Control and Perception of Robotic Movement with Styles

A Dissertation

Presented to

the faculty of the School of Engineering and Applied Science

University of Virginia

in partial fulfillment
of the requirements for the degree

Doctor of Philosophy

by

Lin Bai

August 2019

APPROVAL SHEET

This Dissertation
is submitted in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

Author Signature: Lin Bai

This Dissertation has been read and approved by the examining committee:

Advisor: Luke Dahl

Committee Member: Scott Acton

Committee Member: John Lach

Committee Member: Malathi Veeraraghavan

Committee Member: Daniel Quinn

Committee Member: ______

Accepted for the School of Engineering and Applied Science:

Craig H. Benson, School of Engineering and Applied Science

August 2019

Abstract

Robots are usually thought of as tools that carry out different tasks with functional movements. Robots that need to operate in human-facing environments will require complex options for modifying their movements to communicate changing state information. However, current robotic platforms are limited in reproducing the variation of movement in human perception, such as Baxter Research Robot from Rethink Robotics and Nao from Aldebaran Robotics. This research focuses on improving the feature variation of robotic movement to make robots more accessible, engaging, and collaborative when interacting with humans.

In this research, approaches are designed for solving the problem that occurs when the generated trajectories in high-level controller exceed the physical limits of a particular robotic platform. This work aims to guarantee the trajectories generated by prior high-level controllers are executable on physical robotic platforms. Two of the approaches are used as part of a web-based application that leverages ROS, MATLAB, and onboard low-level controllers to show how the methods can be applied with the technical details abstracted away for a user. This system is implemented on Rethink Robotics Baxter Research Robot with different selections of quality parameters to demonstrate the methods.

To make the human-robot interaction more intuitive and more effective for intent recognition, perceptually meaningful sound is supplemented to the robotic movement. Human vocalization responses to the videos of different simulated robotic movements were recorded and these recordings were analyzed to study how sonic features map to features of movement in human perception. The mapping of features of sound and movement enables us to create appropriate sounds to accompany robotic movement to help convey movement qualities

and make it more expressive. To further improve the variations in robotic movement, a new method for generating the movement reference trajectories with exaggerated variations is proposed. This method is designed based on the affinities between Effort and Space in Laban/Bartenieff Movement System (LBMS).

A user study is carried out for testing whether accompanying movement with sound and the method of generating movement trajectories with spatial affinities are helpful in improving the expressivity of robotic movement. This user study consists of surveys of perceiving the qualities in robotic movement in stick figure animations and videos of Baxter Research Robot. The user study data was analyzed quantitatively using statistical tests.

The existing perceived variations in robotic movement were increased by generating physically feasible movement trajectories for a robot to carry out, accompanying the movements with appropriate sounds and generating more various reference trajectories for a robot to track.

Acknowledgments

I would like to take this opportunity to acknowledge the people who have helped, supported and encouraged me during my Ph.D. studies at the University of Virginia.

I would like to give my deep thanks to my advisor Professor Luke Dahl for his continuous support, guidance, motivation and patience and for being always available to help whenever I seek advice in my research. I would like to express my sincere gratitude to my dissertation committee members, Professor Scott Acton, Professor John Lach, Professor Malathi Veeraraghavan and Professor Daniel Quinn for their time and support and valuable advice of my research which helped me improve this work.

I am grateful to Professor Amy LaViers and Dr. Jon Bellona for the collaboration in this interdisciplinary work. I would like to thank Clay Ford at the StatLab of University of Virginia Library for his helpful suggestions in the data analysis. I would also like to thank Dr. Kelly Cunningham in Engineering Graduate Writing Lab for providing resources to help with my writing.

I gratefully appreciate the University of Virginia Data Science Institute, Office of the President, and the Office of Graduate and Postdoctoral Affairs for awarding me 2015-2016 Presidential Fellowship in Data Science, which partially funded this work.

I would like to thank my colleagues and friends for their friendship, support and company and for sharing the incredible journey with me. I want to thank my parents for their love, support and trust in me. Finally, I would like to thank my husband who has always been there for me. I am grateful to him for his love, support, care and understanding. This dissertation is dedicated to my family.

Contents

1 Introduction **1**

1.1 Research Motivation 1

1.2 Background 2

1.2.1 Laban/Bartenieff Movement System 3

1.2.2 Style-based Robotic Motion Control 5

1.2.3 Relationships between Movement Qualities and Sound Qualities . . . 8

1.3 Research Outline 10

2 Style-based Robotic Motion Control with System Constraints **11**

2.1 Considering Physical Limitations of a Given Robotic Platform 11

2.1.1 Method A: Trimming 12

2.1.2 Method B: Scaling 13

2.1.3 Method C: Optimal Control with System Constraints 13

2.2 Computational Implementation 19

2.3 Application on Rethink Robotics Baxter Research Robot 21

3 Increasing the Perceived Variation of Movement Qualities **27**

3.1 Accompanying Movement with Sound 27

3.1.1 Experiment Design 27

3.1.2 Mapping between Qualities of Movement and Sound 32

3.1.2.1 Qualitative Analysis of the Recorded Vocalization 32

3.1.2.2 Quantitative Analysis of the Recorded Vocalization 36

3.2 Utilizing Affinities between Effort and Space for Reference Trajectories . . . 41

4 Evaluating the Methods with User Study 46

4.1 Experiment Design of the User Study 46

4.2 Quantitative Analysis of the Survey Data of Stick Figure Animations 52

4.2.1 Accuracies of Perception of the Quality of Space in Stick Figure Animations 53

4.2.2 Accuracies of Perception of the Quality of Weight in Stick Figure Animations 56

4.2.3 Accuracies of Perception of the Quality of Time in Stick Figure Animations 60

4.2.4 Comparison of Accuracies of Perception of Movement Qualities in Stick Figure Animations Across Four Methods 63

4.2.5 Conclusions of the Survey of Stick Figure Animations 66

4.3 Quantitative Analysis of the Survey Data of Baxter Research Robot 68

4.3.1 Accuracies of Perception of the Quality of Space of Baxter's Motion . 68

4.3.2 Accuracies of Perception of the Quality of Weight in Baxter's Motion 72

4.3.3 Accuracies of Perception of the Quality of Time in Baxter's Motion . 75

4.3.4 Comparison of Accuracies of Perception of Movement Qualities in Baxter's Motion Across Four Methods 78

4.3.5 Conclusions of the Survey of Baxter Research Robot 81

5 Conclusions and Future Work 84

List of Figures

1 This figure shows the controller architecture for Methods A and B of dealing with the system constraints: trimming and scaling. In both cases, we have a bi-layer structure where user requests ($[Q, R, P, S]$) are treated independently of the system constraints ($x_{lb} \leq x \leq x_{ub}$, $u_{lb} \leq u \leq u_{ub}$). 13

2 This figure shows the controller architecture for Method C presented. The constrained optimal control method is used to generate feasible trajectories as the reference signal for low-level controller of the robot. 14

3 Screenshot of the Robot Control/Choreography Center (RCC) graphical user interface. Users select a desired motion sequence from the menu on the left. Then they pick the qualities for execution of these movements, which are mapped to the weights Q, R, P, and S in the cost function. This is done through the number entry on the right side [1]. 20

4 The reference signal from State 1 to State 2 in Figure 3 for the shoulder twist joint on the right arm of the robot and two different styles of movement with two different sets of quality parameters. The ranges of the angles and the slopes of the curves varies with different qualities. This means the joint positions and velocities are different, which leads to diverse styles of movement. 24

5 We implement Methods A and B proposed in Section 2.1 on the robot to show the effectiveness of our solutions. The original velocity trajectory $y(t)$ is out of the system constraints. Both methods produce a y_{feas} that is within the bound. The quality parameters used here are $[100, 1, 1, 0.1]$. The upper and lower bound of the velocity on this joint are +/-1.5 rad/s. The scaling factor X_{scale} is found to be 4. 25

6 The traces of two different styles of movement on the robot from State 1 to State 2 in Figure 3 for feasible trajectories with two different sets of quality parameters. The figure on the left corresponds to quality parameters $[0.1, 0.1, 0.1, 1]$ (the green line in Figure 4), and the one on the right corresponds to quality parameters $[10, 0.1, 1, 100]$ (the red line in Figure 4). We can see the significant difference in the movement traces in the two cases. The one on the left with a free Flow does not hit the end pose exactly; while the one on the right with a bound Flow hit the end pose more precisely. The one on the right tracks the path more aggressively than the flexible one on the left. 26

7 Picture of the visual stimulus provided to subjects [2]. 30

8 Label counts of the encodings for Weight Effort: Light vs. Strong [2]. 34

9 Label counts of the encodings for Time Effort: Sudden vs. Sustained [2]. . . 35

10 Label counts of the encodings for Space Effort: Direct vs. Indirect [2]. 36

11 Spectral Rolloff for the Space, Weight, and Time Effort Factors [3]. 38

12 Demonstrations of interpolation curves for reference trajectory when initial position y_0 is lower than the final position y_1 [2]. 45

13 Demonstrations of interpolation curves for reference trajectory when initial position y_0 is higher than the final position y_1 [2]. 45

14 Different reference trajectories with different choices of parameter k [2]. . . . 45

15 Screenshot of stick figure animation shown in the survey. 48

16 Screenshot of video of Baxter Research Robot shown in the survey. 48

17 The Dynamosphere: the space that one can reach with Effort factors: Space, Weight, and Time. The eight vertices show the eight Basic Effort Actions. Each BEA consists of a combination of the three Effort factors [4]. 49

18 Multiple comparisons of accuracies of subjects picking the right videos which matches the qualities of Space (Q), Weight (R) and Time (P) for the survey of stick figure animations. 64

19 Multiple comparisons of accuracies of subjects picking the right videos which matches the qualities of Space (Q), Weight (R) and Time (P) for the survey of Baxter Research Robot. 79

List of Tables

1 Joint limits of Baxter Research Robot [5]. 22

2 Laban's Eight Basic Effort Actions [2] 28

3 The values of Q, R, P, and S utilized in the study [2] 30

4 P values for significant findings comparing Effort factors to sound qualities [3]. 39

5 Affinities between Effort and Space [6]. 42

6 Affinities between Effort and Space for each BEA 43

7 Accuracy of subjects' perception of the quality of Space in stick figure animations 53

8 Accuracy of subjects' perception of Dabbing vs. Flicking in stick figure animations 54

9 Accuracy of subjects' perception of Slashing vs. Thrusting in stick figure animations 54

10 Accuracy of subjects' perception of Gliding vs. Floating in stick figure animations 55

11 Accuracy of subjects' perception of Wringing vs. Pressing in stick figure animations 56

12 Accuracy of subjects' perception of the quality of Weight in stick figure animations 57

13 Accuracy of subjects' perception of Dabbing vs. Thrusting in stick figure animations 57

14 Accuracy of subjects' perception of Slashing vs. Flicking in stick figure animations 58

15 Accuracy of subjects' perception of Gliding vs. Pressing in stick figure animations . . . 58
16 Accuracy of subjects' perception of Floating vs. Wringing in stick figure animations . . . 59
17 Accuracy of subjects' perception of the quality of Time in stick figure animations 60
18 Accuracy of subjects' perception of Dabbing vs. Gliding in stick figure animations . . . 61
19 Accuracy of subjects' perception of Flicking vs. Floating in stick figure animations . . . 61
20 Accuracy of subjects' perception of Slashing vs. Wringing in stick figure animations . . . 62
21 Accuracy of subjects' perception of Thrusting vs. Pressing in stick figure animations . . . 62
22 Accuracy of subjects' perception of the quality of Space in Baxter's motion . 69
23 Accuracy of subjects' perception of Dabbing vs. Flicking by Baxter Research Robot . . . 69
24 Accuracy of subjects' perception of Slashing vs. Thrusting by Baxter Research Robot . . . 70
25 Accuracy of subjects' perception of Gliding vs. Floating by Baxter Research Robot . . . 70
26 Accuracy of subjects' perception of Wringing vs. Pressing by Baxter Research Robot . . . 71
27 Accuracy of subjects' perception of the quality of Weight in Baxter's motion 72

28 Accuracy of subjects' perception of Dabbing vs. Thrusting by Baxter Research Robot . . . 73

29 Accuracy of subjects' perception of Slashing vs. Flicking by Baxter Research Robot . . . 73

30 Accuracy of subjects' perception of Gliding vs. Pressing by Baxter Research Robot . . . 74

31 Accuracy of subjects' perception of Floating vs. Wringing by Baxter Research Robot . . . 74

32 Accuracy of subjects' perception of the quality of Time in Baxter's motion . . . 75

33 Accuracy of subjects' perception of Dabbing vs. Gliding by Baxter Research Robot . . . 76

34 Accuracy of subjects' perception of Flicking vs. Floating by Baxter Research Robot . . . 77

35 Accuracy of subjects' perception of Slashing vs. Wringing by Baxter Research Robot . . . 77

36 Accuracy of subjects' perception of Thrusting vs. Pressing by Baxter Research Robot . . . 78

1 Introduction

1.1 Research Motivation

In this research we design methods to increase expressive variety in robotic movement, in order to help people to more easily identify differences in movement style and thus perceive the intent of a robots movement. These robotic movements can be generated to carry out different tasks or to realize the same task in different ways. The improved expressivity enables a robot to carry out more tasks that were not possible without expressive movements. For example, robots that interact with people may require a capacity to expressively indicate hospitality in a restaurant setting, or to be authoritative in an emergency response setting. The same action of inviting people to follow along in both settings may need to be modulated to engender delight in the first setting, and to indicate urgency of the request in the second setting. In another circumstance, on a manufacturing floor, a collaborative robot may need to lift or heave car components in various ways for car assembly. The differences in style for each of these movements correspond to the different function of each movement, and thus people working with these robots need to be able to perceive these stylistic differences in order to understand the intent of a movement.

This research is motivated by helping people identify the differences in the styles of robotic movements and improving people's perception of the intents of movements, despite the limited capabilities of robotic systems. The goal is to make human-robot interactions more intuitive and more effective for intent recognition in robotic movements and to increase both the expressive power and functional capacity of a robot, making it a more useful tool in human-facing applications.

To solve the problem that occurs when the generated trajectories exceed the physical limits of a particular robotic platform, approaches are designed to generate feasible trajectories that the low-level controllers may successfully track. In addition, leveraging the fact that humans often express motivations and intentions by changing both movement quality (body language) as well as sound quality (via non-linguistic aspects of vocal communication such as speech prosody), we investigate how a perceptually meaningful sound can accompany a robot's movement in order to increase the perceived expressivity of the movement. Through a study of human movement and sound correspondences, the expertise of trained musicians were leveraged to develop sonifications that give viewers additional cues to the intent of robotic actions. Furthermore, to add more diversity in robotic movement, a new method for generating the reference trajectories for a robot to track is designed based on the affinities of Space and Effort in Laban/Bartenieff Movement System (LBMS), a system for human movement description and analysis. The new reference trajectories have exaggerated variations that are meaningful in terms of the spatial aspect in LBMS.

1.2 Background

In this section, we will present an overview of research on Laban/Bartenieff Movement System (LBMS) and how it can help us to understand human movement and design robotic movement. We will then present background work on style-based robotic motion control, a method that can be used to generate robotic movement with different styles by changing movement quality parameters. Lastly, we will discuss background work on correspondences between movement qualities and sound qualities. This background knowledge provides solid technical foundation for the research described in this dissertation.

1.2.1 Laban/Bartenieff Movement System

In order to increase the expressivity of robotic movement and make a robot easier to use in human-robot interaction, we need to first understand how human perceives movements and then apply these perceptions into the design of robotic movement. Research has shown that body movements convey intentions and emotions during communication [7, 8, 9]. Our position that a greater variety in movement available to perform tasks in different ways makes robots more useful is informed by LBMS. In LBMS, it is believed that the expression and function of a movement are related to each other [6] and movement quality or Effort (describing how a movement is carried out) is important in both the expression and function of a movement. It is shown in [10] that creating more variation and thoughtful choices in movement quality, workers in a manufacturing setting were able to improve their efficiency and comfort. Similarly, a tool which can easily modify robotic movements and execute these movements on physical platforms will improve the effectiveness of robotic co-workers in these settings. We work on creating such a tool to make robotic movement more expressive and to enable users to quantitatively control the quality of the movement. Considerable effort has been devoted to the development of supervisory control for motion planning algorithm in [11, 12, 13, 14]. In this work, we develop a control scheme to generate trajectories of robotic movement. These trajectories are generated aiming to make the robotic movement better represent the intention. They can be changed easily with parameters that are mapped to different qualities in movement styles. We utilize the concept of movement styles and motion primitives introduced in [15, 16, 17, 18], in which dynamic systems with details of styles were developed.

The Effort system defined by Rudolf Laban [4, 6, 10] is used to describe and analyze how a movement is done. A movement has qualitative characteristics that are described by Laban with four motion factors: Space, Weight, Time, and Flow. We capitalize these four terms to avoid the possible confusion with the technical notions of space, weight, time and flow, as these technical meanings are different from those meant in LBMS. Space Effort describes the attention paid to the environment while a person is moving. A movement can be Direct or Flexible/Indirect in Space Effort, which indicates whether the person is focusing the attention (as in the movement of picking up an object) or diffusing it when moving (as in the movement of spraying air freshener). Weight Effort describes the attitude towards the mover's mass. It may be condensed (like a track athlete) or rarified (like a ballet dancer). Time Effort describes the attitude towards the initiation and finish of a movement. It can be Sudden (like a sprinter) or Sustained (like an elderly person). Flow Effort describes the progression of a series of movements. It can be Free (as in a bungee jumper who cannot easily control what movement comes next) or Bound (as in a diver who exhibits great control over the fall). The four components of Laban's Effort theory are mapped to the weight parameters in the cost function of an optimal control problem in [19]. Optimal control is widely used in the field of robotics [20, 21, 22, 23]. This mapping of Laban Effort factors to the weight parameters in a cost function is an alternative to the prior mapping in [24, 25] and the following mappings in [26, 27].

1.2.2 Style-based Robotic Motion Control

After presenting the qualitative language in LBMS to describe the differences in movement in Section 1.2.1, we review a quantitative interpretation for the varieties in the movement trajectories in this section. During communication, the same “movement” can be expressed in different ways for different intentions. For example, we can “point” in many different ways, with different functional and expressive objectives achieved in each. To ask for a cookie, a child will point to the jar in a very different way than a mother who is angry about a messy room will point to a stray sock. These are two distinct examples where the action is expressive. We can also “paint” in different ways. The brush stroke required to cover an object with a desired paint is very different depending on the viscosity of the particular paint. This is an example that the focus of the movement variation is functional. Research has been conducted in studying expressive robotic motion. It was shown in [26] that flying robots could use their locomotion paths to communicate affective information. In [28], the authors designed a method to generate familiarity of a robot with its expressive behavior. Work has been done to find how the varying movement impact human's feelings. In [29], the relationship of cues in dance movement with dancers expressive intentions was studied. Apart from the studies of how body movement impact human feelings, a framework based on LBMS is proposed to find robot-to-human expression of emotion and intention with robot behavior in [30]. The impact of abstract robotic movement has on human emotions was explored in [31]. The association between the expressivity parameters and dimensional representations of affect is shown in [32].

We now review a mapping between an optimal control problem and this qualitative description of movement provided in [19]. Consider a system with an input vector $u = [u_1, u_2, ..., u_m]^T$, a state vector $x = [x_1, x_2, ..., x_n]^T$, and an output vector $y = [y_1, y_2, ..., y_l]^T$, which tracks a reference signal $r = [r_1, r_2, ..., r_l]^T$. In [19], a quadratic cost function is established to describe a total cost required to generate a movement

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

where the four parameters $Q \in \mathbf{R}^{l\times l}, R \in \mathbf{R}^{m\times m}, P \in \mathbf{R}^{n\times n}, S \in \mathbf{R}^{l\times l}$ correspond to Space, Weight, Time and Flow in Effort System of LBMS respectively.

The problem of generating robotic movement trajectory is formulated and solved as an optimal control problem. The goal for the optimal control problem is to find an input u which minimizes the cost function J in (1) using the parameters Q, R, P, S for generating trajectories with different styles. We solve the optimal control problem

$$\min_u J \tag{2}$$

$$s.t.$$

$$\dot{x} = Ax + Bu \tag{3}$$

$$y = Cx \tag{4}$$

where $A \in \mathbf{R}^{n\times n}$, $B \in \mathbf{R}^{n\times m}$, and $C \in \mathbf{R}^{l\times n}$. The constraints indicate that the state x depends on itself and control vector u and the output y depends on the state x. Solving the optimal control problem with different choices of the four parameters Q, R, P, S produces outputs of different movement trajectories with various styles.

The solution of this problem has a closed algebraic form as in [19]. In general, the dimension of the input, output, and state vectors and the form of the reference trajectory are adjustable parameters, which may vary on different platforms and in different contexts.

In Section 2, we extend the prior problem in [19] by solving the problem of exceeding physical limits on robotic platforms and making it a generally applicable framework across multiple platforms. We implement our framework on Rethink Robotics Baxter Research Robot employing the sequencing in [33]. The control system developed here interfaces with a web application, which is designed to make it possible for users to control robotic movement without the knowledge of the control methods and programs. Similar human-robot interfaces have been developed by researchers and manufacturers. An icon-based programming interface LEGO MINDSTORMS EV3 Software was developed for users to design missions on a LEGO robot [34]. A sketching interface which allows users to control robots by posing stroke gestures on a computer screen was proposed in [35]. An interface for monitoring and controlling multiple robots was investigated in [36]. These interfaces, including the one used here, were designed for different control objectives and made it easier for users to understand and control robots. The web interface used here is not platform-specific, and it can be utilized to generate various robotic motions on different platforms.

We also extend the prior design in [19] by exaggerating the variations in movement trajectories in Section 3.2. In LBMS, there are close relationships among the four components of a movement: Body, Effort, Space, and Shape. We utilize the affinities between the qualities (in Effort) and the directions (in Space) of movement to improve the expression of a movement's qualities. The limitation of capabilities of a given robotic platform might make it unable to change the dynamics of each actuator or generate the desired speeds of motion, but it can alter a spatial pathway. Similar research of using the relationship between Effort and Shape in LBMS is described in [24].

1.2.3 Relationships between Movement Qualities and Sound Qualities

Humans communicate expressive intent through facial expressions, speech, vocalizations and movements [9, 37]. The sound we hear gives us images of movement in our minds [38]. Meaningful combinations of movement and sound are used in various human communications [39]. A sound added to a movement for different purposes may enhances our perception of the movement we see. For example, journalists reporting serious issues at work uses very different tones to the ones they use when talking to family at home, along with the different postures. A study in [40] shows that adding convergent auditory information can improve the accuracy of perception and reproduction of sports movement. Similarly, adding appropriate sound along with robotic movement may improve the interaction between robots and humans.

In order to generate appropriate sound to accompany robotic movement, the mapping relationships between the movement qualities and sound features need to be studied. Research has been done on the relationships between movement and music by studying movement responses to musical sound. For example, in [41], people's unconstrained spontaneous responses of body movements when listening to samba and chacarera music were analyzed. In [42], motion capture data of subjects' free hand movements while listening to short sound examples was analyzed to study the relationship between gesture and sound in music performances. Apart from these studies of unconstrained body movements, studies of how people mime the performance of musical instruments while listening to the music were also conducted, such as making conducting gestures when listening to classical music [43], doing air-drumming gestures when listening to the sound of drum [44], and miming a violin performance while listening to violin recordings [45]. Quantitative analysis of body movements of clarinet performers show that there are correlations of movement patterns to musical features [46].

Instead of studying the movement responses to preexisting sound, this research moves the other way around by asking participants to generate sounds for preexisting movements and corresponding descriptive terms. With the given qualities of the robotic movement stimuli, the sound recordings were analyzed to differentiate the sound features that correspond to a given movement quality. The description of a movement used in this study includes the dynamic qualities of Effort factors in LBMS. Prior studies have been conducted to develop methods to vocalize an appropriate sound for the Effort factors in LBMS and the vocalizations are used in dance pedagogy to help dancers learn new choreography [47]. In [47], the participants perform both movement and sound simultaneously and their goal is to train a system that generates sound from dancers movements in real time. However, our participants perform vocally in responses to viewing movement and our work aims to generate perceptually meaningful sound from the movement quality parameters.

An experiment to identify potential sound-movement correlations via vocalizations of trained musicians and composers was conducted in this research, presented in Section 3.1. Appropriate sound was supplemented to robotic movement to improve the perception of the qualities of movement. This study was also used to suggest sound-design choices for a user study in Section 4 which tested people's ability to distinguish movements based on varying qualities of movement and sound.

1.3 Research Outline

This dissertation is organized as follows. In Section 1, the background knowledge of describing movement in Laban's theory, the prior mapping of movement qualities to parameters in an optimal control problem and background of the relationships between movement qualities and sound qualities are presented. In Section 2, three methods are designed to make sure the robotic motion trajectories generated by high-level controllers are executable on physical robotic platforms. The scheme interfaces with a web-based robotic application to create various trajectories in a user-friendly way and is implemented on Rethink Robotics Baxter Research Robot. In Section 3, a data driven design is developed to improve the feature variation of robotic movement by adding appropriate sound to accompany the movement and by changing the directions of the trajectories to highlight the qualities of the movement. In Section 4, a user study is carried out for testing whether the added sound and changed trajectories help improve people's perception of the qualities of robotic movement. The quantitative analysis of the survey data are presented. Concluding remarks and future topics of the research are provided in Section 5.

2 Style-based Robotic Motion Control with System Constraints

2.1 Considering Physical Limitations of a Given Robotic Platform

This prior framework in [19] can be used to generate varied movement on many platforms. However, there is no consideration of the physical configuration and limitations of a particular robot in [19]. It is a shortcoming in that the joint limits (on position, velocity, and acceleration) can be exceeded, making the robot unable to track some requested movement trajectories with desired qualities. We need to find a way to generate safe motions for the robot with joint limits taken into consideration to avoid the damaging of the mechanical components of the robot. This allows the system to execute the user's request with the optimized trajectory subject to the constraints of the joint limits. In other words, this system ensures that the high-level requests are translated to feasible actions with respect to the low-level controllers.

Let the degree of freedom of the physical platform be given by p. Our output has the dimension of p, i.e. $l = p$. As the degree of freedom corresponds to the number of joint angles, this output will be the movement trajectory of the robot over time. Our state vector is $x = [y_1, \dot{y}_1, ..., y_p, \dot{y}_p]^T$, and $n = 2 * p$. Thus, we can keep track of the joint angles and their velocities in the state vector. The input vector is comprised of the accelerations of these joint angles, i.e. $u = [\ddot{y}_1, ..., \ddot{y}_p]^T$.

Let Y represent a given set of poses in our output space from which, in this section, we will linearly interpolate between the poses to construct a family of reference trajectories

$r = [r_1, r_2, ..., r_p]^T$. These poses and trajectories will comprise our library of primitives. Now, let the physical limits of the robot be given by

$$x_{lb} \leq x \leq x_{ub} \tag{5}$$

$$u_{lb} \leq u \leq u_{ub}, \tag{6}$$

where x_{lb} and x_{ub} are the lower and upper bounds of the state respectively; u_{lb} and u_{ub} are the lower and upper bounds of the input respectively. The methods below will take the resulting, optimized $y(t)$ from the output of the problem solved in [19] to produce a new, feasible trajectory, $y_{feas}(t)$. We pose three possible solutions to the exceeded limit problem: trimming, scaling and solving the optimal control problem with system constraints. Block diagrams of these methods are shown in Figure 1 and Figure 2 respectively.

2.1.1 Method A: Trimming

If any of the limits of one joint in robot is exceeded, we can simply trim the exceeded points to the maximum value of the range, x_{ub} and u_{ub}, or the minimum value of the range, x_{lb} and u_{lb}. This method is easy to implement, but it may significantly change the style of movement. The exceeded points on the trajectory are replaced with the maximum or minimum values, which results in a nonsmooth, flat, or 'clipped' trajectory. However, this is a computationally cheap solution that preserves greater movement variation in the feasible part of the trajectory.

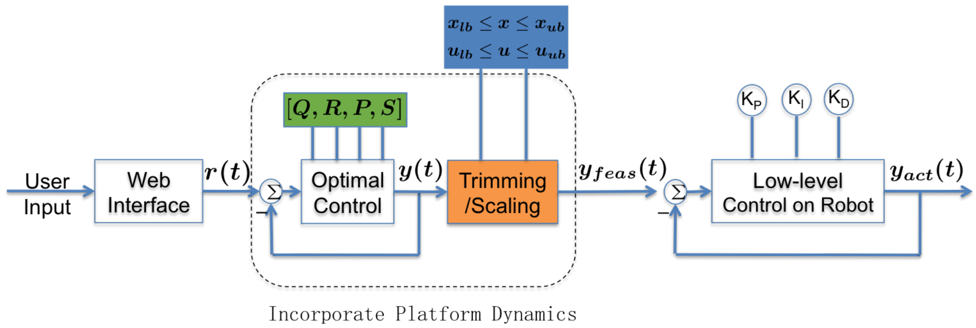


Figure 1: This figure shows the controller architecture for Methods A and B of dealing with the system constraints: trimming and scaling. In both cases, we have a bi-layer structure where user requests ($[Q, R, P, S]$) are treated independently of the system constraints ($x_{lb} \leq x \leq x_{ub}$, $u_{lb} \leq u \leq u_{ub}$).

2.1.2 Method B: Scaling

We can scale the trajectory of the joint to make it stay in the safe range. A scaling factor X_{scale} is needed for scaling the trajectories. We can tune X_{scale} to generate safe trajectories without affecting the overall trajectory shape. Scaling will reduce the range of the movement variation as all movement expression gets 'damped' or 'shrunk' in this method. This may not be desirable when a robot needs to complete a task which asks for precise execution. However, this is a computationally cheap solution that ensures some movement variation along the whole trajectory. A proper X_{scale} is calculated through an iteration to generate a feasible trajectory that is closest to the original trajectory.

2.1.3 Method C: Optimal Control with System Constraints

In order to avoid the problem of sharp change in the trajectory (as in Method A) and avoid reducing the range of movement (as in Method B), we solve the optimal control problem

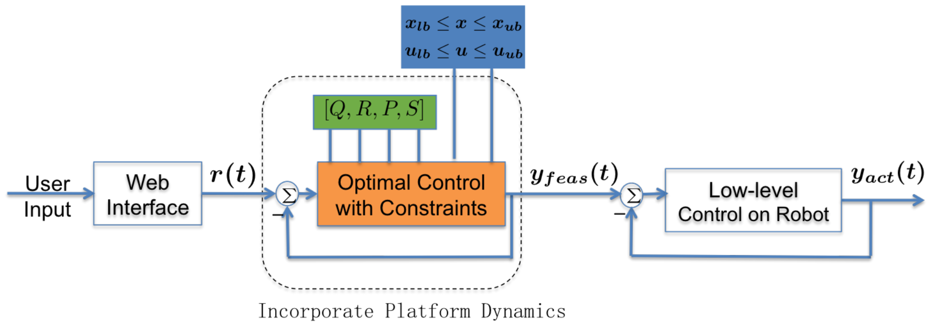


Figure 2: This figure shows the controller architecture for Method C presented. The constrained optimal control method is used to generate feasible trajectories as the reference signal for low-level controller of the robot.

with the constraints on control and state variables. The trajectory generated by solving the constrained optimal control problem will be smooth, which is better than the trajectory in Method A and will not be shrunk, which is better than Method B. The optimal trajectory will be generated to realize different functional and expressive objectives. This will enable us to create more variation in robotic movement and generate safe and feasible trajectories for the robot to execute. Constrained optimal control has been studied in [48, 49, 50, 51, 52] and used in robotic application in [53, 54, 55]. Next, we will show the procedures for solving the optimal control problem with control constraints and state constraints.

Optimal control problem with control constraints

The control u has the constraint as in (6). By absorbing the magnitude of the bounds into the matrix B, the control constraint can be written as $|u(t)| \leq 1$. As the control variable $u(t)$ is constrained in the set U and the $u(t)$ calculated by $\dfrac{\partial H}{\partial u} = 0$ may lie outside the range of the allowed control input, we can no longer take $\dfrac{\partial H}{\partial u} = 0$ [51]. For the constraints on the control inputs, we solve the optimal control problem following the steps as below.

Step 1: Find the Hamiltonian.

$$H(x(t), u(t), t) = \frac{1}{2}u^T(t)R(t)u(t) + \lambda^T(t)(Ax(t) + Bu(t)) \tag{7}$$

where the $\lambda^T(t)$ is a costate variable.

Step 2: List the state and costate equations.

Assume optimal values are $u^*(t)$, $x^*(t)$, $\lambda^*(t)$,

$$\dot{x}^*(t) = (\frac{\partial H}{\partial \lambda})^* = A(t)X^*(t) + B(t)u^*(t) \tag{8}$$

$$\dot{\lambda}(t) = -(\frac{\partial H}{\partial x})^* = -A^T(t)\lambda^*(t) \tag{9}$$

with boundary conditions $X(t_0) = X(t_0)$, $X(t_f) = 0$, t_f is either fixed or free.

Step 3: List the optimal conditions.

Using Pontryagin's minimum Principle, we have

$$\begin{aligned} H(x^*(t), \lambda^*(t), u^*(t)) &\leq H(x^*(t), \lambda^*(t), u(t)) \\ &= \min_{|u(t)| \leq 1} H(x^*(t), \lambda^*(t), u(t)) \end{aligned} \tag{10}$$

$$\begin{aligned} \Rightarrow \quad & \frac{1}{2}u^{*T}(t)R(t)u^*(t) + \lambda^{*T}(t)B(t)u^*(t) \\ & \leq \frac{1}{2}u^T(t)R(t)u(t) + \lambda^{*T}(t)B(t)u(t). \\ & = \min_{|u(t)| \leq 1} \{\frac{1}{2}u^T(t)R(t)u(t) + \lambda^{*T}(t)B(t)u(t)\} \end{aligned} \tag{11}$$

Step 4: Solve for the optimal control.

Denote $q^*(t) = R^{-1}(t)B^T(t)\lambda^*(t)$

$$u^*(t) = -SAT\{q^*(t)\} = -SAT\{R^{-1}(t)B^T(t)\lambda^*(t)\} \tag{12}$$

where

$$SAT\{f_i\} = \begin{cases} f_i, & \text{if } |f_i| \leq 1 \\ \\ sgn\{f_i\}, & \text{if } |f_i| > 1. \end{cases} \tag{13}$$

and $sgn\{f_i\}$ gives the sign of f_i.

With the optimal control (12), the state and costate equations (8) and (9) become a set of nonlinear differential equations, which can be solved by numerical simulations [51].

Optimal control problem with state constraints

We introduce slack variables to transform the optimal control problem with inequality state constraints into an unconstrained optimal control problem on state variables [51] [56].

Consider the optimal control problem which is to find $u(t)$ that minimizes the quadratic cost function

$$J = \frac{1}{2}\int_0^{T_f} [(y-r)^T Q (y-r) + u^T R u + \dot{x}^T P \dot{x}] dt + \frac{1}{2}(y-r)^T S (y-r)\Big|_{T_f} \tag{14}$$

s.t.

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{15}$$

$$S(x(t), t) \leq 0 \tag{16}$$

In this case, the constraint S is of first order. This means that the control $u(t)$ is explicitly present in the first derivative of S.

We introduce a slack variable $\alpha(t)$ as

$$S(x(t), t) + \frac{1}{2}\alpha^2(t) = 0. \tag{17}$$

Differentiating (17) with respect to time, we have

$$S'(x(t), t) + \alpha(t)\dot{\alpha}(t) = 0. \tag{18}$$

As the control $u(t)$ is explicitly present in the first derivative of S, we can solve for $u(t)$ as

$$u(t) = g(x(t), \alpha(t), \dot{\alpha}(t), t). \tag{19}$$

Treat the slack variable $\alpha(t)$ as an additional state variable, the system becomes

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bg(x(t), \alpha(t), \dot{\alpha}(t), t), \\ \dot{\alpha}(t) &= v(t), \quad v(t = t_0) = v(t_0). \end{aligned} \tag{20}$$

The new initial condition $\alpha(t_0)$ is required to satisfy (17), so we have

$$\alpha(t_0) = \pm\sqrt{-2S(x(t), t)}. \tag{21}$$

With this new initial condition (21), the original boundary conditions (17) and (18) are satisfied for any control function v. Thus the original optimal control problem with inequality constraints on state variable is transformed into an unconstrained optimal control problem.

We define the new state vector as

$$z(t) = [z_1(t), z_2(t)]^T = [x(t), \alpha(t)]^T = [M_1 \quad M_2]^T z(t). \tag{22}$$

where $M_1 = [I_{n\times n} \quad 0_{n\times 1}]$ and $M_2 = [0_{1\times n} \quad 1]$.

Thus the new system becomes

$$\dot{z}(t) = F(z(t), v(t), t) \tag{23}$$

So the cost function becomes

$$\begin{aligned} J = \frac{1}{2}\int_0^{T_f} & [(CM_1 z - r)^T Q (CM_1 z - r) \\ & + g^T(x(t), \alpha(t), v(t), t) R g(x(t), \alpha(t), v(t), t) \\ & + (M_2 z v)^T P M_2 z v] dt + \frac{1}{2}(CM_1 z - r)^T S (CM_1 z - r)\big|_{T_f}. \end{aligned} \tag{24}$$

The Hamiltonian is

$$H = \frac{1}{2}[(CM_1z - r)^T Q(CM_1z - r) + g^T(x(t), \alpha(t), v(t), t)Rg(x(t), \alpha(t), v(t), t) + (M_2zv)^T PM_2zv] + \lambda F, \tag{25}$$

where λ is a Lagrange multiplier and $F = \dot{z}$ is from (23).

Differentiating the Hamiltonian (25) with respect to the control v, the state z and Lagrange multiplier λ gives the first order necessary conditions of the control for optimality, costate equation and state equation.

The first order necessary condition can be solved by

$$\frac{\partial H}{\partial v} = 0. \tag{26}$$

The costate equation is

$$-\dot{\lambda} = \frac{\partial H}{\partial z}, \tag{27}$$

with the transversality condition

$$\lambda^T(T_f) = \frac{\partial F}{\partial x}. \tag{28}$$

The state equation is

$$\dot{z} = \frac{\partial H}{\partial \lambda}, \quad z(t_0). \tag{29}$$

The equations of the states (29) and the costates (27) and their initial and final conditions form a two-point boundary value problem. This problem can be solved with specialized software [51].

2.2 Computational Implementation

The procedures of trimming and scaling have been implemented in MATLAB and they were used as part of a computational framework named the Robot Control/Choreography Center (RCC). A screenshot of the RCC is shown in Figure 3 [1]. This is a robotic supervisory control system and graphical user interface that helps users create different feasible motion sequences, modulate the movement quality, and execute them on the robot. Part of this work resulted in the publication [1], and is included in this dissertation with some modifications and augmentations.

Controlling the robot using the RCC does not require knowledge about programming or the robot's configuration. The RCC provides the user with a graphical depiction of the robot's motion model that the user can use to pick their desired motions [1].

In this framework, the user picks the desired motion sequence and quality parameters in the web page. The web page sends the sequence and the corresponding values for $[Q, R, P, S]$ to the MATLAB implementation of the optimal control problem outlined in Section 1.2.2. The trajectory generated by solving the optimal control problem is then trimmed or scaled as described in Sections 2.1.1 and 2.1.2, to make sure the state and control constraints for the platform are satisfied. Then, the feasible trajectory is loaded through Robot Operating System (ROS) into the low-level proportional-integral-derivative (PID) controller as the reference signal and executed on robot. The PID gains are configured to get a satisfactory dynamic response. The actual output of the low-level control part is collected as the negative feedback to realize the tracking of the desired reference signal. All of the implementations in MATLAB and ROS are hidden behind the web page, and they are accessed through the web page by the user. The web page part was developed by collaborator Masoud Bashiri.

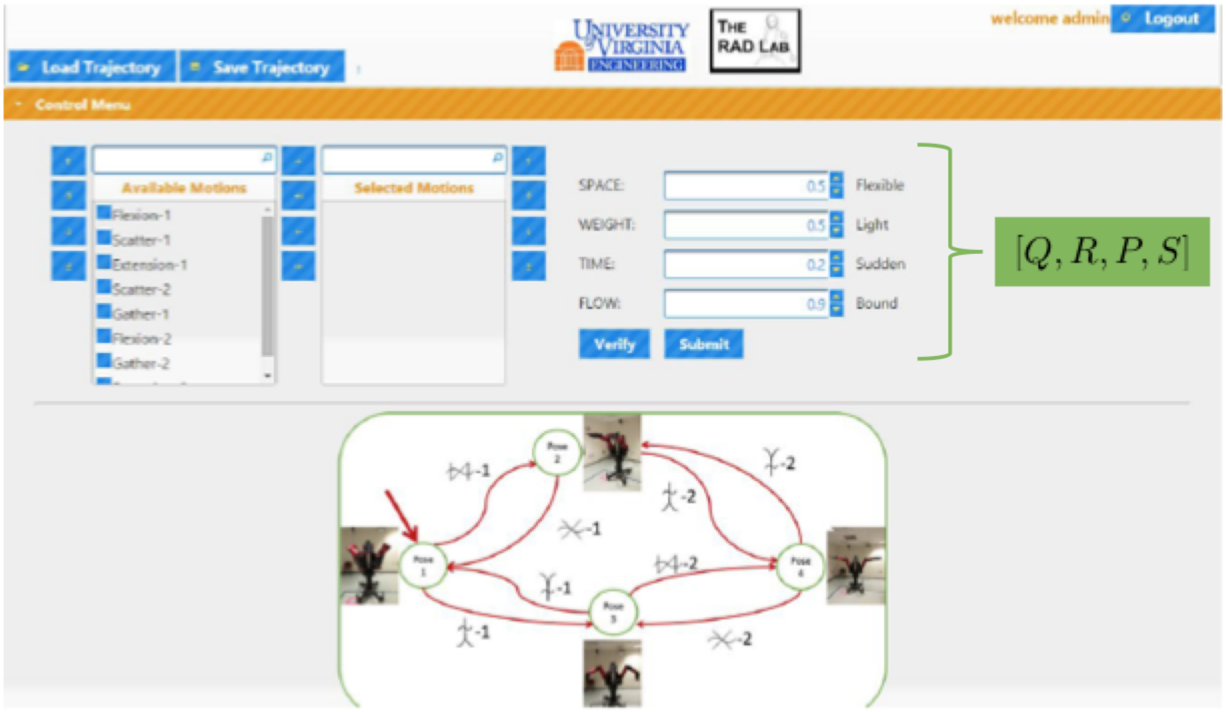


Figure 3: Screenshot of the Robot Control/Choreography Center (RCC) graphical user interface. Users select a desired motion sequence from the menu on the left. Then they pick the qualities for execution of these movements, which are mapped to the weights Q, R, P, and S in the cost function. This is done through the number entry on the right side [1].

Robot Operating System (ROS) is an open source framework for robot software development. ROS comes with client libraries for C++, Python, LISP and Java. The RCC takes advantage of ROS to communicate with different robotic platforms. By removing the Linux command line interface, it provides a user-friendly interface that interacts with separate programs declared in the application's configuration.

The finite state machine (FSM) used in the RCC shows the allowed motion transitions. The set of states consists of feasible postures that are specific to the platform to be controlled.

In this example, the system supports four different types of transitions of movement for Baxter Research Robot: Flexion, Extension, Gathering and Scattering, which are annotated by Motif symbols. The user can also manipulate the quality of movement by changing the control parameters of the controller, which correspond to Q, R, P, S in Equation (1). These four control parameters are associated with words describing specific aspects of LBMS Effort system: Space, Weight, Time and Flow. The users pose transition selections are first checked on the FSM to see if they are feasible. Next, the sequence of motion and the quality of movement are passed as matrices to the MATLAB program that implements the methods proposed in Sections 2.1.1 and 2.1.2 to generate feasible trajectories for the robot to track.

The framework for generating feasible motion for the robot was implemented in MATLAB R2014b. This implementation consists of two levels. The lower-level MATLAB function solves the optimal control problem for one step of movement transition, which generates the optimal trajectory for the robot moving from one pose to another based on the starting pose, ending pose and the choice of quality weights. The upper-level MATLAB function calls the lower-level function for each step in the desired sequence of motion. The output of the framework is the trajectory generated in the upper-level MATLAB function, which the robot will track.

2.3 Application on Rethink Robotics Baxter Research Robot

The framework was implemented on Rethink Robotics Baxter Research Robot to generate various movement in the robot with different selections of weight parameters Q, R, P and S through the RCC. The Baxter Research Robot has seven joints on each arm. The velocity limit, torque limit and range of motion are listed in Table 1. The FSM implemented in the RCC for Baxter Research Robot is shown in Figure 3. The four poses are recorded on the robot to make them physically feasible and are shown in the images surrounding the FSM in Figure 3.

Table 1: Joint limits of Baxter Research Robot [5].

Arm Joint	Torque Limit (N·m)	Velocity Limit (rad/s)	Range of Motion (rad): +limit,-limit: total movement
S0	50	1.5	+0.890,-2.461:3.351
S1	100	1.5	+1.047,-2.147:3.194
E0	50	1.5	+3.028,-3.028:6.056
E1	50	1.5	+2.618,-0.052:2.67
W0	15	4	+3.059,-3.059:6.117
W1	15	4	+2.094,-1.571:3.665
W2	15	4	+3.059,-3.059:6.117

An example of different styles of motion illustrated by joint angle positions for shoulder twist joint on the right arm of the robot is provided in Figure 4. Significantly different styles of motion of the robot arm are generated with distinct selection of quality parameters of movement. According to the prior mapping to Laban's description of movement, we can interpret the joint position trajectory for $[Q, R, P, S] = [0.1, 0.1, 0.1, 1]$ to be flexible, strong, sudden and free, while the trajectory for $[Q, R, P, S] = [10, 0.1, 1, 100]$ to be direct, strong, bound and sudden. When the physical limit is exceeded, due to the choice of start and end pose, our method can be utilized to generate feasible trajectories. Figure 5 shows the differences between the outputs (y_{feas}) of our Methods A and B in adjusting the movement trajectory to make it feasible on a particular platform. To make the robot move along the computed y_{feas}, a PID controller is used to generate the commands for the motors. In our system, the PID gains have been tuned to $K_P = 1000$, $K_I = 500$, $K_D = 10$. The traces of real movement on our robot for the two cases, where both trajectories are feasible, are provided in Figure 6.

Note that the output of the optimization may not always result in movement. For example, for the set of weights $[Q, R, P, S] = [0.1, 10, 10, 0.1]$, we are setting a cost where both terms that involve tracking the references are weighted by a tiny number (the coefficients of $y - r$) and the terms that penalize motion are relatively very large (the coefficients of $\dot{x}$ and u). This result is trivially physically feasible, given that the library of poses is feasible by design, and the particular interpretation of Equation (1) of Laban's motion factors. These weights request a very Strong and Sustained motion, and our mapping determines that the best way to express this is to not move. We pose that for limited platforms, this is the best option as such movements are hard for existing platforms to indicate differently.

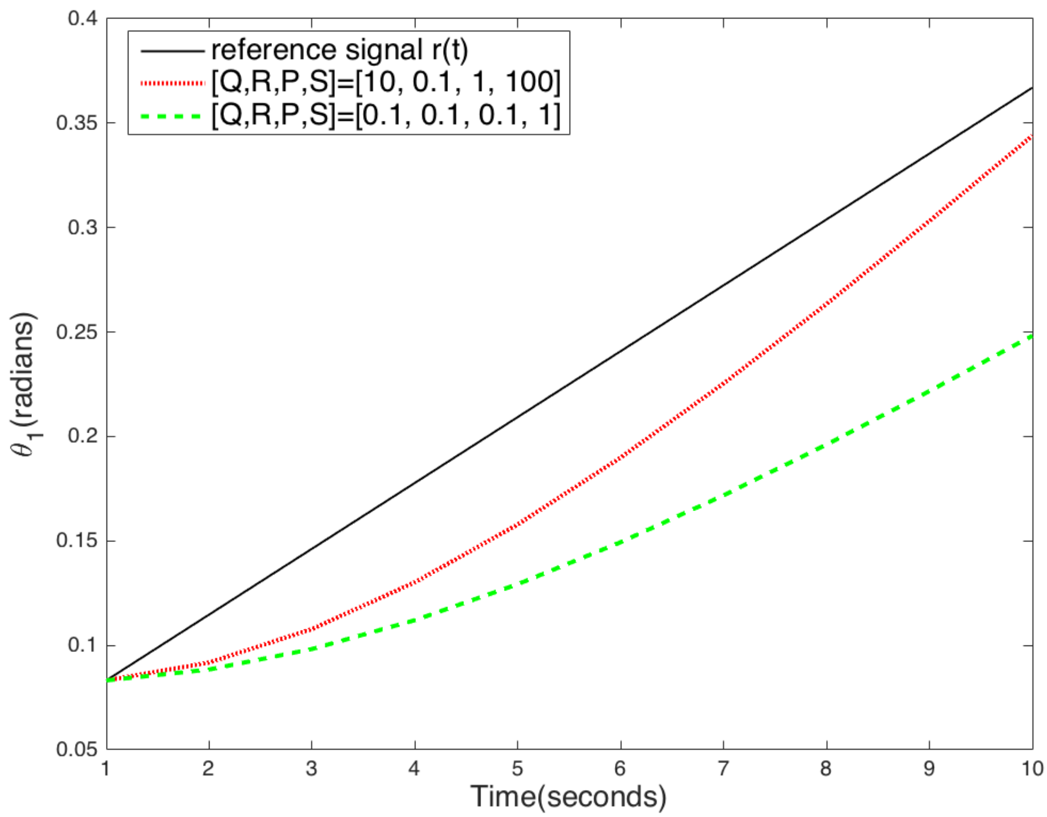


Figure 4: The reference signal from State 1 to State 2 in Figure 3 for the shoulder twist joint on the right arm of the robot and two different styles of movement with two different sets of quality parameters. The ranges of the angles and the slopes of the curves varies with different qualities. This means the joint positions and velocities are different, which leads to diverse styles of movement.

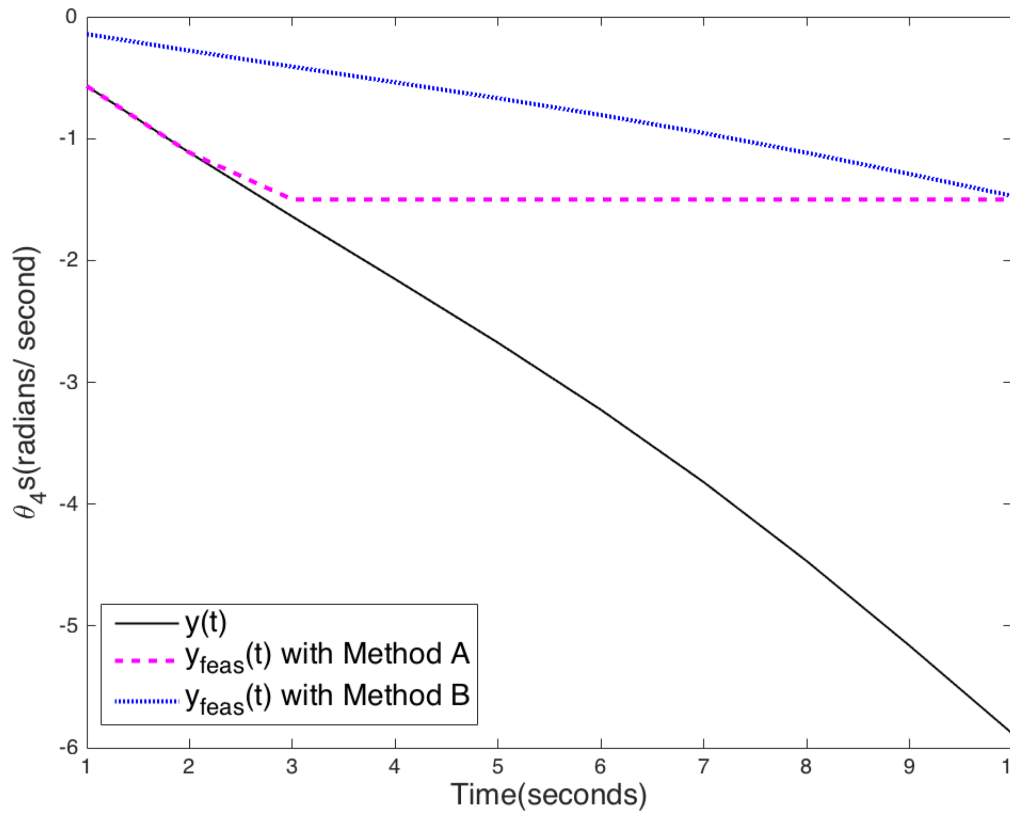
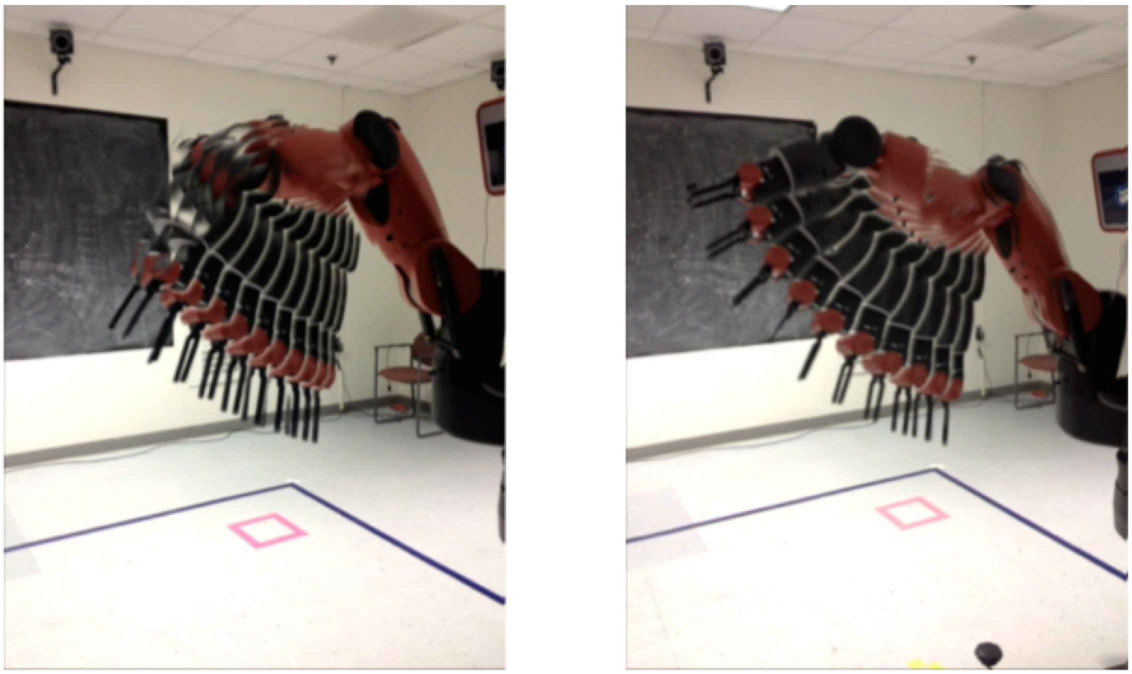


Figure 5: We implement Methods A and B proposed in Section 2.1 on the robot to show the effectiveness of our solutions. The original velocity trajectory $y(t)$ is out of the system constraints. Both methods produce a y_{feas} that is within the bound. The quality parameters used here are $[100, 1, 1, 0.1]$. The upper and lower bound of the velocity on this joint are +/-1.5 rad/s. The scaling factor X_{scale} is found to be 4.



[Q, R, P, S]=[0.1, 0.1, 0.1, 1] [Q, R, P, S]=[10, 0.1, 1, 100]

Figure 6: The traces of two different styles of movement on the robot from State 1 to State 2 in Figure 3 for feasible trajectories with two different sets of quality parameters. The figure on the left corresponds to quality parameters $[0.1, 0.1, 0.1, 1]$ (the green line in Figure 4), and the one on the right corresponds to quality parameters $[10, 0.1, 1, 100]$ (the red line in Figure 4). We can see the significant difference in the movement traces in the two cases. The one on the left with a free Flow does not hit the end pose exactly; while the one on the right with a bound Flow hit the end pose more precisely. The one on the right tracks the path more aggressively than the flexible one on the left.

3 Increasing the Perceived Variation of Movement Qualities

To further improve the variation that is algorithmically defined in a robot's movement in Section 1.2.2, we utilize the harmonic relationships between the Effort element and Spatial aspect in LBMS to generate movement trajectories with exaggerated variations. Apart from changing the trajectories, we also try to supplement the movement with perceptually driven sounds in order to improve the perception of movement qualities. These advancements will make robotic movements more expressive and related to human perception, enabling the possibility for robots to carry out more functions in diverse fields, thus, making robots better tools. Part of the work resulted in the publications [2, 3, 57], and is included in this dissertation with some modifications and augmentations.

3.1 Accompanying Movement with Sound

In this section, we study how sonic features map in accordance with features of movement to enhance the perception of quality of robotic movement. Similar work of trying to find the correspondence between movement and sound is shown in [47, 58, 59].

3.1.1 Experiment Design

This section describes an experiment we designed to find the mapping between the qualities of movement and sound. These mappings were used to create sounds intended to accompany the robotic movement and improve the perception of the qualities of these movements. Expert musicians performed non-verbal vocal sounds in time to animations of the eight Basic Effort

Actions (BEAs). The eight BEAs, dabbing, flickling, floating, gliding, pressing, slashing, thrusting, and wringing, are formed with the pairing of three of the fundamental movement qualities (Effort factors): Space Effort, Weight Effort, and Time Effort [10], as shown in Table 2. Flow Effort is not used in this study, as the movement we use can be considered a single motion. We choose the BEAs as representatives of movements because they include the typical combinations of the three Effort factors. The animations of the eight BEAs were presented to the musicians along with the BEA labels. The musicians were asked vocalize a sound that they thought matched the quality of the movement and the BEA label. Their vocalizations were recorded for analysis described in Sections 3.1.2.1 and 3.1.2.2.

Table 2: Laban's Eight Basic Effort Actions [2]

Movement	Time	Space	Weight
Gliding	Sustained	Direct	Light
Pressing	Sustained	Direct	Strong
Floating	Sustained	Indirect	Light
Wringing	Sustained	Indirect	Strong
Dabbing	Sudden	Direct	Light
Thrusting	Sudden	Direct	Strong
Flicking	Sudden	Indirect	Light
Slashing	Sudden	Indirect	Strong

These recorded sounds were given qualitative labels by the research team in order to gather initial broad correspondences between the Laban Effort factors and various sonic qualities. Then signal analysis of the recordings were performed in order to quantify various sonic qualities. This analysis informed the design of an automated sonification process that was tested with a broader audience in a user study presented in Section 4.

Five graduate students majoring in music composition and two music professionals at the University of Virginia took part in the study. They were recruited due to their experience in performing and improvising music. The participants ranged in age from 24 to 46, with a median age of 32. This study was approved by the Institutional Review Board for Social and Behavioral Research at the University of Virginia.

Different shapes of robot and the position of the core in the robot might impact people's impression of the robotic movement. As our focus of work is not on studying how the changes of shape and core affect people's feelings of robots, we used stick figure animations of human shaped robot to eliminate the possible influences of shape and core. The animations of the stick figure (Figure 7) perform a gesture of extending its arms to the side and then bringing them back to the center. This simple gesture is intended to show the different qualities of the eight BEAs.

By recording and analyzing sound vocalizations that mimic the qualities in these movements, we hope to isolate sound characteristics and accentuate how they vary according to movement qualities. The values of the parameters used in the cost function in Equation (1) for generating the eight BEAs are shown in Table 3.

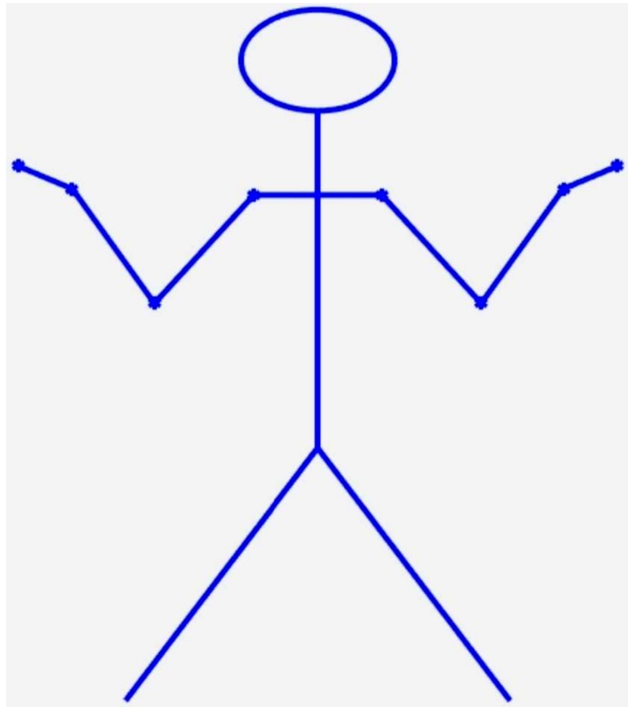


Figure 7: Picture of the visual stimulus provided to subjects [2].

Table 3: The values of Q, R, P, and S utilized in the study [2]

Movement	Q	R	P	S
Gliding	1000	100	100	1000
Pressing	100	0.1	100	100
Floating	0.1	100	100	1000
Wringing	0.1	0.1	100	100
Dabbing	100	100	0.1	1000
Thrusting	100	0.1	0.1	100
Flicking	0.1	100	0.1	1000
Slashing	1	1	1	100

To generate the animation of a stick figure moving its arms with different styles, we implemented in MATLAB the optimal control problem of minimizing the cost function (1) with different values of Q, R, P and S, subject to the system (3) and (4) in Section 1.2.2. The stick figure is designed to have shoulders, elbows and wrists, so we consider a 6-dimensional system described in Equations (3) and (4) with the state $x = [\theta_1, \dot{\theta}_1, \theta_2, \dot{\theta}_2, \theta_3, \dot{\theta}_3]$ and the input $u = [u_{\theta_1}, u_{\theta_2}, u_{\theta_3}]$. These are related by standard double integrator linear matrices (A, B, C). The parameters θ_1, θ_2 and θ_3 are the joint angles in the shoulder, elbow and wrist in each arm of the stick figure, shown in Figure 7.

$$\dot{x} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} u \tag{30}$$

$$y = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} x \tag{31}$$

Participants were shown each of the eight animations, as well as the label of the BEA for the animation (i.e. gliding, pressing, etc.). Then participants were asked to vocalize a sound for each animation, such that their vocalization began at the start of the movement, and lasted the duration of the movement. The videos have consistent durations and the recorded vocalizations were aligned with the videos in time. A three-second countdown was added before each movement to help with timing. Participants were allowed to practice before recording and record up to three takes, indicating the take that they felt best represented the movement.

A custom software interface with Max/MSP was built by collaborator Dr. Jon Bellona to display the animations and record the participants vocalizations. Participants were recorded in an isolated studio environment using a Neumann TLM103 microphone with pop filter and a Focusrite Scarlett 2i2 audio interface for microphone pre-amplification and analog-to-digital conversion. This set up provided a high-quality sound recording for analysis.

3.1.2 Mapping between Qualities of Movement and Sound

In this section, the qualitative and quantitative analysis of the musicians' vocalization recordings are shown to find the mapping relationship between the qualities of movement and the qualities of sound.

3.1.2.1 Qualitative Analysis of the Recorded Vocalization

In order to understand the sound-movement relationships embedded in the musicians' vocalizations, we performed a qualitative encoding of various qualities in the audio recordings. Four members of the research team listened to and applied labels to each of the 56 recordings (7 musicians $\times$ 8 BEAs).

These labels are used to describe the following qualities in each sound. The overall Pitch of the sound is described using the labels *very low, low, medium, high, very high, none.* The overall Amplitude of the sound uses the labels *very soft, soft, medium, loud, very loud.* The overall Timbre of the sound uses *dark tone, dark noise, medium tone, medium noise, bright* *tone, bright noise.* We also apply labels to the shape of how these qualities vary over the duration of the sound. The qualities of Pitch Curve, Amplitude Curve, and Timbre Curve are all described using the labels *start emphasis, middle emphasis, end emphasis, linear increase, linear decrease, sustained, oscillating.*

The movements and sounds were organized according to the BEAs. However we are interested in understanding how the sonic qualities vary with the Effort factors. Thus we organized the labels according to the dimensions of Weight Effort, Time Effort, and Space Effort, and plotted the label counts for each Effort factor in the form of bar graphs, as shown in Figures 8, 9, and 10. For example, the label counts for Sustained Time are the sum of the labels for the first four BEAs, and the label counts for Sudden Time are from the last four BEAs, as seen in Table 2.

Weight: By examining Figure 8, we see interesting relationships between the movement quality of Weight Effort (Light vs. Strong), and the sonic qualities of amplitude and timbre. For example, sounds that corresponded to BEAs with Light Weight Effort were more often labeled with soft amplitude, whereas sounds corresponding to BEAs with Strong Weight Effort were more often labeled medium or loud, suggesting a Weight Effort to amplitude correlation.

For timbre, Strong Weight sounds contained more dark tone and dark noise labels, whereas Light Weight Effort sounds had more mid and bright tone labels. In addition, from the three curves we see that Strong Weight Effort sounds tend to have more end-emphasis, while Light Weight Effort sounds tend to have more sustained and middle emphasis labels.

Time: For the motion factor of Time Effort (Sudden vs. Sustained), a few qualitative correlations were observed. The vocalizations for Sustained Time Effort movements received

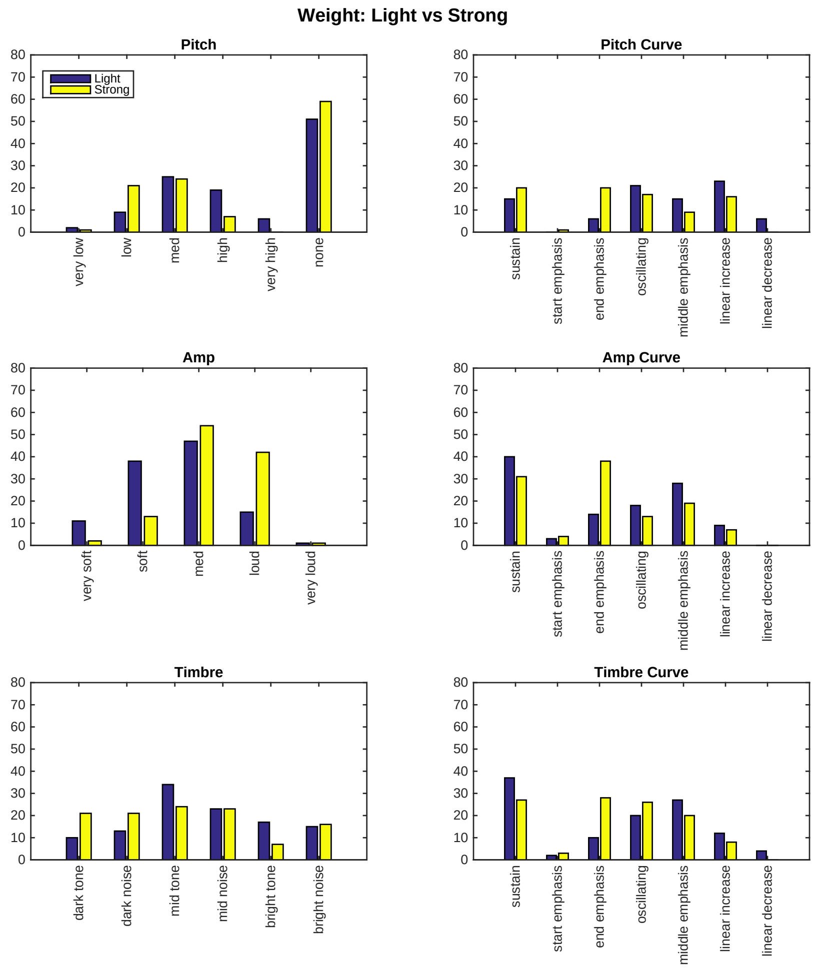


Figure 8: Label counts of the encodings for Weight Effort: Light vs. Strong [2].

more low and medium pitch labels, while Sudden Time Effort movement vocalizations received more labels for "none" pitch, suggesting that sounds of Sudden Time Effort may be less pitched. This is corroborated by the timbre labels where we see that vocalizations for Sustained movements more often have mid tone, whereas Sudden movements are more often vocalized with bright noise.

For sound amplitude, we see that Sudden movement vocalizations were more often labeled as loud, whereas Sustained movement vocalizations had a higher number of soft and medium amplitude labels. In addition, Sudden Time Effort movement vocalizations tend to have

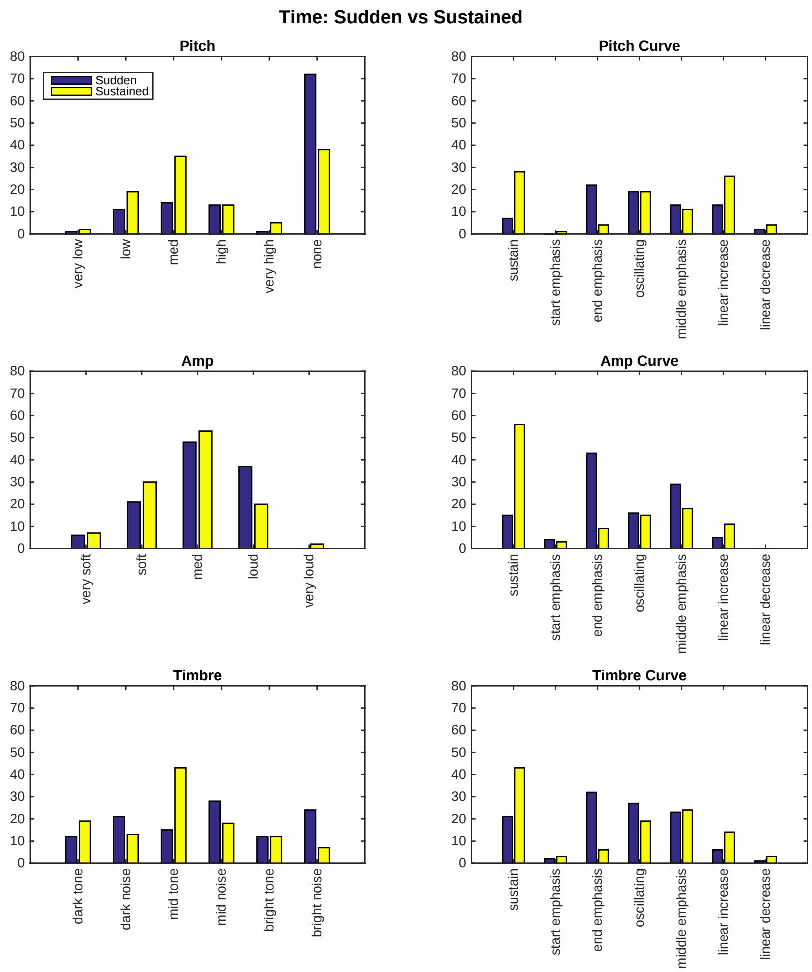


Figure 9: Label counts of the encodings for Time Effort: Sudden vs. Sustained [2].

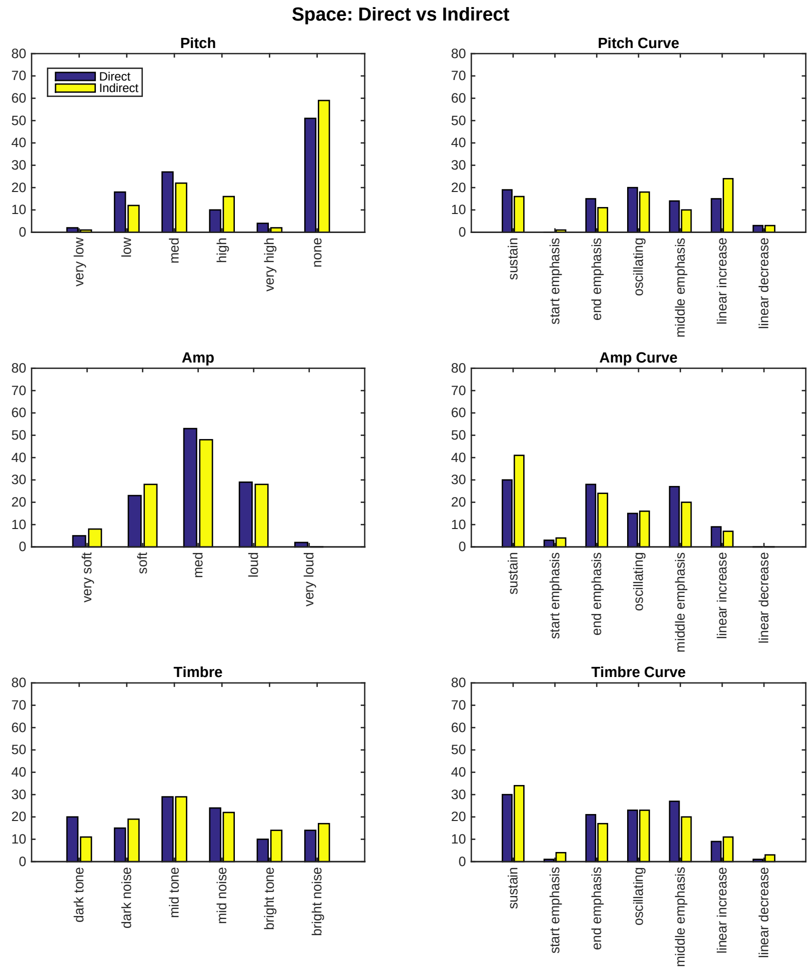


Figure 10: Label counts of the encodings for Space Effort: Direct vs. Indirect [2].

more end- or middle-emphasis labels whereas Sustained Time Effort movement vocalizations contained more sustained curve labels. See Figure 9.

Space: We were not able to discern any clear differences between the sound labels for Direct and Indirect Space Effort movements. The results are plotted in Figure 10. We will discuss this lack of a meaningful result in Section 3.1.2.2.

3.1.2.2 Quantitative Analysis of the Recorded Vocalization

Next we performed a quantitative analysis of the recorded vocalization using MIRtoolbox [60] and studied the mapping of the qualities between movement and sound. The correspondences were used to design a system for parametrically generating sounds to accompany a movement according to its values for Space Effort, Weight Effort and Time Effort [2, 57].

To extract audio features from each recording, the MIRtoolbox [60] library was used. The following audio features were extracted: amplitude envelope, spectral brightness, spectral centroid, spectral roll-off, spectral flatness, and zero-crossing rate. The amplitude envelope is roughly analogous to the changing loudness of a sound over time [3]. Spectral brightness, spectral centroid and spectral roll-off are features that characterize the distribution of energy in the spectrum of a sound. Brightness calculates the percentage of energy above 1500 Hz, centroid calculates the centroid of the magnitude spectrum and roll-off calculates the frequency below which 85% of the energy lies [3]. Spectral flatness is also a measure of the amount of noise in a digital signal, since the magnitude spectrum of white noise is flat [3]. The time-domain zero-crossing rate is used in speech processing to distinguish voiced sound from unvoiced sound [3], and is considered a measurement of the "noisiness" of a signal [61]. The value of each feature for each recording was calculated. As the vocal range of each participant is different, we subtracted the mean of the feature across all recordings of the participant from the recording means we calculated. This subtraction applies to each participant.

The recordings were grouped according to the Effort factors: Space Effort, Weight Effort, and Time Effort to test how the qualities of sound vary with the Effort factors. For example, in Figure 11, it shows that the mean spectral rolloff for movements with Sudden Time Effort is higher than for movements with Sustained Time Effort. To test whether this difference is meaningful, we conducted T-tests between the two values of each Effort Factor (i.e. between Direct and Flexible Space Effort, between Strong and Light Weight Effort, and between

Sudden and Sustained Time Effort). This analysis includes the following sonic features: spectral brightness, roll-off, and centroid, zero-crossing rate and spectral flatness, amplitude envelope and peak, and entropy of envelope and flux envelopes. The details of the analysis is shown in [3]. The p values for statistically significant findings in comparing Effort factors to sound qualities are shown in Table 4. Our significant findings are for Time Effort and Weight Effort.

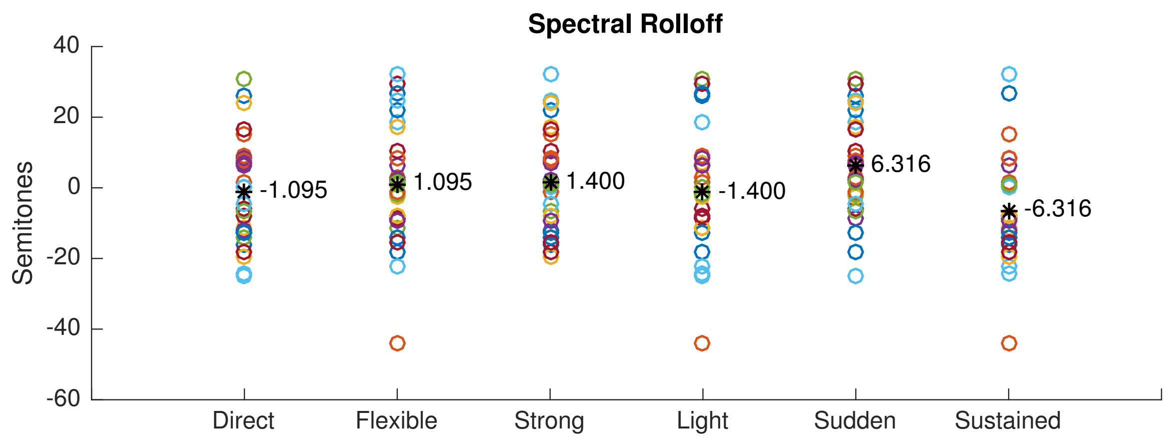


Figure 11: Spectral Rolloff for the Space, Weight, and Time Effort Factors [3].

The correspondences of the qualities of movement and sound are described in detail in [3] and are summarized here.

For Time Effort, both qualitative and quantitative analyses find that: Sudden movements are associated with brighter sounds, whereas sounds for Sustained have darker timbres. Sudden movements are associated with noisier sounds, whereas sounds for Sustained movements are more pitched. The sounds for Sustained movements have smooth amplitude envelopes and smoothly varying timbre, whereas the sounds for Sudden movements tend to have moments of strong emphasis both in amplitude and timbre. Qualitative analyses also find that sounds for Sudden movements tended to be louder than for Sustained movements, but it was

Table 4: P values for significant findings comparing Effort factors to sound qualities [3].

Feature	Effort Factor	P value	Result Summary
Amp. Envelope Entropy	Time	< 0.001	Sudden sounds tend to contain strong peaks
Spectral Flux Entropy	Time	< 0.001	Sudden sounds tend to contain peaks of intense change
Brightness	Time	< 0.01	Sudden sounds tend to be brighter
Spectral Centroid (Log Hz)	Time	< 0.01	Sudden sounds tend to contain higher frequencies
Spectral Flatness	Time	< 0.01	Sudden sounds tend to be noisier
Spectral Roll-off (Log Hz)	Time	< 0.01	Sudden sounds tend to contain more high frequencies
Zero-crossing Rate (Log Hz)	Time	0.0198	Sudden sounds may be noisier
Envelope Peak (dB)	Weight	0.0261	Strong sounds may have louder peak values

neither confirmed nor contradicted by quantitative analysis.

For Weight Effort, both qualitative and quantitative analyses find that Strong movements are associated with sounds that are louder and have higher peak amplitudes, whereas sounds for Light movements are quieter and with smaller peak amplitudes. The qualitative analysis suggested that sounds for Light movements tend to be brighter and more pitched, compared to those for Strong movements which are darker and more noisy, but it was neither confirmed nor contradicted by quantitative analysis. Qualitative analyses also suggested that Strong movements have more end-emphasis, while Light movements have more sustained amplitude envelopes and more middle-emphasis.

However, we did not find any conclusive results with respect to Space Effort and sound qualities. The reason could be that the concept of space in movement does not map precisely onto the qualities of sounds we studied here. Our study focuses on monophonic and spatially static sound, instead of the spatial placement of sound which is a creative dimension of music composition [3].

Based on the qualitative analysis of labeling the sonic features that may map to the movement qualities and the quantitative analysis of the recordings of vocalizations presented here, a sound synthesis application was developed using Max/MSP software by collaborator Dr. Jon Bellona. This application was designed to generate the perceptually meaningful sound to accompany the robotic movement with matching qualities [57]. The mappings of the qualities of Weight and Time to the sound features were used in the application. As the Space Effort indicates the mover's attitude toward the surroundings, which is related to the spatial domain, the stereo field and polyphonic sound were included in the application for the quality of Space.

Our initial experimental design was to use videos of an actual robot, the Baxter Research Robot. However, this platform has limits of maximum velocity and torque which can be used to generate quality of motion (and which the simulation does not suffer from). The torque limit, velocity limit and range of motion are listed in Table 1. The "muted" quality visible on the physical platform, shown in Figure 6, was found to be not appropriate for the experiment described here. Thus, the following section describes an improvement to the method for generating movement. In particular, it creates more complex spatial pathways which may create more visible qualitative differences in the robotic movement trajectories.

3.2 Utilizing Affinities between Effort and Space for Reference Trajectories

In the field of Laban Movement Analysis, there are close relationships between the three Effort elements (Space, Weight and Time) and Spatial aspect of a movement (3 dimensions of Space) [6]. The Effort quality of Flow affects how we change the energy in movement progression from one movement to another and it does not depend on Space. These harmonic relationships between Effort and Space[1] come from the idea of Part/Whole duality in LBMS through people's experiences of moving [6]. For example, a Light Weight Effort is usually linked with a high level in Space – imagine waving to a friend or lifting something off of a very high shelf, both of these examples typically require a rarefaction of Weight Effort. In this section, the Space system and the affinities between the Effort and Space are presented. In this work, we leveraged these affinities to improve the motion design for different styles

[1]Not to be confused with Space Effort, which is a subcomponent of the Effort category, while Space is its own category dealing with patterns in the space around a mover.

in robotic movement.

The Space component of LBMS describes the direction, level and pathway of a particular movement. There are 26 directions of the actions from Place Middle, which consist of combinations of three different levels (high, middle and low), three horizontal directions (left, middle and right) and three sagittal directions (forward, place and backward). The affinities between the qualities in Effort and the directions in Space are listed in Table 5. The Weight Effort Factor shares an affinity with the vertical dimension; the Time Effort Factor shares an affinity with the sagittal dimension; and the Space Effort Factor shares an affinity with the horizontal dimension [6].

Table 5: Affinities between Effort and Space [6].

Effort Element	Spatial Aspect
Light Weight	High
Strong Weight	Low
Sustained Time	Forward
Sudden Time	Backward
Indirect Space	Side Open
Direct Space	Side Across

Effort and Space are two components of movement. Effort describes the qualities of the movement and Space describes the relationship between the movement and the environment [6]. The harmonic relationships between Effort and Space help people perceive movement as a whole [6] and improves people's perception of the qualities of movement. The affinities between the qualities of each BEA and the directions in Space are listed in Table 6.

Table 6: Affinities between Effort and Space for each BEA

BEAs	Effort Element	Spatial Aspect
Dabbing	Direct Space	Side Across
	Light Weight	High
	Sudden Time	Backward
Flicking	Indirect Space	Side Open
	Light Weight	High
	Sudden Time	Backward
Floating	Indirect Space	Side Open
	Light Weight	High
	Sustained Time	Forward
Gliding	Direct Space	Side Across
	Light Weight	High
	Sustained Time	Forward
Slashing	Indirect Space	Side Open
	Strong Weight	Low
	Sudden Time	Backward
Thrusting	Direct Space	Side Across
	Strong Weight	Low
	Sudden Time	Backward
Pressing	Direct Space	Side Across
	Strong Weight	Low
	Sustained Time	Forward
Wringing	Indirect Space	Side Open
	Strong Weight	Low
	Sustained Time	Forward

In Equation (1), the reference trajectory r is the linear interpolation between the initial and final position. To further improve the perceived variation created with the method described in Section 1.2.2, we utilize the affinity relationships between the elements of Effort and Space in LBMS. For example, if the movement has Light Weight Effort, we may adjust the trajectory by making it higher in Space to enhance the perception of the Light quality. The interpolation equations we use to generate a new r for Equation (1) are as follows, demonstrated in Figures 12 and 13.

Acceleration interpolation:

$$r = y_0 - k(y_0 + y_1)t + (k-1)y_0t^2 + (k+1)y_1t^2 \tag{32}$$

where y_0 is the initial position of the trajectory, y_1 is the final position of the trajectory, and k is the coefficient that is associated with the quality in Effort. For example, bigger R in Equation (1) gives Lighter Weight Effort, which corresponds to higher vertical position in Space, that needs bigger k in Equation (32).

Deceleration interpolation:

$$r = y_0 - k(y_0 + y_1)t + (k-1)y_0[1 - (1-t)^2] + (k+1)y_1[1 - (1-t)^2] \tag{33}$$

where y_0 is the initial position of the trajectory, y_1 is the final position of the trajectory, and k is the coefficient that is associated with the quality in Effort.

Figure 14 shows trajectories of the reference signal $r(t)$ with different choices of parameter k. We can see that with larger k, the reference trajectory has exaggerated spatial differences. The evaluation of whether the differences can be perceived by human viewers will be presented in Section 4.

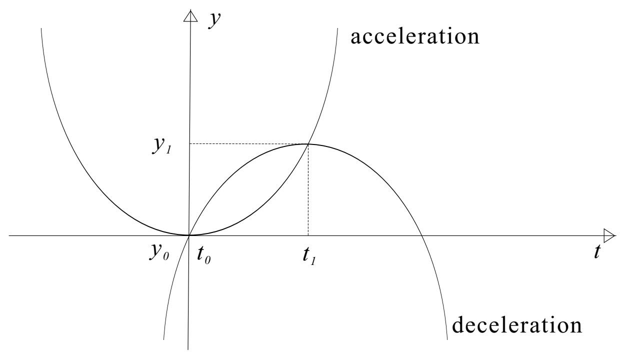


Figure 12: Demonstrations of interpolation curves for reference trajectory when initial position y_0 is lower than the final position y_1 [2].

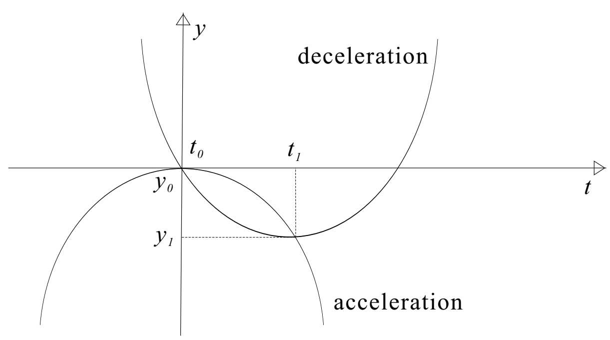


Figure 13: Demonstrations of interpolation curves for reference trajectory when initial position y_0 is higher than the final position y_1 [2].

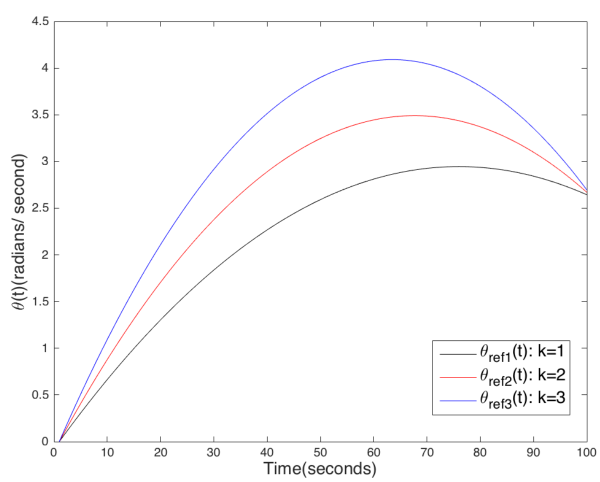


Figure 14: Different reference trajectories with different choices of parameter k [2].

4 Evaluating the Methods with User Study

User studies have been conducted in various applications of human-robot-interaction [62, 63, 64]. In order to test whether the greater variety we created in robotic movement using the spatial affinities of LBMS can be noticed by people and whether the differences people notice help them perceive the robotic movements in terms of the quality of the movements, we conducted a user study. The user study also tested whether the sounds we created to accompany the movement makes it easier for people to perceive the qualities of robotic movements. As discussed in Section 3.1, the sounds added to the movements were generated based on our mapping of qualities between movement and sound.

The participants of the user study were recruited broadly from the engineering school and music department at the University of Virginia and only those above age 18 qualify. The user study was approved by the Institutional Review Board for Social and Behavioral Research at the University of Virginia.

This section describes the experiment design which consists of three hypotheses for the user study to evaluate the results and analysis in Sections 3.1 and 3.2.

4.1 Experiment Design of the User Study

The participants are asked to take two surveys: one survey shows videos of animated stick figure and the other one shows videos of Baxter Research Robot. Both surveys consist of 8 pre-survey questions about the age, gender and background of music and dancing training of the participants and 48 questions of perceiving the qualities of robotic movements. These questions are two-alternative forced choice. In each question, participants were asked to watch two videos and indicate which one is stronger or more direct or more sudden. Some of the videos are with computer-generated sound and some are without sound. Each video of stick figure animations lasts $3s$ and each video of Baxter Research Robot lasts $4s$. The durations of videos are designed to be between $0.5s$ and $5s$, which is the optimal duration for retaining in people's short-term memory [39].

The videos of stick figure animations were created in MATLAB. The start and end poses of the movement is defined by specifying the angles of each joint in joint space. As there are no physical limitations for the stick figure in the animations, the trajectories in between was generated with the optimal control method described in [19]. The videos of Baxter Research Robot were recorded with the robot moving. We first manually moved the robot arm to start and end poses, and recorded the joint angles for these poses. Then we loaded these recorded joint angles in MATLAB and generated the trajectories between the poses of each joint with the method of scaling in Section 2. These trajectories were then loaded to ROS for the robot to execute. Screen shots of videos of stick figure animations and Baxter research robot are shown in Figures 15 and 16.

The process repeats for each pair of the BEAs on each edge of the cube of Effort factors. In the cube of Effort factors in Figure 17, the two BEAs on each edge of the cube share two of the Effort factors and the third Effort factor is different. For example, on one edge of the cube, Dabbing and Flicking, share the same Weight Effort and Time Effort (Light Weight and Sudden Time) and have opposite Space Effort: Dabbing has Direct Space while Flicking has Indirect Space. Our survey is designed to test how well subjects can distinguish the differences of the Effort factors and whether the methods of adding sound and adding spatial affinities help subjects better distinguish these differences.

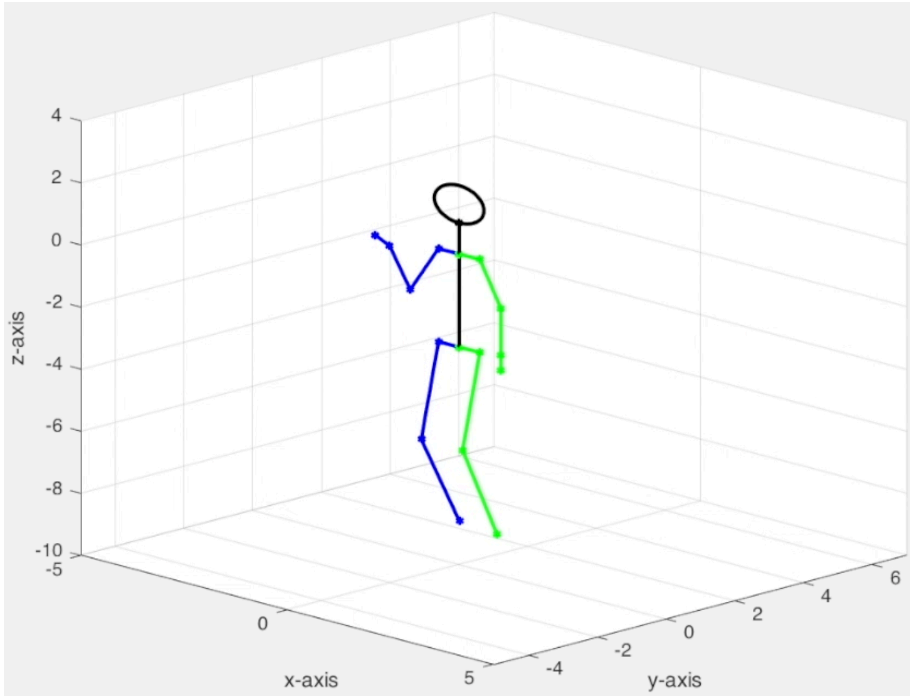


Figure 15: Screenshot of stick figure animation shown in the survey.

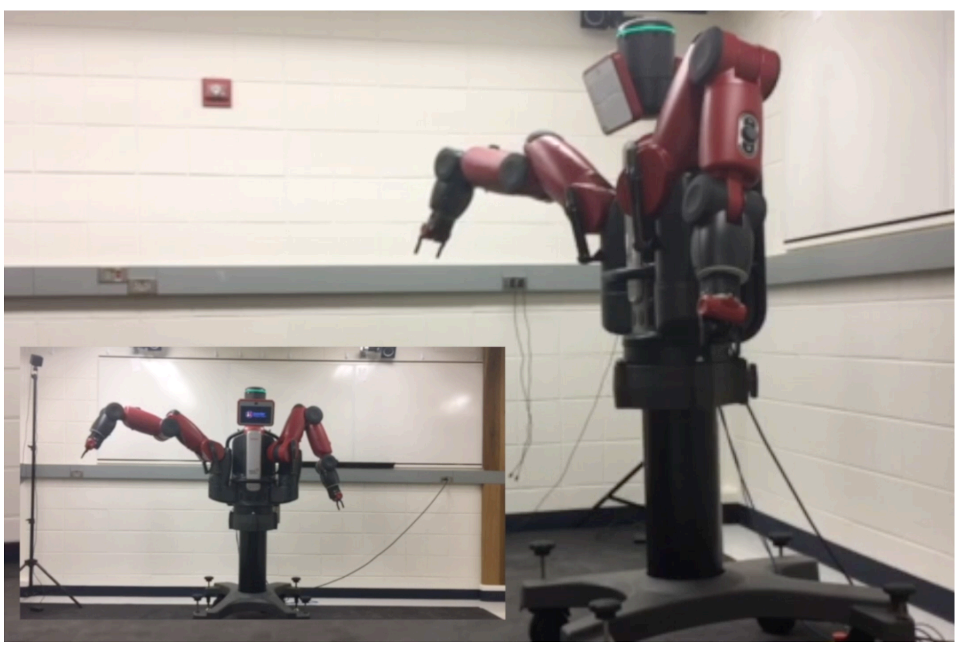


Figure 16: Screenshot of video of Baxter Research Robot shown in the survey.

In our survey, subjects were asked to watch videos of the 12 pairs of BEAs, which correspond to the 12 edges of the dynamosphere in Figure 17. As the two BEAs on one edge share two same Effort factors and have the third one different, we ask the question to test people's ability to distinguish the difference of the third Effort factor. The 12 questions can be divided into 3 categories: 4 questions for Space Effort (corresponding to the pair of BEAs on the four edges that are horizontal in orange in Figure 17), 4 questions for Weight Effort (corresponding to the pair of BEAs on the four edges that are vertical in blue in Figure 17) and 4 questions for Time Effort (corresponding to the pair of BEAs on the four edges that are sagittal in green in Figure 17).

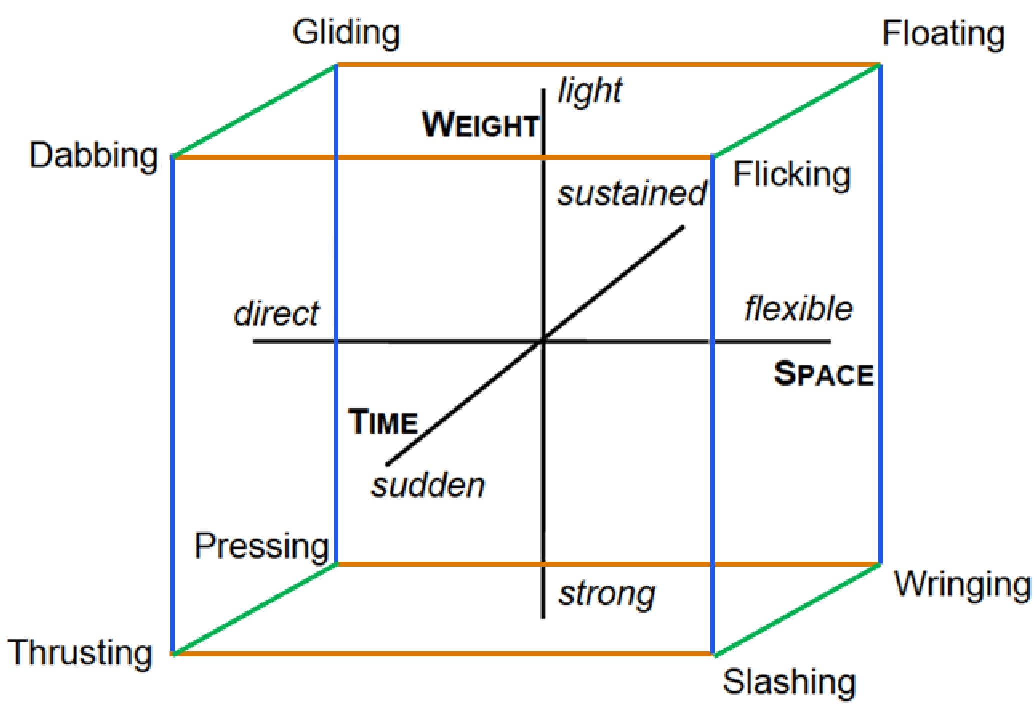


Figure 17: The Dynamosphere: the space that one can reach with Effort factors: Space, Weight, and Time. The eight vertices show the eight Basic Effort Actions. Each BEA consists of a combination of the three Effort factors [4].

This process repeats for the following four cases: (1) the case of using linear interpolation for the reference trajectory of the robotic movements, (2) the case of using linear interpolation for the reference trajectory accompanying with sound, (3) the case of using nonlinear interpolation for the reference trajectory with spatial affinities of the robotic movements, and (4) the case of using nonlinear interpolation for the reference trajectory with spatial affinities and accompanying with sound. There are 12 questions for each case, which gives 48 questions for the four cases. The involvement for each participant takes around 20-30 minutes for each survey.

The research hypotheses [65] we tested in the surveys are as follows.

Hypothesis I: Utilizing the spatial affinities in generating the reference trajectories of robotic movement helps people perceive the qualities of the movement, compared to the case of using linear interpolation for the reference trajectory.

Hypothesis II: Adding appropriate sound to the movement makes it easier to identify movement qualities, compared to the cases without sound. We assume that our mappings of qualities of movement and sound are reasonable.

Hypothesis III: The work that has been done in this dissertation (adding perceptually meaningful sounds and utilizing the spatial affinities to generate the reference trajectories with more varieties) makes it easier to identify the movement qualities.

The experiment of the user study is designed as follows. We show the subjects two videos of movements with different qualities of Space Effort / Weight Effort / Time Effort. For the questions aiming to test subjects' perception of the quality of Space, one of the two videos shows a stick figure moving with one set of qualities with Direct Space Effort. The other video shows a movement that has Indirect Space Effort. The qualities of Weight and Time are the same for the two videos. Some of the videos are generated using linear interpolation, while others using nonlinear interpolation with spatial affinities. Some of the videos have appropriate sounds accompanied to it, while others have no sound accompanied. Then the subjects are asked to pick which video matches better to the described quality of Space in the question (Stronger/Lighter). The questions aiming to test subjects' perception of the quality of Weight and Time are designed with similar process. The sequence of the questions to show the subjects is randomized. We carried out the survey for both the videos of stick figure animations and Baxter Research Robot.

It is worth pointing out that the context in which a movement is carried out might affect how people feel about the movement. However, the questions asked in this user study do not contain information of the context of the movements, as our work does not focus on finding the impact of context in peoples perception of the qualities of robotic movement.

The analysis includes comparing subjects responses across the cases in which there is no sound or there is sound and across the cases of using linear interpolation or nonlinear interpolation with spatial affinities.

Twenty two students in the Engineering School and the Department of Music at the University of Virginia took part in the study. The participants have the ages ranging from 19 to 39, with a median age of 27. Five participants currently play an instrument and two participants have formal training in dance.

4.2 Quantitative Analysis of the Survey Data of Stick Figure Animations

As the questions in our survey are two-alternative forced choices, the chance level of answering each question correctly is 50%. To see whether the subjects were randomly choosing the answers and whether the accuracies of using the different methods are higher than chance level (50%), we calculated the proportions of getting the correct answer (accuracies of perception) and 95% confidence interval of each method for each question in the survey. First, we calculated the proportion for each question to get the observed accuracies in our survey. Then we used a binomial test to compare these proportions with the chance level (50%) to determine whether the accuracies are higher than random choices. If the test result shows that the accuracy of answering a question is higher than random choice, we may conclude that the method used in the question helped subjects to perceive the quality of movement.

Next, we used generalized linear mixed-effects model [66, 67] to test the comparisons across the methods to see whether the proportions (accuracies of perception) are statistically different and then to determine whether the methods of using spatial affinities and adding the sound we designed are more effective in improving the accuracies of perception of qualities of robotic movement than the methods of linear interpolation and moving without sound. The reason we used the generalized linear mixed-effects model is that this test allows much flexibility. The flexibility is needed in our study since the data in the survey is not continuous, and it is not from independent populations, as one participant responds to each of 48 questions in the survey, instead of one participant providing a response to only one question.

4.2.1 Accuracies of Perception of the Quality of Space in Stick Figure Animations

In this section, we analyze the data of the questions aimed to test the perception of the quality of Space in stick figure animations: Dabbing vs. Flicking, Slashing vs. Thrusting, Gliding vs. Floating and Wringing vs. Pressing.

The accuracies of the responses of these four questions are shown in Table 7. All four methods have observed proportions (accuracies) higher than 50% in our survey.

Table 7: Accuracy of subjects' perception of the quality of Space in stick figure animations

Method	Correct	Incorrect	Proportion
Linear interpolation	51	37	0.58
Linear interpolation + sound	46	42	0.52
Spatial affinities	56	32	0.64
Spatial affinities + sound	47	41	0.53

To see whether the estimated accuracies are higher than chance level, we need to calculate the confidence intervals for each question and check whether they have lower bounds above 50%. If a lower bound of confidence interval is above 50%, we may conclude that the corresponding method helped the participants of our survey perceive the quality of movement tested in that question. The detailed results of each question aimed to test the perception of Space in the stick figure animations are shown in Tables 8, 9, 10 and 11.

Table 8: Accuracy of subjects' perception of Dabbing vs. Flicking in stick figure animations

Method	Correct	Incorrect	Proportion	95% confidence interval
Linear interpolation	13	9	0.59	[0.36 0.79]
Linear interpolation + sound	12	10	0.55	[0.32 0.76]
Spatial affinities	13	9	0.59	[0.36 0.79]
Spatial affinities + sound	8	14	0.36	[0.17 0.59]

For the proportions of the four methods for "Dabbing vs. Flicking" in Table 8, the perception of the quality of Space with the first three methods have observed accuracies (the fourth column in the table: Proportion) higher than 50%, but the confidence intervals all cross 50%. So we can't conclude whether the method of spatial affinities and adding sound are effective in improving the perception of the quality of Space or not for the pair of "Dabbing vs. Flicking".

Table 9: Accuracy of subjects' perception of Slashing vs. Thrusting in stick figure animations

Method	Correct	Incorrect	Proportion	95% confidence interval
Linear interpolation	11	11	0.5	[0.28 0.72]
Linear interpolation + sound	4	18	0.18	[0.05 0.40]
Spatial affinities	10	12	0.45	[0.24 0.68]
Spatial affinities + sound	8	14	0.36	[0.17 0.59]

For the proportions of the four methods for "Slashing vs. Thrusting" in Table 9, the perception of the quality of Space of all four methods have observed accuracies lower than or equal to 50%. The confidence interval of "Linear interpolation + sound" has a higher bound lower than 50%. Thus we might conclude that adding sound made it more difficult for subjects to pick the correct one for the pair of "Slashing vs. Thrusting". As the other three confidence intervals cross 50%, which means that both higher and lower than 50% are possible. So we can't conclude whether the method of spatial affinities made the perception of the quality of Space easier or more difficult for the pair of "Slashing vs. Thrusting".

Table 10: Accuracy of subjects' perception of Gliding vs. Floating in stick figure animations

Method	Correct	Incorrect	Proportion	95% confidence interval
Linear interpolation	12	10	0.55	[0.32 0.76]
Linear interpolation + sound	15	7	0.68	[0.45 0.86]
Spatial affinities	16	6	0.73	[0.50 0.89]
Spatial affinities + sound	12	10	0.55	[0.32 0.76]

For the proportions of the four methods for "Gliding vs. Floating" in Table 10, the perception of the quality of Space of all four methods have observed accuracies higher than 50%. For the method of spatial affinities, the confidence interval (shown in red) has a lower bound equal to 50%, which means that the estimated accuracy is at least 50%. We might conclude that the method of spatial affinities is effective in improving the perception of the quality of Space for the pair of "Gliding vs. Floating".

Table 11: Accuracy of subjects' perception of Wringing vs. Pressing in stick figure animations

Method	Correct	Incorrect	Proportion	95% confidence interval
Linear interpolation	15	7	0.68	[0.45 0.86]
Linear interpolation + sound	15	7	0.68	[0.45 0.86]
Spatial affinities	17	5	0.77	[0.55 0.92]
Spatial affinities + sound	19	3	0.86	[0.65 0.97]

For the proportions of the four methods for "Wringing vs. Pressing" in Table 11, the observed accuracies of all groups are higher than chance levels (50%). Two groups (shown in red) have confidence intervals with a lower bound higher than 50%. Thus we might conclude that the methods of spatial affinities with and without sound are both effective in improving the perception of the quality of Space for the pair of "Wringing vs. Pressing".

After comparing with chance level, we will compare the accuracies of perceiving the quality of Space across the four methods (linear interpolation, linear interpolation with sound, spatial affinities and spatial affinities with sound) to see whether adding sound and spatial affinities improve the accuracies. This analysis will be presented in Section 4.2.4.

4.2.2 Accuracies of Perception of the Quality of Weight in Stick Figure Animations

Next, we analyze the data of the questions aimed to test the perception of the quality of Weight in stick figure animations: Dabbing vs. Thrusting, Slashing vs. Flicking, Gliding vs. Pressing and Floating vs. Wringing.

The accuracies of the responses of these four questions are shown in Table 12. All four methods have observed proportions (accuracies) much higher than 50% in our survey, which might mean that the quality of Weight is easy for people to perceive in the stick figure animations.

Table 12: Accuracy of subjects' perception of the quality of Weight in stick figure animations

Method	Correct	Incorrect	Proportion
Linear interpolation	72	16	0.82
Linear interpolation + sound	74	14	0.84
Spatial affinities	74	14	0.84
Spatial affinities + sound	77	11	0.88

To see whether the estimated accuracies are higher than chance level, we need to calculate the confidence intervals for each question and check whether they have lower bounds above 50%. The detailed results of each question aimed to test the perception of Weight are shown in Tables 13, 14, 15 and 16.

Table 13: Accuracy of subjects' perception of Dabbing vs. Thrusting in stick figure animations

Method	Correct	Incorrect	Proportion	95% confidence interval
Linear interpolation	21	1	0.95	[0.77 1.00]
Linear interpolation + sound	20	2	0.91	[0.71 0.99]
Spatial affinities	19	3	0.86	[0.65 0.97]
Spatial affinities + sound	21	1	0.95	[0.77 1.00]

For the proportions of the four methods for the pair of "Dabbing vs. Thrusting" in Table

13, all estimated accuracies are statistically higher than chance levels (50%) with confidence intervals with a lower bound higher than 50% (shown in red). So we might conclude that all methods are effective in improving the perception of the quality of Weight for the pair of "Dabbing vs. Thrusting".

Table 14: Accuracy of subjects' perception of Slashing vs. Flicking in stick figure animations

Method	Correct	Incorrect	Proportion	95% confidence interval
Linear interpolation	19	3	0.86	[0.65 0.97]
Linear interpolation + sound	19	3	0.86	[0.65 0.97]
Spatial affinities	18	4	0.82	[0.60 0.95]
Spatial affinities + sound	19	3	0.86	[0.65 0.97]

For the proportions of the four methods for the pair of "Slashing vs. Flicking" in Table 14, we found that all estimated accuracies are statistically higher than chance levels (50%) with confidence intervals with a lower bound higher than 50% (shown in red). So we might conclude that all methods are effective in improving the perception of the quality of Weight for the pair of "Slashing vs. Flicking".

Table 15: Accuracy of subjects' perception of Gliding vs. Pressing in stick figure animations

Method	Correct	Incorrect	Proportion	95% confidence interval
Linear interpolation	13	9	0.59	[0.36 0.79]
Linear interpolation + sound	16	6	0.73	[0.50 0.89]
Spatial affinities	17	5	0.77	[0.55 0.92]
Spatial affinities + sound	18	4	0.82	[0.60 0.95]

For the proportions of the four methods for the pair of "Gliding vs. Pressing" in Table 15, the estimated accuracies of all groups are higher than chance levels (50%), but only three groups have confidence intervals with a lower bound higher than 50% (shown in red). We might conclude that these three methods are effective in improving the perception of the quality of Weight for the pair of "Gliding vs. Pressing".

Table 16: Accuracy of subjects' perception of Floating vs. Wringing in stick figure animations

Method	Correct	Incorrect	Proportion	95% confidence interval
Linear interpolation	19	3	0.86	[0.65 0.97]
Linear interpolation + sound	19	3	0.86	[0.65 0.97]
Spatial affinities	20	2	0.91	[0.71 0.99]
Spatial affinities + sound	19	3	0.86	[0.65 0.97]

For the proportions of the four methods for the pair of "Floating vs. Wringing" in Table 16, all estimated accuracies are statistically higher than chance levels (50%) with confidence intervals with a lower bound higher than 50% (shown in red). So we might conclude that all methods are effective in improving the perception of the quality of Weight for "Floating vs. Wringing".

After comparing with chance levels, we will compare the accuracies of perceiving the quality of Weight across the four methods (linear interpolation, linear interpolation with sound, spatial affinities and spatial affinities with sound) to see whether adding sound and spatial affinities improve the accuracies. This analysis will be presented in Section 4.2.4.

4.2.3 Accuracies of Perception of the Quality of Time in Stick Figure Animations

In this section, we analyze the data of the questions aimed to test the perception of the quality of Time in stick figure animations: Dabbing vs. Gliding, Flicking vs. Floating, Slashing vs. Wringing and Thrusting vs. Pressing.

The accuracies of the responses of these four questions are shown in Table 17. All four methods have observed proportions (accuracies) much higher than 50% in our survey. This might mean that the quality of Time is easy for people to perceive in stick figure animations.

Table 17: Accuracy of subjects' perception of the quality of Time in stick figure animations

Method	Correct	Incorrect	Proportion
Linear interpolation	78	10	0.89
Linear interpolation + sound	83	5	0.94
Spatial affinities	79	9	0.90
Spatial affinities + sound	87	1	0.99

To see whether the estimated accuracies are higher than chance level, we need to calculate the confidence intervals for each question and check whether they have lower bounds above 50%. The detailed results of each question aimed to test the perception of Time are shown in Tables 18, 19, 20 and 21.

Table 18: Accuracy of subjects' perception of Dabbing vs. Gliding in stick figure animations

Method	Correct	Incorrect	Proportion	95% confidence interval
Linear interpolation	17	5	0.77	[0.55 0.92]
Linear interpolation + sound	21	1	0.95	[0.77 1.00]
Spatial affinities	17	5	0.77	[0.55 0.92]
Spatial affinities + sound	22	0	1.00	[0.85 1.00]

For the proportions of the four methods for the pair of "Dabbing vs. Gliding" in Table 18, we found that adding sound improved people's perception of the quality of Time by about 20%, while the effect of spatial affinities didn't show improvement or impairment of perception. We can also see that all estimated accuracies are statistically higher than chance levels (50%) with confidence intervals with a lower bound higher than 50% (shown in red). So we might conclude that all methods are effective in improving the perception of the quality of Time for "Dabbing vs. Gliding".

Table 19: Accuracy of subjects' perception of Flicking vs. Floating in stick figure animations

Method	Correct	Incorrect	Proportion	95% confidence interval
Linear interpolation	20	2	0.91	[0.71 0.99]
Linear interpolation + sound	19	3	0.86	[0.65 0.97]
Spatial affinities	21	1	0.95	[0.77 1.00]
Spatial affinities + sound	22	0	1.00	[0.85 1.00]

For the proportions of the four methods for the pair of "Flicking vs. Floating" in Table 19, all estimated accuracies are statistically higher than chance levels (50%) with confidence intervals with a lower bound higher than 50% (shown in red). So we might conclude that all methods are effective in improving the perception of the quality of Time for “Flicking vs. Floating”.

Table 20: Accuracy of subjects’ perception of Slashing vs. Wringing in stick figure animations

Method	Correct	Incorrect	Proportion	95% confidence interval
Linear interpolation	21	1	0.95	[0.77 1.00]
Linear interpolation + sound	22	0	1.00	[0.85 1.00]
Spatial affinities	20	2	0.91	[0.71 0.99]
Spatial affinities + sound	22	0	1.00	[0.85 1.00]

For the proportions of the four methods for the pair of “Slashing vs. Wringing” in Table 20, all estimated accuracies are statistically higher than chance levels (50%) with confidence intervals with a lower bound higher than 50% (shown in red). So we might conclude that all methods are effective in improving the perception of the quality of Time for “Slashing vs. Wringing”.

Table 21: Accuracy of subjects’ perception of Thrusting vs. Pressing in stick figure animations

Method	Correct	Incorrect	Proportion	95% confidence interval
Linear interpolation	20	2	0.91	[0.71 0.99]
Linear interpolation + sound	21	1	0.95	[0.77 1.00]
Spatial affinities	21	1	0.95	[0.77 1.00]
Spatial affinities + sound	21	1	0.95	[0.77 1.00]

For the proportions of the four methods for the pair of “Thrusting vs. Pressing”, all estimated accuracies are statistically higher than chance levels (50%) with confidence intervals with a lower bound higher than 50% (shown in red). So we might conclude that all methods are effective in improving the perception of the quality of Time for “Thrusting vs. Pressing”.

4.2.4 Comparison of Accuracies of Perception of Movement Qualities in Stick Figure Animations Across Four Methods

After comparing with chance levels, we are interested to see whether the methods of “spatial affinities” and “adding sound” are helpful in improving people’s perception of the qualities of movement in terms of the Effort factors: Space Effort, Weight Effort and Time Effort. We compared the accuracies of perception across the four methods for each Effort Factor. We categorized the data based on the Effort factors. For each Effort Factor, we did the data analysis to see whether there are differences in the effectiveness of the four methods. The R packages **readxl** [68] was used to load the data of our survey in R [69]. We used the generalized linear mixed-effect model and carried out the test of the general linear hypothesis with the R packages **lme4** [70] and **multcomp** [71] to compare the accuracies across the methods. The confidence intervals for the differences in log odds were calculated between each of the following pairs of methods:

- Linear interpolation vs. Linear interpolation + Sound;
- Linear interpolation vs. Nonlinear interpolation with Spatial Affinities;
- Linear interpolation vs. Nonlinear interpolation with Spatial Affinities + Sound
- Linear interpolation + Sound vs. Nonlinear interpolation with Spatial Affinities + Sound

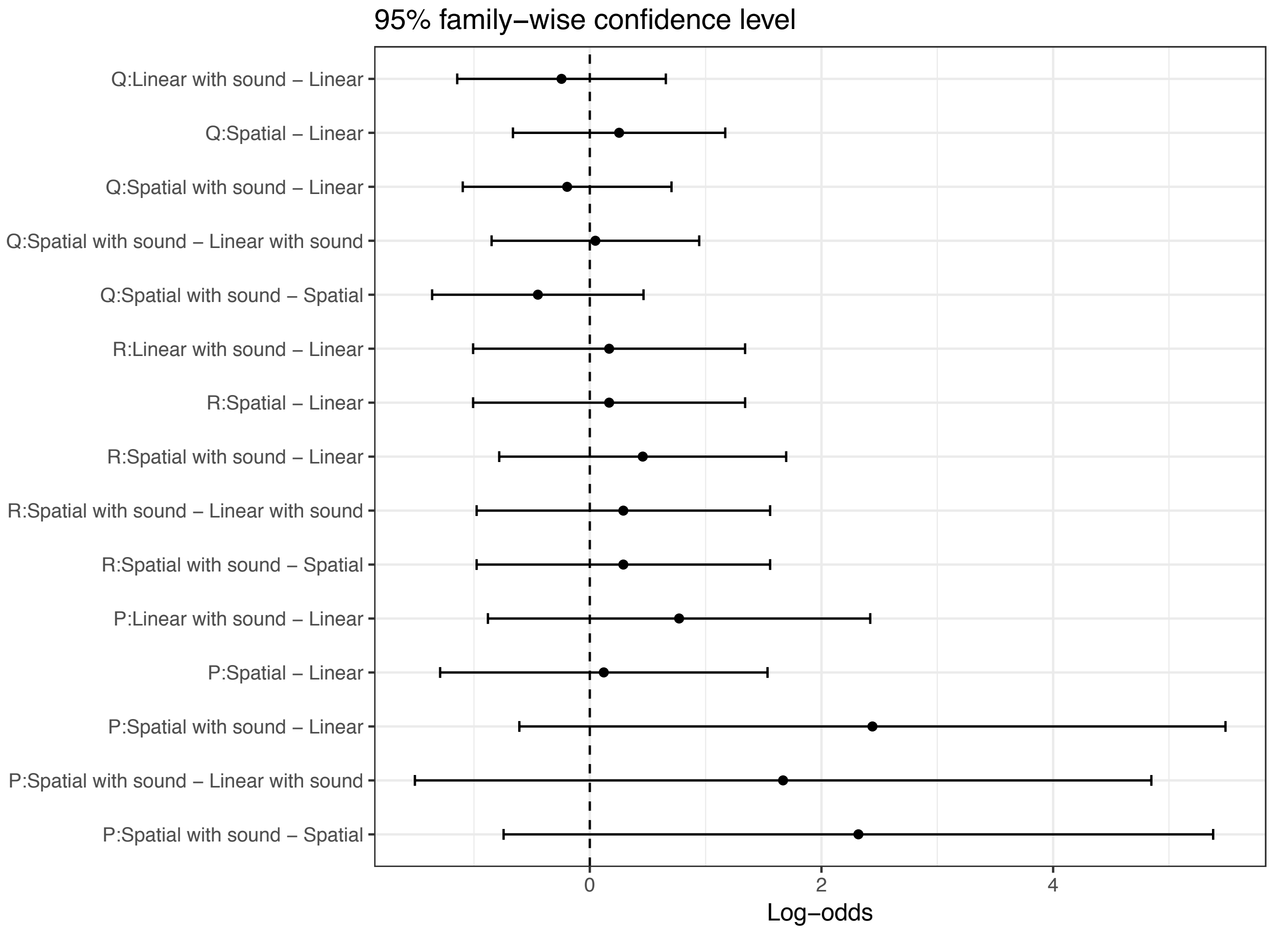


Figure 18: Multiple comparisons of accuracies of subjects picking the right videos which matches the qualities of Space (Q), Weight (R) and Time (P) for the survey of stick figure animations.

- Nonlinear interpolation with Spatial Affinities vs. Nonlinear interpolation with Spatial Affinities + Sound

The plots of the confidence intervals were generated using the R package **ggplot2** [72].

In Figure 18, we showed the 95% of confidence interval by carrying out the general linear hypothesis test for multiple comparisons of the accuracies of subjects picking the correct videos which match the qualities of movement in the question. The first five confidence intervals (in log-odds) are for the accuracies of the perception of the quality of Space; the next five confidence intervals are for the quality of Weight and the last five confidence intervals are for the quality of Time. In the figure, “Linear” is short for “Linear interpolation” and “Spatial” is short for “Nonlinear interpolation with Spatial Affinities”.

For the quality of Space (Q), the actual observed difference in our model is positive for “Spatial - Linear”, which means the observed accuracies of perception using the method “Spatial” is higher than the observed accuracies of perception with the method of “Linear”. The other observed differences in our model are either negative or zero for the quality of Space. This means that the sound added to accompany the movement might make it more difficult for subjects to perceive the quality of Space. However, as all of the confidence of intervals crossed zero (the dashed line), we are not sure whether the difference between the methods are positive or negative. Thus we can’t determine if sound and spatial affinities made a difference in improving people’s perception of the quality of Space in stick figure animations.

For the qualities of Weight (R) and Time (P), the actual observed differences in our model are positive for all comparison pairs, which means the observed accuracies of perception using the method “Linear with sound” is higher than the observed accuracies of perception using the method of “Linear”. We can draw similar conclusions to other four comparisons: the observed accuracies of perception using the method “Spatial” is higher than the observed accuracies of perception with the method of “Linear”, etc. However, as all of the confidence of intervals crossed zero (the dashed line), we are not sure whether the difference between the methods are positive or negative, even though our observed differences are positive. Thus we can’t determine if sound and spatial affinities made a difference in improving people’s perception of the qualities of Weight and Time in stick figure animations.

4.2.5 Conclusions of the Survey of Stick Figure Animations

Based on the data analysis in the previous sections, we have the following discussions and conclusions:

1. Conclusions regarding whether the accuracies are higher than chance level:

(1). All observed accuracies of perception in our survey are higher than 50% in terms of each of the movement qualities: Space, Weight and Time. This means that people are able to perceive the different movement qualities in our survey of stick figure animations.

(2). For the questions aimed to test the perception of the quality of Space, most 95% confidence intervals cover both sides of 50% (see Tables 8, 9, 10 and 11). Thus we can't conclude whether these methods are effective in improving the perception of the quality of Space. But three confidence intervals are above 50% (highlighted in red in Tables 10 and 11). Thus the method of using spatial affinities for generating the reference trajectory is effective for perceiving the difference for the pair of "Gliding vs. Floating". The methods of using spatial affinities with and without sound are both effective for perceiving the difference for the pair of "Wringing and Pressing". But we didn't find evidence that the sound added to accompany the movement made a difference in improving the perception of the quality of Space in robotic movement. This suggests that our sound design choices for conveying the quality of Space were not effective, and possibly detrimental. The result is consistent with our findings in the mapping of the quality of Space between movement and sound in Sections 3.1.2.1 and 3.1.2.2 that there is no precise mapping between the concept of space in movement onto the qualities of the monophonic and spatially static sound in our study. Future work could include exploring ways to generate sounds that have features better mapped to the

the quality of Space in robotic movement.

(3). For the questions aimed to test the perception of the quality of Weight, all 95% confidence intervals have a lower bound above 50% (see Tables 13, 14, 15 and 16), except for one group in Table 15. Thus we might conclude that most accuracies of perception are statistically higher than chance level and all four methods are effective in improving the perception of the quality of Weight. This means that the quality of Weight is easy for people to perceive, no matter which of the four methods we use.

(4). For the questions aimed to test the perception of the quality of Time, all 95% confidence intervals have a lower bound above 50% (see Tables 18, 19, 20 and 21). Thus we might conclude that all accuracies of perception are statistically higher than chance level and all four methods are effective in improving the perception of the quality of Time. This means that the quality of Time is also easy for people to perceive, no matter which of the four methods we use.

2. Conclusions regarding whether the methods of adding sound and spatial affinities are better than no sound and linear interpolation in improving people's perception of the qualities of movement:

(1). The observed data in our survey of stick figure animations shows that the sound we added to accompany the movement and the method of using spatial affinities in generating the reference trajectories improved the accuracies of perception of the qualities of Weight and Time, but not of the quality of Space.

(2). However, all of the confidence intervals crossed zero. Due to the limitation of the sample size, we don't have enough evidence to see if the differences of accuracies are positive or negative precisely. Thus, we can't determine if adding sound and using spatial affinities really made a difference in improving people's perception of the qualities of movement in stick figure animations, even though our observed data shows an improvement in the perception of the qualities of Weight and Time.

4.3 Quantitative Analysis of the Survey Data of Baxter Research Robot

We did similar analysis of data for the survey of Baxter Research Robot. First, we checked whether the subjects were randomly choosing the answers and whether the accuracies of using the different methods are higher than chance level (50%) by calculating the proportions and 95% confidence interval of each method for each question. Next, we tested the comparisons across the methods to see whether the proportions (accuracies of perception) are statistically different and then to determine whether the methods are effective in improving the accuracies of perception of qualities of Baxter's motion.

4.3.1 Accuracies of Perception of the Quality of Space of Baxter's Motion

In this section, we analyze the data of the questions aimed to test the perception of the quality of Space in Baxter's motion: Dabbing vs. Flicking, Slashing vs. Thrusting, Gliding vs. Floating and Wringing vs. Pressing.

The accuracies of the responses of these four questions are shown in Table 22. All four methods have observed proportions (accuracies) higher than 50% in our survey.

Table 22: Accuracy of subjects' perception of the quality of Space in Baxter's motion

Method	Correct	Incorrect	Proportion
Linear interpolation	54	34	0.61
Linear interpolation + sound	49	39	0.56
Spatial affinities	61	27	0.69
Spatial affinities + sound	48	40	0.55

To see whether the estimated accuracies are higher than chance level, we need to calculate the confidence intervals for each question and check whether they have lower bounds above 50%. The detailed results of each question aimed to test the perception of Space are shown in Tables 23, 24, 25 and 26.

Table 23: Accuracy of subjects' perception of Dabbing vs. Flicking by Baxter Research Robot

Method	Correct	Incorrect	Proportion	95% confidence interval
Linear interpolation	16	6	0.73	[0.50 0.89]
Linear interpolation + sound	13	9	0.59	[0.36 0.79]
Spatial affinities	20	2	0.91	[0.71 0.99]
Spatial affinities + sound	16	6	0.73	[0.50 0.89]

For the proportions of the four methods for "Dabbing vs. Flicking" in Table 23, the perception of the quality of Space of all four methods have accuracies higher than 50%, and the confidence intervals of three methods have lower bounds higher than 50%. We can conclude that the three methods with data highlighted in red are effective in improving the perception of the quality of Space in Baxter's motion for the pair of "Dabbing vs. Flicking".

Table 24: Accuracy of subjects' perception of Slashing vs. Thrusting by Baxter Research Robot

Method	Correct	Incorrect	Proportion	95% confidence interval
Linear interpolation	20	2	0.91	[0.71 0.99]
Linear interpolation + sound	19	3	0.86	[0.65 0.97]
Spatial affinities	19	3	0.86	[0.65 0.97]
Spatial affinities + sound	8	14	0.36	[0.17 0.59]

For the proportions of the four methods for "Slashing vs. Thrusting" in Table 24, the perception of the quality of Space of three methods have accuracies statistically higher than 50%, with the confidence intervals having lower bounds higher than 50%. We may conclude that the methods of linear interpolation with and without sound, and the method of spatial affinities are effective in improving the perception of the quality of Space in Baxter's motion for the pair of "Slashing vs. Thrusting".

Table 25: Accuracy of subjects' perception of Gliding vs. Floating by Baxter Research Robot

Method	Correct	Incorrect	Proportion	95% confidence interval
Linear interpolation	13	9	0.59	[0.36 0.79]
Linear interpolation + sound	12	10	0.55	[0.32 0.76]
Spatial affinities	17	5	0.77	[0.55 0.92]
Spatial affinities + sound	17	5	0.77	[0.55 0.92]

For the proportions of the four methods for "Gliding vs. Floating" in Table 25, the perception of the quality of Space with the methods of spatial affinities with and without sound have accuracies higher than 50% and confidence intervals (in red) with a lower bound greater than 50%. We might conclude that the methods of spatial affinities with and without sound are both effective in improving the perception of the quality of Space in Baxter's motion for the pair of "Gliding vs. Floating".

Table 26: Accuracy of subjects' perception of Wringing vs. Pressing by Baxter Research Robot

Method	Correct	Incorrect	Proportion	95% confidence interval
Linear interpolation	5	17	0.23	[0.08 0.45]
Linear interpolation + sound	5	17	0.23	[0.08 0.45]
Spatial affinities	5	17	0.23	[0.08 0.45]
Spatial affinities + sound	7	15	0.32	[0.14 0.55]

For the proportions of the four methods for "Wringing vs. Pressing" in Table 26, the estimated accuracies of all groups are lower than chance levels (50%). Three groups have confidence intervals with a higher bound lower than 50%, which means we might conclude that the first three methods impaired the perception of the quality of Space in Baxter's motion for the pair of "Wringing vs. Pressing".

After comparing with chance level, we will compare the accuracies of perceiving the quality of Space across the four methods (linear interpolation, linear interpolation with sound, spatial affinities and spatial affinities with sound) to see whether the sound we added to accompany the movement and the method of using spatial affinities improve the accuracies. This analysis will be presented in Section 4.3.4.

4.3.2 Accuracies of Perception of the Quality of Weight in Baxter's Motion

In this section, we analyze the data of the questions aimed to test the perception of the quality of Weight in Baxter's motion: Dabbing vs. Thrusting, Slashing vs. Flicking, Gliding vs. Pressing and Floating vs. Wringing.

The accuracies of the responses of these four questions are shown in Table 27. All four methods have observed proportions (accuracies) higher than 50% in our survey.

Table 27: Accuracy of subjects' perception of the quality of Weight in Baxter's motion

Method	Correct	Incorrect	Proportion
Linear interpolation	52	36	0.59
Linear interpolation + sound	62	26	0.70
Spatial affinities	56	32	0.64
Spatial affinities + sound	63	24	0.72

To see whether the estimated accuracies are higher than chance level, we need to calculate the confidence intervals for each question and check whether they have lower bounds above 50%. The detailed results of each question aimed to test the perception of Weight are shown in Tables 28, 29, 30 and 31.

Table 28: Accuracy of subjects' perception of Dabbing vs. Thrusting by Baxter Research Robot

Method	Correct	Incorrect	Proportion	95% confidence interval
Linear interpolation	20	2	0.91	[0.71 0.99]
Linear interpolation + sound	19	3	0.86	[0.65 0.97]
Spatial affinities	18	4	0.82	[0.60 0.95]
Spatial affinities + sound	18	3	0.86	[0.64 0.97]

For the proportions of the four methods for the pair of "Dabbing vs. Thrusting" in Table 28, we found that all estimated accuracies are statistically higher than chance levels (50%) with confidence intervals with a lower bound higher than 50%, which are highlighted in red. So we might conclude that all methods are effective in improving the perception of the quality of Weight in Baxter's motion for the pair of "Dabbing vs. Thrusting".

Table 29: Accuracy of subjects' perception of Slashing vs. Flicking by Baxter Research Robot

Method	Correct	Incorrect	Proportion	95% confidence interval
Linear interpolation	17	5	0.77	[0.55 0.92]
Linear interpolation + sound	19	3	0.86	[0.65 0.97]
Spatial affinities	16	6	0.73	[0.50 0.89]
Spatial affinities + sound	20	2	0.91	[0.71 0.99]

For the proportions of the four methods for the pair of "Slashing vs. Flicking" in Table 29, we found that all estimated accuracies are statistically higher than chance levels (50%) with confidence intervals with a lower bound higher than 50%. So we might conclude that all methods are effective in improving the perception of the quality of Weight in Baxter's motion for the pair of "Slashing vs. Flicking".

Table 30: Accuracy of subjects' perception of Gliding vs. Pressing by Baxter Research Robot

Method	Correct	Incorrect	Proportion	95% confidence interval
Linear interpolation	3	19	0.14	[0.03 0.35]
Linear interpolation + sound	7	15	0.32	[0.14 0.55]
Spatial affinities	3	19	0.14	[0.03 0.35]
Spatial affinities + sound	9	13	0.41	[0.21 0.64]

For the proportions of the four methods for the pair of "Gliding vs. Pressing" in Table 30, the estimated accuracies of all groups are lower than chance levels (50%), and two groups have confidence intervals with a higher bound lower than 50%. We might conclude that these two methods of linear interpolation and spatial affinities without sound impaired the perception of the quality of Weight in Baxter's motion for the pair of "Gliding vs. Pressing".

Table 31: Accuracy of subjects' perception of Floating vs. Wringing by Baxter Research Robot

Method	Correct	Incorrect	Proportion	95% confidence interval
Linear interpolation	12	10	0.55	[0.32 0.76]
Linear interpolation + sound	17	5	0.77	[0.55 0.92]
Spatial affinities	19	3	0.86	[0.65 0.97]
Spatial affinities + sound	16	6	0.73	[0.50 0.89]

For the proportions of the four methods for the pair of "Floating vs. Wringing" in Table 31, the estimated accuracies of three methods are statistically higher than chance levels (50%) with confidence intervals with a lower bound higher than 50%. So we might conclude that all three methods of linear interpolation with sound, and the methods of spatial affinities with and without sound are effective in improving the perception of the quality of Weight.

After comparing with chance level, we will compare the accuracies of perceiving the quality of Weight across the four methods (linear interpolation, linear interpolation with sound, spatial affinities and spatial affinities with sound) to see whether adding sound and spatial affinities improve the accuracies. This analysis will be presented in Section 4.3.4.

4.3.3 Accuracies of Perception of the Quality of Time in Baxter's Motion

Next, we analyze the data of the questions aimed to test the perception of the quality of Time in Baxter's motion: Dabbing vs. Gliding, Flicking vs. Floating, Slashing vs. Wringing and Thrusting vs. Pressing.

The accuracies of the responses of these four questions are shown in Table 32. All four methods have observed proportions (accuracies) much higher than 50% in our survey.

Table 32: Accuracy of subjects' perception of the quality of Time in Baxter's motion

Method	Correct	Incorrect	Proportion
Linear interpolation	81	7	0.92
Linear interpolation + sound	85	3	0.97
Spatial affinities	64	24	0.73
Spatial affinities + sound	77	11	0.88

To see whether the estimated accuracies are higher than chance level, we need to calculate the confidence intervals for each question and check whether they have lower bounds above

50%. The detailed results of each question aimed to test the perception of Time are shown in Tables 33, 34, 35 and 36.

Table 33: Accuracy of subjects' perception of Dabbing vs. Gliding by Baxter Research Robot

Method	Correct	Incorrect	Proportion	95% confidence interval
Linear interpolation	21	1	0.95	[0.77 1.00]
Linear interpolation + sound	21	1	0.95	[0.77 1.00]
Spatial affinities	6	16	0.27	[0.11 0.50]
Spatial affinities + sound	14	8	0.64	[0.41 0.83]

For the proportions of the four methods for the pair of "Dabbing vs. Gliding" in Table 33, the estimated accuracies of the two methods of interpolation with and without sound are both statistically higher than chance levels (50%) with confidence intervals with a lower bound higher than 50%. So we might conclude that these two methods are effective in improving the perception of the quality of Time for the pair of "Dabbing vs. Gliding". But the estimated accuracy of the method of spatial affinities is statistically lower than chance levels (50%) with confidence intervals with a higher bound lower than 50%. So the method of spatial affinities showed impairment of the perception of the quality of Time in Baxter's motion.

Table 34: Accuracy of subjects' perception of Flicking vs. Floating by Baxter Research Robot

Method	Correct	Incorrect	Proportion	95% confidence interval
Linear interpolation	18	4	0.82	[0.60 0.95]
Linear interpolation + sound	21	1	0.95	[0.77 1.00]
Spatial affinities	18	4	0.82	[0.60 0.95]
Spatial affinities + sound	19	3	0.86	[0.65 0.97]

For the proportions of the four methods for the pair of "Flicking vs. Floating" in Table 34, all estimated accuracies are statistically higher than chance levels (50%) with confidence intervals with a lower bound higher than 50%. So we might conclude that all methods are effective in improving the perception of the quality of Time in Baxter's motion for the pair of "Flicking vs. Floating".

Table 35: Accuracy of subjects' perception of Slashing vs. Wringing by Baxter Research Robot

Method	Correct	Incorrect	Proportion	95% confidence interval
Linear interpolation	21	1	0.95	[0.77 1.00]
Linear interpolation + sound	22	0	1.00	[0.85 1.00]
Spatial affinities	18	4	0.82	[0.60 0.95]
Spatial affinities + sound	22	0	1.00	[0.85 1.00]

For the proportions of the four methods for the pair of "Slashing vs. Wringing" in Table 35, all estimated accuracies are statistically higher than chance levels (50%) with confidence intervals with a lower bound higher than 50%. So we might conclude that all methods are effective in improving the perception of the quality of Time in Baxter's motion for the pair of “Slashing vs. Wringing”.

Table 36: Accuracy of subjects’ perception of Thrusting vs. Pressing by Baxter Research Robot

Method	Correct	Incorrect	Proportion	95% confidence interval
Linear interpolation	21	1	0.95	[0.77 1.00]
Linear interpolation + sound	21	1	0.95	[0.77 1.00]
Spatial affinities	22	0	1.00	[0.85 1.00]
Spatial affinities + sound	22	0	1.00	[0.85 1.00]

For the proportions of the four methods for the pair of “Thrusting vs. Pressing” in Table 36, all estimated accuracies are statistically higher than chance levels (50%) with confidence intervals with a lower bound higher than 50%. So we might conclude that all methods are effective in improving the perception of the quality of Time in Baxter’s motion for the pair of “Thrusting vs. Pressing”.

4.3.4 Comparison of Accuracies of Perception of Movement Qualities in Baxter’s Motion Across Four Methods

After comparing with chance levels, we compared the accuracies of perception across the four methods for each Effort Factor to see whether the methods of “spatial affinities” and “adding sound” are helpful in improving people’s perception of the qualities of movement in terms of the Effort factors: Space Effort, Weight Effort and Time Effort. Similarly to the analysis for the data of the survey for the stick figure animations, we categorized the data based on the Effort factors. For each Effort Factor, we did the data analysis to see whether there are differences in the effectiveness of the four methods. The results are shown in Figure

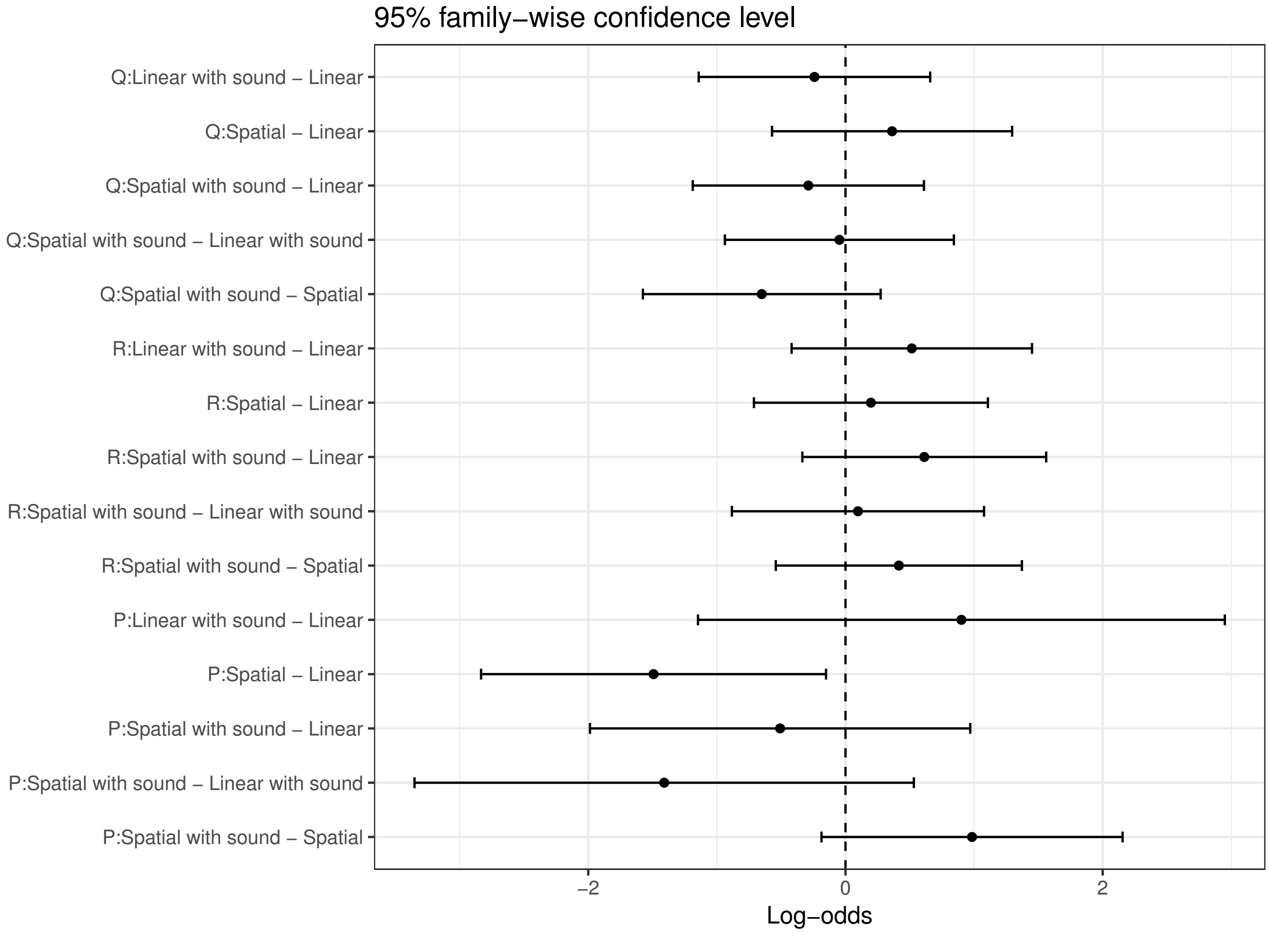


Figure 19: Multiple comparisons of accuracies of subjects picking the right videos which matches the qualities of Space (Q), Weight (R) and Time (P) for the survey of Baxter Research Robot.

19.

In Figure 19, we showed the 95% of confidence interval by carrying out the general linear hypothesis test for multiple comparisons of the accuracies of subjects picking the correct videos which match the qualities of movement in the question. The first five confidence intervals (in log-odds) are for the accuracies of the perception of the quality of Space; the next five confidence intervals are for the quality of Weight and the last five confidence intervals are for the quality of Time. In this figure, "Linear" is short for "Linear interpolation" and "Spatial" is short for "Nonlinear interpolation with Spatial Affinities".

For the quality of Space (Q), the actual observed difference in our model is positive for "Spatial - Linear", which means the observed accuracies of perception using the method "Spatial" is higher than the observed accuracies of perception with the method of "Linear". The other observed differences in our model are negative or near zero for the quality of Space. This means that the sound added to accompany the movement might make it more difficult for subjects to perceive the quality of Space. However, as all of the confidence of intervals crossed zero (the dashed line), we are not sure whether the difference between the methods are positive or negative. Thus we can't determine if sound and spatial affinities made a difference in improving people's perception of the quality of Space in Baxter's motion.

For the quality of Weight (R), the actual observed differences in our model are positive for all five comparisons, which means the observed accuracies of perception using the method "Linear with sound" is higher than the observed accuracies of perception using the method of "Linear". We can draw similar conclusions to other four comparisons: the observed accuracies of perception using the method "Spatial" is higher than the observed accuracies of perception with the method of "Linear", etc. However, as all of the confidence of intervals crossed zero (the dashed line), we are not sure whether the difference between the methods are positive or negative, even though our observed differences are positive. Thus we can't determine if sound and spatial affinities made a difference in improving people's perception of the quality of Weight in Baxter's motion.

For the quality of Time (P), the actual observed difference in our model is positive for "Linear with sound - Linear" and "Spatial with sound - Spatial', which means the observed accuracies of perception using the method "Linear with sound" is higher than the observed accuracies of perception with the method of "Linear" and the observed accuracies of perception using the method "Spatial with sound" is higher than the observed accuracies of perception with the method of "Spatial". This means that the sound added to accompany the movement might make it easier for subjects to perceive the quality of Time. The other three observed differences in our model are negative for the quality of Time. This means that the method of spatial affinities might make it more difficult for subjects to perceive the quality of Time. However, as all of the confidence of intervals crossed zero (the dashed line), we are not sure whether the difference between the methods are positive or negative. Thus we can't determine if sound and spatial affinities made a difference in improving people's perception of the quality of Time in Baxter's motion.

4.3.5 Conclusions of the Survey of Baxter Research Robot

From the data analysis in the previous sections, we have the following discussions and conclusions:

1. Conclusions regarding whether the accuracies are higher than chance level:

(1). All observed accuracies of perception in this survey are higher than 50% in terms of each of the movement qualities: Space, Weight and Time. This means that people are able to perceive the different movement qualities in our survey of Baxter's motion.

(2). For the questions aimed to test the perception of the quality of Space, some 95% confidence intervals cover both sides of 50% (see Tables 23, 24, 25 and 26). Thus we can't conclude whether these methods are effective in improving the perception of the quality of Space. But eight groups have confidence intervals above 50% (highlighted in red in Tables 23, 24 and 25). These highlighted data correspond to the methods that are effective in improving the perception of the quality of Space. Thus, using spatial affinities helped improve the perception of the quality of Space, while adding sound did not. This is consistent with our conclusion in the survey of the stick figure animations. One thing to mention is that for the pair of “Wringing vs. Pressing”, the observed proportions are lower than the chance level. This is due to the scaling of the trajectories when the the joint limit is reached and thus changed the range of the motion.

(3). For the questions aimed to test the perception of the quality of Weight, most 95% confidence intervals have a lower bound above 50% (see Tables 28, 29, 30 and 31), except for the groups in Table 30 and one group in Table 31. Thus we might conclude that most accuracies of perception are statistically higher than chance level and all four methods are effective in improving the perception of the quality of Weight. It means that the quality of Weight is easy for people to perceive, which is consistent with the conclusion in the survey of stick figure animations. For the pair of “Gliding vs. Pressing”, the observed proportions are lower than the chance level. This is due to the scaling of the trajectories when the the joint limit is reached and thus changed the range of the motion.

(4). For the questions aimed to test the perception of the quality of Time, most 95% confidence intervals have a lower bound above 50% (see Tables 33, 34, 35 and 36), except for two groups in Table 33. Thus we might conclude that most accuracies of perception are statistically higher than chance level and all four methods are effective in improving the perception of the quality of Time. This indicates that the quality of Time is easy for people to perceive, which is consistent with the conclusion in the survey of stick figure animations.

2. Conclusions regarding whether the methods of adding sound and spatial affinities are better than no sound and linear interpolation in improving people’s perception of the qualities of movement:

(1). The observed data in our survey of Baxter's motion shows that both of the methods of adding sound and using spatial affinities improved the accuracies of perception of the quality of Weight, while the method of adding sound improved the accuracies of perception of the quality of Time. No improvement is observed for the quality of Space with the methods of adding sound or spatial affinities.

(2). However, the confidence intervals crossed zero. Due to the limitation of the sample size, we don't have enough evidence to see if the differences of accuracies are positive or negative. Thus, we can't determine if sound and spatial affinities really made a difference in improving people's perception of the qualities of Baxter's motion, even though our survey data indicates an improvement in the perception of the qualities of Weight and Time.

5 Conclusions and Future Work

This dissertation studied methods for improving human perception of the qualities of robotic movement and thus making robots better tools in human-facing applications. The additional variety created in our design makes robotic movements more expressive and functional. Three methods were developed to solve the limit-exceeding problem in robotic platforms while generating feasible trajectories for robot to track. The framework was implemented on Rethink Robotics Baxter Research Robot and realized different trajectories on the robot through a highly abstracted web-based computational framework. To further improve the variations in robotic movement, a new method for generating the reference trajectories which utilizes the affinities between the Effort and Space in LBMS was designed. To increase the perceived variation in robotic movement, appropriate sounds were supplemented to the robotic movements to help convey the expression and function of the movement. An experiment was conducted to suggest possible mappings of qualities of movement and sound. A user study was conducted to evaluate the effectiveness of this work for increasing the perceived variations in robotic movement.

The contributions of the dissertation include:

- Solved the limit-exceeding problem and guaranteed the trajectories generated by prior method are executable on physical robotic platforms;
- Supplemented appropriate sounds to robotic movements to improve people's perception of the qualities of robotic movement in human-robot interaction;
- Found the mapping between the qualities of Weight and Time in movement and the

sonic features, which is meaningful for research in other applications;

- Designed a new method for generating movement trajectories with greater variations utilizing the affinities between Effort element and spatial aspect in LBMS, which can be used in designing robotic movement in various scenarios;

- Conducted a user study to evaluate the effectiveness of the methods of linear interpolation with and without sound, and nonlinear interpolation using spatial affinities with and without sound. Data analysis results show that

  - Using spatial affinities helped improve the perception of the quality of Space, but the sound added to accompany the movement did not help with the perception of the quality of Space.

  - All four methods are effective in improving the perception of the movement qualities of Weight and Time.

- Tested whether adding sound and spatial affinities are more effective in improving the perception of the movement qualities than linear interpolation and moving without sound. Test results show that

  - Due to the limitation of sample size, we don't have enough evidence to determine whether adding sound and spatial affinities are more effective.

  - However, the observed data in our user study shows that both the methods of adding sound and spatial affinities improved the perception of the qualities of Weight and Time for stick figure animations and improved the perception of the quality of Weight for Baxter's motion. For the quality of Time in Baxter's motion, the sound we added to accompany the movement improved the perception of the quality.

Future work of this research involves implementing the constrained optimal control problem in different robotic platforms and finding better correlations between the Space Effort of movement quality and the features of sound.

References

[1] A. LaViers, L. Bai, A. Bashiri, G. Heddy, and Y. Sheng, "Abstractions for design-by-humans of heterogeneous behaviors," in *Dance Notations and Robot Motion*. Springer Tracts in Advanced Robotics (STAR), 2015.

[2] L. Bai, J. Bellona, L. Dahl, and A. LaViers, "Design of perceptually meaningful quality in robotic motion," in *Workshop on Artistically Skilled Robots, IEEE/RSJ International Conference on Intelligent Robots and Systems*, Daejeon, Korea, October 2016.

[3] L. Dahl, J. Bellona, L. Bai, and A. LaViers, "Data-driven design of sound for enhancing the perception of expressive robotic movement," in *Proceedings of the 4th International Conference on Movement Computing*, London, United Kingdom, June 2017, pp. 16:1–16:8.

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

[5] (2015, July). [Online]. Available: http://sdk.rethinkrobotics.com/wiki/Hardware_Specifications

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

[7] H. Wallbott, "Bodily expression of emotion," *European Journal of Social Psychology*, vol. 28, no. 6, pp. 879–896, 1998.

[8] M. L. Knapp, J. A. Hall, and T. G. Horgan, *Nonverbal Communication in Human Interaction*. Cengage Learning, 2013.

[9] M. Karg, A.-A. Samadani, R. Gorbet, K. Kuhnlenz, J. Hoey, and D. Kulic, “Body movements for affective expression: A survey of automatic recognition and generation,” *IEEE Transactions on Affective Computing*, vol. 4, no. 4, pp. 341–359, 2013.

[10] R. Laban and F. C. Lawrence, *Effort: Economy of Human Movement*, 2nd ed. Macdonald and Evans, June 1974.

[11] D. Orin, “Supervisory control of a multilegged robot,” *International Journal of Robotics Research*, vol. 1, no. 1, pp. 79–91, Spring 1982.

[12] T. B. Sheridan, “Human supervisory control of robot systems,” in *Procceeding of the IEEE Conference of Robotic and Automation*, San Francisco, CA, USA, April 1986.

[13] N. Sarkar and T. K. Podder, “Coordinated motion planning and control of autonomous underwater vehicle-manipulator systems subject to drag optimization,” *IEEE Journal of Oceanic Engineering*, vol. 26, no. 2, pp. 228–239, Apr 2001.

[14] J.-H. Hwang, R. C. Arkin, and D.-S. Kwon, “Mobile robots at your fingertip: Bezier curve on-line trajectory generation for supervisory control,” in *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)*, vol. 2, Oct 2003, pp. 1444–1449.

[15] M. Brand and A. Hertzmann, “Style machines,” in *Proc. of the 27th Annual Conference on Computer Graphics and Interactive Techniques.* ACM Press/Addison-Wesley Publishing Co., 2000, pp. 183–192.

[16] O. C. Jenkins and M. J. Mataric, “Automated derivation of behavior vocabularies for autonomous humanoid motion,” in *Proc. of the Second International Joint Conference*

on Autonomous Agents and Multiagent Systems. Melbourne, Australia: ACM, July 2003, pp. 225–232.

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

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

[19] A. LaViers and M. Egerstedt, "Style-based robotic motion," in *American Control Conference*, Montreal, QC, Canada, June 2012, pp. 4327–4332.

[20] J. E. Bobrow, S. Dubowsky, and J. Gibson, "Time-optimal control of robotic manipulators along specified paths," *The International Journal of Robotics Research*, vol. 4, no. 3, pp. 3–17, 1985.

[21] S. Kuindersma, R. Deits, M. Fallon, A. Valenzuela, H. Dai, F. Permenter, T. Koolen, P. Marion, and R. Tedrake, "Optimization-based locomotion planning, estimation, and control design for the atlas humanoid robot," *Autonomous Robots*, vol. 40, no. 3, pp. 429–455, 2016.

[22] M. Giftthaler, F. Farshidian, T. Sandy, L. Stadelmann, and J. Buchli, "Efficient kinematic planning for mobile manipulators with non-holonomic constraints using optimal control," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*. Singapore: IEEE, May 2017, pp. 3411–3417.

[23] R. M. Murray, *A mathematical introduction to robotic manipulation.* CRC press, 2017.

[24] D. Chi, M. Costa, L. Zhao, and N. Badler, "The EMOTE model for effort and shape," in *Conference on Computer Graphics and Interactive Techniques*, New Orleans, LA, USA, July 2000, pp. 173–182.

[25] 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, June 2002.

[26] M. Sharma, D. Hildebrandt, G. Newman, J. E. Young, Newman, J. E. Young, and R. Eskicioglu, "Communicating affect via flight path exploring use of the laban effort system for designing affective locomotion paths," in *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, Tokyo, Japan, March 2013, pp. 293–300.

[27] H. Knight and R. Simmons, "Expressive motion with x, y and theta: Laban effort features for mobile robots," in *the 23rd IEEE International Symposium on Robot and Human Interactive Communication*, Edinburgh, UK, August 2014, pp. 267–273.

[28] L. Sentis and O. Khatib, "Synthesis of whole-body behaviors through hierarchical control of behavioral primitives," *International Journal of Humanoid Robotics*, vol. 2, no. 04, pp. 505–518, 2005.

[29] A. Camurri, I. Lagerlöf, and G. Volpe, "Recognizing emotion from dance movement: Comparison of spectator recognition and automated techniques," *The International Journal of Human-Computer Studies*, vol. 59, no. 1-2, pp. 213–225, July 2003.

[30] T. Nakata, T. Sato, and T. Mori, “Expression of emotion and intention by robot body movement,” in *Proceedings of the 5th International Conference on Autonomous Systems*, 1998, pp. 352–359.

[31] J. Harris and E. Sharlin, “Exploring the affect of abstract motion in social human-robot interaction,” in *2011 RO-MAN*, Atlanta, GA, USA, July 2011, pp. 441–448.

[32] L. Malatesta, S. Asteriadis, G. Caridakis, A. Vasalou, and K. Karpouzis, “Associating gesture expressivity with affective representations,” *Engineering Applications of Artificial Intelligence*, vol. 51, pp. 124–135, May 2016.

[33] A. LaViers and M. Egerstedt, “The ballet automaton: a formal model for human motion,” in *American Control Conference, Proceedings of the 2011*, San Francisco, CA, USA, June 2011, pp. 3837–3842.

[34] (2015, September). [Online]. Available: http://www.lego.com/en-us/mindstorms/learn-to-program

[35] D. Sakamoto, K. Honda, M. Inami, and T. Igarashi, “Sketch and run: a stroke-based interface for home robots,” in *Proc. CHI*, Boston, MA, USA, April 2009, pp. 197–200.

[36] R. Parasuraman, S. Galster, P. Squire, H. Furukawa, and C. Miller, “A flexible delegation-type interface enhances system performance in human supervision of multiple robots: empirical studies with roboflag,” *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, vol. 35, no. 4, pp. 481 – 493, July 2005.

[37] M. Tomasello, *Origins of Human Communication*. MIT press, 2010.

[38] Z. Eitan and R. Y. Granot, “How music moves: : Musical parameters and listeners images of motion,” *Music Perception: An Interdisciplinary Journal*, vol. 23, no. 3, pp. 221–248, 2006.

[39] R. I. Godøy and M. L. (Eds.), *Musical Gestures: Sound, Movement, and Meaning*. New York: Routledge, 2010.

[40] A. O. Effenberg, “Movement sonification: effects on perception and action,” *IEEE MultiMedia*, vol. 12, no. 2, pp. 53–59, 2005.

[41] L. Naveda, I. C. Martínez, J. Damesón, A. Pereira Ghiena, R. Herrera, and M. A. Ordás, “Cross-cultural comparisons of unconstrained body responses to argentinian and afro-brazilian music,” in *Music, Mind, and Embodiment*. Springer International Publishing, 2016, pp. 464–482.

[42] B. Caramiaux, F. Bevilacqua, and N. Schnell, “Towards a gesture-sound cross-modal analysis,” in *Proceedings of the 8th International Conference on Gesture in Embodied Communication and Human-Computer Interaction*. Berlin, Heidelberg: Springer-Verlag, 2010, pp. 158–170.

[43] A. Sarasúa and E. Guaus, “Beat tracking from conducting gestural data: A multi-subject study,” in *Proceedings of the 2014 International Workshop on Movement and Computing*. New York, NY, USA: ACM, 2014, pp. 118:118–118:123.

[44] L. Dahl, “Comparing the timing of movement events for air-drumming gestures,” in *Music, Mind, and Embodiment*, R. Kronland-Martinet, M. Aramaki, and S. Ystad, Eds. Cham: Springer International Publishing, 2016, pp. 3–21.

[45] F. Visi, E. Coorevits, R. Schramm, and E. R. Miranda, "Analysis of mimed violin performance movements of neophytes," in *Music, Mind, and Embodiment*. Cham, Switzerland: Springer International Publishing, 2016, pp. 88–108.

[46] M. M. Wanderley, "Quantitative analysis of non-obvious performer gestures," in *Gesture and Sign Language in Human-Computer Interaction*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002, pp. 241–253.

[47] J. Françoise, S. F. Alaoui, T. Schiphorst, and F. Bevilacqua, "Vocalizing dance movement for interactive sonification of laban effort factors," in *Proc. of the 2014 Conference on Designing interactive systems*. Vancouver, BC, Canada: ACM, June 2014, pp. 1079–1082.

[48] D. Jacobson, M. Lele, and J. Speyer, "New necessary conditions of optimality for control problems with state-variable inequality constraints," *Journal of Mathematical Analysis and Applications*, vol. 35, no. 2, pp. 255–284, 1971.

[49] R. F. Hartl, S. P. Sethi, and R. G. Vickson, "A survey of the maximum principles for optimal control problems with state constraints," *SIAM review*, vol. 37, no. 2, pp. 181–218, June 1995.

[50] P. Scokaert and J. Rawlings, "Constrained linear quadratic regulation," *IEEE Transactions on Automatic Control*, vol. 43, no. 8, pp. 1163–1169, August 1998.

[51] D. S. Naidu, *Optimal Control Systems*. CRC Press, 2002.

[52] R. F. Stengel, *Optimal Control and Estimation*. Courier Corporation, 2012.

[53] Feng Lin and R. D. Brandt, "An optimal control approach to robust control of robot manipulators," *IEEE Transactions on Robotics and Automation*, vol. 14, no. 1, pp. 69–77, Feb 1998.

[54] N. Correll, "Parameter estimation and optimal control of swarm-robotic systems: A case study in distributed task allocation," in *2008 IEEE International Conference on Robotics and Automation*. Pasadena, CA, USA: IEEE, May 2008, pp. 3302–3307.

[55] D. J. Braun, F. Petit, F. Huber, S. Haddadin, P. Van Der Smagt, A. Albu-Schäffer, and S. Vijayakumar, "Robots driven by compliant actuators: Optimal control under actuation constraints," *IEEE Transactions on Robotics*, vol. 29, no. 5, pp. 1085–1101, 2013.

[56] D. H. Jacobson and M. M. Lele, "A transformation technique for optimal control problems with a state variable inequality constraint," *IEEE Transactions on Automatic Control*, vol. 14, no. 5, pp. 457–464, October 1969.

[57] J. Bellona, L. Bai, L. Dahl, and A. LaViers, "Empirically informed sound synthesis application for enhancing the perception of expressive robotic movement," in *Proceedings of the International Conference on Auditory Display*, Pennsylvania State University, PA, USA, June 2017, pp. 73–80.

[58] J. Françoise, N. Schnell, R. Borghesi, and F. Bevilacqua, "Probabilistic models for designing motion and sound relationships," in *Proceedings of the 2014 International Conference on New Interfaces for Musical Expression*, Goldsmiths, University of London, UK, June 2014, pp. 287–292.

[59] K. Nymoen, R. I. Gody, A. R. Jensenius, and J. Torresen, “Analysing correspondence between sound objects and body motion,” *ACM Transactions on Applied Perception (TAP)*, vol. 10, no. 2, pp. 9:–9:22, 2013.

[60] O. Lartillot and P. Toiviainen, “A matlab toolbox for musical feature extraction from audio,” in *Proc. of the 10 th Int. Conference on Digital Audio Effects*, Bordeaux, France, September 2007, pp. 237–244.

[61] E. Scheirer and M. Slaney, “Construction and evaluation of a robust multifeature speech/music discriminator,” in *1997 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 2, April 1997, pp. 1331–1334.

[62] E. Pacchierotti, H. I. Christensen, and P. Jensfelt, “Human-robot embodied interaction in hallway settings: a pilot user study,” in *ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, 2005.*, Nashville, TN, USA, 2005, pp. 164–171.

[63] H. Huttenrauch and K. S. Eklundh, “Fetch-and-carry with cero: Observations from a long-term user study with a service robot,” in *Proceedings. 11th IEEE International Workshop on Robot and Human Interactive Communication.* Berlin, Germany: IEEE, 2002, pp. 158–163.

[64] T. L. Chen, M. Ciocarlie, S. Cousins, P. M. Grice, K. Hawkins, K. Hsiao, C. C. Kemp, C. King, D. A. Lazewatsky, A. E. Leeper, H. Nguyen, A. Paepcke, C. Pantofaru, W. D. Smart, and L. Takayama, “Robots for humanity: using assistive robotics to empower

people with disabilities," *IEEE Robotics Automation Magazine*, vol. 20, no. 1, pp. 30–39, 2013.

[65] D. C. Howell, *Statistical Methods for Psychology*, 7th ed. Cengage Learning, 2009.

[66] R. H. Myers, *Classical and Modern Regression with Applications*, 2nd ed. Duxbury Thomson Learning, 1990.

[67] D. C. Montgomery, *Design and Analysis of Experiments*, 9th ed. John Wiley & Sons, 2017.

[68] H. Wickham and J. Bryan, *readxl: Read Excel Files*, 2019.

[69] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2019.

[70] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015.

[71] T. Hothorn, F. Bretz, and P. Westfall, "Simultaneous inference in general parametric models," *Biometrical Journal*, vol. 50, no. 3, pp. 346–363, 2008.

[72] H. Wickham, *ggplot2: Elegant Graphics for Data Analysis.* Springer-Verlag New York, 2016.