Implementation of a Multiple Switch Time Approach to Style-Based Motion Segmentation

A Thesis

Presented to

the Faculty of the School of Engineering and Applied Science
University of Virginia

In Partial Fulfillment
of the Requirements for the Degree

Master of Science
in Systems Engineering

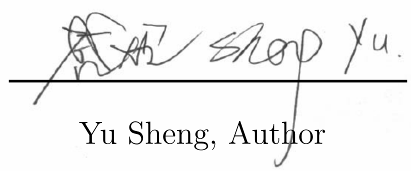
by

Yu Sheng

August
2014

APPROVAL SHEET

The thesis
is submitted in partial fulfillment of the requirements
for the degree of
Master of Science



Yu Sheng, Author

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

Prof. Amy LaViers, Thesis Advisor

Prof. Matthew Gerber , Committee Chairperson

Prof. Peter Beling, Committee Member

Accepted for the School of Engineering and Applied Science:

Dean, School of Engineering and Applied Science

August

2014

Acknowledgment

First, I would like to thank my advisor Professor LaViers, for her continuous support and help on my study and research; I could not finish my MS Proposal and Thesis without her patience and effort. Second, I would like to thank my committee members, Professor Gerber, Professor Beling and Professor Lin, for those precious questions and advices. Then, I would like to thank my cousin LeiNaiguang and his wife DongXiaojie for all the help and encouragement during the past years. And thanks to the members of RAD Lab, my friends and family, Jerry, Yuxiang, Dad, Mom and Imon.

The data used in this paper was obtained from mocap.cs.cmu.edu. The database was created with funding from NSF EIA-0196217.

Abstract

This thesis presents progress on segmenting human movement based on a notion of movement quality. This research is an extension of a style-based motion classification where, here, this classification is used to segment long motion phrases into smaller, discrete motion snippets. In particular, this thesis presents a given trajectory that is segmented into three shorter trajectories that each has their own length and quality. The objective of the thesis is to refine this segmentation, extend it to an arbitrary number of segmentation points, apply it to motion capture data and explore other extensions. A key novel contribution of this thesis is the analytical derivation of first order necessary conditions for optimality. The research may be used to build a library of motion primitives and aid the study of motion recognition in automation.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Motion, which is talked about, performed and planned by almost everybody in everyday life, is so familiar to us as human beings that sometimes we do not realize we execute so many movements in our life. For example, to get a mug on the dinner table, we have to separate it into several smaller tasks: stand up from the chair, open the door of the study, avoid the running cat, and finally grab the mug. This example suggests the fact that the way we execute movements and the way we think about them are not the same. We make decisions step-by-step with finite operations in our mind but movements are coherent and in a space of infinite states. It is obvious that our nervous system has a way of mapping between discreteness and continuousness by abstracting from a higher dimensional space to a lower one. Since this thesis is in the discipline of engineering, we are more interested in the state of arts in artificial systems, where the situation is not the same as natural systems.

As mathematical theories and computation tools developed, more and more researches have been devoted to the study of human motion, progress has been made in recognizing human gesture, controlling humanoid robots and other applications. The similarity of these previous works is they started from observed data and ended in certain functions. There is no doubt that these data-driven, or bottom-up, research efforts have solved many problems. However, as these progress expressed through mathematical or technical terms, like control laws, it is difficult for ordinary people without engineering training to benefit from them directly. This obstacle might be removed if computational systems are able to abstract motions to lower dimensional space of commands qualitatively, or implement the experience of movement to higher dimensional space of actions.

The difficulty of relating the qualitative experience of movement to a quantitative analysis – a top-down approach – lies in the gap between the nature of observable human motion and human cognition. The motions are executed in continuous time and have uncountable states; whereas, the state space of human decisions is finite, and they are made discretely. Since the machines are designed by human, it is easier to do experiments and collect data on them than modeling human or testing human systems. As a result of this relatively low barrier to entry, the bottom-up approach is common. The top-down approach, which is the focus of this thesis, offers significant progress that breaks with convention.

Before introducing our top-down framework, remember the human motion example in the first paragraph. People do not pick up a mug directly; the negative feedback loop in our body keeps modifying the motion trajectory until the goal is achieved. This intriguing claim could rise another question of why robots look so different from human when they execute a similar motion and with a feedback loop as well? Similarly, it is easy to discern a human from a human-like toy or Micky the Mouse. Whereas, it would not be so obvious what's the difference from a human to a zombie or Marvin. The first step of this top-down approach is modeling motion with mapping between the discreteness and continuousness.

The work of style-based motion classification by LaViers [1] provided a way of connecting the quantitative notion of motion to qualitative data, which combines the motion system of effort from Laban's dance theory [2] and inverse optimal control method from engineering discipline [3]. This research of motion segmentation offers a way of applying the classification progress in [1], which focus on single movement, to a series of movements by segmenting large dataset to smaller pieces. This segmentation work also makes it possible to build up a library of motion primitives, that have distinctive characteristic from each other.

This thesis begins with an introduction of the fundamental theory of the research and the parallel researches on motion with different approaches in Ch. 2, with a formal statement of the problem to be presented in Ch. 3. After that, we present the analytical and numerical solution to the problem in Ch. 4 & 5. A segmentaition method built on Fourier Transformation will be introduced in CH. 6. The comparison with other methods will be discussed in Ch. 7. And in Ch. 8 the expected contribution of the thesis to robots control and motion recognition will be bestowed.

Chapter 2

Background

The research presented in this thesis is a direct extension of the style-based motion classification by LaViers [1] and the optimal control of switch system [3]. Further, the work of LaViers [1] was built on the theory of Laban [2] and inverse optimal control [4]. Figure 2.1 shows the relationship of the fundamental theories.

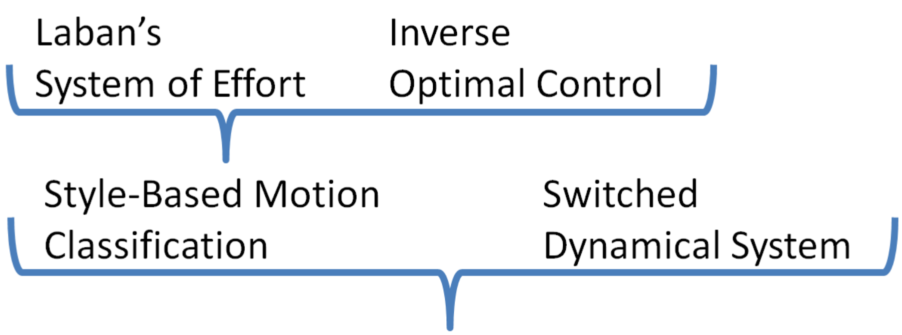


Figure 2.1: Theory foundation for the thesis.

In this chapter, we will first discuss the former approaches of classifying or segmenting motions in Sec. 2.1. And then the theory of Laban and style-based motion classification will be explained in Sec. 2.2. After that, the literature of inverse optimal control (Sec. 2.3) and will be reviewd.

2.1 Previous Attempts at Classification and Segmentation

The method in [1] is similar to many other inverse optimal control approaches [5–8] and other robotic control research [9–11]. And former researchers also have attempted to segment dynamic motion primitives (*movemes* [12]) which could be assembled into consistent motions [13–16] and used to generate orders for robots [17–19]. However, the method presented here incorporates a more corporal sense of movement by employing notions from dance theory; namely, movement *quality* as defined by dance scholar Rudolf Laban. Laban's set of codified *motion factors* measures quality, a quantity observed by mover and viewer that describes the nature of any given movement. His eight basic *efforts* illustrate the variety of possibly qualities.

Engineers have come up with some ideas of assigning certain style for motions. The work of [20] presented an inverse kinematics system based on a learned model of human poses. A real-time motion generation technique that allows us to generate the motions of a particular individual performing parameterized displacement activities was presented in [21]. Hauser et al. presented a method of computing efficient and natural looking motions for humanoid robots walking on varied terrain in [22], which used a small set of high-quality motion primitives.

It is necessary to point out that the majority method of human or robot motion studies is different from the one mentioned in Section 2.1. A mathematical representation for characterizing human arm motions was presented in [23], and the method was very simple for implementation and generated a human-like posture for an arbitrary arm configuration. The work of [24] presented two incremental teaching approaches to transfer gestures and associated constraints to a humanoid robot without using historical data, and compared the results with a batch training procedure. [25] presented an implementation of a PCA/ICA/HMM-based system to encode, generalize, recognize and reproduce gestures, and demonstrates the usefulness of using a stochastic method to encode the characteristic elements of a gesture and the organization of these elements. [26] presented a robotics-based approach for the synthesis of human motion using task-level control, it characterizes effort expenditure in terms of musculoskeletal parameters, rather than just skeletal parameters. [27] showed that the parametric description is qualitatively more appropriate to model the key features extracted from the trajectories. Li et al. surveys an idea of group motion segmentation in [28] and tries to build up the foundation for solving the problem of multi-object activity recognition. Rao et al. studied the problem of segmenting tracked feature point trajectories of multiple moving objects in an image sequence in [29]. These research focus on separating motion from the background or extracting the trajectory; they share the same term of motion segmentation as us by are not working on time series segmenting.

2.2 Style-Based Motion Classification and Laban

In order to achieve the desired shift from top-down to bottom-up methods, we introduce the concept of quality. This idea is not an invention from engineering but of dance theorist Rudolf von Laban. Laban developed a system of effort to describe the movement quality. He used four factors to describe a motion [2], which are space, weight, time and flow. The dynamosphere, Fig. 2.2, shows the relationship between three of them.

The dynamosphere is powerful for dancers, choreographers, and movement practitioners and observers, in general. However, this qualitative model does not indicate how to solve an engineering problem with data. A quantitative model then given by

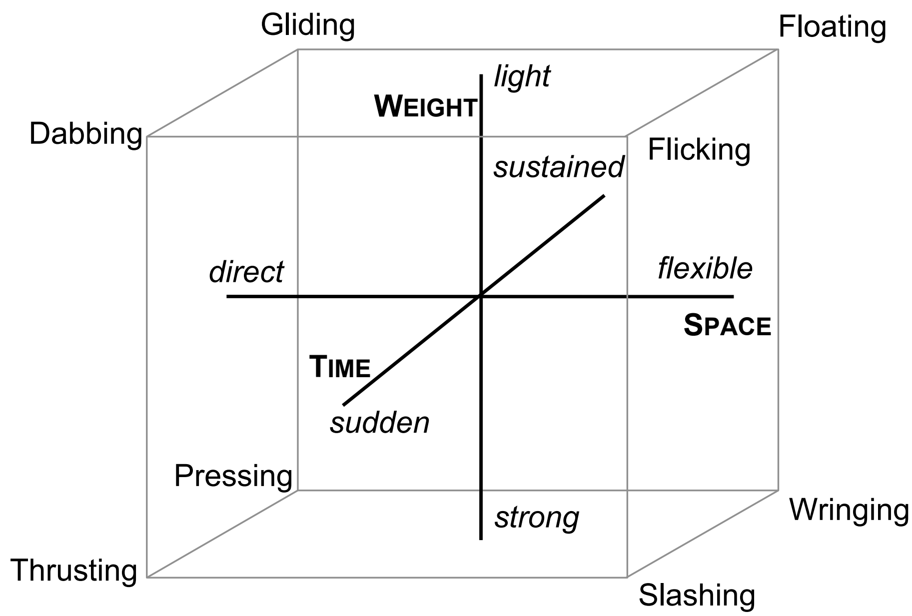


Figure 2.2: The dynamosphere. Laban's arrangement of eight basic efforts according to the axes of space, weight and time. The forth factor flow is not included in this sphere and the use of flow may be bound or free.

LaViers [30]. In the work an association was made as below,

$$Q \sim direct \tag{2.1}$$

$$R \sim light \tag{2.2}$$

$$P \sim sustained \tag{2.3}$$

$$S \sim bound. \tag{2.4}$$

Then a quadratic cost function was established as

$$J_u = \frac{1}{2}\int_0^T [(y-\sigma)'Q(y-\sigma) + u'Ru + \dot{x}'P\dot{x}]\,dt + \frac{1}{2}(y-\sigma)'S(y-\sigma)\bigg|_T \tag{2.5}$$

where $Q \in \mathbb{R}^{l\times l}$, $R \in \mathbb{R}^{m\times m}$, $P \in \mathbb{R}^{n\times n}$, and $S \in \mathbb{R}^{l\times l}$.

In order to generate a trajectory with desired quality, she then solved the optimal control problem as below,

$$\min_u J_u \tag{2.6}$$

$$s.t. \begin{cases} \dot{x} = Ax + Bu \quad x(0) = x_0 \\ y = Cx \end{cases}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, and $C \in \mathbb{R}^{l\times n}$. The optimal x (and thus y) are elected by the weights in Eq. 2.5. Based on the notion of style-based motion, the method of style-based motion classification was constructed by solving an inverse optimal problem [1].

2.3 Inverse Optimal Control

In [6] a way of transferring biological motions to robots was presented with some numerical method, but the analytical solution for the inverse optimal control problem was not been solved. Hatz etc. tried to estimate system parameter with inverse optimal control method in [31]. [32] shows an approach of obtaining physics-based representation of realistic character motion with inverse optimization method.

When researchers mentioned inverse optimal control in the 1950s, what they meant is to design regulators with optimization method. Fujii et al. presented a complete optimality condition for the multi-input inverse optimal control problem in [33]. Krstic and Tsiotras [34] use inverse optimal control to reconstruct optimal controllers from knowledge of a control Lyapunov function and a particular stabilizing control policy. Li, et al. [35] present an inverse optimal control method for nonlinear systems based on computing an approximate value function given a control policy.

The area became active again in the new millennium and was named inverse reinforcement learning by many researchers. In [36], a new algorithm was presented which allowed the researcher to solve the inverse problem without solving the forward problem. A optimality principle for human walking was studied in [37] with biological motor system.

Chapter 3

Statement of Problem

In this chapter, we will first formulate the problem we dealing with; and then the objectives and hypothesis of the research will be presented and discussed.

As shown in Fig. 3.1, we will segment a trajectory into $n+1$ pieces with n switching points. Note that the number of τ and π are not the same, we have one

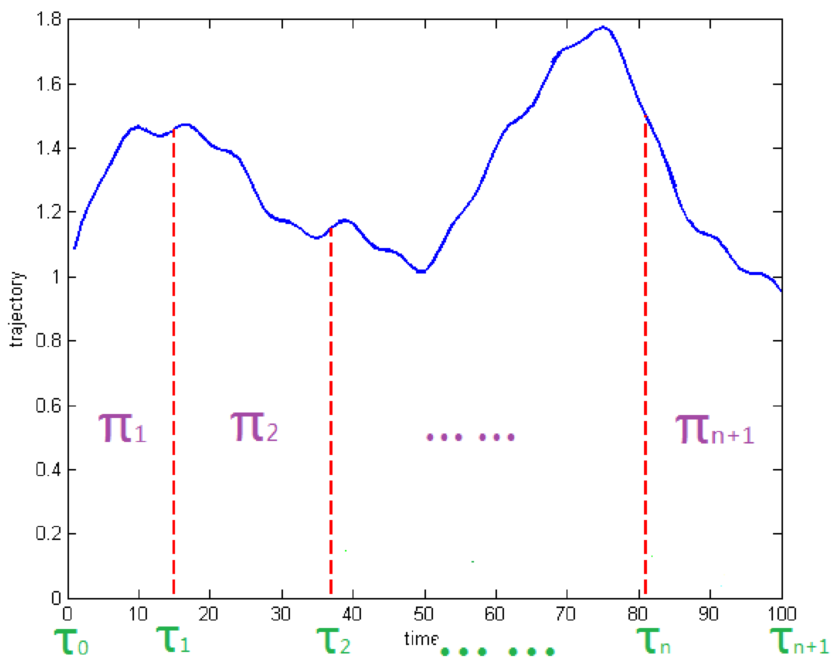


Figure 3.1: The n switching time problem with n segmenting points and n+1 qualities.

more element of quality than switch time. The problem of this research is to find a way to obtain proper π and τ for any trajectory. Here, we define

$$\mu = [\pi; \tau] \tag{3.1}$$

as our desired parameter. We set up a minimization problem to find μ, which is

$$\min_{\mu} \int L(x, \rho)dt + \Psi(x, \rho) \tag{3.2}$$

where L is the cost function and Ψ is the terminal cost. Since x is dependent on μ, we have to solve the control vector u for the dynamical system

$$\dot{x} = f(x, u) \tag{3.3}$$

Thus an optimal control problem will be established as

$$u = \arg\min \int F(x, u, \mu)dt + \psi(x, \rho). \tag{3.4}$$

Notice that there is a μ in F, which presents the dependence of u on μ.

This thesis is based on the following hypothesis: the previous notion of motion quality in 1-dimension provides a way to pose the segmentation problem by incorporating a human-generated single segment; in this thesis the success of this solution will be measured by analytical rigor.

The objective of this thesis is three-fold:

1. Segment a trajectory with single switch time by using an inverse optimal control problem (Ch. 4) ,

   (a) dealing with discontinuities that arise analytically,

   (b) implementing the solution.

2. Segment that trajectory into n pieces and classify the segments simultaneously (Ch. 5) ,
   (a) dealing with discontinuities that arise analytically and deduct the first order necessary optimal condition for the problem (Sec. 5.1) ,
   (b) implementing the solution with two segment points (Sec. 5.2) .
3. Compare the novel method to existing data analysis techniques (Ch. 6 & 7) .

Chapter 4

Single Switch Time Problem

To solve the problem step by step, we started at a single switch time style-based motion segmentation problem as

$$\min_{\tau} \int_0^T L(x,\rho)dt + \Psi_1(x(\tau),\rho(\tau)) + \Psi_2(x(T),\rho(T)) \tag{4.1}$$

$$\text{s.t.} \begin{cases} u_1 = \arg\min \int_0^{\tau} F(x,u_1,\sigma)dt + \psi_1(x(\tau),\sigma(\tau)) \\ u_2 = \arg\min \int_{\tau}^{T} F(x,u_2,\sigma)dt + \psi_2(x(T),\sigma(T)) \\ \dot{x} = \begin{cases} f_x(x,u_1) & [0,\tau) \quad x(0) = x_0 \\ f_x(x,u_2) & (\tau,T] \quad x(\tau) = x_\tau \end{cases} \end{cases} \tag{4.2}$$

where $\tau \in \mathbb{R}$ is a switch time variable that segments the novel data, $\rho \in \mathbb{R}^l$, into two movements; $x \in \mathbb{R}^n$ is the state; x_τ is given by integrating $f_x(x,u_1)$ on $[0,\tau)$; $u_1, u_2 \in \mathbb{R}^m$ are the inputs; $y \in \mathbb{R}^l$ is the output; and $\sigma \in \mathbb{R}^l$ is the reference signal representing a nominal movement.

4.1 Analytical Solution

Theorem 4.1.1. *The first order optimality conditions on τ with respect to the cost given in are*

$$\kappa_\tau(0) = 0 \tag{4.3}$$

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

$$\begin{cases} \dot{\kappa}_\tau = -\lambda_x \dfrac{\partial f_x}{\partial \tau} - \lambda_\xi \dfrac{\partial f_\xi}{\partial \tau} - \lambda_\sigma \dfrac{\partial f_\sigma}{\partial \tau} \\ \kappa_\tau(T) = 0 \end{cases}, \tag{4.4}$$

and with

$$\dot{\lambda}_x = -\frac{\partial L}{\partial x} - \lambda_x \frac{\partial f_x}{\partial x} - \lambda_\xi \frac{\partial f_\xi}{\partial x} \tag{4.5}$$

$$\dot{\lambda}_\xi = -\lambda_x \frac{\partial f_x}{\partial \xi} - \lambda_\xi \frac{\partial f_\xi}{\partial \xi} \tag{4.6}$$

$$\dot{\lambda}_\sigma = -\lambda_x \frac{\partial f_x}{\partial \sigma} - \lambda_\xi \frac{\partial f_\xi}{\partial \sigma} - \lambda_\sigma \frac{\partial f_\sigma}{\partial \sigma} \tag{4.7}$$

$$\textit{with} \begin{cases} \lambda_x(\tau^-) = \frac{\partial \Psi_1}{\partial x}(x(\tau)) - \lambda_\xi(\tau^-)\frac{\partial G_1}{\partial x}(x(\tau)) + \lambda_x(\tau^+) \\ \lambda_x(T) = \frac{\partial \Psi_2}{\partial x}(x(T)) - \lambda_\xi(T)\frac{\partial G_2}{\partial x}(x(T)) \\ \lambda_\xi(0) = 0 \\ \lambda_\xi(\tau^+) = 0 \\ \lambda_\sigma(0) = 0 \\ \lambda_\sigma(\tau^+) = \lambda_\sigma(\tau^-) + \lambda_\xi(\tau^-)\frac{\partial G_1}{\partial \sigma} \end{cases} \tag{4.8}$$

Since the results for this single switch time problem has been published in [38] and this single switch time problem is a special case of the problem we will discuss in the following sections, we do not present the proof for the analytical solution and the process we obtain the numerical solution in detail. Please see [38] for this detail.

4.2 Numerical Results

The results of the simulation are shown in Fig.4.1. The red trajectory is the original data ρ. The green line is the reference data σ. The blue trajectory $y(u^*)$ is generated as described in Sec. 2.2 from the original data ρ, quality π and segmenting time τ. Initial intuition might say that the segment time τ should be at the neighbor of a extremum point of ρ.

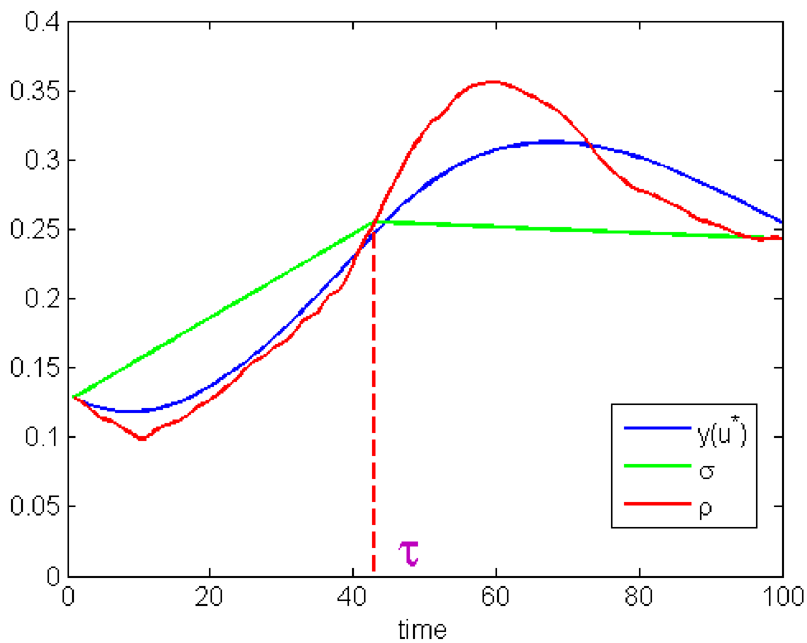
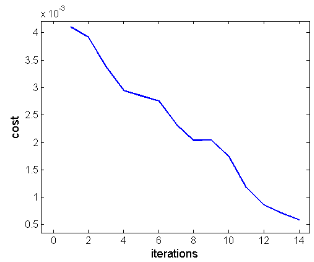
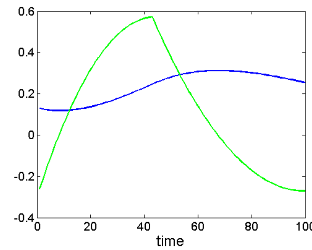


Figure 4.1: This figure shows the result of simulation: the output y (in blue) tracks the original data ρ (in red), the green line is the reference signal σ and the purple τ with red dash line indicates the segmentation point.

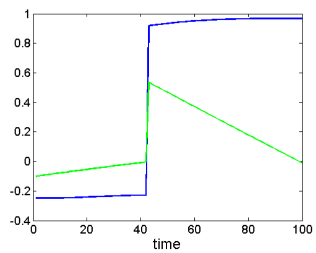
Figure 4.2(a) shows the cost converges when performing gradient descent with fixed step on τ. The plot of $x(t)$ over time is shown in Fig. 4.2(b) where the blue line is θ (the joint angle over time) and the green line is $\dot{\theta}$ (the velocity of the joint angle). They are both continuous and the sharp turn of $\dot{\theta}$ on τ reflects the fact that the control vector u is not continuous on τ. Various costates involved are also shown in Fig. 4.2; note their discontinuities at τ.



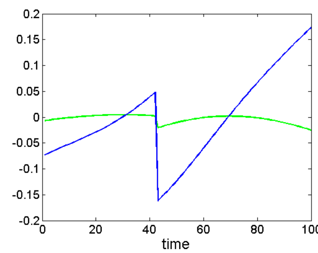
(a) Converging cost during numerical implementation.



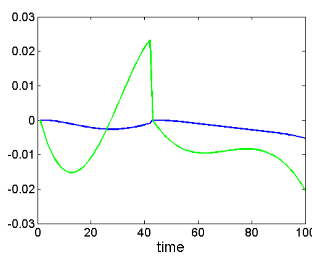
(b) $x(t) \in \mathbb{R}^2$.



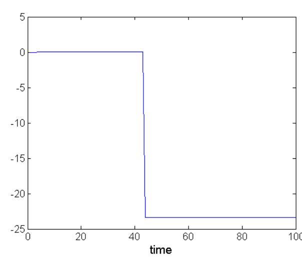
(c) $\xi(t) \in \mathbb{R}^2$.



(d) $\lambda_x(t)^T \in \mathbb{R}^2$.



(e) $\lambda_\xi(t)^T \in \mathbb{R}^2$.



(f) $\lambda_\sigma \in \mathbb{R}^2$.

Figure 4.2: These figures provide insight into the optimization process; from here we can see that the cost converges and that the x and λ_σ are continuous, and the costates involved have discontinuities.

Chapter 5

Multiple Switch Time Problem

We now extend the results from previous chapter to a multiple switch time problem with simultaneous classification. Let n be the number of switching times $\tau_1, \tau_2, ..., \tau_n$. Recalling that in the single switch time problem [38], we segmented the original trajectory into two shorter trajectories with the same quality; similarly we will segment the original trajectory into $n+1$ segments distinguished by qualities $\pi_1, \pi_2, ..., \pi_{n+1}$ by n switch time.

Setting $\pi = [\pi_1; \pi_2; ...; \pi_{n+1}]$, $\tau = [\tau_1; \tau_2; ...; \tau_n]$ and $\mu = [\tau; \pi]$, the problem is

$$\min_{\mu} \sum_{i=1}^{n+1} \left[\int_{\tau_{i-1}}^{\tau_i} L(x, \rho) dt + \Psi_k(x(\tau_k), \rho(\tau_k)) \right] \quad (5.1)$$

$$\text{s.t.} \begin{cases} u_i = \arg\min \int_{\tau_{i-1}}^{\tau_i} F(x, u_i, \sigma, \pi_i) dt + \psi_i(x(\tau_i), \rho(\tau_i)) \\ \text{s.t.} \\ \dot{x} = f(x, u_i) \quad x(\tau_{i-1}) = \begin{cases} x_0 & i = 1 \\ x(\tau_{i-2}) + \int_{\tau_{i-2}}^{\tau_{i-1}} \dot{x} dt & i = 2, 3, ..., n+1 \end{cases} \\ \dot{\sigma} = f_\sigma(\sigma, \tau, \rho) \quad \sigma(\tau_{i-1}) = \rho(\tau_{i-1}) \\ \qquad \sigma(\tau_i) = \rho(\tau_i) \end{cases}, \quad (5.2)$$

where $\tau_1, \tau_2, ..., \tau_n \in \mathbb{R}$ are switch time variables which segment the novel data $\rho \in \mathbb{R}^l$, into shorter "movements", $\tau_0 = 0$ and $\tau_{n+1} = T$ are the initial and terminal times, $x \in \mathbb{R}^n$ is the state, $u_1, u_2, ..., u_{n+1} \in \mathbb{R}^m$ are the inputs, $y \in \mathbb{R}^l$ is the output, $\sigma \in \mathbb{R}^l$ is a reference signal.

First, we derive the first order necessary optimal condition as the analytical solution for the problem. Then, we provide numerical implementation.

5.1 Analytical Solution

5.1.1 Forward Probelm

The first task at hand is to solve the "forward" optimization problem for u (which is the same as that used to generate quality-endowed trajectories from the previous section).

Thus, we solve

$$\min_{u_i} \int_{\tau_{i-1}}^{\tau_i} F(x, u_i, \sigma)dt + \psi_i(x(\tau_i), \sigma(\tau_i)) \quad (5.3)$$

$$\text{s.t.} \begin{cases} \dot{x} = f(x, u_i) & x(\tau_{i-1}) = \begin{cases} x_0 & i = 1 \\ x(\tau_{i-2}) + \int_{\tau_{i-2}}^{\tau_{i-1}} \dot{x} dt & i = 2, 3, ..., n+1 \end{cases} \\ \dot{\sigma} = f_\sigma(\sigma, \tau, \rho) & \sigma(\tau_{i-1}) = \rho(\tau_{i-1}) \\ & \sigma(\tau_i) = \rho(\tau_i) \end{cases} . \quad (5.4)$$

Lemma 5.1.1. *The first order necessary optimality conditions on u with respect to the cost in Eq. 5.3 subject to the constraints in Eq. 5.4 are given by*

$$\delta \hat{J}(u_i; \nu) = \kappa_{u_i}(\tau_{i-1}) = 0 \quad (5.5)$$

where

$$\dot{\kappa}_{u_i} = -\frac{\partial F}{\partial u_i} - \lambda\frac{\partial f_x}{\partial u_i} \tag{5.6}$$

$$\kappa_{u_i}(\tau_i) = 0 \tag{5.7}$$

and with

$$\dot{\lambda} = -\frac{\partial F}{\partial x} - \lambda\frac{\partial f_x}{\partial x} \tag{5.8}$$

$$\lambda(\tau_i^-) = \frac{\partial \psi_i}{\partial x}(x(\tau_i), \sigma(\tau_i)) \tag{5.9}$$

Proof. The augmented cost is given by

$$\hat{J}_{u_i} = \int_{\tau_{i-1}}^{\tau_i} \left[F(x, u_i, \sigma) + \lambda\left(f - \dot{x}\right)\right] dt + \psi_i(x(\tau_i)).$$

Consider a variation in u such that

$$u_i \mapsto u_i + \epsilon\nu.$$

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

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

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

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

$$= \int_{\tau_{i-1}}^{\tau_i}\left[\left(\frac{\partial F}{\partial u_i} + \lambda\frac{\partial f}{\partial u_i}\right)\nu + \left(\frac{\partial F}{\partial x} + \lambda\frac{\partial f}{\partial x}\right)\eta - \lambda\dot{\eta}\right] dt + \frac{\partial \psi_i}{\partial x}(x(\tau_i))\eta(\tau_{i-1}^-).$$

Integrating by parts gives

$$= \int_{\tau_{i-1}}^{\tau_i} \left[\left(\frac{\partial F}{\partial u_i} + \lambda \frac{\partial f}{\partial u_i} \right) \nu + \left(\frac{\partial F}{\partial x} + \lambda \frac{\partial f}{\partial x} - \dot{\lambda} \right) \eta \right] dt$$

$$+\lambda(\tau_{i-1})\eta(\tau_{i-1}) - \lambda(\tau_i^-)\eta(\tau_i^-) + \frac{\partial \psi_i}{\partial x}(x(\tau_i))\eta(\tau_i^-)).$$

To fill in the boundary conditions, note that $\eta(\tau_{i-1}) = 0$ because η starts at $x_{\tau_{i-1}}$, which is fixed. The following substitutions

$$\dot{\lambda} = -\frac{\partial F}{\partial x} - \lambda \frac{\partial f}{\partial x}$$
$$\lambda(\tau_i^-) = \frac{\partial \psi_i}{\partial x}(x(\tau_i))$$

reduce the derivative to

$$\delta \hat{J}_{u_i}(u_i;\nu) = \int_{\tau_{i-1}}^{\tau_i} \left(\frac{\partial F}{\partial u_i} + \lambda \frac{\partial f}{\partial u_i} \right) \nu \; dt,$$

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

$$\kappa_{u_i} = \int_t^{\tau_i} \left(\lambda \frac{\partial f}{\partial u_i} + \lambda \frac{\partial f}{\partial u_i} \right) dt.$$

Differentiating κ_u with respect to time gives the expressions in Eq. 5.7, which concludes the proof. ∎

Then the optimizer, u_i^*, can be expressed as a function of x, ξ, π_i and σ, where is the costate satisfying

$$\dot{\xi} = -\frac{\partial F^T}{\partial x} - \frac{\partial f^T}{\partial x}\xi \tag{5.10}$$

$$\xi(\tau_i) = \frac{\partial \psi}{\partial x}(x(\tau_i), \sigma(\tau_i)). \tag{5.11}$$

Plugging in the optimal u_i^* into the equations for x and ξ gives the expression for the Hamiltonian dynamics on interval (τ_{i-1}, τ_i):

$$\dot{x} = f_x(x, u_i^*(x, \xi, \sigma, \tau_{i-1}, \tau_i, \pi_i)) \tag{5.12}$$

$$\dot{\xi} = f_\xi(\xi, u_i^*(x, \xi, \sigma, \tau_{i-1}, \tau_i, \pi_i)), \tag{5.13}$$

which we denote with

$$\dot{x} = f_{x,u_i}(x, \xi, \sigma, \tau_{i-1}, \tau_i, \pi_i) \tag{5.14}$$

$$\dot{\xi} = f_{\xi,u_i}(x, \xi, \sigma, \tau_{i-1}, \tau_i, \pi_i). \tag{5.15}$$

5.1.2 Inverse Problem

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

$$\min_{\mu} \sum_{i=1}^{n+1} \left[\int_{\tau_{i-1}}^{\tau_i} L(x, \rho) dt + \Psi_i(x(\tau_i), \rho(\tau_i)) \right] \tag{5.16}$$

$$\text{s.t.} \begin{cases} \dot{x} = f_{x,u_i}(x, \xi, \sigma, \mu) & x(\tau_{i-1}) \\ & = \begin{cases} x_0 & i = 1 \\ x(\tau_{i-2}) + \int_{\tau_{i-2}}^{\tau_{i-1}} \dot{x} dt & i = 2, 3, ..., n+1 \end{cases} \\ \dot{\xi} = f_{\xi,u_i}(x, \xi, \sigma, \mu) & \xi(\tau_i^-) = G_i(x(\tau_i), \sigma(\tau_i)) \\ \dot{\sigma} = f_\sigma(\sigma, \tau_{i-1}, \tau_i, \rho) & \sigma(\tau_{i-1}) = \rho(\tau_{i-1}) \\ & \sigma(\tau_i) = \rho(\tau_i) \end{cases}, \tag{5.17}$$

where ρ is the empirical data we wish to mimic and thus segment with $\tau = [\tau_1, \tau_2, ..., \tau_n]^T$. The function G_i is place-holder for the conditions derived in the Eq. 5.9; namely,

$G_i = \frac{\partial \psi_i}{\partial x}(x(\tau_i))$.

Theorem 5.1.1. *The first order optimality conditions on μ with respect to the cost given in are*

$$\kappa_\mu(0) = 0 \tag{5.18}$$

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

$$\dot{\kappa}_\mu = -\lambda_x \frac{\partial f_x}{\partial \mu} - \lambda_\xi \frac{\partial f_\xi}{\partial \mu} - \lambda_\sigma \frac{\partial f_\sigma}{\partial \mu} \tag{5.19}$$

$$\kappa_\mu(T) = 0, \tag{5.20}$$

and with

$$\dot{\lambda}_x = -\frac{\partial L}{\partial x} - \lambda_x \frac{\partial f_x}{\partial x} - \lambda_\xi \frac{\partial f_\xi}{\partial x} \tag{5.21}$$

$$\dot{\lambda}_\xi = -\lambda_x \frac{\partial f_x}{\partial \xi} - \lambda_\xi \frac{\partial f_\xi}{\partial \xi} \tag{5.22}$$

$$\dot{\lambda}_\sigma = -\lambda_x \frac{\partial f_x}{\partial \sigma} - \lambda_\xi \frac{\partial f_\xi}{\partial \sigma} - \lambda_\sigma \frac{\partial f_\sigma}{\partial \sigma} \tag{5.23}$$

$$\textit{with} \begin{cases} \lambda_x(\tau_i^-) = \frac{\partial \Psi_i}{\partial x}(x(\tau_i)) - \lambda_\xi(\tau_i^-)\frac{\partial G_i}{\partial x}(x(\tau_i)) + \lambda_x(\tau_i^+) \\ \lambda_x(T) = \frac{\partial \Psi_{n+1}}{\partial x}(x(T)) - \lambda_\xi(T)\frac{\partial G_{n+1}}{\partial x}(x(T)) \\ \lambda_\xi(0) = 0 \\ \lambda_\xi(\tau_i^+) = 0 \\ \lambda_\sigma(0) = 0 \\ \lambda_\sigma(\tau_i^+) = \lambda_\sigma(\tau_i^-) + \lambda_\xi(\tau_i^-)\frac{\partial G_i}{\partial \sigma} \\ i = 1, 2, ..., n \end{cases} \tag{5.24}$$

Proof. We begin by augmenting the cost,

$$\hat{J}_\mu = \int_0^T L(x,\rho)dt$$
$$+\sum_{i=1}^{n+1}\left\{\int_{\tau_{i-1}}^{\tau_i}\left[\lambda_x(f_{x,u_{i-1}}-\dot{x})+\lambda_\xi(f_{\xi,u_{i-1}}-\dot{\xi})+\lambda_\sigma(f_\sigma-\dot{\sigma})\right]dt\right.$$
$$\left.+\Psi_i(x(\tau_i),\rho(\tau_i))\right\}.$$

Now consider a variation in τ such that

$$\mu \mapsto \mu + \epsilon\theta.$$

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

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

The variation in τ also disturbs the costate's boundary condition, $\xi(\tau_i^-)$, which is given by $\xi(\tau_i^-) \mapsto \xi(\tau_i^-) + \epsilon\nu(\tau_i^-)$,but since $\xi(\tau_i^-) = G_i(x(\tau_i),\sigma(\tau_i))$, the relation can be written as

$$\xi(\tau_i^-) \mapsto \xi(\tau_i^-) + \epsilon\left(\frac{\partial G_i}{\partial x}\eta(\tau_i^-) + \frac{\partial G_i}{\partial \sigma}\omega(\tau_i^-)\right), \tag{5.25}$$

thus we obtain a expression for $\nu(\tau_i^-)$ as

$$\nu(\tau_i^-) = \left(\frac{\partial G_i}{\partial x}\eta(\tau_i) + \frac{\partial G_i}{\partial \sigma}\omega(\tau_i)\right) \tag{5.26}$$

The difference in the augmented cost and its variation is given by:

$$\hat{J}_\mu(\mu + \epsilon\theta) - \hat{J}_\mu(\mu) =$$

$$\sum_{i=1}^{n+1} \left\{ \int_{\tau_{i-1}}^{\tau_i} \left[\epsilon \frac{\partial L}{\partial x}\eta - \epsilon\lambda_x \dot{\eta} - \epsilon\lambda_\xi \dot{\nu} - \epsilon\lambda_\sigma \dot{\omega} \right.\right.$$

$$+\epsilon\lambda_x \left(\frac{\partial f_{x,u_i}}{\partial x}\eta + \frac{\partial f_{x,u_i}}{\partial \xi}\nu + \frac{\partial f_{x,u_i}}{\partial \sigma}\omega + \frac{\partial f_{x,u_i}}{\partial \mu}\theta \right)$$

$$\left. +\epsilon\lambda_\xi \left(\frac{\partial f_{\xi,u_i}}{\partial x}\eta + \frac{\partial f_{\xi,u_i}}{\partial \xi}\nu + \frac{\partial f_{\xi,u_i}}{\partial \sigma}\omega + \frac{\partial f_{\xi,u_i}}{\partial \mu}\theta \right) +\epsilon\lambda_\sigma \left(\frac{\partial f_\sigma}{\partial \sigma}\omega + \frac{\partial f_\sigma}{\partial \mu}\theta \right) \right] dt$$

$$\left. +\epsilon \frac{\partial \Psi_i}{\partial x}(x(\tau_i))\eta(\tau_i) \right\} + \mathcal{O}(\epsilon).$$

$$+\sum_{i=1}^{n} \int_{\tau_i^-}^{\tau_i^- + \epsilon\theta_i} \left[\epsilon \frac{\partial L}{\partial x}\eta - \epsilon\lambda_x \dot{\eta} - \epsilon\lambda_\xi \dot{\nu} - \epsilon\lambda_\sigma \dot{\omega} \right.$$

$$+\epsilon\lambda_x \left(\frac{\partial f_{x,u_i}}{\partial x}\eta + \frac{\partial f_{x,u_i}}{\partial \xi}\nu + \frac{\partial f_{x,u_i}}{\partial \sigma}\omega + \frac{\partial f_{x,u_i}}{\partial \mu}\theta \right)$$

$$\left. +\epsilon\lambda_\xi \left(\frac{\partial f_{\xi,u_i}}{\partial x}\eta + \frac{\partial f_{\xi,u_i}}{\partial \xi}\nu + \frac{\partial f_{\xi,u_i}}{\partial \sigma}\omega + \frac{\partial f_{\xi,u_i}}{\partial \mu}\theta \right) + \epsilon\lambda_\sigma \left(\frac{\partial f_\sigma}{\partial \sigma}\omega + \frac{\partial f_\sigma}{\partial \mu}\theta \right) \right] dt$$

$$-\sum_{i=1}^{n} \int_{\tau_i^+}^{\tau_i^+ + \epsilon\theta_i} \left[\epsilon \frac{\partial L}{\partial x}\eta - \epsilon\lambda_x \dot{\eta} - \epsilon\lambda_\xi \dot{\nu} - \epsilon\lambda_\sigma \dot{\omega} \right.$$

$$+\epsilon\lambda_x \left(\frac{\partial f_{x,u_i}}{\partial x}\eta + \frac{\partial f_{x,u_i}}{\partial \xi}\nu + \frac{\partial f_{x,u_i}}{\partial \sigma}\omega + \frac{\partial f_{x,u_i}}{\partial \mu}\theta \right)$$

$$\left. +\epsilon\lambda_\xi \left(\frac{\partial f_{\xi,u_i}}{\partial x}\eta + \frac{\partial f_{\xi,u_i}}{\partial \xi}\nu + \frac{\partial f_{\xi,u_i}}{\partial \sigma}\omega + \frac{\partial f_{\xi,u_i}}{\partial \mu}\theta \right) + \epsilon\lambda_\sigma \left(\frac{\partial f_\sigma}{\partial \sigma}\omega + \frac{\partial f_\sigma}{\partial \mu}\theta \right) \right] dt$$

Now the mean value theorem is applied on $[\tau_i, \tau_i + \epsilon\theta_i]$, causing the integrals there to vanish. Computing the directional derivative of the augmented cost produces:

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

$$= \sum_{i=1}^{n+1} \left\{ \int_{\tau_{i-1}}^{\tau_i} \left[\frac{\partial L}{\partial x}\eta - \lambda_x \dot{\eta} - \lambda_\xi \dot{\nu} - \lambda_\sigma \dot{\omega} \right.\right.$$

$$+\lambda_x \left(\frac{\partial f_{x,u_i}}{\partial x}\eta + \frac{\partial f_{x,u_i}}{\partial \xi}\nu + \frac{\partial f_{x,u_i}}{\partial \sigma}\omega + \frac{\partial f_{x,u_i}}{\partial \mu}\theta \right)$$

$$+\lambda_\xi \left(\frac{\partial f_{\xi,u_i}}{\partial x}\eta + \frac{\partial f_{\xi,u_i}}{\partial \xi}\nu + \frac{\partial f_{\xi,u_i}}{\partial \sigma}\omega + \frac{\partial f_{\xi,u_i}}{\partial \mu}\theta \right) + \lambda_\sigma \left(\frac{\partial f_\sigma}{\partial \sigma}\omega + \frac{\partial f_\sigma}{\partial \mu}\theta \right) \Bigg] \, dt$$

$$+\frac{\partial \Psi_i}{\partial x}(x(\tau_i))\eta(\tau_i) \Bigg\}$$

Integrating by parts and collecting terms by variation gives

$$= \int_0^T \left(\frac{\partial L}{\partial x} + \lambda_x \frac{\partial f_x}{\partial x} + \lambda_\xi \frac{\partial f_\xi}{\partial x} + \dot{\lambda}_x \right) \eta \; dt + \sum_{i=1}^{n+1} \left[\lambda_x(\tau_{i-1}^+)\eta(\tau_{i-1}) - \lambda_x(\tau_i^-)\eta(\tau_i) \right]$$

$$+ \int_0^T \left(\lambda_x \frac{\partial f_x}{\partial \xi} + \lambda_\xi \frac{\partial f_\xi}{\partial \xi} + \dot{\lambda}_\xi \right) \nu \; dt + \sum_{i=1}^{n+1} \left[\lambda_\xi(\tau_{i-1}^+)\nu(\tau_{i-1}^+) - \lambda_\xi(\tau_i^-)\nu(\tau_i^-) \right]$$

$$+ \int_0^T \left(\lambda_x \frac{\partial f_x}{\partial \sigma} + \lambda_\xi \frac{\partial f_\xi}{\partial \sigma} + \lambda_\sigma \frac{\partial f_\sigma}{\partial \sigma} + \dot{\lambda}_\sigma \right) \omega \; dt + \sum_{i=1}^{n+1} \left[\lambda_\sigma(\tau_{i-1}^+)\omega(\tau_{i-1}) - \lambda_\sigma(\tau_i^-)\omega(\tau_i) \right]$$

$$+ \int_0^T \left(\lambda_x \frac{\partial f_x}{\partial \mu} + \lambda_\xi \frac{\partial f_\xi}{\partial \mu} + \lambda_\sigma \frac{\partial f_\sigma}{\partial \mu} \right) \theta \; dt + \sum_{i=1}^{n+1} \frac{\partial \Psi_i}{\partial x}(x(\tau_i))\eta(\tau_i)$$

To unravel the boundary conditions, remember that $\eta(0) = 0$ since η starts at x_0, which is fixed. Likewise, $\sigma(0)$, $\sigma(x_{\tau_i})$ and $\sigma(T)$ are fixed to be the value of ρ at each respective time; thus, $\omega(0) = \omega(T) = 0$. It has also be shown that an expression in terms of $\eta(\tau_i)$ and $\omega(\tau_i)$ for $\nu(\tau_i)$ can be obtained as in Eq. 5.26.

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

$$\dot{\lambda}_x = -\frac{\partial L}{\partial x} - \lambda_x \frac{\partial f_x}{\partial x} - \lambda_\xi \frac{\partial f_\xi}{\partial x}$$
$$\dot{\lambda}_\xi = -\lambda_x \frac{\partial f_x}{\partial \xi} - \lambda_\xi \frac{\partial f_\xi}{\partial \xi}$$
$$\dot{\lambda}_\sigma = -\lambda_x \frac{\partial f_x}{\partial \sigma} - \lambda_\xi \frac{\partial f_\xi}{\partial \sigma} - \lambda_\sigma \frac{\partial f_\sigma}{\partial \sigma},$$

then the derivative has reduced to

$$= \sum_{i=1}^{n} \left(-\lambda_x(\tau_i^-) + \lambda_x(\tau_i^+) + \frac{\partial \Psi_i}{\partial x}(x(\tau_i)) + \lambda_\xi(\tau_i^-) \frac{\partial G_i}{\partial x}(x(\tau_i)) \right) \eta(\tau_i)$$

$$+ \left(\frac{\partial \Psi_{n+1}}{\partial x}(x(T)) - \lambda_x(T) - \lambda_\xi(T) \frac{\partial G_{n+1}}{\partial x} \right) \eta(T)$$

$$+ \sum_{i=1}^{n} \lambda_\xi(\tau_i^+)\nu(\tau_i^+) + \lambda_\xi(0)\nu(0)$$

$$+ \sum_{i=1}^{n} \left(\lambda_\sigma(\tau_i^+) - \lambda_\sigma(\tau_i^-) - \lambda_\xi(\tau_i^-) \frac{\partial G_i}{\partial \sigma} \right) \omega(\tau_i)$$

$$+ \int_0^T \left(\lambda_x \frac{\partial f_x}{\partial \mu} + \lambda_\xi \frac{\partial f_\xi}{\partial \mu} + \lambda_\sigma \frac{\partial f_\sigma}{\partial \mu} \right) \theta \; dt.$$

Setting

$$\lambda_x(\tau_i^-) = \frac{\partial \Psi_i}{\partial x}(x(\tau_i)) - \lambda_\xi(\tau_i^-) \frac{\partial G_i}{\partial x}(x(\tau_i)) + \lambda_x(\tau_i^+) \tag{5.27}$$

$$\lambda_x(T) = \frac{\partial \Psi_{n+1}}{\partial x}(x(T)) - \lambda_\xi(T) \frac{\partial G_{n+1}}{\partial x}(x(T)) \tag{5.28}$$

$$\lambda_\xi(0) = 0 \tag{5.29}$$

$$\lambda_\xi(\tau_i^+) = 0 \tag{5.30}$$

$$\lambda_\sigma(0) = 0 \tag{5.31}$$

$$\lambda_\sigma(\tau_i^+) = \lambda_\sigma(\tau_i^-) + \lambda_\xi(\tau_i^-) \frac{\partial G_i}{\partial \sigma} \tag{5.32}$$

$$i = 1, 2, ..., n$$

leaves

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

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

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

Differentiating κ_μ with respect to time gives the expressions in Eq. 5.20, concluding the proof. ■

5.1.3 Remarks

To express this result more explicitly, we will rewrite the κ_μ to κ_π and κ_τ in the following Remarks.

Remark 5.1.1.

$$\begin{cases} \dot{\kappa}_{\pi_i} = \begin{cases} -\lambda_x \dfrac{\partial f_{x,u_i}}{\partial \pi_i} - \lambda_\xi \dfrac{\partial f_{x,u_i}}{\partial \pi_i} - \lambda_\sigma \dfrac{\partial f_\sigma}{\partial \pi_i} & (\tau_{i-1}, \tau_i) \\ 0 & else \end{cases} \\ \kappa_{\pi_i}(T) = 0 \end{cases} \tag{5.33}$$

Remark 5.1.2.

$$\begin{cases} \dot{\kappa}_{\tau_i} = \begin{cases} -\lambda_x \dfrac{\partial f_{x,u_i}}{\partial \tau_i} - \lambda_\xi \dfrac{\partial f_{x,u_i}}{\partial \tau_i} - \lambda_\sigma \dfrac{\partial f_\sigma}{\partial \tau_i} & (\tau_{i-1}, \tau_i) \\ -\lambda_x \dfrac{\partial f_{x,u_{i+1}}}{\partial \tau_i} - \lambda_\xi \dfrac{\partial f_{x,u_{i+1}}}{\partial \tau_i} - \lambda_\sigma \dfrac{\partial f_\sigma}{\partial \tau_i} & (\tau_i, \tau_{i+1}) \\ 0 & else \end{cases} \\ \kappa_{\tau_i}(T) = 0 \end{cases} \tag{5.34}$$

Eq. 5.33 indicates that the quality of a segmented shorter trajectory only depends the information in the in terval. Similarly , Eq. 5.34 indicates that a certain segmenting time relies on the information of the two intervals beside it. However, these do not mean that we get same analytical solution with the one interval situation. Because the costates λ are influenced by every interval.

5.2 Numerical Results

In this section, the way of obtaining a numerical solution will be explained.

5.2.1 Necessary Derivations

We use a similar cost function as Eq.2.5

$$F = \frac{1}{2}\int_0^T \left[(y-\sigma)^T Q (y-\sigma) + u^T R u + \dot{x}^T P \dot{x}\right] dt \tag{5.35}$$

$$\psi_i = \frac{1}{2}(y-\sigma)^T S (y-\sigma)\Big|_{\tau_i} \tag{5.36}$$

$$L(x,\rho) = \frac{1}{2}\|y-\rho\|^2 \tag{5.37}$$

$$\Psi_i = \frac{1}{2}\|y-\sigma\|^2\Big|_{\tau_i}. \tag{5.38}$$

And we choose the following metries to set up the comparison of opur simulated trajectory with real data. We choose the same linear system as in Eq. 2.6, which is

$$\begin{cases} \dot{x} = Ax + Bu \quad x(0) = x_0 \\ y = Cx \end{cases} \tag{5.39}$$

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

Differentiating the Hamiltonian of the forward probelm in Sec. 5.1.1 :

$$H = \frac{1}{2}[(y-r)^T Q(y-r) + u^T Ru + \dot{x}^T P\dot{x}] + \lambda f \quad (5.40)$$

$$\frac{\partial H}{\partial u} = u^T(R + B^T PB) + x^T A^T PB + \lambda B = 0 \quad (5.41)$$

from which it follows that

$$u = -(R + B^T PB)^{-1}(B^T PAx - B^T \lambda^T) \quad (5.42)$$

and

$$\frac{\partial H}{\partial x} = -\dot{\lambda} = x^T(CTQC + ATPA) + u^T B^T PA + \lambda A - \sigma^T QC \quad (5.43)$$

from which it follows that

$$\begin{aligned}\dot{\xi} = -\dot{\lambda}^T &= [A^T PB(R + B^T PB)^{-1}B^T PA - C^T QC - A^T PA]x \\ &+[A^T PB(R + B^T PB)^{-1}B^T - A^T]\xi + C^T Q\sigma \quad (5.44)\\ \dot{x} &= (A - B(R + B^T PB)^{-1}B^T PA)x - B(R + B^T PB)^{-1}B^T \xi \quad (5.45)\end{aligned}$$

The reference signal we choose is

$$\dot{\sigma} = \frac{\rho(\tau_i) - \rho(\tau_{i-1})}{\tau_i - \tau_{i-1}} \quad (5.46)$$

The following derivatives will be necessary for the optimal conditions:

$$\frac{\partial \psi_i}{\partial x}(x(\tau_i)) = C^T SCx(\tau_i) - C^T S\sigma(\tau_i) \quad (5.47)$$

$$\frac{\partial^2 \psi_i}{\partial x^2}(x(\tau_i)) = C^T SC \quad (5.48)$$

$$\frac{\partial L}{\partial x} = x^T C^T C - \rho^T C \quad (5.49)$$

$$\frac{\partial \Psi_i}{\partial x}(x(\tau_i)) = C^T Cx(\tau_i) - C^T \rho(\tau_i) \quad (5.50)$$

$$\frac{\partial f_x}{\partial x} = A - B(R + B^T PB)^{-1} B^T PA \tag{5.51}$$

$$\frac{\partial f_\xi}{\partial x} = A^T PB(R + B^T PB)^{-1} B^T PA - C^T QC - A^T PA) \tag{5.52}$$

$$\frac{\partial f_x}{\partial \xi} = -B(R + B^T PB)^{-1} B^T \tag{5.53}$$

$$\frac{\partial f_\xi}{\partial \xi} = A^T PB(R + B^T PB)^{-1} B^T - A^T \tag{5.54}$$

$$\frac{\partial f_x}{\partial \sigma} = 0 \tag{5.55}$$

$$\frac{\partial f_\xi}{\partial \sigma} = C^T Q \tag{5.56}$$

$$\frac{\partial f_\sigma}{\partial \sigma} = 0 \tag{5.57}$$

$$\frac{\partial f_x}{\partial \tau_i} = 0 \tag{5.58}$$

$$\frac{\partial f_\xi}{\partial \tau_i} = 0 \tag{5.59}$$

$$\frac{\partial f_\sigma}{\partial \tau_i} = \begin{cases} -\dfrac{\sigma(\tau_i) - \sigma(\tau_{i-1})}{(\tau_i - \tau_{i-1})^2} & (\tau_{i-1}, \tau) \\ \dfrac{\sigma(\tau_{i+1}) - \sigma(\tau_i)}{(\tau_{i+1} - \tau_i)^2} & (\tau_i, \tau_{i+1}) \\ 0 & else \end{cases} . \tag{5.60}$$

Let

$$\Lambda = R + B^T PB \tag{5.61}$$

and

$$\pi_i = [\pi_{i1}; \pi_{i2}; \pi_{i3}; \pi_{i4}], \tag{5.62}$$

then the derivatives associated with π are:

$$\frac{\partial f_x}{\partial \pi_i} = \left[\frac{\partial f_x}{\partial \pi_{i1}}, \frac{\partial f_x}{\partial \pi_{i2}}, \frac{\partial f_x}{\partial \pi_{i3}}, \frac{\partial f_x}{\partial \pi_{i4}}\right] \tag{5.63}$$

$$\frac{\partial f_x}{\partial \pi_{i1}} = 0_{n\times l} \tag{5.64}$$

$$\frac{\partial f_x}{\partial \pi_{i2}} = B\Lambda^{-2}B^T(\pi_{i3}Ax + \xi) \tag{5.65}$$

$$\frac{\partial f_x}{\partial \pi_{i3}} = -B\Lambda^{-1}B^TAx + BB^TBH^{-1}B^T(\pi_{i3}Ax + \xi) \tag{5.66}$$

$$\frac{\partial f_x}{\partial \pi_{i4}} = 0_{n\times l} \tag{5.67}$$

$$\frac{\partial f_\xi}{\partial \pi_i} = \left[\frac{\partial f_\xi}{\partial \pi_{i1}}, \frac{\partial f_\xi}{\partial \pi_{i2}}, \frac{\partial f_\xi}{\partial \pi_{i3}}, \frac{\partial f_\xi}{\partial \pi_{i4}}\right] \tag{5.68}$$

$$\frac{\partial f_\xi}{\partial \pi_{i1}} = C^T(\sigma - Cx) \tag{5.69}$$

$$\frac{\partial f_\xi}{\partial \pi_{i2}} = -\pi_{i3}A^TB\Lambda^{-2}B^T(\pi_{i3}Ax + \xi) \tag{5.70}$$

$$\begin{aligned}\frac{\partial f_\xi}{\partial \pi_{i3}} = \quad & A^TB\Lambda^{-1}B^T(\pi_{i3}Ax + \xi)\pi_{i3}A^TBB^TB\Lambda^{-2}B^T(\pi_{i3}Ax + \xi) \\ & +(2\pi_{i2}A^TB\Lambda^{-1}B^TA - A^TA)x\end{aligned} \tag{5.71}$$

$$\frac{\partial f_\xi}{\partial \pi_{i4}} = 0_{n\times l} \tag{5.72}$$

5.2.2 Solving the Boundary Condition

The Hamiltonian dynamics are given by

$$\begin{bmatrix} \dot{x} \\ \dot{\xi} \\ \dot{w}_x \\ \dot{w}_\xi \end{bmatrix} = \begin{bmatrix} M & \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \\ \begin{matrix} -C^TC & 0 \\ 0 & 0 \end{matrix} & -M^T \end{bmatrix} \begin{bmatrix} x \\ \xi \\ w_x \\ w_\xi \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ C^TQ & 0 \\ 0 & C^T \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \sigma \\ \rho \end{bmatrix} \tag{5.73}$$

The entries of M are

$$M_{11} = A - B(R + b^T PB)^{-1} B^T PA \tag{5.74}$$

$$M_{12} = -B(R + B^T PB)^{-1} B^T \tag{5.75}$$

$$M_{21} = A^T PB(R + B^T PB)^{-1} B^T PA - C^T QC - A^T PA \tag{5.76}$$

$$M_{22} = A^T PB(R + B^T PB)^{-1} B^T - A^T \tag{5.77}$$

Set

$$z = [x, \xi, w_x, w_\xi]^T, \tag{5.78}$$

then we rewrite Eq.5.73 as

$$\dot{z} = \mathcal{M} z + \mathcal{N} \zeta, \tag{5.79}$$

where

$$\mathcal{M} = \begin{bmatrix} \multicolumn{2}{c}{\multirow{2}{*}{M}} & 0 & 0 \\ & & 0 & 0 \\ -C^T C & 0 & \multicolumn{2}{c}{\multirow{2}{*}{-M^T}} \\ 0 & 0 & & \end{bmatrix}, \tag{5.80}$$

$$\mathcal{N} = \begin{bmatrix} 0 & 0 \\ C^T Q & 0 \\ 0 & C^T \\ 0 & 0 \end{bmatrix}, \tag{5.81}$$

$$\zeta = \begin{bmatrix} \sigma \\ \rho \end{bmatrix}. \tag{5.82}$$

The solutions of z on $\tau_i 6-$ is

$$z(\tau_i^-) = e^{\mathcal{M}(\tau_i - \tau_{i-1})} z_0 + \int_{\tau_{i-1}}^{\tau_i} e^{\mathcal{M}(\tau_i - t)} \mathcal{N} \zeta(t) dt, \tag{5.83}$$

which can be written as

$$z(\tau_i^-) = \Omega_i z_{\tau_{i-1}} + \Gamma_i, \tag{5.84}$$

where

$$\Omega_i = e^{\mathcal{M}(\tau_i - \tau_{i-1})}, \tag{5.85}$$

$$\Gamma_i = \int_{\tau_{i-1}}^{\tau_i} e^{\mathcal{M}(\tau_i - t)} \mathcal{N} \zeta(t) dt. \tag{5.86}$$

We can write the jump condition at each switch time or segmenting point of z as

$$z(\tau_i^+) = \alpha_i z(\tau_i^-) + \beta_i, \tag{5.87}$$

where

$$\alpha_i = \begin{bmatrix} I_n & 0 & 0 & 0 \\ C^T SC & 0 & 0 & 0 \\ -C^T SC & 0 & I_n & C^T SC \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{5.88}$$

$$\beta_i = \begin{bmatrix} 0 \\ -\sigma(\tau)C^T \\ \rho(\tau)C^T \\ 0 \end{bmatrix}. \tag{5.89}$$

Lemma 5.2.1. *The expression of* $z(\tau)$ *on* z_0 *are*

$$z(\tau_i^-) = \left(\prod_{k=1}^{i-1} \Omega_k \alpha_k\right) \Omega_i z_0 + \left[\sum_{j=1}^{i-1} (\alpha_j \Gamma_j + \beta_j) \prod_{k=j+1}^{i-1} \Omega_k \alpha_k\right] \Omega_i + \Gamma_i, \tag{5.90}$$

$$z(\tau_i^+) = \left(\prod_{k=1}^{i} \Omega_k \alpha_k\right) z_0 + \sum_{j=1}^{i} (\alpha_j \Gamma_j + \beta_j) \prod_{k=j+1}^{i} \Omega_k \alpha_k, \tag{5.91}$$

$$i = 0, 1, 2, ..., n+1.$$

Proof. We use mathematical induction to prove the lemma.

When $i = 0$, $z(\tau_0) = z(0)$, so the statement holds for $i = 0$.

When $i = 1$, $z(\tau_1^-) = \Omega_1 z_0 + \Gamma_1$ and $z(\tau_1^+) = \Omega_1 \alpha_1 z_0 + \alpha_1 \Gamma_1 + \beta_1$, so the statement holds for $i = 1$.

When $i = 2$,

$$z(\tau_2^-) = \Omega_2 z(\tau_1^+) + \Gamma_2 = (\Omega_1 \alpha_1)\Omega_2 z_0 + (\alpha_1 \Gamma_1 + \beta_1)\Omega_2 + \Gamma_2,$$

$$z(\tau_2^+) = \alpha_2 z(\tau_2^-) + \beta_2 = (\Omega_1 \alpha_1 \Omega_2 \alpha_2) z_0 + (\alpha_1 \Gamma_1 + \beta_1)\Omega_2 \alpha_2 + (\alpha_2 \Gamma_2 + \beta_2),$$

so the statement holds for $i = 2$.

Assume that the statement holds when $i = m$, so

$$z(\tau_m^-) = \left(\prod_{k=1}^{m-1} \Omega_k \alpha_k\right) \Omega_m z_0 + \left[\sum_{j=1}^{m-1} (\alpha_j \Gamma_j + \beta_j) \prod_{k=j+1}^{m-1} \Omega_k \alpha_k\right] \Omega_m + \Gamma_m,$$

$$z(\tau_m^+) = \left(\prod_{k=1}^{m} \Omega_k \alpha_k\right) z_0 + \sum_{j=1}^{m} (\alpha_j \Gamma_j + \beta_j) \prod_{k=j+1}^{m} \Omega_k \alpha_k.$$

When $i = m+1$,

$$z(\tau_{m+1}^{-}) = \Omega_{m+1} z(\tau_m^{+}) + \Gamma_{m+1}$$
$$= \left(\prod_{k=1}^{m} \Omega_k \alpha_k\right) \Omega_{m+1} z_0 + \left[\sum_{j=1}^{m} (\alpha_j \Gamma_j + \beta_j) \prod_{k=j+1}^{m} \Omega_k \alpha_k\right] \Omega_{m+1} + \Gamma_{m+1},$$

$$z(\tau_{m+1}^{+}) = \alpha_{m+1} z(\tau_{m+1}^{-}) + \beta_{m+1}$$
$$= \left(\prod_{k=1}^{m} \Omega_k \alpha_k\right) \Omega_{m+1} \alpha_{m+1} z_0$$
$$+ \left[\sum_{j=1}^{m} (\alpha_j \Gamma_j + \beta_j) \prod_{k=j+1}^{m} \Omega_k \alpha_k\right] \Omega_{m+1} \alpha_{m+1} + (\Gamma_{m+1} \alpha_{m+1} + \beta_{m+1})$$
$$= \left(\prod_{k=1}^{m+1} \Omega_k \alpha_k\right) z_0 + \sum_{j=1}^{m+1} (\alpha_j \Gamma_j + \beta_j) \prod_{k=j+1}^{m+1} \Omega_k \alpha_k,$$

so the statement also holds for $i = m+1$.

Thus the statement holds for i=0, 1, 2, ... , n+1 . ■

Since we set $\tau_{n+1} = T$ and there is no discontinuity at terminal, so

$$z(T) = z(\tau_{n+1}) = z(\tau_{n+1}^{-})$$
$$= \left(\prod_{k=1}^{n} \Omega_k \alpha_k\right) \Omega_{n+1} z_0 + \left[\sum_{j=1}^{n} (\alpha_j \Gamma_j + \beta_j) \prod_{k=j+1}^{n} \Omega_k \alpha_k\right] \Omega_{n+1} + \Gamma_n \tag{5.92}$$

Then the solution of $z(T)$ can be written as

$$z(T) = \Phi z_0 + \eta, \tag{5.93}$$

where

$$\Phi = \left(\prod_{k=1}^{n} \Omega_k \alpha_k \right) \Omega_{n+1}, \tag{5.94}$$

$$\eta = \left[\sum_{j=1}^{n} (\alpha_j \Gamma_j + \beta_j) \prod_{k=j+1}^{n} \Omega_k \alpha_k \right] \Omega_{n+1} + \Gamma_n. \tag{5.95}$$

And from this algebra, we can find the necessary boundary conditions for z, let

$$\Phi = \begin{bmatrix} \Phi_{11} & \Phi_{12} & \Phi_{13} & \Phi_{14} \\ \Phi_{21} & \Phi_{22} & \Phi_{23} & \Phi_{24} \\ \Phi_{31} & \Phi_{32} & \Phi_{33} & \Phi_{34} \\ \Phi_{41} & \Phi_{42} & \Phi_{43} & \Phi_{44} \end{bmatrix}, \tag{5.96}$$

$$\eta = \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix}. \tag{5.97}$$

Next, plug in the following boundary conditions:

$$\xi(T) = C^T S C x(T) - C^T S \sigma(T) \tag{5.98}$$

$$w_x(T) = C^T C x(T) - C^T \rho(T) - C^T S C w_\xi(T) \tag{5.99}$$

$$w_\xi(0) = 0 \tag{5.100}$$

Rearranging the resulting expression gives

$$\begin{bmatrix} x(T) \\ \xi(0) \\ w_x(0) \\ w_\xi(T) \end{bmatrix} = \mathcal{A}^{-1}\left(\begin{bmatrix} \Phi_{11} \\ \Phi_{21} \\ \Phi_{31} \\ \Phi_{41} \end{bmatrix} x_0 + \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix} + \mathcal{B}\begin{bmatrix} \sigma(T) \\ \rho(T) \end{bmatrix}\right), \tag{5.101}$$

where

$$\mathcal{A} = \begin{bmatrix} I_n & -\Phi_{12} & -\Phi_{13} & 0 \\ C^TSC & -\Phi_{22} & -\Phi_{23} & 0 \\ -C^TC & -\Phi_{32} & -\Phi_{33} & C^TSC \\ 0 & -\Phi_{42} & -\Phi_{43} & I_n \end{bmatrix}, \tag{5.102}$$

$$\mathcal{B} = \begin{bmatrix} 0 & 0 \\ C^TS & 0 \\ 0 & -C^T \\ 0 & 0 \end{bmatrix} \tag{5.103}$$

and where the invertibility of $\mathcal{A}$ is guaranteed by the complete controllability of the system [1]. Thus the desired dynamics, $\dot{z}$ and z_0, where y recreates ρ optimally according to the structure of J_μ as best as the structure of J_u will allow are given.

5.2.3 Simulation Results

For this implementation, we choose

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 \end{bmatrix}.$$

And the weight matrices in J_u are set as

$$Q = qI_l$$
$$R = rI_m$$
$$P = pI_n$$
$$S = sI_l$$

where I_n is an $n \times n$ identity matrix.

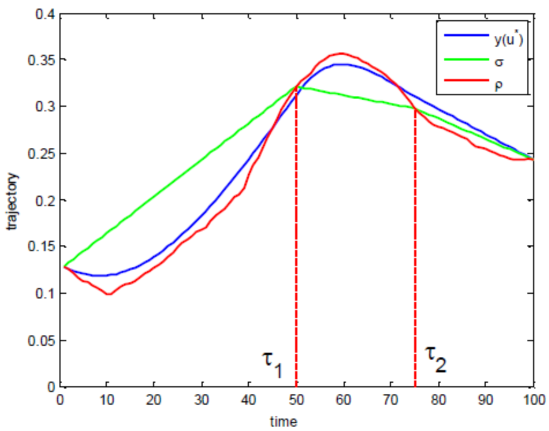
Thus the quality parameter could be presented as

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

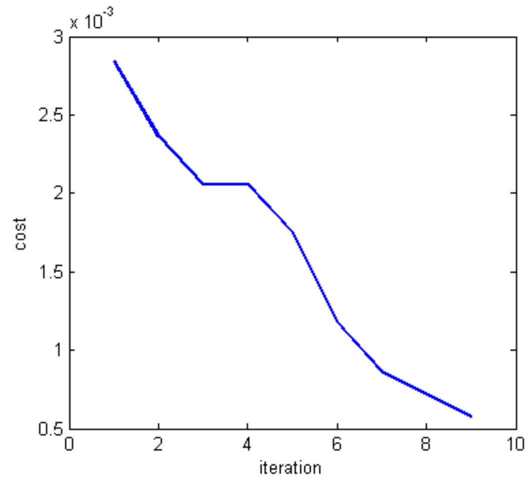
With the setup, we can implement the segmentation method on a single joint angle, like the motion of the elbow. However, as it is an initial process, the implementation showed in this section worked on some simulated data rather than the real data captured from human motion.

First, we will generate an original trajectory ρ with an initial segmentation point $\bar{\tau}$. The quality of the trajectory before and after $\bar{\tau}$ are different. What should be pointed out is that, this is the only parameter that helps to generate ρ and the solution of the minimization to find τ^* might not equal $\bar{\tau}$. Then, we initialize the states of the linear system in Eq.2.6 as $x_0 = [\rho(0), \rho'(0)]^T$. Performing gradient descent with fixed step on τ, we can get the converging cost as shown in Fig.5.1(b) and obtain the optimal value of τ as shown in Figure 5.1(a).

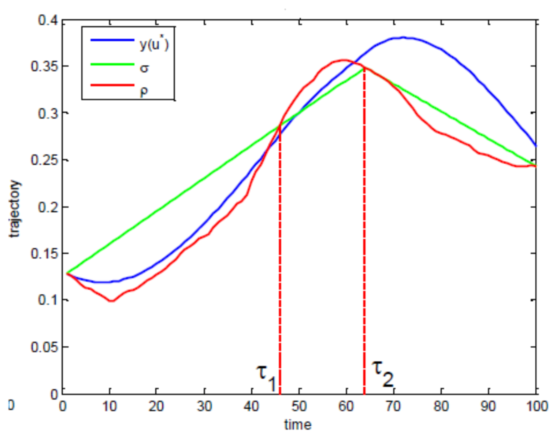
Since we only achieve first order necessary optimality condition for the problem, a global optimal is not guaranteed during numerical process. Figure. 5.1(c) is a good example for this. The blue trajectory goes apart from the red one after τ_2 and have a similar shape as it. This is because the numerical solution we get in this instance is a local optimal or saddle point, so the quality we get for the second and third pieces are not good enough. As the generation of the trajectory is influenced by the whole information on $[0, T]$, there is a maximum point after τ_2 rather than between τ_1 and τ_2. Figure. 5.1(d) shows another type of unsatisfied result we got. It is caused by the reference signal σ we choose.



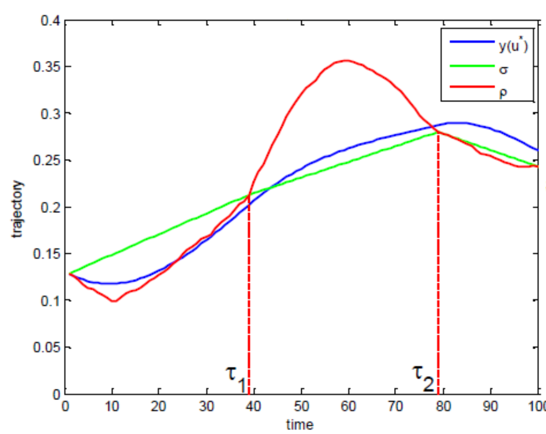
(a) ρ is the original trajectory needed to be segmented. σ is a reference ,. $y(u*)$ is the trajectory with a certain quality generated by our algorithm.



(b) Converging cost during numerical implementation.



(c) ρ is the original trajectory needed to be segmented. σ is a reference ,. $y(u*)$ is the trajectory with a certain quality generated by our algorithm.



(d) ρ is the original trajectory needed to be segmented. σ is a reference ,. $y(u*)$ is the trajectory with a certain quality generated by our algorithm.

Figure 5.1: The numerical results for a segmentation problem with 2 segmenting points.

Chapter 6

Segmentation Method Based on Fourier Transformation

In this chapter, we present a segmentation method building on Fourier Transformation [39], more specifically we use Fast Fourier Transform algorithm [40] to analyze the frequency characteristic of data, which is the movement trajectory in our research.

We established a utility function to compute the difference in frequency spectrum between each segmented pieces. As shown in Fig. 6.1, we choose the couple of segmenting points where the utility reaches its maximum. The segmentation result and the spectrum associated with each segmented piece are as in 6.2.

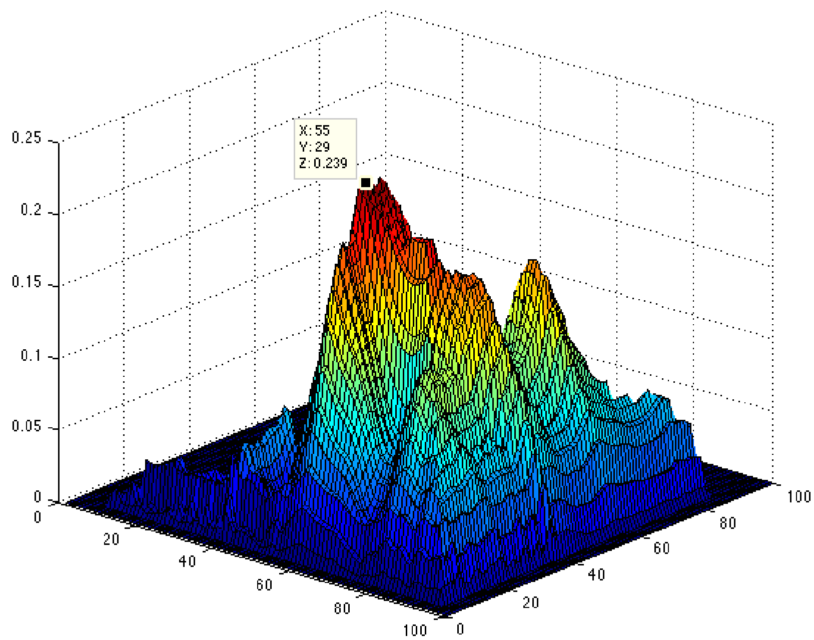


Figure 6.1: Searching for segmenting points combination with maximum utility.

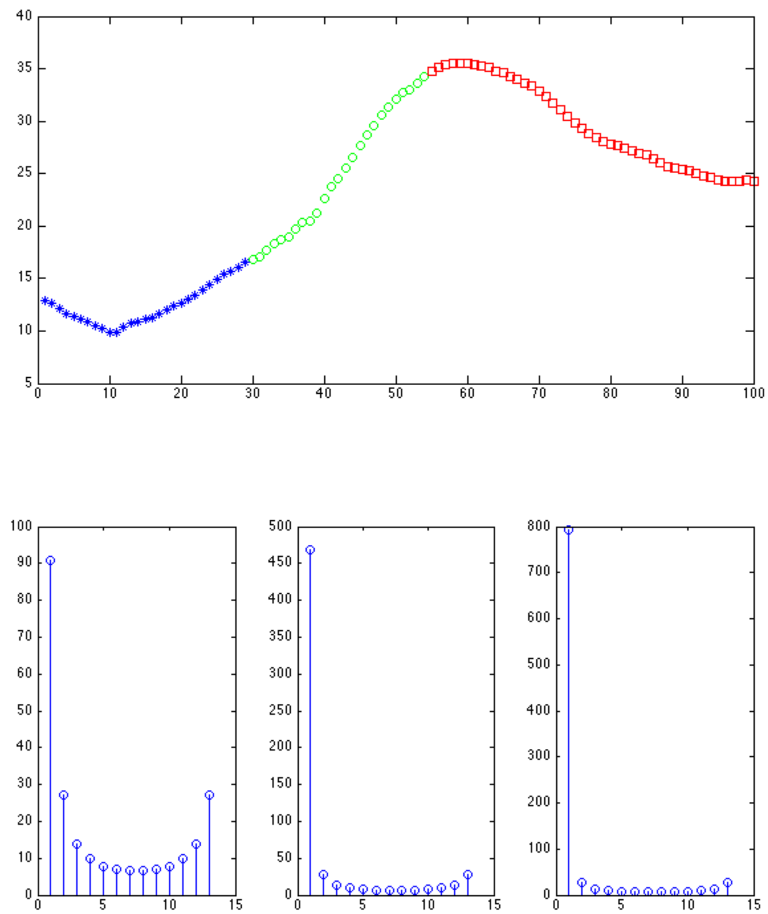


Figure 6.2: Segmentation result with Fourier transformation and the spectrum associated with each segmented piece.

Chapter 7

Comparison to Existing Data Analysis Methods

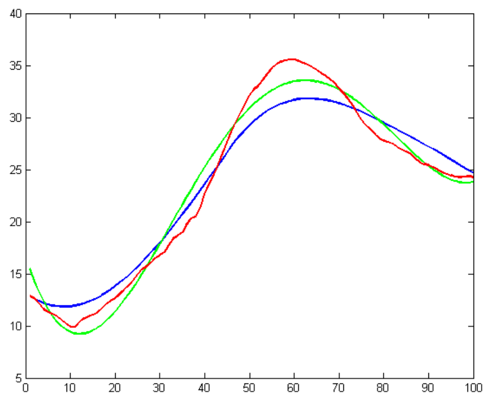
In this chapter, we will compare our method of style-based segmentation with other existing data analysis methods. The main point to note here is how our methods fundamentally different from these common data analysis techniques.

7.1 Polynomial Curve Fitting

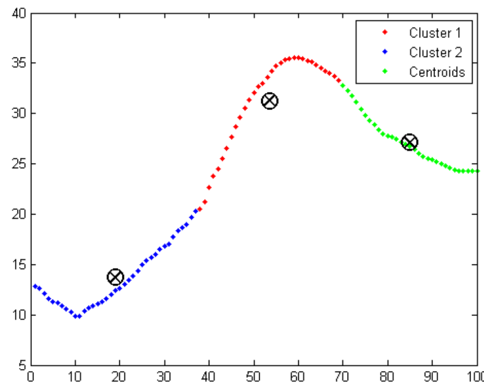
As we aim to segment the original trajectory into 3 shorter ones, we have 12 parameter for the polynomial curve fitting. As shown in 7.1(a), polynomial curve fitting has a smaller cost than our classification method. However, as explained in Section 2.3, our classification method provides a mapping between the qualitative description and the quantitative data.

7.2 Segmenting by Clustering Methods

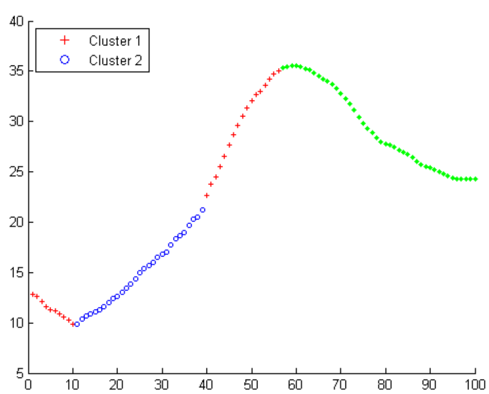
K-means clustering [41] tends to segment the trajectory in to three shorter ones with almost same number of points as in Figure 7.1(b). It is not a proper way of segmentation because the segments are not temporally contiguous. With the Gaussian distribution model, EM-clustering [42] can achieve better result that K-means clustering. Sometimes the algorithm even clusters points not close to each other into one set, the red dots in Figure 7.1(c). Figure 7.1(d) shows that the optimal cluster sets number is 4.



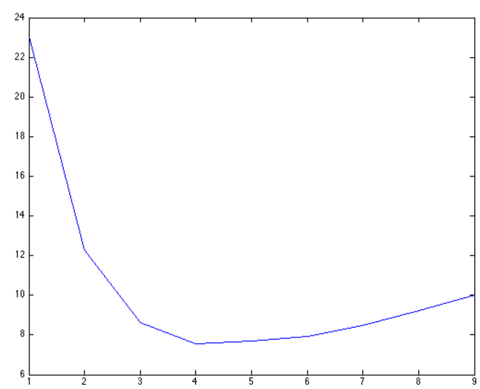
(a) Polynomial Curve Fitting



(b) K-Means Clustering



(c) Expectation-maximization Algorithm



(d) Selecting the Number of Clusters

Figure 7.1: Segmenting results with data clustering methods.

7.3 Summary

The shown in Table 7.1, our method has a better performance than other methods.

Table 7.1: A summary of Qualitative Comparison

	Robust Approximation	Contiguous Segmentation
K-means	N	Y
EM	Y	N
Style-based Segmentation	Y	Y

Chapter 8

Conclusion

In summary, a novel approach to the segmentation of human movement has been posed; in particular, the ability to incorporate a user defined notion of a single motion as a template for segmenting a longer trajectory makes this approach distinctive. This thesis gives an analytical solution and design a numerical method to solve the problem. And, the study of human motion – particularly as it may be incorporated into robotics and automation – will benefit from the tool proposed here as discussed in this chapter.

If human motions could be segmented into movements with certain quality, then it would be possible to use a combination of movements to represent a motion. And the once continuous trajectory would become discrete and the redundant information would be eliminated by 'sampling'. Also, the state space of motion would become finite instead of infinite and this would be much easier for human to analyze something discrete and finite. With these primitives, it will be possible to build up a library of motion primitives. We could achieve better understanding of human motion based on these primitives. The traditional way of motion recognition as comparing the captured data with database could cost a lot both in time and storage space. And at most occasions, the data of the motion is redundant for machine to recognize. With the segmentation technique, the motion could be representing as a string, whose letters represent a certain motion primitive in the motion library. Then the comparing process could be executed very fast. Or the gesture could be spoken out directly, since the string might be meaningful in a dictionary for human motion.

Bibliography

[1] A. LaViers and M. Egerstedt, "Style-based abstractions for human motion classification," in *5th International Conference on Cyber-Physical Systems*, 2014.

[2] R. v. Laban and L. Ullmann, *The language of movement : a guidebook to choreutics.* Boston: Plays, inc, 1976.

[3] F. Delmotte, E. I. Verriest, and M. Egerstedt, "Optimal impulsive control of delay systems," *ESAIM: Control, Optimisation and Calculus of Variations*, vol. 14, no. 04, pp. 767–779, 2008.

[4] D. P. Bertsekas, D. P. Bertsekas, D. P. Bertsekas, and D. P. Bertsekas, *Dynamic programming and optimal control.* Athena Scientific Belmont, MA, 1995, vol. 1, no. 2.

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

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

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

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

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

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

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

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

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

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

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

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

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

[18] E. Drumwright, O. Jenkins, and M. Mataric, "Exemplar-based primitives for humanoid movement classification and control," in *Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE International Conference on*, vol. 1. IEEE, 2004, pp. 140–145.

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

[20] K. Grochow, S. L. Martin, A. Hertzmann, and Z. Popović, "Style-based inverse kinematics," in *ACM Transactions on Graphics (TOG)*, vol. 23, no. 3. ACM, 2004, pp. 522–531.

[21] R. Urtasun, P. Glardon, R. Boulic, D. Thalmann, and P. Fua, "Style-based motion synthesis," in *Computer Graphics Forum*, vol. 23, no. 4. Wiley Online Library, 2004, pp. 799–812.

[22] K. Hauser, T. Bretl, K. Harada, and J.-C. Latombe, “Using motion primitives in probabilistic sample-based planning for humanoid robots,” in *Algorithmic Foundation of Robotics VII.* Springer, 2008, pp. 507–522.

[23] S. Kim, C. H. Kim, and J. H. Park, “Human-like arm motion generation for humanoid robots using motion capture database,” in *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on.* IEEE, 2006, pp. 3486–3491.

[24] S. Calinon and A. Billard, “Incremental learning of gestures by imitation in a humanoid robot,” in *Proceedings of the ACM/IEEE international conference on Human-robot interaction.* ACM, 2007, pp. 255–262.

[25] ——, “Recognition and reproduction of gestures using a probabilistic framework combining pca, ica and hmm,” in *Proceedings of the 22nd international conference on Machine learning.* ACM, 2005, pp. 105–112.

[26] O. Khatib, E. Demircan, V. De Sapio, L. Sentis, T. Besier, and S. Delp, “Robotics-based synthesis of human motion,” *Journal of physiology-Paris*, vol. 103, no. 3, pp. 211–219, 2009.

[27] S. Calinon and A. Billard, “Stochastic gesture production and recognition model for a humanoid robot,” in *Intelligent Robots and Systems, 2004.(IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on*, vol. 3. IEEE, 2004, pp. 2769–2774.

[28] R. Li and R. Chellappa, “Group motion segmentation using a spatio-temporal driving force model,” in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on.* IEEE, 2010, pp. 2038–2045.

[29] S. Rao, R. Tron, R. Vidal, and Y. Ma, “Motion segmentation in the presence of outlying, incomplete, or corrupted trajectories,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 10, pp. 1832–1845, 2010.

[30] A. LaViers and M. Egerstedt, "Style based robotic motion," in *American Control Conference (ACC), 2012*, June 2012, pp. 4327–4332.

[31] K. Hatz, J. P. Schloder, and H. G. Bock, "Estimating parameters in optimal control problems," *SIAM Journal on Scientific Computing*, vol. 34, no. 3, pp. A1707–A1728, 2012.

[32] C. K. Liu, A. Hertzmann, and Z. Popović, "Learning physics-based motion style with nonlinear inverse optimization," in *ACM Transactions on Graphics (TOG)*, vol. 24, no. 3. ACM, 2005, pp. 1071–1081.

[33] T. Fujii and M. Narazaki, "A complete optimally condition in the inverse problem of optimal control," in *Decision and Control, 1984. The 23rd IEEE Conference on*, vol. 23, Dec 1984, pp. 1656–1659.

[34] M. Krstic and P. Tsiotras, "Inverse optimal stabilization of a rigid spacecraft," *Automatic Control, IEEE Transactions on*, vol. 44, no. 5, pp. 1042–1049, 1999.

[35] W. Li, E. Todorov, and D. Liu, "Inverse optimality design for biological movement systems," in *IFAC*, vol. 18, no. 1, 2011, pp. 9662–9667.

[36] K. Dvijotham and E. Todorov, "Inverse optimal control with linearly-solvable mdps," in *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, 2010, pp. 335–342.

[37] G. Arechavaleta, J.-P. Laumond, H. Hicheur, and A. Berthoz, "An optimality principle governing human walking," *Robotics, IEEE Transactions on*, vol. 24, no. 1, pp. 5–14, 2008.

[38] Y. Sheng and A. LaViers, "Style-based human motion segmentation," in *Systems, Man, and Cybernetics (SMC), 2014 IEEE International Conference on*, 2014 (Accepted).

[39] L. Debnath and D. Bhatta, *Integral transforms and their applications.* CRC press, 2010.

[40] A. V. Oppenheim, R. W. Schafer, J. R. Buck *et al.*, *Discrete-time signal processing.* Prentice-hall Englewood Cliffs, 1989, vol. 2.

[41] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics.* Berkeley, Calif.: University of California Press, 1967, pp. 281–297. [Online]. Available: http://projecteuclid.org/euclid.bsmsp/1200512992

[42] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Statist. Soc. Ser. B*, vol. 39, no. 1, pp. 1–38, 1977, with discussion.