}^{disco}) \tag{32}$$

and

$$pose(trans(\mathcal{G}_{arm_1}^{cheer} \times \mathcal{G}_{arm_2}^{cheer}, O_{infeas}^{NAO}), X_{unaesth}^{cheer}). \tag{33}$$

Finally, the framework presented here is able to scale the timing of these trajectories with the stylistic weights or knobs we outlined previously. Namely, we consider a

4-dimensional system with double integrator dynamics where $x = [\theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2]^T$, $u = [u_{\theta_1}, u_{\theta_2}]^T$, and

$$\dot{x} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} u \tag{34}$$

$$y = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} x. \tag{35}$$

Note that $y(t)$ describes the pose of the arm over time. We then select the following weight matrices for Eq. 8:

$$\begin{cases} Q_{disco} = 0.1 \cdot I \\ R_{disco} = 0.1 \cdot I \\ P_{disco} = I \\ S_{disco} = 100 \cdot I \end{cases} \tag{36}$$

and

$$\begin{cases} Q_{cheer} = I \\ R_{cheer} = I \\ P_{cheer} = 10 \cdot I \\ S_{cheer} = 100 \cdot I \end{cases} \tag{37}$$

where I is the identity matrix. The nominal movement reference signal, r, which Eq. 8 encourages the system to track, is simply the linear interpolation between the desired end poses given by the motion sequence.

The relative scaling of these matrices reflects the fact that we'd like the disco behavior to be loose but confident; hence, we employ a small weight on trajectory following (Q_{disco}) and a large one on end point matching (S_{disco}) to produce a trajectory that is, in Laban's terms, flexible and bound. We would also like this behavior to be more energetic. Thus, we

employ a small weight on the input (R_{disco}) and change in state (P_{disco}) creating trajectories which are also strong and sudden. On the other hand, the cheer behavior is more rigid but still confident (which implies a larger Q_{cheer} and large S_{cheer}) and energetic but somewhat sustained (larger R_{cheer} and P_{cheer}). Thus, in Laban's framework these trajectories are direct, bound, strong, and sustained.

In order to test the effectiveness of the chosen rules and weights (as these behaviors which we define are subjective and require a human eye to validate – as in Sec. 4.3), random allowable sample paths for both behaviors were generated. The results are animated in simulation and on the Aldebaran NAO humanoid robotic platform; snapshots are provided in Fig. 16.

Significant changes take place between the distinct cases of systems that were animated. Namely, the disco style enforces movements which are largely below the shoulders, as most casual social dancing styles exhibit. The notable exception to this generalization is the distinctive motion when the arm raises up, away from the body, and moves contralaterally and down towards the hips which occurs frequently in this behavior. This classic disco motion recurs frequently when the NAO is moving according to the disco task. Furthermore, the dynamic quality of the simulated movements are loose, imprecise, and give the impression of someone who is carefree and having a good time.

On the other hand, when in the cheerleader mode, the NAO displays not only poses associated with cheerleaders, but also moves frequently through positions with the hands on hips or clasped in front of the chest. In other words, it is not just the shape of the robot that calls to mind a cheerleader association but also the motions the system is performing as it moves between poses. The dynamic quality in this case is more rigid and purposeful; this gives the feeling that the mover is imparting great, specific effort into the movements.

$\diamond$

In this chapter, a quantitative definition of style was introduced. The definition is comprised of three distinct mathematical objects, or *stylistic* parameters, which drive a two

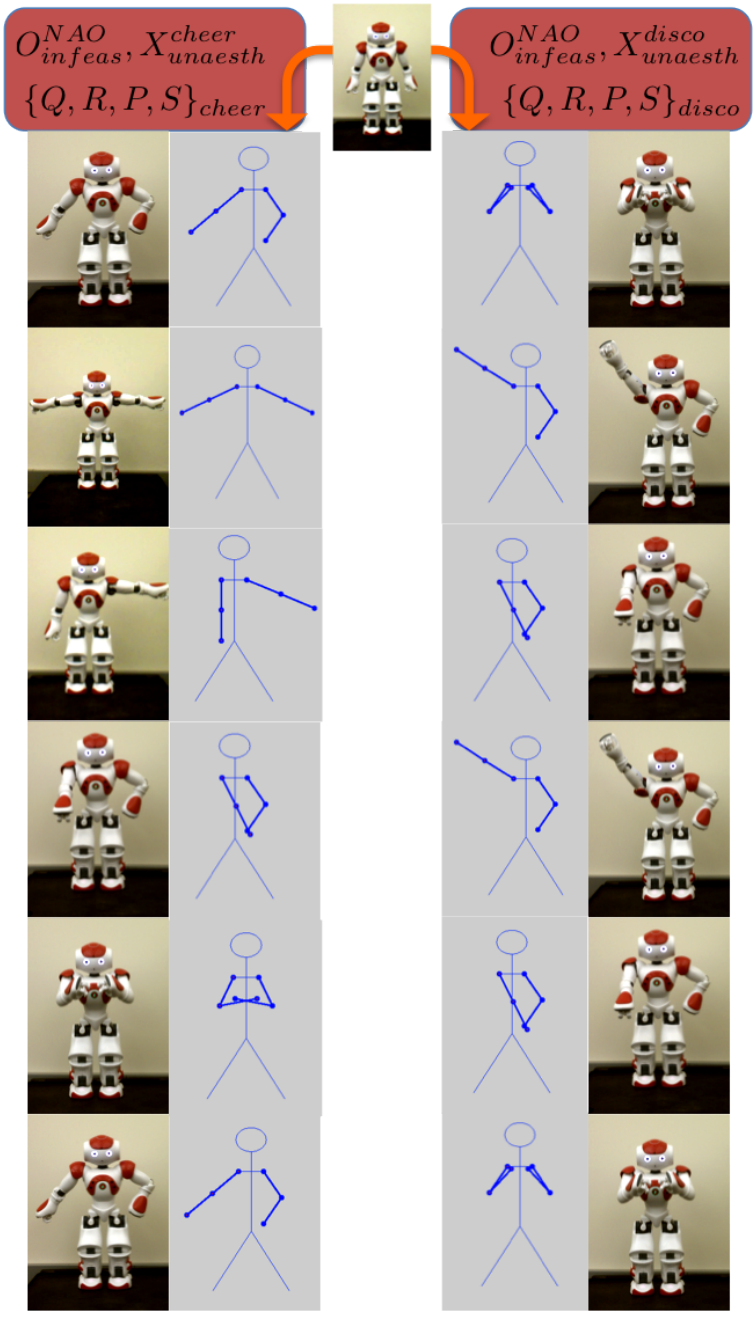


Figure 16. Two example (partial) sequences demonstrate the result of our control method. The left-hand sequence is an example of the system evolving under the "cheer" constraint while the sequence on the right is illustrating the "disco" style.

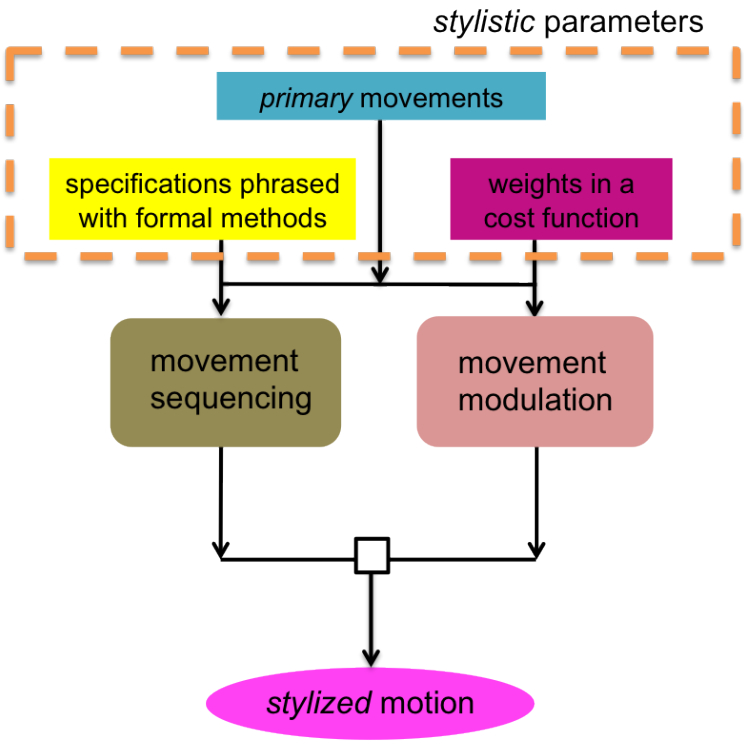


Figure 17. The figure shows a conceptual overview of how the three stylistic parameters and the two-part movement scheme fit together.

part generation scheme. The first, an automaton, lists basic, primary movements, which are represented abstractly by the transitions of the automaton; the states in the automaton are selected body poses. These primary movements are then recombined and sequenced according to knobs on a supervisory controller, which are scripted either as sets or LTL specifications. The primary movements are, in parallel, modulated via an optimal control problem where weights in the cost function determine the *quality* of each movement; this scheme generates specific trajectories which are not explicitly part of the automaton. The overall scheme is illustrated in Fig. 17.

The definition of "style" of movement presented in this chapter is precise such that it can be implemented on robots. This final section has presented an example of exactly that. Note that the robot is able to move in the crafted style of motion *automatically* and the sequences presented are one possible instantiation of movement that satisfies the definition.

CHAPTER 3

THE INVERSE PROBLEM: A STYLISTIC INTERPRETATION OF HUMAN MOVEMENT

The next question this thesis will consider is: "Given raw data of a human moving, can the *quality* of distinct movements contained therein be extracted?" These (two) questions can be formulated as inverse optimal control problems.

Thus, the topic of this chapter is the solutions to two optimal control problems defined by a nested pair of cost functions. The solutions derived here allows for the notion of style defined in the previous chapter to be used as a tool for analyzing data. These nested cost functions take advantage of the useful abstractions set up in Ch. 2 for the goal of extracting a specific kinesthetic concept of motion quality. The first cost function sets up the definition of quality as it applies to movement from Chapter 2, and the second describes a metric for success toward recreating a novel data signal. Solving for optimality conditions simultaneously and gathering three costates for the original state, an algebraic solution to the posed question can be derived and applied to real motion capture data.

This formulation provides a natural way to define a motion *classification* (in Sec. 3.2) and a *segmentation* (in Sec. 3.3) problem. Instead of constraining the problem via statistical metrics or arbitrary mathematical constructs, the idea is to have a human mover define a motion, seeding the problem with one known "movement," which the machinery can then interpret stylistically in order to segment a longer series of movements from this mover. An application to real data is presented in Sec. 3.4.

3.1 Style-based Abstractions for Human Motion Analysis

To manage the complexity associated with engineered systems, abstractions are typically needed to express and analyze characteristics of these systems. Such an abstraction-based approach is often taken when describing human motion, e.g., by segmenting the motion

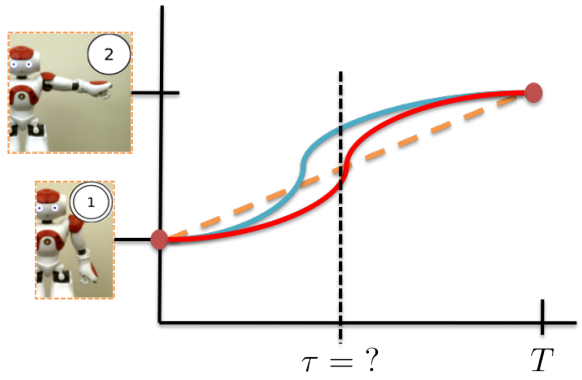


Figure 18. This segmentation problem uses an output (in red) to try and best recreate some real data (in blue) – say joint angle motion over time – with a simulated signal as produced in Sec. 2.3. The variables that generate this output will be the minimizer for optimal control problems presented here.

into primitive movements or so-called movemes. Indeed, there are several definitions of what the appropriate abstraction for human motion analysis might be as discussed in Ch. 1. The aim here is to build a system where the levels of abstraction are defined *kinesthetically* through movement theory and the use of an individual's actual movement as a tracked reference. To do this, human motion styles must be accounted for in a principled way. The tool presented here has the desirable quality that it is flexible enough to incorporate the notion of style from Ch. 2. This work introduces a general method for approximating any empirical signal using optimal control, and such extensions will be discussed in Ch. 5.

The solution presented here consists of two cost functions. The first is called the "template cost function" and is a template for simulating a candidate matching trajectory. The second is referred to as the "metric cost function" and is a metric on how well the simulation is approximating the empirical data. The simulated trajectory will indeed be different than the empirical one, but according to the measure defined by the metric cost function, the simulation will be the best possible recreation of the data. The optimal parameters, a seeming by-product of the process, that define the template cost function can be used to make comparisons between signals and recreate new trajectories that are somehow of the same ilk or style. Fig. 18 displays a conceptual image of this idea.

The benefit of this method is its ability to describe trajectories in a principled way. Here,

the type of trajectories used to motivate this flexible approach are those of humans dancing. Human movement is a rich and varied phenomenon of a complex system that is of interest in robotics. In particular, roboticists often aim to have robots "move in a certain way" or, in other words, move in a certain style. So far in this dissertation, the case has been made for a precise definition of stylistic movement that is rooted in dance theory. The method outlined in this chapter applies this definition to real data and extracts style parameters that can then be used to automatically generate robotic movements that are stylistically similar.

3.2 Quality Matching: The Inverse Problem

Now the precise optimal control framework that will enable the analysis of human movement described above is presented. In Section 3.4, the problem is specialized for the stylistic template cost function where the two reference trajectories employed, $\rho \in \mathbb{R}^l$ and $\sigma \in \mathbb{R}^l$, correspond to real, empirical data of human movement and the nominal movement between the same endpoints, respectively.

Consider the general optimal control problem

$$\min_u J_u = \int_0^T F(x, \sigma, \pi)dt + \psi(x(T), \sigma(T), \pi) \tag{38}$$

$$s.t. \begin{cases} \dot{x} = f(x, u) \quad x(0) = x_0 \\ y = h(x) \end{cases} \tag{39}$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the input, $y \in \mathbb{R}^l$ is the output, σ is a reference, and π is a vector of weighting parameters.

The maximum principle states that the optimizer, u^*, can be expressed as a function of x, ξ, σ, and π. In this context, ξ is the costate satisfying

$$\begin{aligned} \dot{\xi} &= -\frac{\partial F}{\partial x}^T - \frac{\partial f}{\partial x}^T \xi \\ \xi(T) &= \frac{\partial \psi}{\partial x}(x(T), \sigma(T), \pi). \end{aligned} \tag{40}$$

Plugging in the optimal u^* into the equations for x and ξ gives the expression for the

Hamiltonian dynamics to be

$$\dot{x} = f_x(x, u^*(x, \xi, \sigma, \pi)) \tag{41}$$

$$\dot{\xi} = f_\xi(\xi, u^*(x, \xi, \sigma, \pi)), \tag{42}$$

which is denoted as

$$\dot{x} = f_x(x, \xi, \sigma, \pi) \tag{43}$$

$$\dot{\xi} = f_\xi(x, \xi, \sigma, \pi). \tag{44}$$

Now a second cost function, which needs to be minimized with respect to the weighting parameters π under the constraints imposed by the problem just outlined, is introduced. This optimal control problem can be written as follows:

$$\min_{\pi} J_\pi = \int_0^T L(x, \rho)dt + \Psi(x(T), \rho(T)) \tag{45}$$

$$s.t. \begin{cases} \dot{x} = f_x(x, \xi, \sigma, \pi) & x(0) = x_0 \\ \dot{\xi} = f_\xi(x, \xi, \sigma, \pi) & \xi(T) = \frac{\partial \psi}{\partial x}(x(T), \sigma(T), \pi) \end{cases} \tag{46}$$

where ρ is the empirical data to be mimicked and classified.

Theorem 3.2.1 *The first order optimality conditions on π with respect to the cost given in Equation 45 under the constraints in Equation 74 is*

$$\kappa(0) = 0 \tag{47}$$

where

$$\dot{\kappa} = -\lambda_x \frac{\partial f_x}{\partial \pi} - \lambda_\xi \frac{\partial f_\xi}{\partial \pi}$$
$$\kappa(T) = 0 \tag{48}$$

and with

$$\begin{aligned}
\dot{x} &= f_x(x,\xi,\sigma,\pi) && x(0) = x_0 \\
\dot{\xi} &= f_\xi(x,\xi,\sigma,\pi) && \xi(T) = \frac{\partial \psi}{\partial x}(x(T),\sigma(T),\pi) \\
\dot{\lambda}_x &= -\frac{\partial L}{\partial x} - \lambda_x \frac{\partial f_x}{\partial x} - \lambda_\xi \frac{\partial f_\xi}{\partial x} && \lambda_x(T) = \frac{\partial \Psi}{\partial x}(x(T)) - \lambda_\xi(T)\frac{\partial^2 \psi}{\partial x^2}(x(T)) \\
\dot{\lambda}_\xi &= -\lambda_x \frac{\partial f_x}{\partial \xi} - \lambda_\xi \frac{\partial f_\xi}{\partial \xi} && \lambda_\xi(0) = 0.
\end{aligned} \tag{49}$$

Proof: By including the constraints in the cost, the augmented cost is obtained. This is given by

$$\hat{J}_\pi = \int_0^T \left[L(x,\rho) + \lambda_x(f_x - \dot{x}) + \lambda_\xi(f_\xi - \dot{\xi}) \right] dt + \Psi(x(T),\rho(T)).$$

Now consider a variation in π such that

$$\pi \mapsto \pi + \epsilon\theta.$$

Such a variation also causes a variation in the state and costate, i.e.,

$$\Rightarrow \begin{cases} x \mapsto x + \epsilon\eta \\ \xi \mapsto \xi + \epsilon\nu. \end{cases}$$

The variation in π also disturbs the costate's boundary condition, $\xi(T)$, which is given by

$$\xi(T) \mapsto \xi(T) + \epsilon\nu(T).$$

Computing the directional derivative of the augmented cost produces the following:

$$\delta\hat{J}_\pi(\pi;\theta) = \lim_{\epsilon\to 0} \frac{1}{\epsilon}\left[\hat{J}_\pi(\pi+\epsilon\theta) - \hat{J}_\pi(\pi) \right]$$

$$= \int_0^T \left[\frac{\partial L}{\partial x}\eta + \lambda_x\left(\frac{\partial f_x}{\partial x}\eta + \frac{\partial f_x}{\partial \xi}\nu + \frac{\partial f_x}{\partial \pi}\theta \right) - \lambda_x\dot{\eta} + \lambda_\xi\left(\frac{\partial f_\xi}{\partial x}\eta + \frac{\partial f_\xi}{\partial \xi}\nu + \frac{\partial f_\xi}{\partial \pi}\theta \right) - \lambda_\xi\dot{\nu} \right] dt$$
$$+ \frac{\partial \Psi}{\partial x}(x(T))\eta(T).$$

Integrating by parts and collecting terms by variation gives

$$
\begin{aligned}
= \int_0^T \left(\frac{\partial L}{\partial x} + \lambda_x \frac{\partial f_x}{\partial x} + \lambda_\xi \frac{\partial f_\xi}{\partial x} + \dot{\lambda}_x \right) dt \cdot \eta + \lambda_x(0)\eta(0) - \lambda_x(T)\eta(T) \\
+ \int_0^T \left(\lambda_x \frac{\partial f_x}{\partial \xi} + \lambda_\xi \frac{\partial f_\xi}{\partial \xi} + \dot{\lambda}_\xi \right) dt \cdot \nu + \lambda_\xi(0)\nu(0) - \lambda_\xi(T)\nu(T) \\
+ \int_0^T \left(\lambda_x \frac{\partial f_x}{\partial \pi} + \lambda_\xi \frac{\partial f_\xi}{\partial \pi} \right) dt \cdot \theta + \frac{\partial \Psi}{\partial x}(x(T))\eta(T).
\end{aligned}
$$

To unravel the boundary conditions, remember that $\eta(0) = 0$ since η starts at x_0, which is fixed. Next, consider $\nu(T)$. The boundary condition in Equation 40 dictates the terminal value of the costate giving $\xi(T)$ as a function of $x(T)$. An expression for $\nu(T)$ is obtained through the following computation:

$$
\begin{aligned}
\tilde{\xi}(T) &= \xi(T) + \epsilon\nu(T) \\
\Rightarrow \nu(T) &= \frac{1}{\epsilon}\left[\tilde{\xi}(T) - \xi(T)\right] \\
&= \frac{1}{\epsilon}\left[\frac{\partial \psi}{\partial x}(x(T) + \epsilon\eta(T)) - \frac{\partial \psi}{\partial x}(x(T))\right] \\
&= \frac{1}{\epsilon}\left[\frac{\partial \psi}{\partial x}(x(T)) + \epsilon\frac{\partial^2 \psi}{\partial x^2}(x(T))\eta(T) - \frac{\partial \psi}{\partial x}(x(T))\right] \\
&= \frac{\partial^2 \psi}{\partial x^2}(x(T))\eta(T). \qquad (50)
\end{aligned}
$$

If the dynamics of the new costates are set to be

$$
\dot{\lambda}_x = -\frac{\partial L}{\partial x} - \lambda_x \frac{\partial f_x}{\partial x} - \lambda_\xi \frac{\partial f_\xi}{\partial x} \qquad (51)
$$

$$
\dot{\lambda}_\xi = -\lambda_x \frac{\partial f_x}{\partial \xi} - \lambda_\xi \frac{\partial f_\xi}{\partial \xi}, \qquad (52)
$$

then the derivative has reduced to

$$
\begin{aligned}
&\delta \hat{J}_\pi(\pi; \theta) = \\
&= -\lambda_x(T)\eta(T) + \lambda_\xi(0)\nu(0) - \lambda_\xi(T)\frac{\partial^2 \psi}{\partial x^2}(x(T))\eta(T) \\
&+ \int_0^T \left(\lambda_x \frac{\partial f_x}{\partial \pi} + \lambda_\xi \frac{\partial f_\xi}{\partial \pi} \right) dt \cdot \theta + \frac{\partial \Psi}{\partial x}(x(T))\eta(T).
\end{aligned}
$$

Setting

$$\lambda_x(T) = \frac{\partial \Psi}{\partial x}(x(T)) - \lambda_\xi(T)\frac{\partial^2 \psi}{\partial x^2}(x(T)) \tag{53}$$

$$\lambda_\xi(0) = 0 \tag{54}$$

leaves

$$\delta \hat{J}_\pi(\pi;\theta) = \int_0^T \left(\lambda_x \frac{\partial f_x}{\partial \pi} + \lambda_\xi \frac{\partial f_\xi}{\partial \pi}\right) dt \cdot \theta,$$

which should equal zero for all values of θ to achieve optimality. To keep track of this in the compact way presented in the theorem, let

$$\kappa = \int_t^T \left(\lambda_x \frac{\partial f_x}{\partial \pi} + \lambda_\xi \frac{\partial f_\xi}{\partial \pi}\right) dt. \tag{55}$$

Differentiating κ with respect to time gives the expressions in Equations 48 and 47, which concludes the proof. ■

3.3 A Style-based Segmentation Problem

Our mapping to Laban's description of movement *quality* is now leveraged in order to segment human motion into basic movements. Similar to the optimization set up in the last section, here the goal is to use a simulated trajectory, which is of a desirable or useful form, in order to "best" match real data. Instead of weights determining the quality of the simulated trajectory, the optimization parameter now is a switch time, signaling the start of a new "movement," which may be considered an alternate description of a *moveme*.

In particular, the problem is now

$$\min_\tau \int_0^T L(x,\rho)dt + \Psi_1(x(\tau),\rho(\tau)) + \Psi_2(x(T),\rho(T)) \tag{56}$$

$$\text{s.t.} \begin{cases} u = \arg\min \int_0^T F(x,u,\sigma)dt + \psi_1(x(\tau),\sigma(\tau)) + \psi_2(x(T),\sigma(T)) \\ \qquad\qquad \text{s.t.} \begin{cases} \dot{x} = f_x(x,u) \\ x(0) = x_0 \end{cases} \end{cases} \tag{57}$$

where $\tau \in \mathbb{R}$ is a switch time variable which segments the novel data, $\rho \in \mathbb{R}^l$, into two "movements," $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the input, $y \in \mathbb{R}^l$ is the output, $\sigma \in \mathbb{R}^l$ is a reference signal.

The first task at hand is to solve the "forward" optimization problem for u (which is the same as that used to generate quality-endowed trajectories from the previous section now split into two time intervals by a variable τ) in terms of the new optimization variable τ.

Thus, we solve

$$\min_u \int_0^T F(x,u,\sigma)dt + \psi_1(x(\tau),\sigma(\tau)) + \psi_2(x(T),\sigma(T)) \tag{58}$$

$$\text{s.t.} \begin{cases} \dot{x} = f(x,u) \\ x(0) = x_0 \end{cases}. \tag{59}$$

Lemma 3.3.1 *The first order necessary optimality conditions on u with respect to the cost in Eq. 58 subject to the constraints in Eqs. 59-59 are given by*

$$\kappa_u(0) = 0 \tag{60}$$

where

$$\begin{aligned} \dot{\kappa}_u &= -\frac{\partial F}{\partial u} - \lambda \frac{\partial f_x}{\partial u} \\ \kappa_u(T) &= 0 \end{aligned} \tag{61}$$

and with

$$\dot{\lambda} = -\frac{\partial F}{\partial x} - \lambda \frac{\partial f_x}{\partial x} \tag{62}$$

$$\lambda(\tau^-) - \lambda(\tau^+) = \Delta\lambda(\tau) = \frac{\partial \psi_1}{\partial x}(x(\tau),\sigma(\tau)) \tag{63}$$

$$\lambda(T) = \frac{\partial \psi_2}{\partial x}(x(T),\sigma(T)) \tag{64}$$

Proof: The augmented cost is given by

$$\hat{J}_u = \int_0^T F(x,u,\sigma) + \lambda\,(f - \dot{x})\,dt + \psi_1(x(\tau)) + \psi_2(x(T)).$$

Consider a variation in u such that

$$u \mapsto u + \epsilon v.$$

This variation also causes a variation in the state, i.e.,

$$\Rightarrow x \mapsto x + \epsilon\eta.$$

Computing the directional derivative of the augmented cost produces the following:

$$\delta\hat{J}_\pi(\pi;\theta) = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[\hat{J}_u(u+\epsilon\nu) - \hat{J}_u(u)\right]$$

$$= \int_0^T \left(\frac{\partial F}{\partial u} + \lambda\frac{\partial f}{\partial u}\right)\nu + \left(\frac{\partial F}{\partial x} + \lambda\frac{\partial f}{\partial x}\right)\eta - \lambda\dot{\eta}\,dt + \frac{\partial \psi_1}{\partial x}(x(\tau))\eta(\tau) + \frac{\partial \psi_2}{\partial x}(x(T))\eta(T).$$

Integrating by parts on the intervals $[0,\tau]$ and $[\tau,T]$ in order to catch the discontinuity at τ gives

$$= \int_0^T \left(\frac{\partial F}{\partial u} + \lambda\frac{\partial f}{\partial u}\right)\nu + \left(\frac{\partial F}{\partial x} + \lambda\frac{\partial f}{\partial x} - \dot{\lambda}\right)\eta\,dt$$
$$+\lambda(0)\eta(0) - \lambda(\tau^-)\eta(\tau) + \lambda(\tau^+)\eta(\tau) - \lambda(T)\eta(T)$$
$$+\frac{\partial \psi_1}{\partial x}(x(\tau))\eta(\tau) + \frac{\partial \psi_2}{\partial x}(x(T))\eta(T).$$

To fill in the boundary conditions, note that $\eta(0) = 0$ because η starts at x_0, which is fixed. The following substitutions

$$\dot{\lambda} = -\frac{\partial F}{\partial x} - \lambda\frac{\partial f}{\partial x} \tag{65}$$

$$\lambda(\tau^-) - \lambda(\tau^+) = \frac{\partial \psi_1}{\partial x}(x(\tau)) \tag{66}$$

$$\lambda(T) = \frac{\partial \psi_2}{\partial x}(x(T)) \tag{67}$$

reduce the derivative to

$$\delta\hat{J}_u(u;\nu) = \int_0^T \left(\frac{\partial F}{\partial u} + \lambda\frac{\partial f}{\partial u}\right)\nu\,dt,$$

which should equal zero for all values of ν to achieve optimality. To keep track of this in the compact way presented in the lemma, let

$$\kappa_u = \int_t^T \left(\lambda\frac{\partial f}{\partial u} + \lambda\frac{\partial f}{\partial u}\right)dt. \tag{68}$$

Differentiating κ_u with respect to time gives the expressions in Eqn. 61, which concludes the proof. ■

This lemma now becomes a constraint set for an inverse problem. Namely, we note that the optimizer, u^*, can be expressed as a function of x, $\lambda^T = \xi$, σ, and τ, i.e.,

$$\dot{x} = f_x(x, u^*(x, \xi, \sigma, \tau)) \tag{69}$$

$$\dot{\xi} = f_\xi(\xi, u^*(x, \xi, \sigma, \tau)), \tag{70}$$

which we denote with

$$\dot{x} = f_x(x, \xi, \sigma, \tau) \tag{71}$$

$$\dot{\xi} = f_\xi(x, \xi, \sigma, \tau). \tag{72}$$

A second cost function, which is minimized with respect to the timing parameter τ under the constraints imposed by the forward problem (Lemma 3.3.1), is now defined. That is,

$$\min_\tau J_\tau = \int_0^T \mathcal{L}(x, \rho)dt + \Psi_1(x(\tau), \rho(\tau)) + \Psi_2(x(T), \rho(T)) \tag{73}$$

$$s.t. \begin{cases} \dot{x} = f_x(x, \xi, \sigma, \tau) & x(0) = x_0 \\ \dot{\xi} = f_\xi(x, \xi, \sigma, \tau) & \xi(\tau^-) - \xi(\tau^+) = \Delta\xi(\tau) = G_1(x(\tau), \sigma(\tau)) \\ & \xi(T) = G_2(x(T), \sigma(T)) \\ \dot{\sigma} = f_\sigma(\sigma, \tau, \rho) & \sigma(0) = \rho(0) \\ & \sigma(\tau) = \rho(\tau) \\ & \sigma(T) = \rho(T) \end{cases} \tag{74}$$

where ρ is the empirical data we wish to mimic and thus segment with τ. The functions G_1 and G_2 are place-holders for the conditions derived in the lemma; namely, $G_1 = \frac{\partial \psi_1}{\partial x}(x(\tau))$, and $G_2 = \frac{\partial \psi_2}{\partial x}(x(T))$.

Theorem 3.3.1 *The first order optimality conditions on τ with respect to the cost given in are*

$$\frac{\partial J_\tau}{\partial \tau} = 0 \tag{75}$$

where

$$\frac{\partial J_\tau}{\partial \tau} = \kappa_\tau(0) + \mu\Big(f_x(x,\xi(\tau^-),\sigma,\tau) - f_x(x,\xi(\tau^+),\sigma,\tau) + f_\xi(x,\xi(\tau^-),\sigma,\tau) - f_\xi(x,\xi(\tau^+),\sigma,\tau) + \frac{\partial G_1}{\partial \sigma}\dot{\sigma}(\tau) + \frac{\partial \Psi_1}{\partial x}\dot{x}(\tau)\Big), \tag{76}$$

where the value $\kappa(0)$ *is given by the differential equation*

$$\begin{cases} \dot{\kappa}_\tau = -\lambda_x \dfrac{\partial f_x}{\partial \tau} - \lambda_\xi \dfrac{\partial f_\xi}{\partial \tau} - \lambda_\sigma \dfrac{\partial f_\sigma}{\partial \tau} \\ \kappa_\tau(T) = 0 \end{cases}, \tag{77}$$

and with

$$\begin{array}{ll}
\dot{x} = f_x(x,\xi,\sigma,\tau) & x(0) = x_0 \\
\dot{\xi} = f_\xi(x,\xi,\sigma,\tau) & \xi(\tau^-) - \xi(\tau^+) = \Delta\xi(\tau) = G_1(x(\tau),\sigma(\tau)) \\
 & \xi(T) = G_2(x(T),\sigma(T)) \\
\dot{\sigma} = f_\sigma(\sigma,\tau,\rho) & \sigma(0) = \rho(0) \\
 & \sigma(\tau) = \rho(\tau) \\
 & \sigma(T) = \rho(T) \\
\dot{\lambda}_x = -\dfrac{\partial L}{\partial x} - \lambda_x \dfrac{\partial f_x}{\partial x} - \lambda_\xi \dfrac{\partial f_\xi}{\partial x} & \lambda_x(\tau^-) = -\lambda_x(\tau^+) - \dfrac{\partial \Psi_1}{\partial x}(x(\tau)) - \lambda_\xi(\tau^+)\dfrac{\partial G_1}{\partial x}(x(\tau)) \\
 & \lambda_x(T) = \dfrac{\partial \Psi_2}{\partial x}(x(T)) - \lambda_\xi(T)\dfrac{\partial G_2}{\partial x} \\
\dot{\lambda}_\xi = -\lambda_x \dfrac{\partial f_x}{\partial \xi} - \lambda_\xi \dfrac{\partial f_\xi}{\partial \xi} & \lambda_\xi(0) = 0 \\
 & \lambda_\xi(\tau^-) = \lambda_\xi(\tau^+) \\
\dot{\lambda}_\sigma = -\lambda_x \dfrac{\partial f_x}{\partial \sigma} - \lambda_\xi \dfrac{\partial f_\xi}{\partial \sigma} - \lambda_\sigma \dfrac{\partial f_\sigma}{\partial \sigma} & \\
\mu = \lambda_\xi(\tau^+) &
\end{array} \tag{78}$$

Proof: Augmenting the cost, including a constraint for the jump in the costate, ξ yields

$$\hat{J}_\tau = \int_0^T \left[L(x,\rho) + \lambda_x(f_x - \dot{x}) + \lambda_\xi(f_\xi - \dot{\xi}) + \lambda_\sigma(f_\sigma - \dot{\sigma}) \right] dt$$
$$+\Psi_1(x(\tau),\rho(\tau)) + \Psi_2(x(T),\rho(T)) + \mu\left(G_1(x(\tau),\sigma(\tau) - \Delta\xi(\tau)\right).$$

Now consider a variation in τ such that

$$\tau \mapsto \tau + \epsilon\theta.$$

Such a variation also causes a variation in the state, x, costate, ξ, and the dynamics of the "nominal" move, σ, which is interpolated between $\rho(0)$, $\rho(\tau)$, and $\rho(T)$, i.e.,

$$\Rightarrow \begin{cases} x \mapsto x + \epsilon\eta \\ \xi \mapsto \xi + \epsilon\nu \\ \sigma \mapsto \sigma + \epsilon\omega. \end{cases}$$

The variation in τ also disturbs the costate's boundary condition, $\xi(T)$, which is given by

$$\xi(T) \mapsto \xi(T) + \epsilon\nu(T),$$

but since $\xi(T) = G_2(x(T),\sigma(T))$ this can be written as

$$\xi(T) \mapsto \xi(T) + \epsilon\left(\frac{\partial G_2}{\partial x}\eta(T) + \frac{\partial G_2}{\partial \sigma}\omega(T)\right).$$

Due to the jump in the costate ξ, the variation in τ also produces a non-small variation across the interval $[\tau, \tau + \epsilon\theta]$. On this interval, the expression for the disturbed costate as $\tilde{\xi} = \xi + \epsilon\nu$ breaks down (since it assumes a "small" ϵ) so in what follows it will be denoted simply as $\tilde{\xi}$ on this diminishing interval. The difference in the augmented cost and its variation is given by:

$$\hat{J}_\tau(\tau + \epsilon\theta) - \hat{J}_\tau(\tau) = \int_0^\tau \left[\epsilon\frac{\partial L}{\partial x}\eta + \epsilon\lambda_x\left(\frac{\partial f_x}{\partial x}\eta + \frac{\partial f_x}{\partial \xi}\nu + \frac{\partial f_x}{\partial \sigma}\omega + \frac{\partial f_x}{\partial \tau}\theta\right) - \epsilon\lambda_x\dot{\eta} \right.$$
$$\left. +\epsilon\lambda_\xi\left(\frac{\partial f_\xi}{\partial x}\eta + \frac{\partial f_\xi}{\partial \xi}\nu + \frac{\partial f_\xi}{\partial \sigma}\omega + \frac{\partial f_\xi}{\partial \tau}\theta\right) - \epsilon\lambda_\xi\dot{\nu} + \epsilon\lambda_\sigma\left(\frac{\partial f_\sigma}{\partial \sigma}\omega + \frac{\partial f_\sigma}{\partial \tau}\theta\right) - \epsilon\lambda_\sigma\dot{\omega} \right] dt$$

$$+\epsilon\frac{\partial\Psi_1}{\partial x}(x(\tau))\eta(\tau) + \int_{\tau+\epsilon\theta}^{T}\left[\epsilon\frac{\partial L}{\partial x}\eta + \epsilon\lambda_x\left(\frac{\partial f_x}{\partial x}\eta + \frac{\partial f_x}{\partial\xi}\nu + \frac{\partial f_x}{\partial\sigma}\omega + \frac{\partial f_x}{\partial\tau}\theta\right) - \epsilon\lambda_x\dot{\eta}\right.$$
$$\left.+\epsilon\lambda_\xi\left(\frac{\partial f_\xi}{\partial x}\eta + \frac{\partial f_\xi}{\partial\xi}\nu + \frac{\partial f_\xi}{\partial\sigma}\omega + \frac{\partial f_\xi}{\partial\tau}\theta\right) - \epsilon\lambda_\xi\dot{\nu} + \epsilon\lambda_\sigma\left(\frac{\partial f_\sigma}{\partial\sigma}\omega + \frac{\partial f_\sigma}{\partial\tau}\theta\right) - \epsilon\lambda_\sigma\dot{\omega}\right]dt$$
$$+\epsilon\frac{\partial\Psi_2}{\partial x}(x(T))\eta(T) + \int_{\tau}^{\tau+\epsilon\theta}\left[\epsilon\frac{\partial L}{\partial x}\eta + \lambda_x\left(f_x(x,\tilde{\xi},\sigma,\tau) - f_x(x,\xi,\sigma,\tau)\right.\right.$$
$$\left.+\epsilon\frac{\partial f_x}{\partial x}\eta + \epsilon\frac{\partial f_x}{\partial\sigma}\omega + \epsilon\frac{\partial f_x}{\partial\tau}\theta\right) - \epsilon\lambda_x\dot{\eta} + \lambda_\xi\left(f_\xi(x,\tilde{\xi},\sigma,\tau) - f_\xi(x,\xi,\sigma,\tau)\right.$$
$$\left.\left.+\epsilon\frac{\partial f_\xi}{\partial x}\eta + \epsilon\frac{\partial f_\xi}{\partial\sigma}\omega + \epsilon\frac{\partial f_\xi}{\partial\tau}\theta - \dot{\tilde{\xi}} + \dot{\xi}\right) + \epsilon\lambda_\sigma\left(\frac{\partial f_\sigma}{\partial\sigma}\omega + \frac{\partial f_\sigma}{\partial\tau}\theta\right) - \epsilon\lambda_\sigma\dot{\omega}\right]dt \ ...$$
$$... + \epsilon\mu\frac{\partial G}{\partial x}\left(\eta(\tau) + \theta(\tau)\dot{x}(\tau)\right) + \epsilon\mu\frac{\partial G}{\partial\sigma}\left(\omega(\tau) + \theta(\tau)\dot{\sigma}(\tau)\right) - \mu\Delta\tilde{\xi}(\tau) + \mu\Delta\xi(\tau) + O(\epsilon).$$

Now special consideration must be given to the terms where $\tilde{\xi} \neq \xi + \epsilon\nu$. The mean value theorem is applied on the interval $[\tau, \tau + \epsilon\theta]$, and the equality $\tilde{\xi}(\tau^+) - \xi(\tau^+) = \eta(\tau) + \theta\dot{\xi}(\tau^+)$ is used for the jump constraint (μ) terms. Note now that the variation ν may have a discontinuity due to the discontinuity in ξ. Computing the directional derivative of the augmented cost produces the following:

$$\delta\hat{J}_\tau(\tau;\theta) = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[\hat{J}_\tau(\tau+\epsilon\theta) - \hat{J}_\tau(\tau)\right]$$
$$= \int_0^T\left[\frac{\partial L}{\partial x}\eta + \lambda_x\left(\frac{\partial f_x}{\partial x}\eta + \frac{\partial f_x}{\partial\xi}\nu + \frac{\partial f_x}{\partial\sigma}\omega + \frac{\partial f_x}{\partial\tau}\theta\right) - \lambda_x\dot{\eta} + \lambda_\xi\left(\frac{\partial f_\xi}{\partial x}\eta + \frac{\partial f_\xi}{\partial\xi}\nu + \frac{\partial f_\xi}{\partial\sigma}\omega + \frac{\partial f_\xi}{\partial\tau}\theta\right)\right.$$
$$\left.-\lambda_\xi\dot{\nu} + \lambda_\sigma\left(\frac{\partial f_\sigma}{\partial\sigma}\omega + \frac{\partial f_\sigma}{\partial\tau}\theta\right) - \lambda_\sigma\dot{\omega}\right]dt + \frac{\partial\Psi_1}{\partial x}(x(\tau))\eta(\tau) + \frac{\partial\Psi_2}{\partial x}(x(T))\eta(T).$$
$$+\theta\left(\lambda_x(\tau^+)\left(f_x(x,\xi(\tau^-),\sigma,\tau) - f_x(x,\xi(\tau^+),\sigma,\tau)\right)\right.$$
$$\left.+\lambda_\xi(\tau^+)\left(f_\xi(x,\xi(\tau^-),\sigma,\tau) - f_\xi(x,\xi(\tau^+),\sigma,\tau)\right) + \lambda_\xi(\tau^+)\Delta\dot{\xi}(\tau)\right)$$
$$+\mu\frac{\partial G}{\partial x}\left(\eta(\tau) + \theta\dot{x}(\tau)\right) + \mu\frac{\partial G}{\partial\sigma}\left(\omega(\tau) + \theta\dot{\sigma}(\tau)\right) - \mu\Delta\nu(\tau) - \mu\theta\Delta\dot{\xi}(\tau)$$

Integrating by parts, over the intervals $[0, \tau]$ and $[\tau, T]$ – for the costate's variation, ν – in anticipation of the discontinuity in ξ at τ, and collecting terms by variation gives

$$= \int_0^T\left(\frac{\partial L}{\partial x} + \lambda_x\frac{\partial f_x}{\partial x} + \lambda_\xi\frac{\partial f_\xi}{\partial x} + \dot{\lambda}_x\right)\eta\,dt + \lambda_x(0)\eta(0) - \lambda_x(\tau^-)\eta(\tau) + \lambda_x(\tau^+)\eta(\tau) - \lambda_x(T)\eta(T)$$

$$+\int_0^T \left(\lambda_x \frac{\partial f_x}{\partial \xi} + \lambda_\xi \frac{\partial f_\xi}{\partial \xi} + \dot{\lambda}_\xi\right)\nu\, dt + \lambda_\xi(0)\nu(0) - \lambda_\xi(\tau^-)\nu(\tau^-) + \lambda_\xi(\tau^+)\nu(\tau^+) - \lambda_\xi(T)\nu(T)$$

$$+\int_0^T \left(\lambda_x \frac{\partial f_x}{\partial \sigma} + \lambda_\xi \frac{\partial f_\xi}{\partial \sigma} + \lambda_\sigma \frac{\partial f_\sigma}{\partial \sigma} + \dot{\lambda}_\sigma\right)\omega\, dt + \lambda_\sigma(0)\omega(0) - \lambda_\sigma(T)\omega(T)$$

$$+\int_0^T \left(\lambda_x \frac{\partial f_x}{\partial \tau} + \lambda_\xi \frac{\partial f_\xi}{\partial \tau} + \lambda_\sigma \frac{\partial f_\sigma}{\partial \tau}\right)\theta\, dt + \left(\frac{\partial \Psi_1}{\partial x}(x(\tau)) + \mu \frac{\partial G}{\partial x}\right)\eta(\tau) + \frac{\partial \Psi_2}{\partial x}(x(T))\eta(T)$$

$$+\theta\left(\lambda_x(\tau^+)\left(f_x(x, \xi(\tau^-), \sigma, \tau) - f_x(x, \xi(\tau^+), \sigma, \tau)\right)\right.$$

$$\left.+\lambda_\xi(\tau^+)\left(f_\xi(x, \xi(\tau^-), \sigma, \tau) - f_\xi(x, \xi(\tau^+), \sigma, \tau)\right) + \lambda_\xi(\tau^+)\Delta\dot{\xi}(\tau)\right)$$

$$+\theta\mu \frac{\partial G}{\partial x}\dot{x}(\tau) + \mu \frac{\partial G}{\partial \sigma}\left(\omega(\tau) + \theta\dot{\sigma}(\tau)\right) - \mu\Delta\nu(\tau) - \mu\theta\Delta\dot{\xi}(\tau).$$

To unravel the boundary conditions, remember that $\eta(0) = 0$ since η starts at x_0, which is fixed. Likewise, $\sigma(0)$, $\sigma(\tau)$, and $\sigma(T)$ are fixed to be the value of ρ at each respective time; thus, $\omega(0) = \omega(\tau) = \omega(T) = 0$. It has also be shown that an expression in terms of $\eta(T)$ and $\omega(T)$ (which is zero) for $\nu(T)$ can be obtained.

Further, if the dynamics of the new costates are set to be

$$\dot{\lambda}_x = -\frac{\partial L}{\partial x} - \lambda_x \frac{\partial f_x}{\partial x} - \lambda_\xi \frac{\partial f_\xi}{\partial x} \tag{79}$$

$$\dot{\lambda}_\xi = -\lambda_x \frac{\partial f_x}{\partial \xi} - \lambda_\xi \frac{\partial f_\xi}{\partial \xi} \tag{80}$$

$$\dot{\lambda}_\sigma = -\lambda_x \frac{\partial f_x}{\partial \sigma} - \lambda_\xi \frac{\partial f_\xi}{\partial \sigma} - \lambda_\sigma \frac{\partial f_\sigma}{\partial \sigma}, \tag{81}$$

then the derivative has reduced to

$$\delta \hat{J}_\tau(\tau; \theta)$$

$$= \left(\lambda_x(\tau^-) + \lambda_x(\tau^+) + \frac{\partial \Psi_1}{\partial x}(x(\tau)) + \mu \frac{\partial G}{\partial x}(x(\tau))\right)\eta(\tau)$$

$$+\left(\frac{\partial \Psi_2}{\partial x}(x(T)) - \lambda_x(T) - \lambda_\xi(T)\frac{\partial F}{\partial x}\right)\eta(T) + \lambda_\xi(0)\nu(0)$$

$$+(\mu - \lambda_\xi(\tau^-))\nu(\tau^-) + (\lambda_\xi(\tau^+) - \mu)\nu(\tau^+) + \int_0^T \left(\lambda_x \frac{\partial f_x}{\partial \tau} + \lambda_\xi \frac{\partial f_\xi}{\partial \tau} + \lambda_\sigma \frac{\partial f_\sigma}{\partial \tau}\right)\theta\, dt$$

$$+\theta\Big[\lambda_x(\tau^+)\left(f_x(x, \xi(\tau^-), \sigma, \tau) - f_x(x, \xi(\tau^+), \sigma, \tau)\right)$$

$$+\lambda_\xi(\tau^+)\left(f_\xi(x, \xi(\tau^-), \sigma, \tau) - f_\xi(x, \xi(\tau^+), \sigma, \tau)\right)$$

$$+\left(\lambda_\xi(\tau^+)-\mu\right)\Delta\dot{\xi}(\tau)+\mu\frac{\partial G}{\partial x}\dot{x}(\tau)+\mu\frac{\partial G}{\partial \sigma}\dot{\sigma}(\tau)\Bigg].$$

Setting

$$\mu = \lambda_\xi(\tau^+) \tag{82}$$

$$\lambda_x(\tau^-) = -\lambda_x(\tau^+) - \frac{\partial \Psi_1}{\partial x}(x(\tau)) - \lambda_\xi(\tau^+)\frac{\partial G}{\partial x}(x(\tau)) \tag{83}$$

$$\lambda_x(T) = \frac{\partial \Psi_2}{\partial x}(x(T)) - \lambda_\xi(T)\frac{\partial F}{\partial x} \tag{84}$$

$$\lambda_\xi(0) = 0 \tag{85}$$

$$\lambda_\xi(\tau^-) = \lambda_\xi(\tau^+) \tag{86}$$

leaves

$$\delta\hat{J}_\tau(\tau;\theta) = \int_0^T \left(\lambda_x\frac{\partial f_x}{\partial \tau}+\lambda_\xi\frac{\partial f_\xi}{\partial \tau}+\lambda_\sigma\frac{\partial f_\sigma}{\partial \tau}\right)\theta\, dt + \theta\mu\Bigg[f_x(x,\xi(\tau^-),\sigma,\tau) - f_x(x,\xi(\tau^+),\sigma,\tau)$$

$$+f_\xi(x,\xi(\tau^-),\sigma,\tau) - f_\xi(x,\xi(\tau^+),\sigma,\tau) + \frac{\partial G}{\partial x}\dot{x}(\tau)+\frac{\partial G}{\partial \sigma}\dot{\sigma}(\tau)\Bigg].$$

which should equal zero for all values of θ to achieve optimality. To keep track of this in the compact way presented in the theorem, let

$$\kappa_\tau = \int_t^T \left(\lambda_x\frac{\partial f_x}{\partial \tau}+\lambda_\xi\frac{\partial f_\xi}{\partial \tau}+\lambda_\sigma\frac{\partial f_\sigma}{\partial \tau}\right)dt. \tag{87}$$

Differentiating κ_τ with respect to time gives the expressions in Eqn. 77, which concludes the proof. ■

3.4 Application to Human Movement

Now consider the case where the template cost function, J_u, is given by Eqn. 8. As motivated in Chs. 1 and 2, this cost function can be thought of as an abstraction that captures the natural variation found in the way humans execute a given movement because it derives from an established kinesthetic theory of movement. In this section, a second cost function, J_π, is used to compare the output of the template cost function to some excerpt of real movement. The optimization of π, the variable containing the stylistic "knobs" found in

the template cost function as enumerated in Ch. 2, over the metric cost function produces the desired analysis of a given movement's *quality*.

For this problem, the weight matrices in J_u are constrained to be of the form

$$Q = qI_l$$

$$R = rI_m$$

$$P = pI_n$$

$$S = sI_l$$

where I_n is an $n \times n$ identity matrix. Thus, the optimization parameter can be written as

$$\pi = \begin{bmatrix} q \\ r \\ p \\ s \end{bmatrix}.$$

The metric cost function, which is comprised of $L(x, \rho)$ and $\Psi(x(T), \rho(T))$ as in Eqn. 45, is chosen to be

$$L(x, \rho) = \frac{1}{2}||y - \rho||^2 \tag{88}$$

$$\Psi(x(T), \rho(T)) = \frac{1}{2}||y - \rho||^2\Big|_T. \tag{89}$$

The state, input, output, and references are, respectively, $x = [\theta_1, \dot{\theta}_1, \theta_2, \dot{\theta}_2, ..., \theta_l, \dot{\theta}_l]'$[1], $u = [u_{\theta_1}, u_{\theta_2}, ..., u_{\theta_l}]'$, $y = [\theta_1, \theta_2, ..., \theta_l]'$, $\sigma = [\sigma_1, \sigma_2, ..., \sigma_l]'$, and $\rho = [\rho_1, \rho_2, ..., \rho_l]'$ where σ is chosen to be a linear interpolation between the endpoints of ρ (this is the nominal move) and ρ is the data to be classified. The system dynamics matrices A, B, and C are the canonical matrices for a completely controllable double integrator driftless system. The ρ that is employed is expected to be motion from a human dancer recorded on a motion capture system. The x, y, and z rotations for each bone represented in the skeleton extracted

[1] In practice, it may be most effective to perform this analysis on each joint angle individually as different joints may move with different qualities, a notion Laban described as simultaneity versus succession.

from raw data are shown with corresponding nominal movements as functions of time for the left leg of a human in Figure 19.

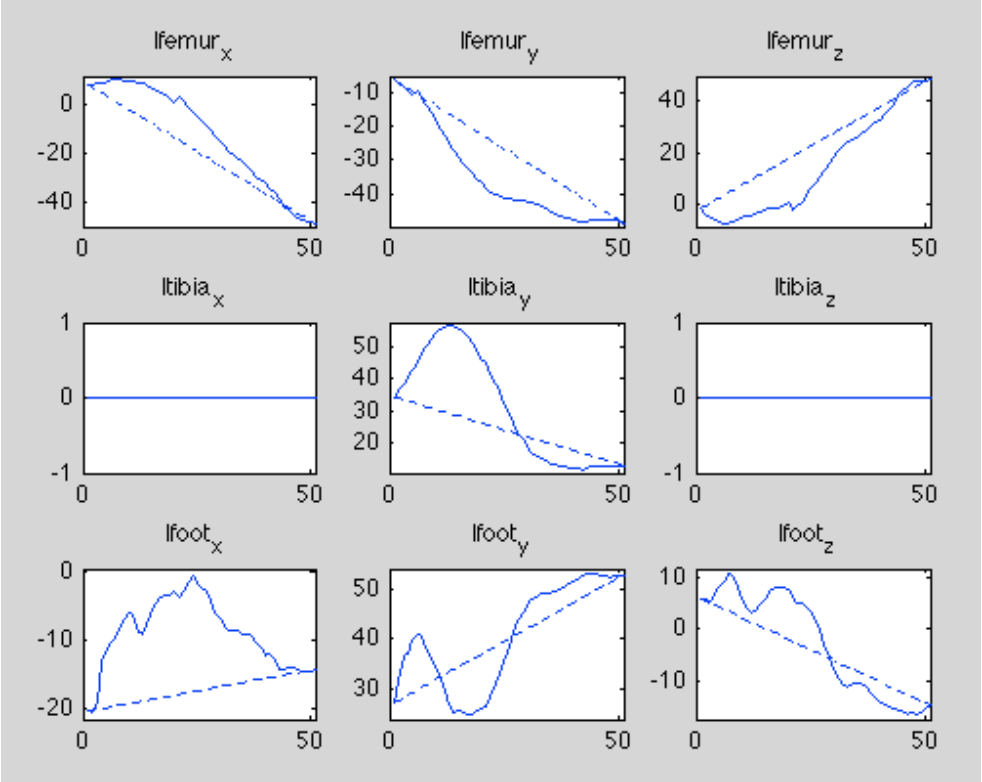


Figure 19. The plots in this figure are examples for ρ, in solid line, and the nominal reference σ, in dashed line. The subfigure titles are the names of skeleton bones in the Acclaim motion capture (.asf/.amc) file framework.

Note that now the notion of system dynamics needed here is a little unusual because the original state x ends up with three co-states. The co-state ξ is used to enforce the output to follow the trajectory specified by Q, R, P, and S. The co-states λ'_x and λ'_ξ are used to enforce π to be the optimal set of weights for the output y to approximate the empirical signal ρ.

Per the optimality conditions in Theorem 3.2.1, the following derivatives are necessary:

$$\frac{\partial \psi}{\partial x}(x(T)) = C'SCx(T) - C'S\sigma(T) \tag{90}$$

$$\frac{\partial^2 \psi}{\partial x^2}(x(T)) = C'SC \tag{91}$$

$$\frac{\partial L}{\partial x} = x'C'C - \rho'C \tag{92}$$

$$\frac{\partial \Psi}{\partial x}(x(T)) = x(T)'C'C - \rho(T)'C \tag{93}$$

$$\frac{\partial f_x}{\partial x} = A - B(R + B'PB)^{-1}B'PA \tag{94}$$

$$\frac{\partial f_x}{\partial \xi} = -B(R + B'PB)^{-1}B' \tag{95}$$

$$\frac{\partial f_\xi}{\partial x} = A'PB(R + B'PB)^{-1}B'PA - C'QC - A'PA \tag{96}$$

$$\frac{\partial f_\xi}{\partial \xi} = A'PB(R + B'PB)^{-1}B' - A' \tag{97}$$

$$\frac{\partial f_x}{\partial \pi} = \begin{bmatrix} \frac{\partial f_x}{\partial \pi_1} & \frac{\partial f_x}{\partial \pi_2} & \frac{\partial f_x}{\partial \pi_3} & \frac{\partial f_x}{\partial \pi_4} \end{bmatrix}$$

$$\frac{\partial f_x}{\partial \pi_1} = 0_{nx1} \tag{98}$$

$$\frac{\partial f_x}{\partial \pi_2} = BH^{-2}B'(\pi_3 Ax + \xi) \tag{99}$$

$$\frac{\partial f_x}{\partial \pi_3} = -BH^{-1}B'Ax + BB'BH^{-1}B'(\pi_3 Ax + \xi) \frac{\partial f_x}{\partial \pi_4} = 0_{nx1} \tag{100}$$

$$\frac{\partial f_\xi}{\partial \pi} = \begin{bmatrix} \frac{\partial f_\xi}{\partial \pi_1} & \frac{\partial f_\xi}{\partial \pi_2} & \frac{\partial f_\xi}{\partial \pi_3} & \frac{\partial f_\xi}{\partial \pi_4} \end{bmatrix}$$

$$\frac{\partial f_\xi}{\partial \pi_1} = C'(\sigma - Cx) \tag{101}$$

$$\frac{\partial f_\xi}{\partial \pi_2} = -\pi_3 A'BH^{-2}B'(\pi_3 Ax + \xi) \tag{102}$$

$$\frac{\partial f_\xi}{\partial \pi_3} = (2\pi_2 A'BH^{-1}B'A - A'A)x + A'BH^{-1}B'(\pi_3 Ax + \xi)\pi_3 A'BB'BH^{-2}B'(\pi_3 Ax + \xi) \tag{103}$$

$$\frac{\partial f_\xi}{\partial \pi_4} = 0_{nx1} \tag{104}$$

where $H = R + B'PB$.

Now, to assemble the Hamiltonian dynamics (x with all three of the costates), let $w_x =$

λ'_x and $w_\xi = \lambda'_\xi$. The dynamics are then given by

$$\begin{bmatrix} \dot{x} \\ \dot{\xi} \\ \dot{w}_x \\ \dot{w}_\xi \end{bmatrix} = \begin{bmatrix} & M & & 0 & 0 \\ & & & 0 & 0 \\ -C'C & 0 & & -M' & \\ 0 & 0 & & & \end{bmatrix} \begin{bmatrix} x \\ \xi \\ w_x \\ w_\xi \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ C'Q & 0 \\ 0 & C' \\ 0 & 0 \end{bmatrix} \begin{bmatrix} r_\theta \\ \rho \end{bmatrix} \tag{105}$$

where the entries for M describe the dynamics of $[x, \xi]'$ under Eqns. 8 and 13. The entries of M are given by

$$M_{11} = A - B(R + B'PB)^{-1}B'PA \tag{106}$$

$$M_{12} = -B(R + B'PB)^{-1}B' \tag{107}$$

$$M_{21} = A'PB(R + B'PB)^{-1}B'PA - C'QC - A'PA \tag{108}$$

$$M_{22} = A'PB(R + B'PB)^{-1}B' - A'. \tag{109}$$

Setting $z = [x, \xi, w_x, w_\xi]'$ and $\zeta = [\sigma, \rho]'$ allows the dynamics to be written in the following form:

$$\dot{z} = \mathcal{M}z + \mathcal{N}\zeta. \tag{110}$$

The solution for $z(T)$ can be written as

$$\begin{aligned} z(T) &= e^{\mathcal{M}T}z_0 + \int_0^T e^{\mathcal{M}(T-t)}\mathcal{N}\zeta(t)dt \\ &= \Phi z_0 + \eta. \end{aligned} \tag{111}$$

Then, let

$$\Phi = \begin{bmatrix} \Phi_{11} & \Phi_{12} & \Phi_{13} & \Phi_{14} \\ \Phi_{21} & \Phi_{22} & \Phi_{23} & \Phi_{24} \\ \Phi_{31} & \Phi_{32} & \Phi_{33} & \Phi_{34} \\ \Phi_{41} & \Phi_{42} & \Phi_{43} & \Phi_{44} \end{bmatrix} \tag{112}$$

$$\eta = \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix}. \tag{113}$$

Next, plug in the following boundary conditions (along with x_0):

$$\xi(T) = C'SCx(T) - C'S\sigma(T) \tag{114}$$

$$w_x(T) = C'SCw_\xi(T) - C'Cx(T) + C'\rho(T) \tag{115}$$

$$w_{\xi 0} = 0. \tag{116}$$

Rearranging the resulting expression gives

$$\begin{bmatrix} x(T) \\ \xi(0) \\ w_x(0) \\ w_\xi(T) \end{bmatrix} = \mathcal{A}^{-1}\left(\begin{bmatrix} \Phi_{11} \\ \Phi_{21} \\ \Phi_{31} \\ \Phi_{41} \end{bmatrix} x_0 + \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix} + \mathcal{B}\begin{bmatrix} \sigma(T) \\ \rho(T) \end{bmatrix}\right), \tag{117}$$

where

$$\mathcal{A} = \begin{bmatrix} I_n & -\Phi_{12} & -\Phi_{13} & 0 \\ C'SC & -\Phi_{22} & -\Phi_{23} & 0 \\ C'C & -\Phi_{32} & -\Phi_{33} & -C'SC \\ 0 & -\Phi_{42} & -\Phi_{43} & I_n \end{bmatrix} \tag{118}$$

$$\mathcal{B} = \begin{bmatrix} 0 & 0 \\ C'S & 0 \\ 0 & C' \\ 0 & 0 \end{bmatrix} \tag{119}$$

and where the invertibility of $\mathcal{A}$ is guaranteed by the complete controllability of the system. Thus, the desired dynamics, $\dot{z}$ and z_0, where y recreates ρ optimally according to the structure of J_π as best as the structure of J_u will allow are given.

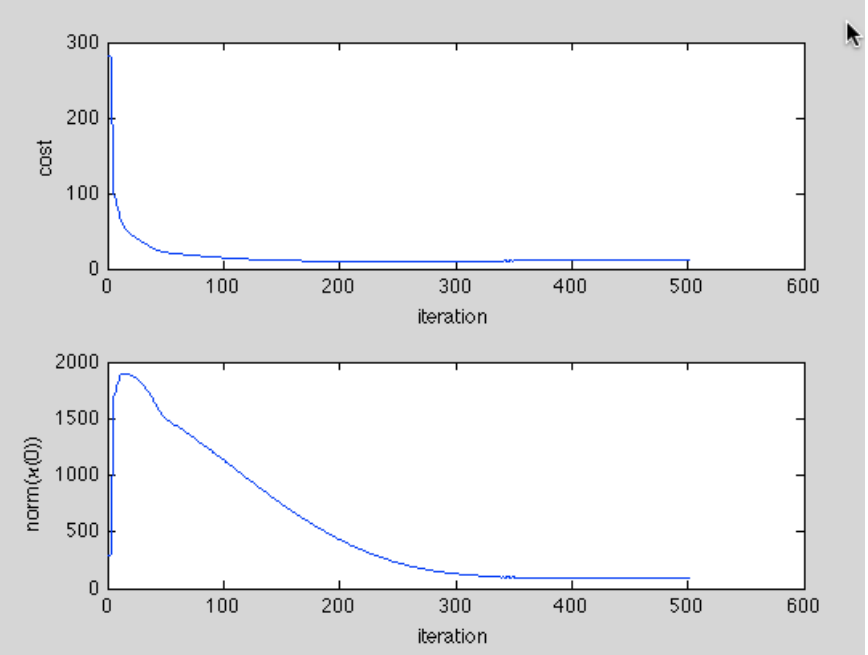


Figure 20. The lower figure shows that the solution satisfies the optimality condition in Eq. 47, and the upper figure shows the corresponding converging cost function.

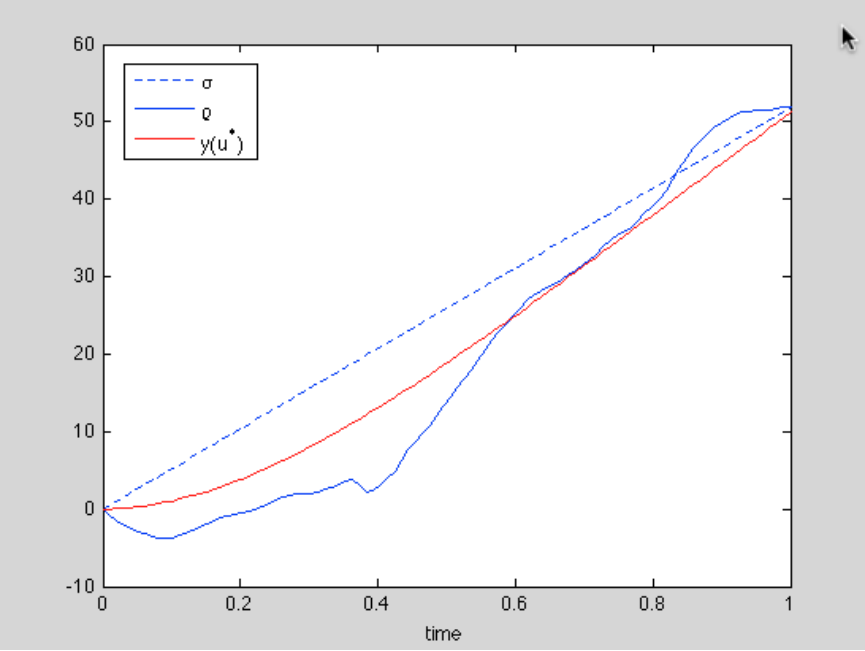


Figure 21. This figure shows the result of the simulation: an output y which tracks data of a subject's femur as closely as possible using the mapping between Laban's motion factors and a cost function.

We now implement this setup on a single joint angle, namely, motion of the left femur in the z-direction that has been smoothed with 3 passes of a sliding average of width 2. This

results in a 4×4 system, which is initialized to match ρ at $t = 0$ (i.e., $x_0 \approx [\rho(0); \rho'(0)]$) and normalized to start at zero and last for one unit of time (i.e., $\rho(0) = 0$ and $T = 1$). Performing gradient descent with Armijo step size on π (using $\alpha = \beta = 0.5$), we get the converging cost (shown in Fig. 20) and an output y that mimics ρ as shown in Fig. 21. The the following values where computed for π: $q = 1.53$, $r = 0.018$, $p = 0.0001$, and $s \triangleq 1$). These values for the optimization parameter π prescribe a relative location for the trajectory ρ in Laban's dynamosphere (Fig. 4); this motion is then considered *direct*, *light*, and *sudden*.

This process can, likewise, be implemented on each joint angle. This was done for the full nine-dimensions of the motion captured leg (with the same parameters as the one-dimensional example). Figure 22 shows the results of this simulation. The simulated motion was animated as shown in Fig. 23 in order to demonstrate how well these approximations are able to capture the essence of the original movement and *recreate* that movement.

$\diamond$

In this chapter we have presented a method for motion analysis using inverse optimal control in concert with the choreographic abstractions for movement *generation* from the previous chapter. This method for *interpretation* can be thought of as a way to close the loop, so to speak, on the movement generation scheme presented in this thesis. A conceptual graphic is provided in Fig. 24.

With the inverse problems solved here, the *quality* of movement can be determined and used to segment longer motion phrases into *primary* movements. The extracted quality takes the form of four matrices that can generate movement as described in Sec. 2.3. The primary movements compose an automaton as outlined in Sec. 2.1. Techniques to determine an automaton's structure based on samples of its movements were reviewed in Sec. 1.1, which can recreate the result of Sec. 2.2. Thus, the *stylistic* parameters from Ch. 2 can be extracted from real data and subsequently implemented on a robot (as in Sec. 2.4).

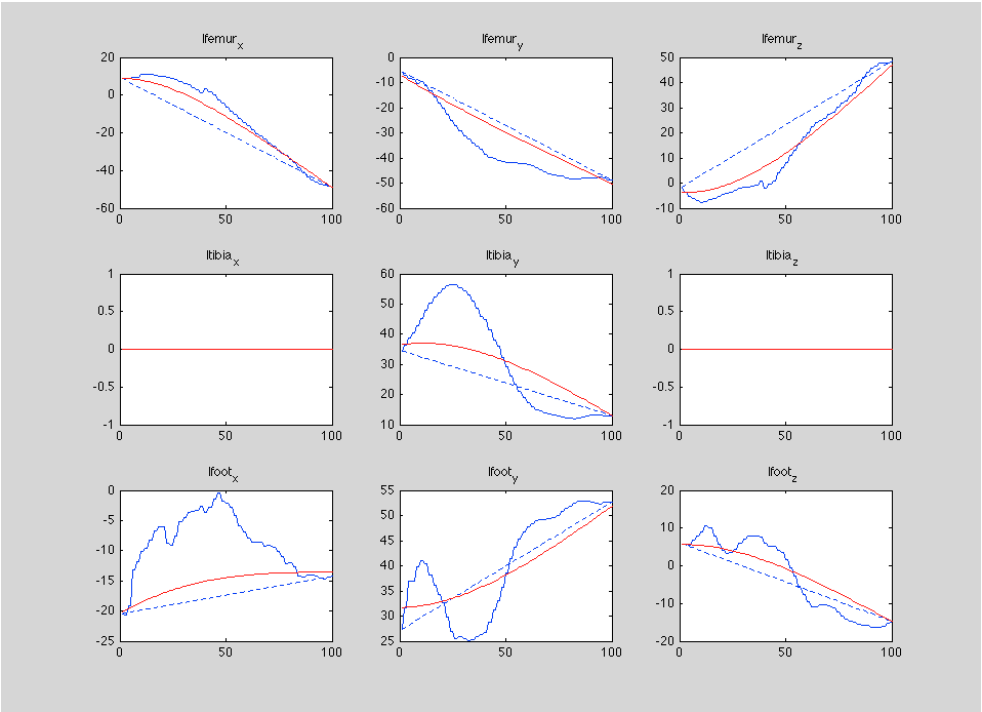


Figure 22. This figure shows the result of the simulation: an output y (in red) which tracks data of a subject's femur (in solid blue) as closely as possible using the mapping between Laban's motion factors and a cost function.

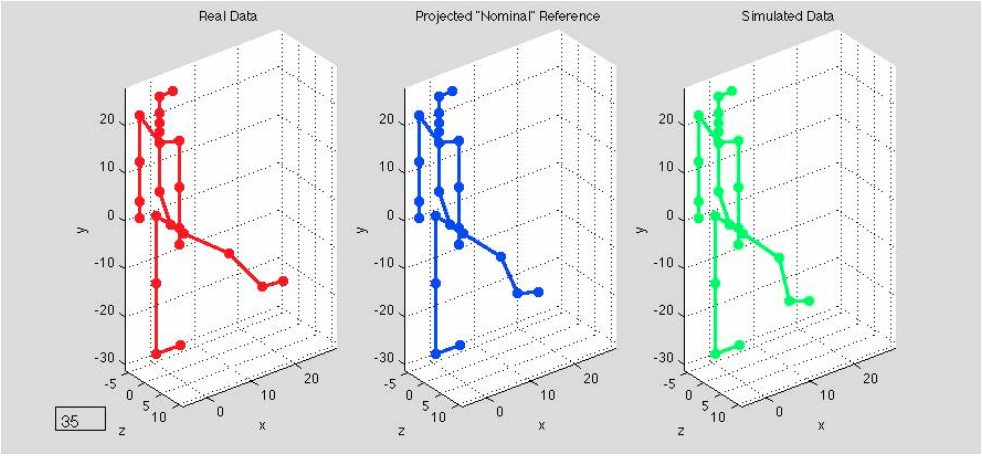


Figure 23. This figure is a snapshot from an animation of the original raw data, ρ, the nominal move, σ, and the optimal output, y.

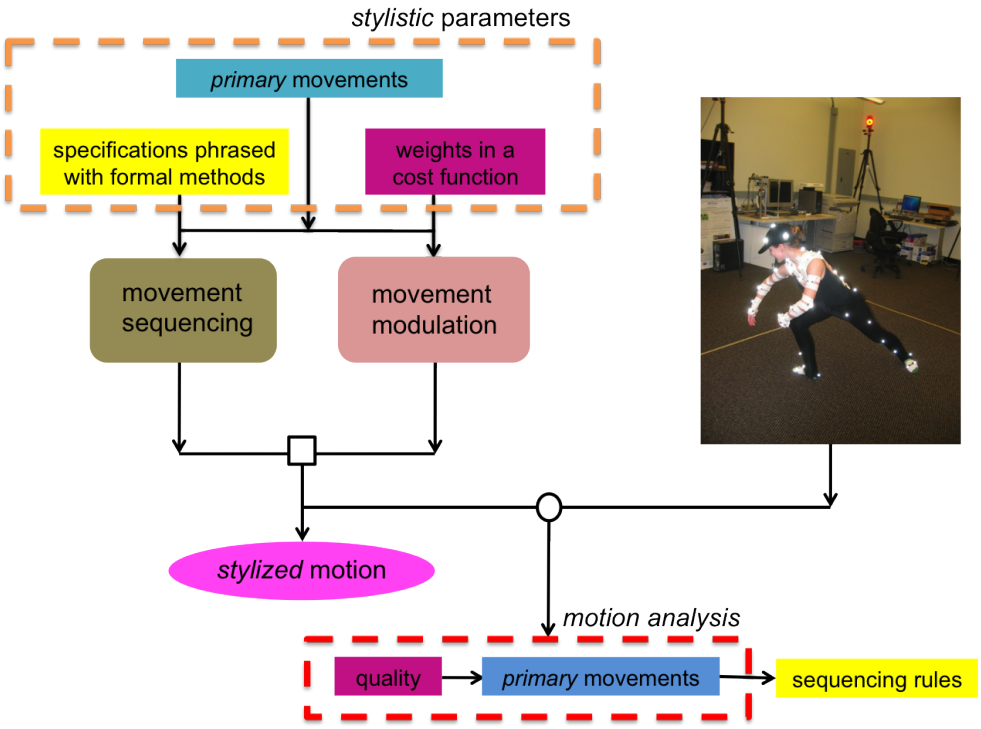


Figure 24. The figure shows a conceptual overview of how the inverse optimal control problems presented in this chapter "close the loop" on the generation scheme presented in Ch. 2, i.e., the three stylistic parameters which feed the movement generation can be extracted from real movement. The parameters circled in red are directly accounted for in this chapter, and techniques to extract the third parameter from the first two have been reviewed in this dissertation. Photo snapped by Amir Rahmani.

CHAPTER 4

AUTOMATON: AN APPLICATION IN ROBOTICS AND DANCE PERFORMANCE

The topic of this thesis might, most simply, be considered "dancing robots." Indeed, a method for taking human movement and animating some version of it on a robot has been provided. This is accomplished not from a direct joint angle mapping from one highly articulated skeleton (the human) to a deficient one (the robot). Instead, the method extracts key stylistic features of human movement (recovered via motion capture). Thus, novel sequences, which are in the same *style*, can be produced rather than simple imitation. However, such a narrative, casting the work as simply a method for some kind of mimicry, short changes the diverse applications this thesis can have. The aim is to use engineering and art to tease out an understanding of human behavior, which is key to applications in robotics, human-centered technology, and the performing arts.

On the part of robotics, this thesis has motivated a vein of research complementary to typical task-based robotic motion generation. Functional tasks have a straightforward interpretation for metrics of success, efficiency, and completion and are, thus, implementable on robots in a variety of shapes. However, robots that are shaped like humans have a greater vocabulary of movement to aspire to as much of human movement serves an as yet unknown function. Dance is my favorite such example, but movement aesthetics and *style* have to be considered in other application areas as well. To cite two popular examples of how a fully autonomous humanoid robot might first appear in complex roles: in-home hospice robots should move as though they are gentle and caring and museum tour guide robots should have a brisk air of authority about their movements. Thus, a quantitative parameterization of high-level aspects of human movement is a necessary component for human-shaped robots that interact with humans.

Technology relies more and more on an interpretation of human movement. Human-centered technology has moved from vending machine style interfaces where users need to figure out the function of various buttons to operate a machine to iPhones where much of the interaction takes place based on the machine's interpretation of the motion of the human fingers. Following this trend, Microsoft's Kinect extracts a human skeleton using computer vision techniques and aims to incorporate full-body motions into the Xbox entertainment system. Being able to interpret the style and other high-level aspects of a user's movement can enhance these efforts. Similarly, physical rehabilitation is an activity that can benefit from the metrics on human movement developed here. For example, the measure of quality introduced in Ch. 3 takes into account a motion's entire trajectory, and the cost function can be tailored to better produce outputs that mimic specific motions associated with various rehabilitative tasks.

The third broad area this work can interface with is art. The analysis and creation of art can, and has, benefited from quantitative tools as discussed in Sec. 1.3.3. On the part of analysis, being able to generate – and even *measure* – stylized movement from three parameters allows for movement from different dancers, choreographers, genres, and sections of the same piece to be compared quantitatively. Such analysis can bolster the body of academic work which *qualitatively* discusses the evolution of and connection between various genres of dance and the differences between the work of two choreographers.

On the part of creation, to back up claims that this thesis has a tight interconnection with the arts and to provide an artistic outlet for the concepts herein, an original work of contemporary dance was choreographed. What follows is a discussion of how principles from this thesis were included in the choreography of the show, *Automaton*.

4.1 Expectations and Reactions: Programmed Behavior

The robot used in the show is immediately engaging to audiences. The robotic figure calls to mind experiences of humanoids in pop culture, film, and fantasy. This presents unique

Figure 25. The four dancers in *Automaton* along with the humanoid robot used in the show. Photo by Christian Moreno.

challenges and opportunities from a choreographic perspective. Because the robot conjures up the audience member's impressions about robots, one of the opportunities is to challenge these preconceived ideas. In particular, I believe most people feel quite differently about a piece of technology that is human-shaped (for example, see Fig. 26). While few people would personify their ice maker or iPhone, everyone immediately asks what the robot's name is (and sometimes gives it one themselves).

Thus, one of the choreographic goals of the piece is to blur this boundary between human-shaped and non-human-shaped technology as well as the boundary between humans and programmed devices. In particular, at the interface where humans interact with technology, their own movements mimic the robotic qualities associated with programmed devices [75]. The confluence of a discrete description of movement, a robot to implement it on, and trained dancers who could also execute that description with a high level of accuracy made *Automaton* a perfect setting to explore the interaction of natural human movement with movement that interfaces with a programmed device.

In the third movement of the piece, two dancers perform with the NAO humanoid robot.

Figure 26. Two dancers depict the robot being aggressively confronted. The extent to which this image conjures up empathy for a plastic container for motors, circuit boards, and a processor is surprising until its human shape is taken into account. Photo by Christian Moreno.

This trio explores the idea that there is some level of expected behavior and reaction from any being or device with which a person interacts. For example, the dancers engage in a game of patty cake. This is a set of movements, a series of alternating hand claps between two people, that are essentially pre-programmed in many children (at least in Western culture). In the trio one dancer breaks the pattern of the movement – in a sense malfunctioning – and causes his partner to have a moment of extreme confusion. The confused partner then turns to the robot and makes it his new patty cake partner. But the robot also breaks the traditional pattern of patty cake. Here both the human and the robot "malfunctioned"– seemingly on purpose – or started playing a new game, causing the human partner, in a twist of irony, to experience a situation he seemed not to know how to navigate, a plight that typically affects robotic systems– particularly in dynamic environments.

This vignette displays a reversal of roles: instead of humans evolving under unknown patterns where next movements are not predictable, they engaged in a child's game with preset movements. Because the robot can easily mimic such preprogrammed movements, it can play along too. But what happens to a human when the rules of this game are slightly different? In this case (and in many others), the human malfunctions just like a robot presented with an unpredicted situation. This is analogous to an autonomous car that was designed for operation in the U.S., where cars drive on the right side of the road, that has just been dropped in the U.K., where cars drive on the left side of the road – it would crash instantly. This metaphor mirrors aspects of human behavior as well.

4.2 Rule-based Movement Generation

Choreographers are constantly seeking sources of inspiration and techniques to generate new movement. Often this involves using measures external to their bodies to generate new constraints for their creative problems – à la Merce Cunningham and his dice. The tools in this thesis afford such an opportunity: style of motion may be *programmed* through the

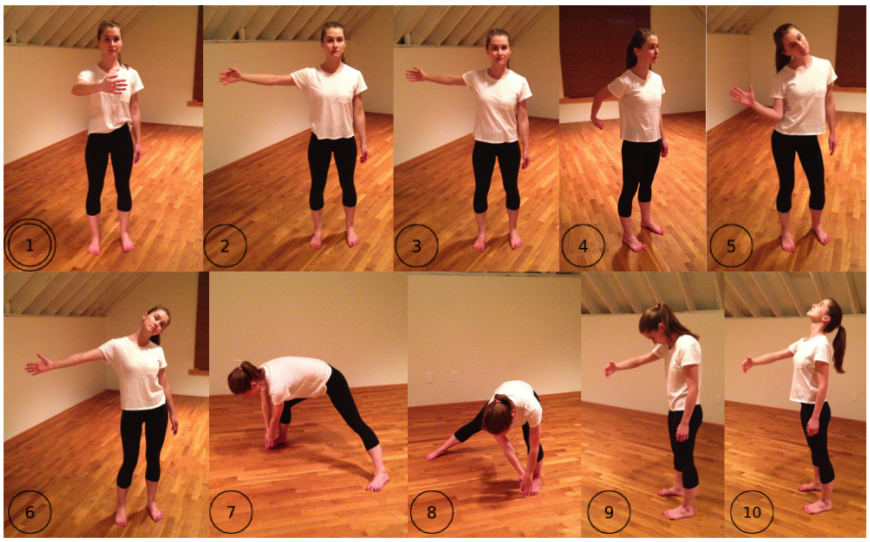


Figure 27. The underlying poses for the five styles of movement performed during *Automaton*. Photos snapped by Jessica Portugal.

automata-theoretic description of style[1] presented in Ch. 2.

Developing movement "styles" using automata inherits the same advantage that automata afford to the technological objects which they are traditionally used to describe. Namely, they allow for discrete changes in state. In the case of a remote control, this means that when the highest channel is reached, the TV interprets the "up" signal from the remote as a command to go *down* to the first or lowest channel. This is to some degree inconsistent and would be extremely inconvenient to capture this desired behavior using continuous mathematics. In the case of dancing, the programmatic advantage is that moving between body poses does not have to be governed by the natural movement patterns of the choreographer, but instead offers an opportunity to produce unusual sequences as drawing arrows between states is a completely disembodied task – just as Cunningham's technique of rolling the dice.

Furthermore, this method allows for a uniquely quantitative comparison between generated styles of movement. Figures 27-31 show the rule sets for each dancer. Just as the movement patterns *look* different to watch, these automata have different structures. For

[1] Audience members were given an experiential tutorial in automata via the folding of their programs. See Appendix A.

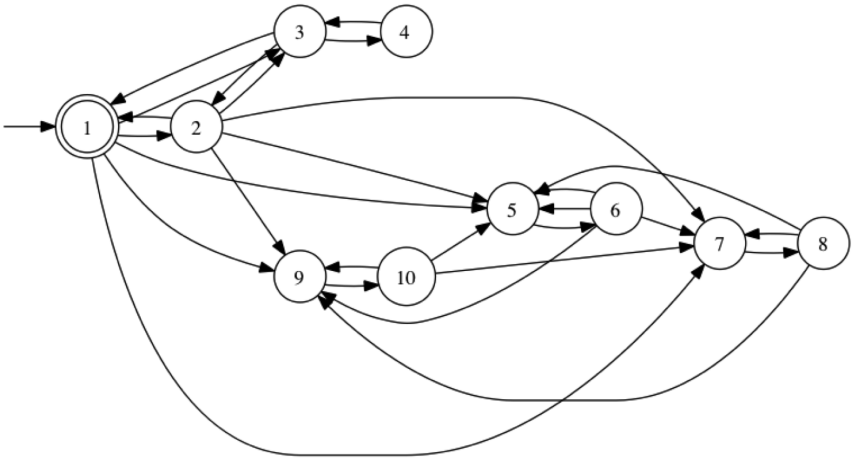


Figure 28. The rule set for the dancer Helen and the humanoid robot.

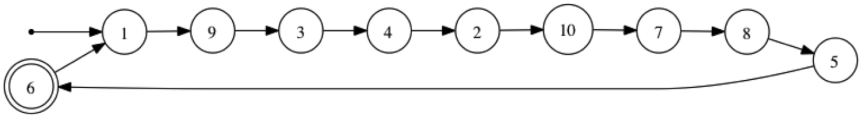


Figure 29. The rule set for the dancer Camille.

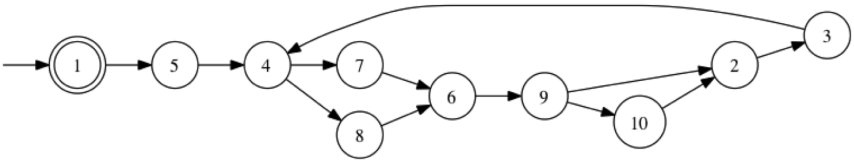


Figure 30. The rule set for the dancer Alex.

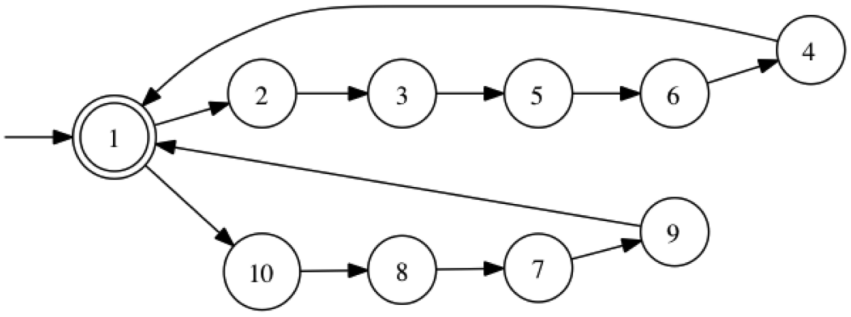


Figure 31. The rule set for the dancer Erik.

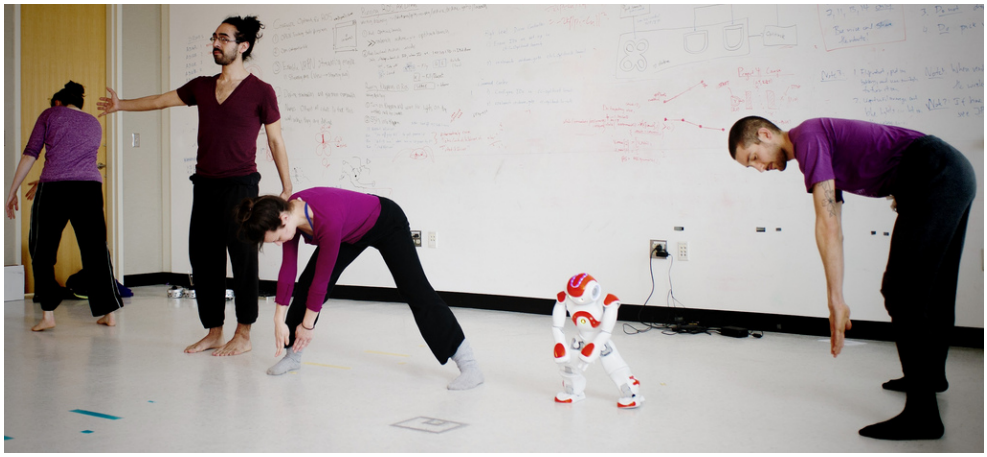


Figure 32. Five styles of movement are performed during *Automaton*. Photo by Rob Felt.

example, one automaton is quite linear and represents the same sequence of movement repeated over and over; whereas others present more options and allow for repeated movement.

Moreover, these descriptions are amenable to robotics, and the styles of movement may be *automated*. Thus, the rule-set pictured in Fig. 28, is also animated on the NAO humanoid robot. By employing different rule sets on each dancer and the robot, the audience is able to see different automata animated, live. A snapshot from this section of the piece is shown in Fig. 32.

4.3 What did the audience see?: A Human Study

The final piece of this thesis will be to provide an initial step toward validating the goal of endowing a robot with a consistent style of movement. The sequencing framework in Ch. 2 is used in *Automaton* as described in the previous section. Audience members were then polled after the performances (held on April 6 and 13, 2013) for their impressions of the robot's movement. Appendix B contains the Institutional Review Board (IRB) forms related to this study.

Between the two showings, about 100 people saw *Automaton*, and 28 audience members volunteered to fill out a survey collecting their impressions of the robot's movement.

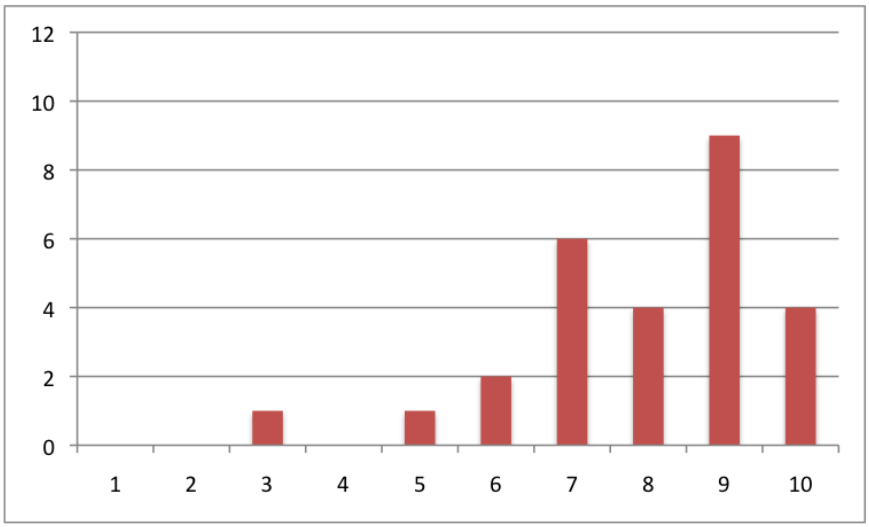


Figure 33. A histogram of survey responses collected between both showings of *Automaton* completed by volunteers from the audience. The question shown here was stated as follows: *Please rate your impression of the overall consistency of the robots movement. A rating of 1 indicates that it seemed to be moving completely randomly. A rating of 10 indicates that it was moving in a very specific way*. Each column counts how many subjects selected 1, 2, 3, ..., 10, respectively, for this question.

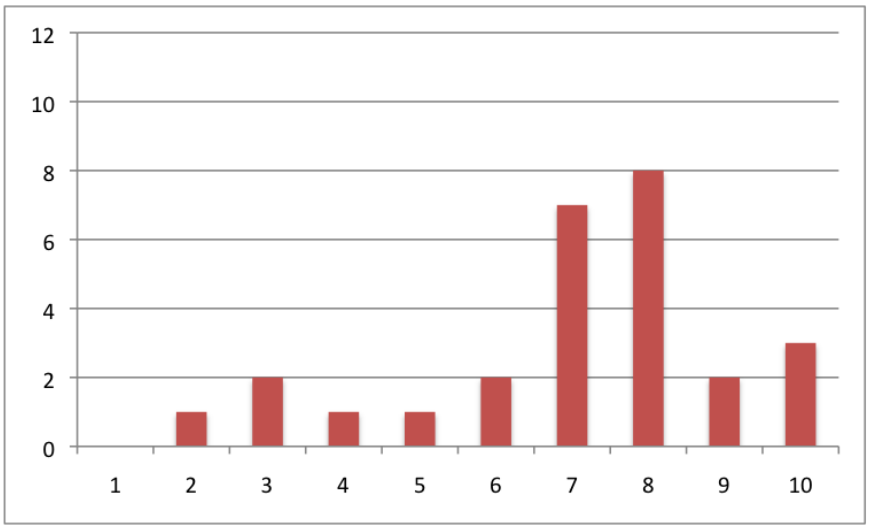


Figure 34. A histogram of survey responses collected between both showings of *Automaton* completed by volunteers from the audience. The question shown here was stated as follows: *Please rate your impression of the style of the robots movement. A rating of 1 indicates that it seemed to be moving without style. A rating of 10 indicates that it was moving in a very stylized way*. Each column counts how many subjects selected 1, 2, 3, ..., 10, respectively, for this question.

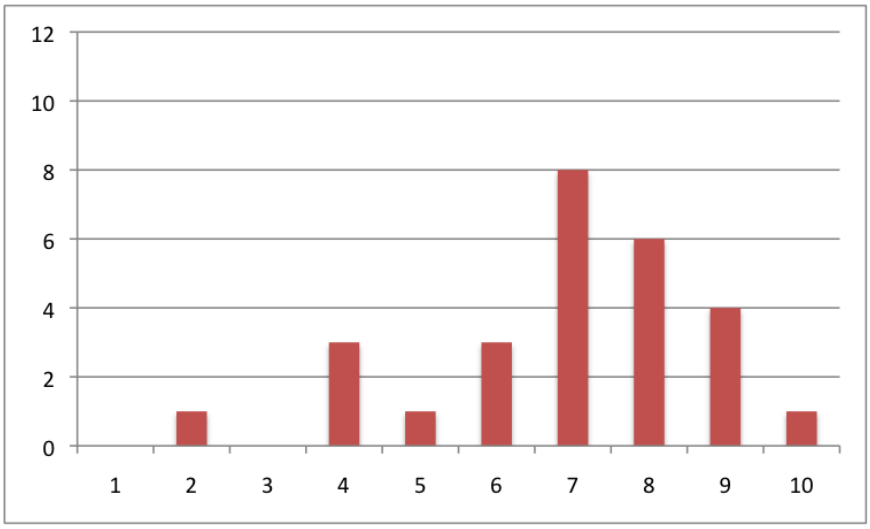


Figure 35. A histogram of survey responses collected between both showings of *Automaton* completed by volunteers from the audience. The question shown here was stated as follows: *Please rate your impression of the virtuosity of the robots movement. A rating of 1 indicates that it seemed to be very bad at moving. A rating of 10 indicates that it was very good at moving*. Each column counts how many subjects selected 1, 2, 3, ..., 10, respectively, for this question.

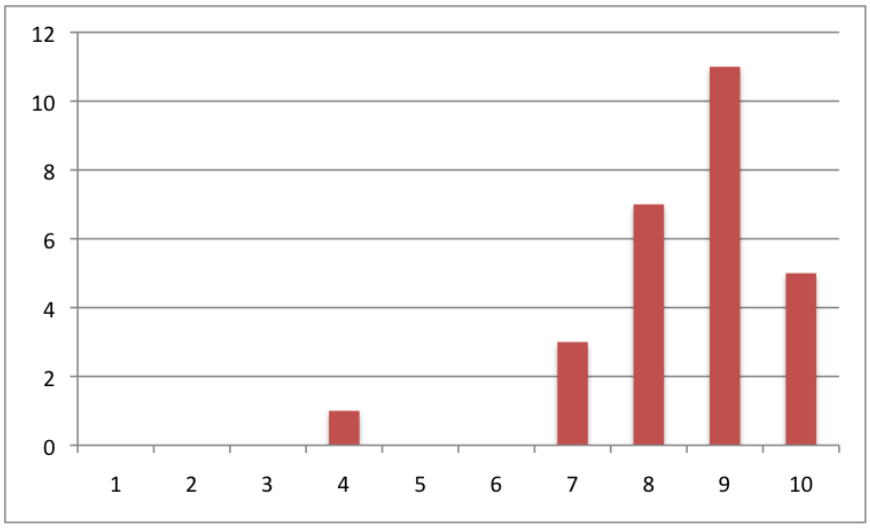


Figure 36. A histogram of survey responses collected between both showings of *Automaton* completed by volunteers from the audience. The question shown here was stated as follows: *Please rate your overall enjoyment of the robots movement. A rating of 1 indicates that you really did not enjoy it. A rating of 10 indicates that you very much enjoyed it*. Each column counts how many subjects selected 1, 2, 3, ..., 10, respectively, for this question.

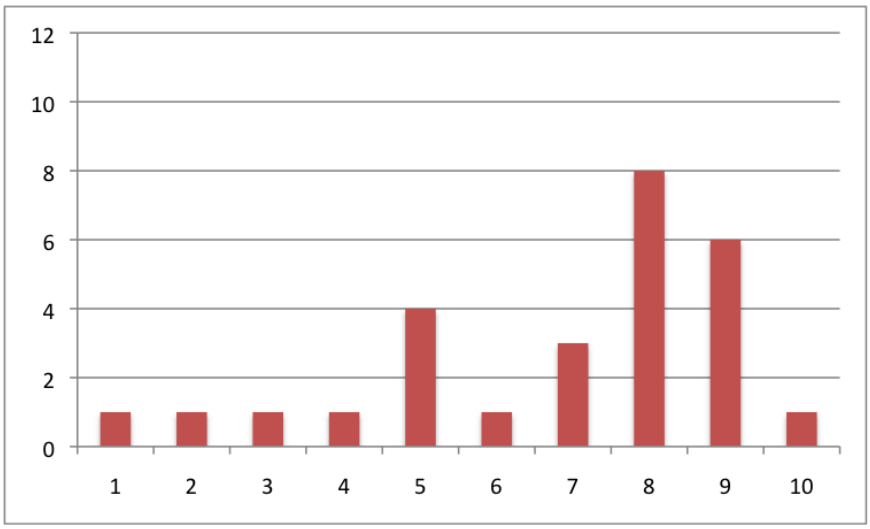


Figure 37. A histogram of survey responses collected between both showings of *Automaton* completed by volunteers from the audience. The question shown here was stated as follows: *Please rate your feelings about the robots movement. A rating of 1 indicates that you felt very little about the movement. A rating of 10 indicates that you were very moved by the movement*. Each column counts how many subjects selected 1, 2, 3, ..., 10, respectively, for this question.

Figures 33-37 show the results of the responses and the averaged values for each question. (The table in Appendix B, Fig. 46, contains the complete responses.) The questions aimed to get at the viewer's impression of a colloquial notion of "style" as it applied to the movement of the robot in the show. The study subjects were asked if their impressions of the movement were that the movement was "specific," "stylized," "virtuosic," "enjoyable," and whether they "felt" anything at all about the movement. These questions were phrased broadly and were not meant to define a subjective or psychological notion of "style." Instead, the questions simply probed whether the audience believed that something consistent was underlying the robot's movement.

The responses indicate the viewers polled did experience the robot's movement as something other than random motions set to music. In particular, the responses to the first two questions support this claim. Question 1 polled the viewer's sense that the movement the robot was executing was very "specific." The average of almost 8 on a scale of 1 to 10, with 1 corresponding to random motion and 10 corresponding to very specific motion, indicates the audience did feel the robot's movement was "specific." Question 2 was similar

except phrased using the term "style"; the average rating of about 7 out of 10 indicates that the viewers polled found the robot's motion to be more "stylized" than not. The responses to the other questions (Figs. 35-37) also indicate the audience sensed some high-level design of the robot's movement. Thus, this initial study indicates we have achieved robotic movement that is *perceived* to be of a certain style.

◊

This chapter reviewed applications for the work in Chs. 2 and 3 and described the performative component of this thesis. The work in this thesis provides a window into the phenomenon of creative movement. An understanding of phenomena such as why – or even how – humans move the way they do is key for the automated systems of the future. In this sense, *Automaton* is a fitting presentation for the applications of this thesis.

CHAPTER 5

EXTENSIONS: BEYOND STYLE

This chapter presents a dynamic clustering method that assigns data to appropriate clusters for data with a time varying component. The segmentation method from the previous chapter is then generalized in order to automatically scale the window over which interactions are considered thus extracting time scale information from the data. These time scales are restricted to be of a certain form in order to make statements about persistent patterns in the data.

There are many available clustering methods, but we review Zelnik-Manor and Perona's self-tuning algorithm [76] leaving the details to the next section as we build directly onto it in there. This algorithm treats data as nodes in a graph and then uses notions of connectedness via spectral properties of the graph Laplacian to determine the data's structure. Clusters are defined as closely related communities within a fully connected graph where edge weights indicate relatedness. A measure of assignment quality determines the optimal cluster number – automatically.

In these examples, data is treated as a single, static set. Further methods are needed to account for the inherent time-varying nature of these data sets – should poses that occur far from each other in time be treated the same as those temporally close? – and even changing movement styles over time. For example, Lahiri uses repeated patterns in clusters assigned at periodic time points as a way to identify resilient clusters [77].

5.1 Dynamic Clustering

The method presented in [78] addresses clustering for time-varying data, building on [76]. Many analyses of data look for patterns that accrue over time. Biologists studying the interactions of social animals often consider the duration and repetition of an animals proximity with other of their species. And, as discussed in the literature review, clustering has been

an important tool in analyzing motion capture data. These are two examples of data that is inherently time-varying while clustering methods may not be taking that into account.

A key feature of the algorithm is that it retains the desirable quality of all spectral clustering techniques in that it partitions data without presuming any direct scale to be characteristic of the data *a priori*. In fact, the final algorithm has only two parameters that can be tuned: a weighting factor on the similarity measure and the length of the time interval over which the dynamic algorithm is applied. Both parameters are robust to the type of changes that will occur as the data evolves in time; this result is theoretically pleasing and has also proven itself useful in practice.

We begin with a finite dataset with n entries that form the set $\mathcal{V}$; hence $n = |\mathcal{V}|$. Onto this dataset we may apply a notion of abstract distance, $\delta : \mathcal{V} \times \mathcal{V} \to \mathcal{R}$, that is defined between every pair of points (i, j) in the dataset $\mathcal{V}$ (let $\delta(i, j) = \delta_{ij}$). This distance is an alternate expression for the similarity measure between data points.

The structure that δ imposes over the data can be encoded in a matrix by defining a complete, weighted graph over the data points, $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$, where $\mathcal{E}$ is the edge set of $\mathcal{G}$ and the edge weights, w_{ij}, are a monotonically decreasing function of the distance measure between data points, δ_{ij}. For a δ that corresponds to Euclidean distance, points which are closest to each other have the biggest weights, are most similar in terms of location, and, thus, will be more likely to be in the same cluster.

The normalized weighted adjacency matrix is defined as

$$L_N = \Delta^{-\frac{1}{2}} A \Delta^{-\frac{1}{2}} = I - \Delta^{-\frac{1}{2}} L \Delta^{-\frac{1}{2}}, \tag{120}$$

where the graph Laplacian is $L = \Delta - A$; $\Delta_{n\times n}$ is the diagonal degree matrix which is defined as

$$\Delta_{ij} = \begin{cases} deg(v_i) & i = j \\ 0 & i \neq j \end{cases}, \tag{121}$$

for our completely connected graph this is simply the identity matrix scaled by the number of nodes in $\mathcal{V}$, $n \cdot \mathbb{I}_{nxn}$; and $A_{n\times n}$ is the weighted adjacency matrix which encodes the

distances between nodes and is defined as

$$A_{ij} = \begin{cases} w_{ij} & (v_i, v_j) \in E \\ 0 & o.w. \end{cases} . \tag{122}$$

Thus, L_N is a direct variant of the the graph Laplacian; we will use the spectral properties of this matrix to perform the clustering.

This matrix L_N contains the complete connectivity information of $\mathcal{G}$ and the similarity metric defined by δ. Since L is symmetric and positive semi-definite, we know that L_N is also symmetric and has eigenvalues belonging to the interval $(-\infty, 1]$. Furthermore, for the completely connected graph, the largest eigenvalue of L_N equals 1 and its associated eigenvector is a vector of ones (perhaps scaled by a constant), $\mathbf{1} \in \mathbb{R}^n$.

For further intuition about this matrix, consider an ideally clustered graph with c clusters where there is an edge-weight of zero between nodes which do not belong to the same cluster. The matrix L_N is now in block diagonal form with c blocks as in Eq. 123.

$$L_N = \begin{bmatrix} L_{N_1} & & & \\ & L_{N_2} & & \\ & & \ddots & \\ & & & L_{N_c} \end{bmatrix} \tag{123}$$

In this disconnected case, L_N has c eigenvalues equal to one. Correspondingly, the c eigenvectors associated with these characteristic eigenvalues define the clustering: if row i of eigenvector j contains a nonzero entry, then node i is in cluster j for $i = \{1, ..., n\}$ and $j = \{1, ..., c\}$. Every other entry of the eigenvector will be a zero and none of the cluster assignments are multiply defined. This structure is shown by:

$$V_{n \times c} = \begin{bmatrix} \mathbf{1}_1 & 0 & & 0 \\ 0 & \mathbf{1}_2 & & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & & \mathbf{1}_c \end{bmatrix}, \tag{124}$$

where the c eigenvectors corresponding to eigenvalues of 1 have been concatenated into a matrix, V and $\mathbf{1}_i$ is a vector of ones with dimension equal to the number of nodes in the ith connected component . We can also think of this ideal case as one where the eigenvectors have no inter-dependence when expressed in terms of the n dimensional space defined by the data points (and nodes of our constructed graph); that is, each basis vector is used once and only once to describe the eigenvectors.

Next, consider our graph $\mathcal{G}$, where $\mathcal{V} = \{1, ..., n\}$ are the nodes of the graph defined by n data points; $\mathcal{E} = \{1, ..., m\}$ is the edge set of a complete graph, K_n (where $m = \frac{n(n-1)}{2}$); and w_{ij} is a weight on edge $(v_i, v_j) \in \mathcal{E}$ that defines a monotonically decreasing function of distance (or monotonically increasing function of similarity) $\delta : \mathcal{E} \rightarrow \mathbb{R}$ between the two nodes.

There is one connected component in this graph; hence one eigenvalue of L_N is equal to 1. A first look at the eigenvectors in the form of Eq. 124 shows a single vector of ones describing one uninformative cluster; however, the data may still have underlying clusters. These clusters are determined by groups of data points which are close to each other (where closeness is defined by δ since some edge weights will be small relative to the others).

Since the eigenvalues and eigenvectors vary continuously with modifications to the matrix entries, the number of relatively large eigenvalues of L_N (close to one) gives a good, initial guess of the number of clusters, c, in the dataset. Again, only one eigenvalue of L_N equals 1 (since there is one connected component in a complete graph), but the other eigenvalues corresponding to highly clustered regions of the graph, which have now drifted, remain close to 1.

Likewise, the eigenvectors that before gave a definite clustering of the graph's clusters are now skewed from the ideal case. The effect of this skew can be reversed or reduced by rotating the new eigenvectors until they are close to the ideal case again. As before, we would like to express these eigenvectors with the smallest possible inter-dependence in the n-dimensional space of the dataset.

Let the first c eigenvectors of L_N compose $V \in \mathcal{R}^{n\times c}$. Rotating V such that each row of the new, rotated matrix Z has only one non-zero entry provides the cluster assignment for each node. Defining this rotation by a matrix $R(\Theta)_{c\times c}$, we want to find a vector $\Theta \in \mathbb{R}^{c(c-1)/2}$ where each entry is an angle in $[-\frac{\pi}{2}, \frac{\pi}{2})$ such that it minimizes a cost function,

$$\min_{\Theta} \mathcal{J} = \sum_{i=1}^{n}\sum_{j=1}^{c}\left(\frac{Z_{ij}}{M_i}\right)^2, \tag{125}$$

subject to the constraint

$$Z_{n\times c} = V_{n\times c}R(\Theta)_{c\times c}. \tag{126}$$

To achieve clear clustering assignments one term on each row of Z should be large relative to the other entries of that row; in the ideal case, this corresponds to only one nonzero entry per row (as in Eq. 124). Thus, we set $M_i = \max_j |Z_{ij}|$ in order for the optimization to achieve the desired structure in Z. This optimization can be solved using the gradient descent approach proposed by [79] with the update rule:

$$\Theta_{k+1} = \Theta_k - \alpha\nabla\mathcal{J}|_{\Theta=\Theta_k}. \tag{127}$$

After this rotation node i belongs to the cluster given by $arg\ max_j|Z_{ij}|$.

Which number of clusters produces the rotation closest to the ideal case? To answer this, we consider the *quality* of each clustering. Quality can be defined as

$$q(c,n) = 1 - \left(\frac{\mathcal{J}}{n} - 1\right). \tag{128}$$

Note that n is the minimum of $\mathcal{J}$ for the ideally clustered case, and in that case q is equal to 1. Thus, clustering quality is correspondingly farther from 1 for poorer cluster assignments. Since the number of nodes $n = |\mathcal{V}|$ is constant, $arg\ max_c q(c,n)$ provides the optimum cluster number. This summarizes [76] and results in a number of clusters chosen independent of human choice or prejudice.

Clustering at each time step and basing the dynamic case on the frequency of each static cluster places undue importance on these intermediate clusterings. In particular, since the

relationships of nodes inside the intermediate clusterings is lost in this method, it is not clear what the interpretation of these intermediate clusters is. That is, information about nodes which are close to each other inside a larger cluster is lost with this type of analysis.

Instead, we look for an algorithm that captures the information contained within these larger clusters at each time point. Our choice of weight - an aggregate, pairwise distance between nodes - tracks the relative positions of each node as it moves between other clusters and among other nodes and retains that information throughout a desired time interval. In other words, our algorithm saves the information about inter-agent relationships over some time rather than considering only the movement of agents between clusters.

We define the distance between two dynamically evolving points in our dataset, $d : \mathcal{V}(t) \times \mathcal{V}(t) \to \mathcal{R}$, as a function of their instantaneous distance over a window of time leading to current time (let $d(i, j) = d_{ij}$). In general:

$$d_{ij}(k) = f(\delta_{ij}(0), \cdots, \delta_{ij}(k)). \tag{129}$$

Specifically, we choose this function to simply sum the distances over time, i.e.

$$d_{ij}(k) = \sum_{m=k-l}^{k} \delta_{ij}(m), \tag{130}$$

over a sliding time period $[k-l, k] \in [t_0, t_f]$. We redefine our graph to contain the aggregated information of inter-agent distance over a period of time as $\mathcal{G}(\mathcal{V}, \mathcal{E}, \hat{w}(k))$. Here we define new weights of the form

$$\hat{w}_{ij}(k) = -\exp\left(\frac{d_{ij}^2(k)}{\sigma_i(k)\sigma_j(k)}\right), \tag{131}$$

where $\sigma_i(k)$ is a local scaling factor for the inter-agent distances and is chosen to be the aggregated distance to the 7th nearest neighbor to node i. This choice of weight is inherited from [76] and will be discussed further below. Now, the data points are described in terms not of their absolute positions (or similarities), but in terms of net movement over the sliding window. However, the algorithm still tries to reduce the inter-dependence of the new, aggregate eigenvectors in the same n-dimensional space.

This choice of weight self-tunes the scale of the algorithm to the scale inherent in the data. Using the 7th nearest neighbor biases the method to assign clusterings where the number of nodes in each cluster is on the order of 7; while it has proven useful for a wide range of data sets of various sizes of clusters [76], if the number of nodes in a cluster should be more on the order of 7 million in an ill defined data set (where the difference in the weights inside and out of a given cluster is small), this parameter may need to be adjusted though it is expected to remain constant for one time varying data set. This implements the second main contribution to spectral clustering in [76].

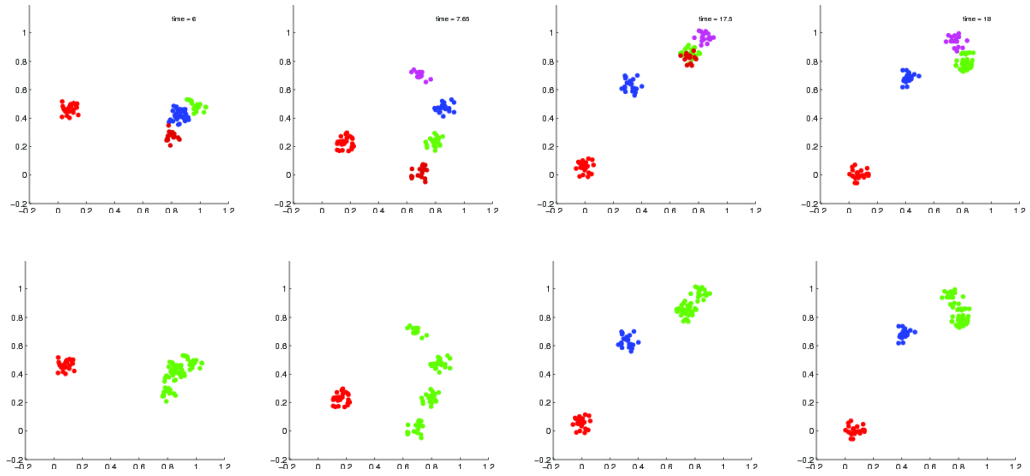
In the dynamic extension presented in this thesis, we have introduced a second parameter, the length of the sliding time window l, which, while invariant to many types of data or small changes in a given dataset, may need to be tweaked for individual dataset. This window represents the algorithm's ability to use the past temporal information to influence clusterings. Thus, it should be sized according to the rate of change of the data and to any notion of memory the system has. That is, the window should be short enough so that quickly changing data is able to form and reform clusters - the lack of any window may result in equal weights across all edges as after some point each node has undergone a similar net overall change. Even for a very slowly evolving system, the window should remain small enough to accurately model the fact that nodes "forget" past interactions after some time.

To compare the performance of our dynamic clustering to that of the static one, we generated a 100 point dataset with groups of points moving in $\mathcal{R}^2$. The distance metric was defined as the Euclidean distance between the position of each two points,

$$\delta_{ij}(k) = \|x_i(k) - x_j(k)\|, \tag{132}$$

summed over a time window of 10 seconds. Figure 5.1 depicts the results of the proposed clustering algorithm for the dynamic case on the top row versus that of the static case on the bottom row.

The figure illustrates our algorithm's ability to keep track of the cluster associations



(a) t=6 (b) t=7.65 (c) t=17.5 (d) t=18

Figure 38. A comparison of previous, static methods and our choice of dynamic weights: spectral clustering using aggregated distances (top) using instantaneous distances (bottom).

over time. Consider the three clusters which converge at time t=6: dynamic clustering shows a blue, burgundy, and green cluster while static clustering shows them as one green cluster. In the case where clusters converge or diverge and remain close or far for some time, the dynamic algorithm identifies this as a cluster merge or split, respectively. These phenomena are visible at times t=18 when the green and burgundy clusters have merged and t=7.65 with the emergence of the pink cluster.

5.2 Automatic Time-scale Extraction

The framework developed in Sec. 3.3 can be applied to this dynamic clustering algorithm in order to automatically extract a time-varying, yet persistent sliding window or time-scale. In Sec. 3.3 it was a notion of movement quality developed in Ch. 2 that motivated a movement segmentation problem. In the case of the dynamic clustering method presented here, a natural question to ask is what is the optimal width of the sliding window, the parameter l such that the quality, $q(c,n)$, is maximized?

To outline how this can be accomplished, consider computing the quality of the clusterings that result for a range of values for l. Some clustering will result in an optimal quality,

l^*, at each time. It is desirable for the resulting signal $l^*(t)$ to provide some measure of the intrinsic time scale of the data in order to answer questions like "Does the data vary quickly or are changes over time persistent?" Thus, it is desirable that $l(t)$ have a constant value – or even better a piecewise constant value that allows for discrete changes in the time scale that must persist over some amount of time.

This can be formulated as an inverse timing problem as presented in Sec. 3.3. In this case, the signal we wish to match, ρ, is $l^*(t)$, and the nominal reference, σ, is a piecewise constant function (with a minimum step width γ). The optimization variable τ will give the value of the time constant (τ is the time constant over the first interval and $T - \tau$ the value over the second).

$\diamond$

This chapter has presented a dynamic clustering method and an application outside of human movement for the material presented in Ch. 3. The final section of this chapter aimed to cast the inverse problems posed in this thesis as tools with general applications. It should also be noted that clustering methods are widely used in human movement analysis as reviewed in Sec. 1.1, and the dynamic method with automatic time-scale extraction presented here may have applications in this arena as well.

APPENDIX A

THE PROGRAM FOR *AUTOMATON*

This appendix presents the program from the performances of *Automaton*. Audience members were given a sheet of unfolded paper with Figs. 39 and 40 printed on either side. This sheet of paper was accompanied with a half sheet of paper with the image in Fig. 41 printed on it. Thus, in folding their programs, the audience experienced the technical definition of an automaton.

A u t o m a t o n

April 6 and 13, 2013
Georgia Institute of Technology
Atlanta, GA

Choreography by Amy LaViers in collaboration with dancers
Alex Abarca, Helen Hale, Camille Jackson, and Erik Thurmond.

Based on research performed under the advisement of
Dr. Magnus Egerstedt.

Brought to you by
The Georgia Robotics and Intelligent Systems (GRITS) Laboratory
and The Department of Electrical and Computer Engineering.

Thank you.

To my adviser, collaborator, and mentor, Magnus Egerstedt, for creating this body of research with me and enabling this show. My collaborators in movement, Alex Abarca, Helen Hale, Camille Jackson, and Erik Thurmond: I thank each of you for bringing your energy, thought, and bodies to this showing and my research. This performance of *Automaton* has been funded by the Electrical and Computer Engineering department. This award has not only produced a show, but funded rehearsals where I have re-thought every component of my thesis from a new angle. I cannot say enough about what their sponsorship has done for the work (in part, because it's hard to quantify and in part, because it is such a big contribution). Thanks to the GA Tech newsroom, in particular Liz Klipp, for creating such a thoughtful and accurate representation of this showing for the GA Tech homepage. Thanks also to the Ferst Center for the Arts, in particular Jack Rogers, for their help in the show's production details. Thanks to Marion Crowder in ECE for finding us microphones! And, thanks to Ann Terka, also in ECE, for handling and the funds for the show. Thanks most of all to my family and friends who have been recruited for many behind the scenes tasks along the way.

Figure 39. The front of the unfolded program for *Automaton*.

THE ARTISTS

Amy LaViers is a PhD candidate in Electrical and Computer Engineering at Georgia Tech. She studies robotics and control theory through the lens of movement and dance. With her PhD, she hopes to push our understanding of dance and develop tools for artists using quantitative methods from engineering. Her research began at Princeton University where she earned a certificate in Dance and degree in Mechanical and Aerospace Engineering. Her senior thesis earned top honors in the MAE department, the School of Engineering and Applied Science, and the Lewis Center for the Arts. At Georgia Tech, she is the recipient of the ECE Graduate Teaching Excellence Award and a finalist for the CETL/BP Outstanding Graduate Teaching Award. *Automaton* explores the ideas of style and automation outlined in her PhD dissertation.

Alex Abarca is a native of Houston. A graduate of the High School for Performing & Visual Arts and the University of Houston; he danced with various companies and independent artists while in Texas. He had the opportunity to perform at the historic Inside/Out stage a Jacob's Pillow with Suchu Dance. He came to Atlanta about four years ago as a company member with CORE performance company. Now as a freelance dance-artist, he dances for Catellier Dance Projects, Ellisorus Rex & the DANCE machine, and The Lucky Penny (Blake Beckham). He was also a featured dancer for the ARTech residency with Sean Curran and Dancers. Alex is excited about working with Amy LaViers in Automaton, and sharing the stage with a robot (which he named Bobb).

Helen Hale is an Atlanta-based modern dance choreographer and performer. Her work has been presented by Dashboard Co-op, Art on the Atlanta BeltLine, WonderRoot Community Arts Center, BurnAway Magazine, Dance Truck, The Lucky Penny, The Hambidge Center, The Layman Group, and the High Museum of Art. Helen received a BFA from Temple University (PA) in 2009 and has performed with companies and choreographers in Philadelphia, New York, and Atlanta including Kun-Yang Lin/Dancers, Staibdance, Coriolis Dance Inc., Duende Dance Theater, Tahni Holt, Meg Foley/Moving Parts, Cynthia St. Clair, Makoto Hirano, Lykion Ton Ellinidon Atlanta/Troupe Hellas, Gathering Wild Dance, Room To Move, Ground Delivery Dance Theater, Corian Ellisor, and Alisa Mittin. Helen is the recipient of a Walthall Artist Fellowship for professional development as well as a recent fellow at the Hambidge Center for Creative Arts and Sciences. Helen continues to work in conjunction with Dashboard Co-op and to perform with Staibdance while developing her current project, *Sanity Ceremonies*, on which she is working closely with the artistic support and brilliance of installation artist/costume designer, Amanda Baumgardner, and motion designer/filmmaker, Mike Boutté.

Camille Jackson Since moving to Atlanta in 2004, Camille has had the opportunity to work with many wonderful choreographers including Susan Eldridge/DENSE, Blake Beckham/Dance Truck, Alisa Mitten, Emily Christianson, Greg Catellier, and Lilli Ransijn to name a few. She currently teaches creative movement at Moving in the Spirit and Creative Learning Atlanta. When not in the studio, Camille can be found gardening or rehearsing with the Emory Gamelan Ensemble.

Erik Thurmond is from Gwinnett County, Georgia. After training at the Alabama School of Fine Arts he moved to New York and studied at the American Musical and Dramatic Academy and The Alvin Ailey School. After spending a few months in Tel Aviv, Isreal last year he has made his way back to Atlanta where he is currently dancing for Staib Dance and CORE Performance Group.

THE PROGRAM

1.
Amy LaViers
Music: Nico Muhly

2.
Helen Hale + Camille Jackson
Music: Abbott + Costello

3.
Alex Abarca + Erik Thurmond

4.
Alex Abarca, Helen Hale, Camille Jackson, + Erik Thurmond
Music: Avicii, Nico Muhly, + Skrillex

Figure 40. The back of the unfolded program for *Automaton*.

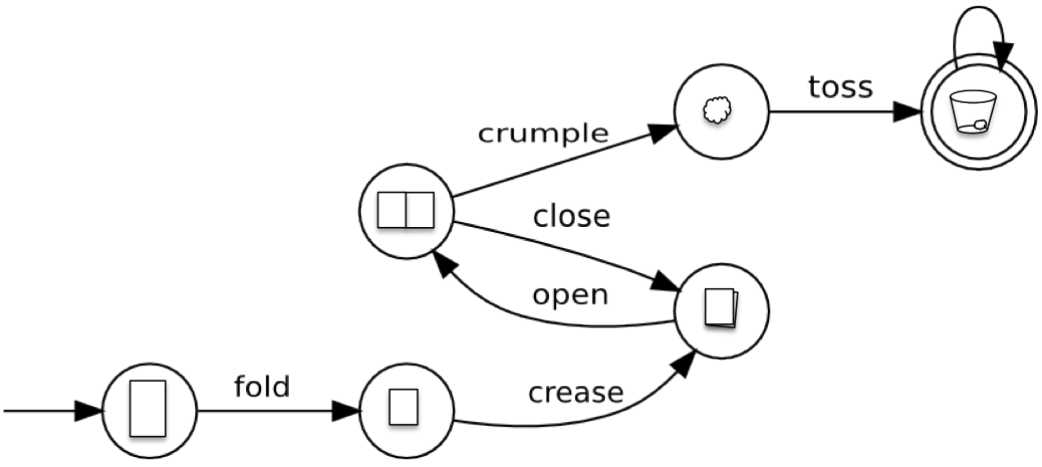


Figure 41. The insert that accompanied the unfolded program for *Automaton*.

APPENDIX B

HUMAN STUDY IRB FORMS

This appendix contains the forms related to and the results of the human study proposed in Ch. 4.

Georgia Tech Office of Research Integrity Assurance

Protocol Number: H13034
Funding Agency: N/A
Review Type: Exempt, Category 2
Title: Human Experience of Style-Based Robotic Movement
Number of Subjects: 50

March 5, 2013

Magnus Egerstedt
ECE
0250

Dear Dr. Egerstedt:

The Institutional Review Board (IRB) has carefully considered the referenced protocol. Your approval is effective as of **February 26, 2013**. The proposed procedures are exempt from further review by the Georgia Tech Institutional Review Board.

Project qualified for exemption status under 45 CFR 46 101b. **2.**

Thank you for allowing us the opportunity to review your plans. If any complaints or other evidence of risk should occur, or if there is a significant change in the plans, the IRB must be notified.

If you have any questions concerning this approval or regulations governing human subject activities,

please feel free to contact Dennis Folds, IRB Chair, at 404/407-7262, or me at 404 / 894-6942.

Sincerely,

Melanie J. Clark, CIP
IRB Compliance Officer

cc: Dr. Dennis Folds, IRB Chair
Amy La Viers

Figure 42. The approval form from Georgia Tech's Institutional Review Board (IRB) after the showings of *Automaton*.

CONSENT DOCUMENT FOR ENROLLING PARTICIPANTS IN A RESEARCH STUDY

Georgia Institute of Technology
Project Title: Human Experience of Style-Based Robotic Movement
Investigators: Magnus Egerstedt, Ph.D and Amy LaViers

Protocol and Consent Title: Human Experience of Style-Based Robotic Movement
Study Consent Form 01/22/2013 version 1

You are being asked to be a volunteer in a research study.

Purpose:

The purpose of this study is to better understand people's perspective of robots as human-like, stylized movers. For example, does our algorithm convince human views that a small humanoid robot is doing the disco? The study will help us understand the effect of a specific movement generation algorithm in the context of a dance show and keeping robotic movement in line with a particular style of movement. We expect to enroll 50 people in this study.

Exclusion/Inclusion Criteria:

Only those persons who are 18 years of age and older are eligible to participate.

Procedures:

If you decide to be in the study, your part will involve a single session. First, you will watch a 1-hour dance show that includes a small humanoid robot. Then you will be given a survey that asks you to rate your impressions of the robot's style of movement. This survey will take no more than twenty minutes; however, you may leave or stop at any time.

Risks or Discomforts:

The risks involved are no greater than those involved in daily activities, such as watching a live show for 1 hour and filling out a questionnaire.

Benefits:

You are not likely to benefit in any way from joining this study. We hope that what we learn will help roboticists control the style of motion of humanoid robots.

Compensation to You:

Consent Form approved by Georgia Tech IRB February 26, 2013

Page 1 of 3

Figure 43. The consent form provided to audience member volunteers after the showings of *Automaton*.

Georgia Institute of Technology
Human Experience of Style-Based Robotic Movement

Data ID: _____ (Note: *This ID does not link this questionnaire to you and your personal information, but rather links the questionnaire to the data collected by the program.*)

Instructions: Please rate each of the following. Circle a value from 1 (inconsistent) to 10 (consistent). You may leave comments about any of the tasks if you wish.

1. Please rate your impression of the overall consistency of the robot's movement. A rating of 1 indicates that it seemed to be moving completely randomly. A rating of 10 indicates that it was moving in a very specific way.

   1 2 3 4 5 6 7 8 9 10

   Comments:

2. Please rate your impression of the style of the robot's movement. A rating of 1 indicates that it seemed to be moving without style. A rating of 10 indicates that it was moving in a very stylized way.

   1 2 3 4 5 6 7 8 9 10

   Comments:

3. Please rate your impression of the virtuosity of the robot's movement. A rating of 1 indicates that it seemed to be very bad at moving. A rating of 10 indicates that it was very good at moving.

   1 2 3 4 5 6 7 8 9 10

   Comments:

4. Please rate your overall enjoyment of the robot's movement. A rating of 1 indicates that you really did not enjoy it. A rating of 10 indicates that you very much enjoyed it.

   1 2 3 4 5 6 7 8 9 10

   Comments:

5. Please rate your feelings about the robot's movement. A rating of 1 indicates that you felt very little about the movement. A rating of 10 indicates that you were very moved by the movement.

   1 2 3 4 5 6 7 8 9 10

   Comments:

Figure 44. The questionnaire provided to audience member volunteers after the showings of *Automaton*.

6. Please give details about your overall impression of the robot's movement in a short paragraph.

Figure 45. The second page of the questionnaire provided to audience member volunteers after the showings of *Automaton*.

Q-1	Q-2	Q-3	Q-4	Q-5	Q-6 (Comments)
10	10	10	10	9	I saw tension, form, and shape.
6	8	2	10	8	Motions were in keeping with dancers' motions.
8	9	7	9	9	
8	7	8	8	8	
9	8	8	7	9	
7	7	7	7	5	
9	8	9	10	8	
9	7	8	8	7	I felt that you could definitely tell what the movements of the robot were supposed to represent. Also, for the balance and joint limitations of the robot I felt the style was pretty similar to an actual human movement.
9	4	8	8	4	
7	6	6	7	6	
9	10	9	10	9	It was interesting and impressive how unrandom the movement was. The movement was more fluid than I expected. There was emotion and a connection, especially when with other dancers. I was surprised and impressed by the movement. Especially how fluid and almost "human" it was. Especially in the last piece. It matched the emotion and beat of the people performers.
9	2	7	4	1	
7	7	7	9	8	
9	3	7	9	8	Amusing and entertaining. Seemed to imitate dancers.
5	6	4	8	7	
8	8	7	9	8	
3	3	6	10	10	
7	7	5	8	5	
9	9	7	9	8	The robot was moving like the other dancers but at a tad different speed, with different sequence of movements and with a different feeling to its movements, which was kind of cool because all the dancers, robot included, were doing the same things but different somehow. It's like the robot knew what needed to be done but was not simply choreographed to do dance, but rather, as was presented, was taught the style.
8	7	8	9	7	The movements of the robot, though limited, but could easily be related to the movement of the other performers. Different motion styles and dance movements were incorporated in the robots performance, which were distinguishable. Overall, in my opinion, the robot was definitely moving in a manner in which dancers move.
10	10	9	9	9	It was impressive how the robot's movements were in harmony with the dancers, yet not hard-coded. It was more impressive with some understanding of what was going on inside.
9	8	9	9	5	
10	5	7	9	2	From my seat, I could not see the actual movement incredibly well. I found that its poses were quite entertaining, but the actual movement seemed limited by the abilities of the robot. I couldn't see much more than a position to position movement. It seemed smooth, but not anything out of the ordinary.
10	8	8	9	5	It was funny when the robot mooned the crowd. Great work!
6	8	6	8	8	The robot movements were very smooth indicating that there was some background optimal control algorithm running. Some robot actions seemed random (which was probably intended?) It was a great experience watching it dance!
7	7	4	8	3	The movement appeared to be structured. At times he could not keep up with the humans. Though his movement was less "robotic" than expected.
7.96	7.04	6.93	8.52	6.85	

Figure 46. The data collected between both showings of *Automaton* of surveys (see Appendix B) completed by volunteers from the audience. The final row of the table contains the averaged response for each question.

REFERENCES

[1] J. E. Hopcroft, *Introduction to Automata Theory, Languages, and Computation, 3/E*. Pearson Education India, 2008.

[2] M. Michalowski, S. Sabanovic, and H. Kozima, "A dancing robot for rhythmic social interaction," in *Proceedings of the ACM/IEEE international conference on Human-robot interaction*, pp. 89–96, ACM, 2007.

[3] A. LaViers, "Learning the primary colors of dance," Master's thesis, Princeton University, 2009.

[4] A. Thomaz and C. Breazeal, "Teachable robots: Understanding human teaching behavior to build more effective robot learners," *Artificial Intelligence*, vol. 172, no. 6-7, pp. 716–737, 2008.

[5] B. Akgun, M. Cakmak, J. Yoo, and A. Thomaz, "Augmenting kinesthetic teaching with keyframes," *Robotics*, pp. 1–8, 2011.

[6] B. Akgun, M. Cakmak, K. Jiang, and A. Thomaz, "Keyframe-based learning from demonstration," *International Journal of Social Robotics*, pp. 1–13, 2012.

[7] S. Jiang, S. Patrick, H. Zhao, and A. D. Ames, "Outputs of human walking for bipedal robotic controller design," in *2012 American Control Conference, Montréal*, 2012.

[8] R. W. Sinnet and A. D. Ames, "Extending two-dimensional human-inspired bipedal robotic walking to three dimensions through geometric reduction," in *2012 American Control Conference, Montréal*, 2012.

[9] A. D. Ames, E. A. Cousineau, and M. J. Powell, "Dynamically stable bipedal robotic walking with nao via human-inspired hybrid zero dynamics," in *Proceedings of the 15th ACM international conference on Hybrid Systems: Computation and Control*, pp. 135–144, ACM, 2012.

[10] S. N. Yadukumar, M. Pasupuleti, and A. D. Ames, "From formal methods to algorithmic implementation of human inspired control on bipedal robots," in *Algorithmic Foundations of Robotics X*, pp. 511–526, Springer, 2013.

[11] M. Gillies, "Learning finite-state machine controllers from motion capture data," *Computational Intelligence and AI in Games, IEEE Transactions on*, vol. 1, no. 1, pp. 63–72, 2009.

[12] L. Kovar, M. Gleicher, and F. Pighin, "Motion graphs," in *ACM Transactions on Graphics (TOG)*, vol. 21:3, pp. 473–482, ACM, 2002.

[13] A. Nerode, "Linear automaton transformations," *Proceedings of the American Mathematical Society*, vol. 9, no. 4, pp. 541–544, 1958.

[14] D. Angluin, "Learning regular sets from queries and counterexamples* 1," *Information and computation*, vol. 75, no. 2, pp. 87–106, 1987.

[15] I. Chattopadhyay, Y. Wen, A. Ray, and S. Phoha, "Unsupervised inductive learning in symbolic sequences via recursive identification of self-similar semantics," in *American Control Conference (ACC), 2011*, pp. 125–130, IEEE, 2011.

[16] B. Bollig, J.-P. Katoen, C. Kern, M. Leucker, D. Neider, and D. Piegdon, "libalf: The automata learning framework," in *Computer Aided Verification*, pp. 360–364, Springer, 2010.

[17] J. Casti, "On the general inverse problem of optimal control theory.," *Journal of Optimization Theory and Applications*, vol. 32, no. 4, pp. 491 – 497, 1980.

[18] B. D. Ziebart, A. Maas, J. A. Bagnell, and A. K. Dey, "Human behavior modeling with maximum entropy inverse optimal control," in *AAAI Spring Symposium on Human Behavior Modeling*, pp. 92–97, 2009.

[19] *Analysis of optimal control models for the human locomotion.*, (Universite Paris-Sud, Laboratoire des signaux et syste'mes, Gif-sur-Yvette, 91192, France), 2010.

[20] A.-S. Puydupin-Jamin, M. Johnson, and T. Bretl, "A convex approach to inverse optimal control and its application to modeling human locomotion," in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pp. 531–536, IEEE, 2012.

[21] K. Mombaur, A. Truong, and J. Laumond, "From human to humanoid locomotion-an inverse optimal control approach.," *AUTONOMOUS ROBOTS*, vol. 28, no. 3, pp. 369 – 383, 2009.

[22] A. Nakazawa, S. Nakaoka, K. Ikeuchi, and K. Yokoi, "Imitating human dance motions through motion structure analysis," in *Intelligent Robots and Systems, 2002. IEEE/RSJ International Conference on*, vol. 3, pp. 2539–2544, IEEE, 2002.

[23] A. Nakazawa, S. Nakaoka, and K. Ikeuchi, "Synthesize stylistic human motion from examples," in *Robotics and Automation, 2003. Proceedings. ICRA'03. IEEE International Conference on*, vol. 3, pp. 3899–3904, IEEE, 2003.

[24] T. Flash and N. Hogan, "The coordination of arm movements: an experimentally confirmed mathematical model," *Journal of Neuroscience*, vol. 5, no. 7, p. 1688, 1985.

[25] S. LaValle, *Planning algorithms*. Cambridge Univ Pr, 2006.

[26] R. A. Brooks, "Behavior-based humanoid robotics," in *Intelligent Robots and Systems' 96, IROS 96, Proceedings of the 1996 IEEE/RSJ International Conference on*, vol. 1, pp. 1–8, IEEE, 1996.

[27] R. Arkin, *Behavior-based robotics*. The MIT Press, 1998.

[28] C. Bregler, "Learning and recognizing human dynamics in video sequences," in *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*, pp. 568–574, IEEE, 2002.

[29] D. Del Vecchio, R. Murray, and P. Perona, "Decomposition of human motion into dynamics-based primitives with application to drawing tasks* 1," *Automatica*, vol. 39, no. 12, pp. 2085–2098, 2003.

[30] L. Wang, W. Hu, and T. Tan, "Recent developments in human motion analysis," *Pattern recognition*, vol. 36, no. 3, pp. 585–601, 2003.

[31] M. Egerstedt, T. Balch, F. Dellaert, F. Delmotte, and Z. Khan, "What are the ants doing? vision-based tracking and reconstruction of control programs," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2005)*, pp. 18–22, 2005.

[32] C. Fanti, L. Zelnik-Manor, and P. Perona, "Hybrid models for human motion recognition," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 1, pp. 1166–1173, IEEE, 2005.

[33] X. Ren, "Finding people in archive films through tracking," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pp. 1–8, IEEE, 2008.

[34] O. Jenkins and M. Mataric, "Automated derivation of behavior vocabularies for autonomous humanoid motion," in *Proceedings of the second international joint conference on Autonomous agents and multiagent systems*, pp. 225–232, ACM, 2003.

[35] S. Nakaoka, A. Nakazawa, K. Yokoi, and K. Ikeuchi, "Leg motion primitives for a dancing humanoid robot," in *Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE International Conference on*, vol. 1, pp. 610–615, IEEE, 2004.

[36] M. Do, J. Romero, H. Kjellstrom, P. Azad, T. Asfour, D. Kragic, and R. Dillmann, "Grasp recognition and mapping on humanoid robots," in *Humanoid Robots, 2009. Humanoids 2009. 9th IEEE-RAS International Conference on*, pp. 465–471, IEEE, 2009.

[37] S. Miller, J. Van Den Berg, M. Fritz, T. Darrell, K. Goldberg, and P. Abbeel, "A geometric approach to robotic laundry folding," *The International Journal of Robotics Research*, vol. 31, no. 2, pp. 249–267, 2012.

[38] D. Kulic, C. Ott, D. Lee, J. Ishikawa, and Y. Nakamura, "Incremental learning of full body motion primitives and their sequencing through human motion observation," *The International Journal of Robotics Research*, vol. 31, no. 3, pp. 330–345, 2012.

[39] L. Lee and W. Grimson, "Gait analysis for recognition and classification," in *Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference on*, pp. 148–155, Ieee, 2002.

[40] E. Surer and A. Kose, "Methods and technologies for gait analysis," *Computer Analysis of Human Behavior*, pp. 105–123, 2011.

[41] D. Del Vecchio, R. Murray, and P. Perona, "Primitives for human motion: a dynamical approach," in *15th IFAC World Congress on Automatic Control*, Citeseer, 2002.

[42] D. Del Vecchio, R. Murray, and P. Perona, "Segmentation of human motion into dynamics based primitives with application to drawing tasks," in *American Control Conference, Proceedings of the 2003*, vol. 2, pp. 1816–1823, 2003.

[43] D. Del Vecchio, R. Murray, and P. Perona, "Classification of human actions into dynamics based primitives with application to drawing tasks," in *European Control Conference, University of Cambridge, UK*, 2003.

[44] T. Wongpiromsarn, U. Topcu, and R. M. Murray, "Receding horizon control for temporal logic specifications," in *Proc. International Conference on Hybrid Systems: Computation and Control*, pp. 101–110, 2010.

[45] M. M. Quottrup, T. Bak, and R. Izadi-Zamanabadi, "Multi-robot motion planning: A timed automata approach," in *Proceedings of the 2004 IEEE International Conference on Robotics and Automation*, (New Orleans, LA), pp. 4417–4422, April 2004.

[46] H. Kress-Gazit, G. Fainekos, and G. Pappas, "Where's waldo? sensor-based temporal logic motion planning," in *IEEE International Conference on Robotics and Automation*, pp. 3116–3121, Citeseer, 2007.

[47] M. Kloetzer and C. Belta, "A fully automated framework for control of linear systems from temporal logic specifications," *IEEE Transactions on Automatic Control*, vol. 53, no. 1, pp. 287–297, 2008.

[48] X. C. Ding, C. Belta, and C. G. Cassandras, "Receding horizon surveillance with temporal logic specifications," in *the 49th IEEE Conference on Decision and Control*, 2010.

[49] E. M. M. Clarke, D. Peled, and O. Grumberg, *Model checking*. MIT Press, 1999.

[50] N. Piterman and A. Pnueli, "Synthesis of reactive(1) designs," in *In Proc. Verification, Model Checking, and Abstract Interpretation (VMCAI'06*, pp. 364–380, Springer, 2006.

[51] P. Gastin and D. Oddoux, "Fast LTL to Buchi automata translation," *Lecture Notes in Computer Science*, pp. 53–65, 2001.

[52] M. Brand and A. Hertzmann, "Style machines," in *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pp. 183–192, ACM Press/Addison-Wesley Publishing Co., 2000.

[53] T. Matsubara, S. Hyon, and J. Morimoto, "Learning stylistic dynamic movement primitives from multiple demonstrations," in *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pp. 1277–1283, IEEE, 2010.

[54] E. Bradley and J. Stuart, "Using chaos to generate variations on movement sequences," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 8, no. 4, pp. 800–807, 1998.

[55] E. Hsu, K. Pulli, and J. Popović, "Style translation for human motion," in *ACM Transactions on Graphics (TOG)*, vol. 24:3, pp. 1082–1089, ACM, 2005.

[56] P. Martin and M. Egerstedt, "Timing control of switched systems with applications to robotic marionettes," *Discrete Event Dynamic Systems*, vol. 20, no. 2, pp. 233–248, 2010.

[57] P. Kingston and M. Egerstedt, "Time and output warping of control systems: Comparing and imitating motions," *Automatica*, 2011.

[58] M. Gielniak, C. Liu, and A. Thomaz, "Stylized motion generalization through adaptation of velocity profiles," in *RO-MAN, IEEE*, pp. 304–309, IEEE, 2010.

[59] G. Warren and S. Cook, *Classical ballet technique*. Gainesville, FL: University of South Florida Press, 1989.

[60] www.intermartialarts.com, "Tai chi - 24 form."

[61] P. Hackney, *Making connections: Total body integration through Bartenieff fundamentals*. Routledge, 1998.

[62] V. Maletic, *Body, space, expression*. Berlin: Walter de Gruyter & Co., 1987.

[63] J. Newlove and J. Dalby, *Laban for All*. Nick Hern Books, 2004.

[64] G. Hoffman and G. Weinberg, "Shimon: an interactive improvisational robotic marimba player," in *Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems*, pp. 3097–3102, ACM, 2010.

[65] C. Huepe, R. F. Cádiz, and M. Colasso, "Generating music from flocking dynamics," in *American Control Conference (ACC), 2012*, pp. 4339–4344, IEEE, 2012.

[66] R. Copeland, *Merce Cunningham: The modernizing of modern dance*. Routledge, 2003.

[67] W. Forsyth, A. Sommer, and R. Sulcas, *William Forsythe: Improvisation Technologies*. ZKM, 1999.

[68] M. Palazzi and N. Z. Shaw, "Synchronous objects for one flat thing, reproduced," in *SIGGRAPH 2009: Talks*, p. 2, ACM, 2009.

[69] A. LaViers and M. Egerstedt, "The ballet automaton: A formal model for human motion," *American Control Conference, Proceedings of the 2011*, 2011.

[70] A. LaViers, Y. Chen, C. Belta, and M. Egerstedt, "Automatic generation of balletic motions," *ACM/IEEE Second International Conference on Cyber-Physical Systems*, pp. 13–21, 2011.

[71] A. LaViers, Y. Chen, C. Belta, and M. Egerstedt, "Automatic sequencing of ballet poses," *Robotics & Automation Magazine, IEEE*, vol. 18, no. 3, pp. 87–95, 2011.

[72] A. LaViers and M. Egerstedt, "Style based robotic motion," *American Control Conference, Proceedings of the 2012*, 2012.

[73] C. Cassandras and S. Lafortune, *Introduction to discrete event systems*. Springer, 2008.

[74] T. Nakata, T. Mori, and T. Sato, "Analysis of impression of robot bodily expression," *Journal of Robotics and Mechatronics*, vol. 14, no. 1, pp. 27–36, 2002.

[75] D. Haraway, "A cyborg manifesto: Science, technology, and socialist-feminism in the late twentieth century," *Simians, Cyborgs and Women: The Reinvention of Nature*, pp. pp.149–181., 1991.

[76] P. Perona and L. Zelnik-Manor, "Self-tuning spectral clustering," *Advances in neural information processing systems*, vol. 17, pp. 1601–1608, 2004.

[77] M. Lahiri and T. Berger-Wolf, "Mining periodic behavior in dynamic social networks," in *Data Mining, 2008. ICDM'08. Eighth IEEE International Conference on*, pp. 373–382, IEEE, 2008.

[78] A. LaViers, A. Rahmani, and M. Egerstedt, "Dynamic spectral clustering," *19th International Symposium on Mathematical Theory of Networks and Systems*, 2010.

[79] V. Goyal and M. Vetterli, "Block transform adaptation by stochastic gradient descent," in *IEEE Digital Signal Processing Workshop*, 1998.