Doctoral Dissertation

Expert-Induced Latent Features as Feedback for Sit-to-Stand Therapy

Bryan Lao

February 10, 2020

Graduate School of Information Science
Nara Institute of Science and Technology

A Doctoral Dissertation
submitted to Graduate School of Information Science,
Nara Institute of Science and Technology
in partial fulfillment of the requirements for the degree of
Doctor of ENGINEERING

Bryan Lao

Thesis Committee:

Professor Kazushi Ikeda	(Supervisor)
Professor Satoshi Nakamura	(Co-supervisor)
Associate Professor Tomoya Tamei	(Co-supervisor)
Professor Ichiro Takeuchi	(Nagoya Institute of Technology)

Expert-Induced Latent Features as Feedback for Sit-to-Stand Therapy*

Bryan Lao

Abstract

Physiotherapy is a rehabilitation activity that restores physical function. Its labor-intensive process, coupled with a severe shortage in trained healthcare providers, have made this service increasingly inaccessible. Telehealth has emerged as an alternative to conventional therapy, but with various limitations. Chief among them is the model choice for the human subject. Due to the complexity of the human motor system, matching the appropriate model to a particular motor task is non-trivial. To address this problem, we propose the use of latent variable models, due to its desirable properties in the context of physiotherapy.

In this dissertation, we propose modifications to two latent variable models and extract meaningful features from the modified latent spaces. The latent features represent both motor strategies as well as feedback mechanisms. First, we introduce an aggregated approach to muscle synergy extraction, where the group synergy set allows the direct comparison between synergy activations. We propose and analyze three metrics to infer some characteristics of intervention. Second, we introduce a reorganization approach to a Gaussian Process embedding of multiple motor task performances. We formulate two forms of augmented feedback from the reorganized latent space for facilitating self-exercise.

The utility of the proposed metrics is demonstrated through the sit-to-stand task, an important whole-body exercise in physiotherapy. By comparing the extracted metrics through various natural and expert-induced conditions, we are able to gain some insight on the expert strategy as well as build tools for providing expert-level feedback.

*Doctoral Dissertation, Graduate School of Information Science,
Nara Institute of Science and Technology, February 10, 2020.

Keywords:

Physiotherapy, latent variable model, muscle synergy, Gaussian process, sit-to-stand, augmented feedback

Contents

List of Figures

1 Introduction

Recent demographic trends have caused a severe shortage of trained medical practitioners. As the world population ages at an ever-increasing rate, the workforce simply cannot keep up using traditional methods. Moving at an equally impressive rate, however, is the technology available to us today. Information and Communication Technologies (ICT) have the potential to improve the quality of healthcare and to make it accessible to more people. Telehealth is technology-driven healthcare that has developed to enhance patient satisfaction, overcome physical barriers, and reduce medical costs. In the age of artificial intelligence, these systems need not be limited to the automation of mundane, repetitive tasks. They can instead leverage on the latest technologies and perform smarter, sophisticated tasks. However, the current state of telehealth has not yet caught up.

Many of the existing telehealth systems are limited in their scope, targeted only at a handful of tasks and a specific set of users. These limitations are caused in no small part by the complexity of the human body, logistical constraints of physiotherapy, and the ethical concerns of safety. These issues continue to challenge the progress of telehealth solutions today. However, one can still aim at making many small meaningful improvements to make a system that is "greater than the sum of its parts". The author believes that the existing telehealth solutions can stand to improve significantly by creating flexible frameworks with the following properties: incorporating therapist motor skill, modeling the user's whole body, and learning in a data-efficient manner. This thesis aims to address the identified limitations of existing telehealth systems, with the hope that the results of our study can benefit the fields of telehealth and physiotherapy.

1.1 Research Motivation

Physiotherapy is a rehabilitation activity that improves and restores physical function. While conventional face-to-face therapy is effective in treating many common injuries, access has become increasingly difficult for many individuals. Among many factors, the shortage of practitioners and physical distance were identified as major contributors [16]. This shortage is expected to worsen as the world population is aging at an unprecedented rate. At present, so-called developed countries are already considered aged, led by Japan (28%) and Italy (23%); developing countries are following this trend at an even faster pace [17]. As a result, countries are shifting long-term elderly care from institutions to home- and community-based services, and remote therapy has emerged as an accessible alternative to conventional therapy [18].

Remote therapy, or telehealth, is defined as the delivery of health-related services and information via telecommunications technologies [19]. Recent developments use immersive technologies, like augmented reality (AR) and virtual reality (VR) to simulate environments close to conventional therapy. These systems tend to use single-purpose models, which work well for a single exercise and a few patients. However, they can be difficult to implement in practice due to the logistical constraints of physiotherapy.

Physiotherapy can be a labor-intensive process, involving multiple face-to-face sessions with a physical therapist (PT) (Figure 1.1). In each therapy session, the PT and patient practice a large variety of movements with a limited amount of time. This amount of practice is an order of magnitude lower than what is expected to induce neural reorganization [20]. Thus, the patient is required to continue the exercises themselves, without the corrective guidance from the PT [3]. The main challenge is to develop a flexible framework that can model arbitrary expert-guided motor behavior with only few available data. These issues could potentially be addressed using a generative model to learn therapist-induced motor behavior.

The long-term goal of this research is to develop a general framework for personalized self-exercise that incorporates the expert therapist's motor skill. The framework needs to accept arbitrary expert-guided exercises and provide the user with intuitive feedback that matches the expert's. The emphasis is on providing

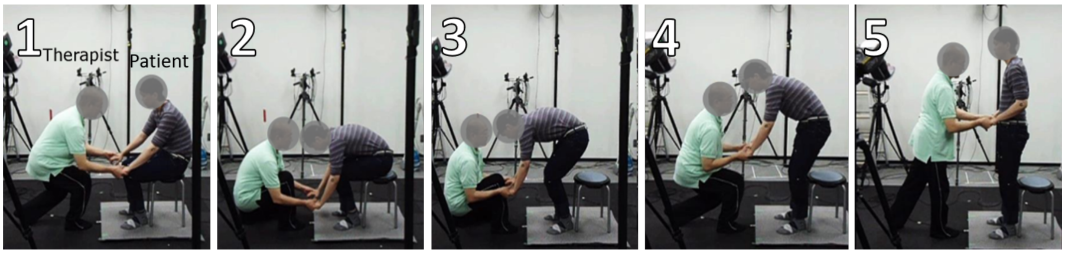

Figure 1.1: Therapist (left) induces proper form of exercise on the patient (right).

expert-level feedback with the limited number of demonstrations available during collocated therapy.

The solutions formulated in this study are inspired by the synergistic concepts found in motor control theory. "Synergy" literally means *working together*, referring to how parts of the human motor system produce a motor behavior. The number of degrees of freedom in the human body far exceeds the minimum number required to finish a task. As a result, a given task can be accomplished by different movement patterns. This is referred to as the "degree of freedom" problem [21]. How the central nervous system (CNS) generates the appropriate muscle patterns to achieve a behavioral goal remains an open question, but it has been hypothesized by many that the CNS simplifies the computational difficulty of motor control using low-dimensional latent controllers called *synergies* [22]. The synergy hypothesis is appealing from a computational point of view, as it represents a physiological analog to the theoretical latent variable model.

1.2 Overview of the Problem

Physiotherapy is a labor-intensive process that has become increasingly inaccessible due to a shortage of trained practitioners [16]. Telehealth has emerged as an accessible alternative to conventional therapy, however with its own sets of limitation. The main challenges are: developing a model that can adapt to the large variety of exercises performed per session and the individual needs of each patient; and from this data-efficient model, extract simple feedback mechanisms to facilitate self-exercise. A promising approach is to use a generative model to learn therapist-induced motor behavior and extract low-dimensional features as feedback.

In recent years, telehealth systems make use of immersive technologies to simulate environments close to conventional therapy. However, these systems tend to use single-purpose frameworks, which can be cumbersome to calibrate for the variety of different exercises and users. We highlight three issues in particular:

1. The physical therapist's motor skills are ignored. Many systems are designed according to the information provided by the expert PT. Expert *knowledge* such as what types of exercise and target angles are useful, but expert *motor skills* play an equally important role during therapy [23].

2. Multi-joint movements are largely unaddressed. Focus has been largely on isolated individual movements, such as finger motion, knee, or shoulder movements. Important whole-body movements, such as the sit-to-stand exercise, cannot be addressed.

3. Patient-specific calibration is difficult to achieve. On the one hand, many systems rely on generic expert templates. The template may differ greatly from the target user's body type and physical condition, potentially suggesting painful postures. On the other hand, automatic calibration would require large amounts of personalized training data, which is impractical with the limited time available during a therapy session. For context, an average post-stroke therapy session was found to be 36 minutes long, requiring patients to perform up to 17 types of movements [24].

The presentation of the feedback mechanisms play an important role in motor skills learning. *Feedback* refers to performance-related information that a learner receives for performing a task, and augmented feedback refers to supplementary information that is not naturally learned. Typically, two types of augmented feedback can be presented to a learner. One conveys knowledge of results (KR) while the other conveys knowledge of performance (KP) [25]. By proposing to use more complex models and higher-dimensional data, dimensionality reduction is necessary. For this purpose, we propose to modify existing latent variable models and extract such meaningful feedback mechanisms.

To work towards a practical implementation of a data-efficient self-exercise framework, the aforementioned issues need to be addressed. Our strategy is to first formulate a data-efficient model that incorporates the expert therapist's motor skills. Second, meaningful features are extracted from the model for presenting simple augmented feedback mechanisms.

1.3 Research Contribution

This thesis aims at a personalized self-exercise framework that incorporates the expert therapist's motor skill. We address several limitations of existing telehealth feedback systems, highlighting three issues in particular: exclusion of expert motor skill, focus on single-joint movements, and high cost of personalized calibration. These issues are addressed through simple metrics extracted from latent variable models. Their utility is demonstrated through the sit-to-stand task, an important whole-body exercise in physiotherapy.

The main contributions of this study are as follows: (1) First, we characterize the changes in motor behavior induced by an expert therapist. The concept of muscle synergy is used to quantify the components of motor strategy. We introduce an aggregated approach to muscle synergy extraction, which results in a novel comparison method that favors the temporal component of synergies. From the extracted synergies, we extract and analyze three metrics to infer some characteristics of intervention. (2) Second, we develop a data-efficient framework for expert-level feedback. The concept of Gaussian Process Dynamical Model (GPDM) is used to encode high-dimensional human motion data efficiently in a low-dimensional latent space. We introduce a reorganization approach to a Gaussian Process embedding of multiple motor task performances. From the reorganized latent space, we extract two types of augmented feedback mechanisms for self-correction.

1.4 Organization of the Thesis

This thesis is divided into five chapters. Chapter 2 gives an overview of related works. Chapter 3 presents the application of a linear latent variable model to quantify neuromusculuar changes during physiotherapy. Chapter 4 presents the application of a non-linear latent variable model to encode expert-induced movements in a data-efficient manner. Lastly, Chapter 5 concludes the thesis with a summary of our key results and an outlook on future research.

2 Related Works

In this thesis, latent variable models are proposed as a natural way of understanding and representing the underlying mechanisms of human motion in physiotherapy. This chapter summarizes some of the fundamental concepts and problems addressed in the succeeding chapters. Each section includes a literature survey and a succinct explanation of the relevant concept. The sections are organized as follows: Section 2.1 discusses conventional physiotherapy practice and presents examples of existing telehealth solutions. Section 2.2 introduces the latent variable model as the common conceptual basis of the mathematical models used in this work. Section 2.3 describes an example of the linear latent variable model, in the context of muscle synergy analysis. Section 2.4 describes an example of the non-linear latent variable model, in the context of modeling high-dimensional data. Finally, Section 2.5 provides a general background and motivation for our chosen motor task, the sit-to-stand motion.

2.1 Modern Physiotherapy

Physiotherapy is a type of intervention that improves and restores physical function, usually after an injury or surgery. For example, after a dislocated shoulder is put back in place, a patient is taught several shoulder exercises to help restore strength and range-of-movement [3]. An essential part of the therapeutic process is continuous and frequent assessments by a healthcare professional, called a physical therapist (PT) [26].

Conventional physiotherapy involves three major activities between patient and PT: *assessment*, *movement guidance*, and *exercise between sessions* [23]. First, the PT diagnoses the patient's conditions and prescribes the appropriate exercises to target the problematic area. Second, the PT guides the patient in performing the exercise, where proper form is emphasized. Finally, without the guidance of the PT, the patient continues performing the exercises regularly until the next session. This cycle repeats indefinitely until the patient's condition improves. In the absence of feedback, patients risk re-aggravating an injury due to incorrect movements [27].

Several studies investigate patient response during sit-to-stand motion. These studies demonstrate two common interactions between patient and healthcare professional: caregiving and physiotherapy. Chugo et al developed a force assistance system for standing up motion based on the "Kamiya scheme" (Figure 2.1A). The system focuses on providing assistance based on the the remaining amount of force necessary to stand up [1]. On the other hand, Kogami et al analyzed the effects of intervention for standing up motion based on the "Bobath concept" (Figure 2.1B). The study focuses on quantifying the effects of *intervention* on hemiplegic patients [2]. Although some actions may appear similar, the main purpose of caregiving is providing support to a patient's activities of daily living, while the goal of physiotherapy is to help the patient regain the ability to stand up. Currently, it is not well-understood how these techniques affect the sit-to-stand strategies, thus the need for further empirical evidence.

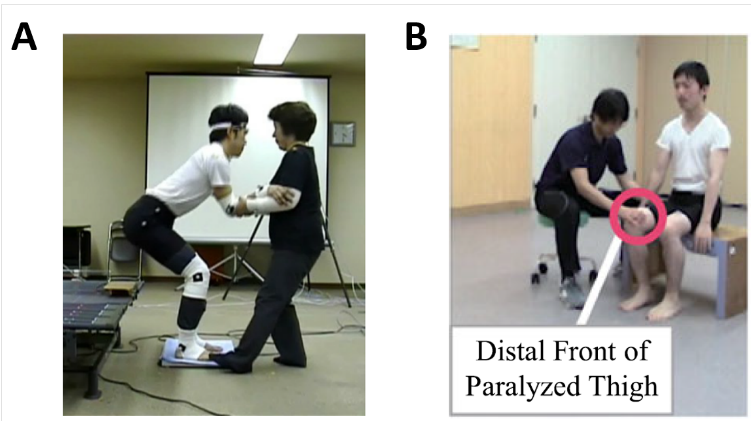

Figure 2.1: Styles of assisted sit-to-stand, **(A)** "Kamiya scheme" used by a nursing specialist [1], **(B)** "Bobath" concept used by a physical therapist [2].

In recent times, remote therapy has emerged as an accessible alternative to conventional therapy, driven by lack of access and an increasing shortage of practitioners [16]. Remote therapy, or telehealth, is defined as the delivery of health-related services and information via telecommunications technologies (ICT) [19]. Recent developments use immersive technologies, like augmented reality (AR) and virtual reality (VR) to simulate environments close to conventional therapy.

These systems usually have sensor to track the user's joints, then present a virtual avatar of his limbs as well as a target location. An AR example is the popular augmented mirror setup that guides users through pre-recorded exercises. For example, Physio@Home demonstrates four shoulder exercises. The system tracks a user's joints and overlays them on the user's virtual body, where a target shoulder angle is presented (Figure 2.2A) [3]. A VR example is the simulation of a 3D environment to facilitate arm-reaching exercises. This system also tracks a user's joint positions, while vibratory feedback is given when task performance is successful (Figure 2.2B) [4]. These systems tend to use simple single-limb models, and calibrated to only a few patients and a single exercise. They can be difficult to implement in practice due to the logistical constraints of physiotherapy.

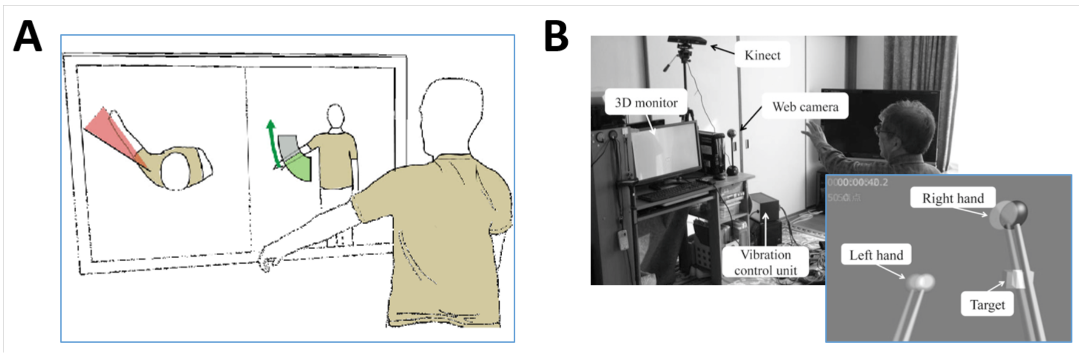

Figure 2.2: Existing telehealth solutions, **(A)** Augmented reality mirror setup [3], **(B)** 3D virtual reality telerehabilitation system [4]

2.2 Latent Variable Models

A common approach to machine learning is to learn in a supervised manner, i.e. find a mapping from a known input to a known output variable. Tasks such as classification and regression can be performed with great success, especially when all relevant factors are observable. Since the elemental organization of the data is not necessarily learned, this type of model is expected to work well only for the type of data used in training. An alternative approach is to learn in an unsupervised manner, taking into consideration that the data generation mechanism could be latent (unobservable).

The Latent Variable Model (LVM) approach assumes that the responses of manifest (observable) variables are controlled by a set of latent (unobservable) variables. LVMs can be used in a variety of machine learning tasks, where the latent variable fulfills different roles. For dimensionality reduction, latent variables are the low-dimensional representations of high-dimensional samples. For clustering, the latent points can represent real cluster membership. For noise reduction, they can represent the noise-free version of the observed data [28].

Many popular machine learning models can be considered LVMs, such as Principal Component Analysis and neural networks. Linear models are preferred in the medical field for their simplicity and simplifying power. However, they can be sometimes limited by the fact that the linearity assumption often does not hold. Thus, the resulting latent variables often become less interpretable after being transformed. By removing the assumption on linearity, more complex types of data can potentially be modeled (Figure 2.3). In the next sections, we discuss natural use-cases for linear models in modeling muscle activation, and a non-linear model used in modeling human motion.

2.3 Muscle Synergy Models

The human body is capable of performing a wide variety of motor tasks with a great deal of flexibility. This highlights the fact that the musculoskeletal system has a large number of degrees of freedom (DoF). Because there are more DoFs than those required to perform most tasks, the inherent dimensionality of performing a task is much lower than what is observed. The central nervous system (CNS) is believed to utilize predefined modules to simplify motor control, referred to as “motor primitives” or “synergies”. Synergies may refer to the working together of different physiological systems, such as joints, muscles, and

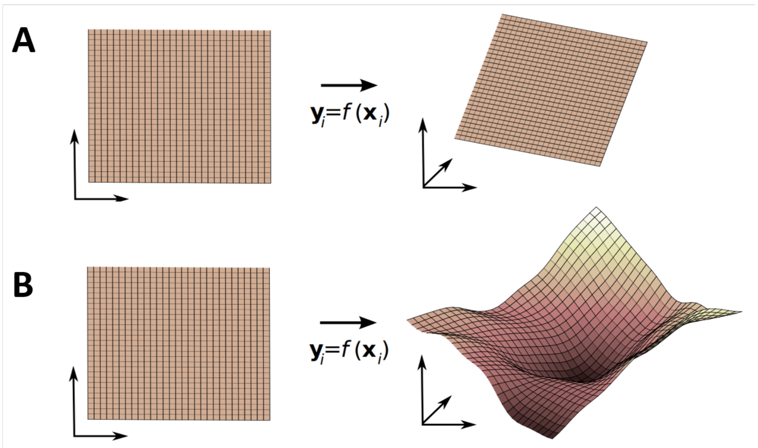

Figure 2.3: Dimensionality reduction mapping, **(A)** Linear, **(B)** Non-linear [5].

neurons.

As an open problem, several interpretations exist for the concept of synergies. However, substantial empirical evidence on animal musculature and neural data analysis point to neural origins of control. The neural strategy of simplifying control is referred to as the muscle synergy hypothesis [6]. A common explanation is that muscle synergies are the CNS's solution to the degrees of freedom problem, as a way to simplify the control of complex motions. Instead of controlling thousands of motor units, control is instead simplified to a few modules [29]. Another explanation is that the CNS adopts a hierarchical control strategy, using synergies as a translation between task-level goals and execution-level commands [30].

Muscle synergies are typically modeled using linear LVMs and extracted from muscle activation data, called electromyography (EMG). In the fields of motor control and neuroscience, two primary techniques are used in the decomposition of EMG signals: principal component analysis (PCA) and non-negative matrix factorization (NMF) [31]. Both are linear decomposition techniques which assume that the measured data is driven by a smaller set of hidden variables. In particular, an EMG channel of a particular observation can be represented as a linear combination:

$$M_j = c_{1j}W_1 + c_{2j}W_2 + \ldots + c_{nj}W_n + \text{error}, \tag{2.1}$$

where vector W_i is a synergy and scalar c_{ij} is a scaling factor. In both cases we assume that the CNS sends signals C to a small set of modules W, wherein each module contains information on the muscles' activation magnitude. The appropriate muscles are then activated, causing flexion/extension of the associated limbs (Figure 2.4).

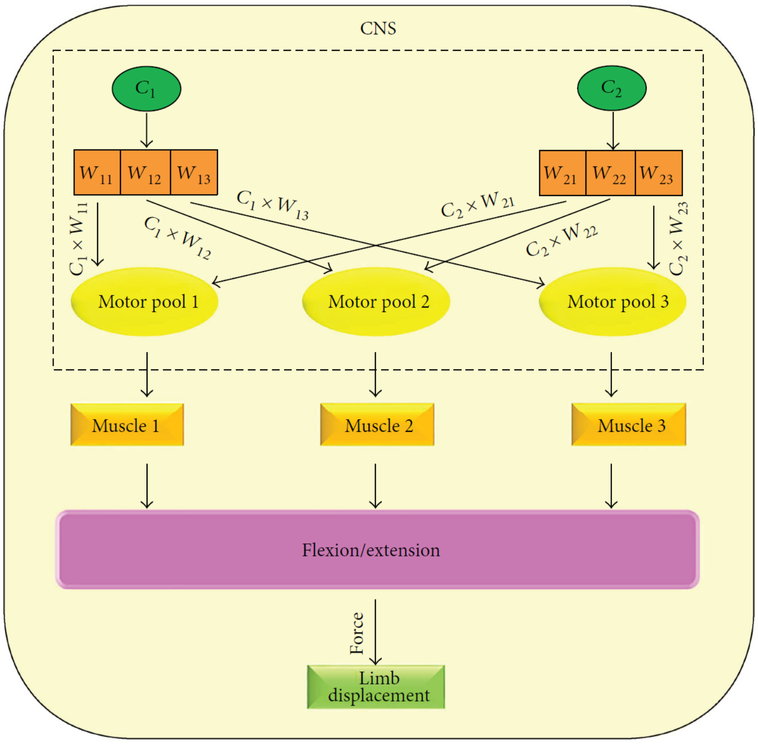

Figure 2.4: Muscle control scheme of the Central Nervous System [6].

Although both techniques are similar in their mathematical representation, their individual constraints cause differences in the extracted components and their respective interpretations. On the one hand, PCA has an orthogonality constraint. The resulting components simply describe the direction of the data, lacking physiological analogs. On the other hand, NMF has a non-negativity constraint. The resulting components are constrained to be non-negative, which is physiologically consistent with the current understanding of neural and muscular output. In particular, neurons are considered to have only a positive state (firing) and a zero state (resting). Thus, NMF is the preferred decomposition technique, for its interpretability.

Several studies investigate the muscle synergy structure of sit-to-stand motion, an important activity of daily living. These studies demonstrate the utility of synergistic features in analyzing motor tasks. Yang et al demonstrated that two distinct sit-to-stand strategies can be generated through a change in the synergy activation timing. In their setup, the subjects were explicitly instructed to perform two different strategies [32]. Kogami et al also report the modulation of the synergy activation timing to produce a different sit-to-stand strategy. In their setup, physical therapists were asked to assist in the standing up of hemiplegic patients [2]. The results of such studies demonstrate that synergistic features may provide new insights or biomarkers for evaluating motor task performance.

2.4 Gaussian Process Latent Variable Model

The main reason for using a non-linear model is that many processes are *inherently* non-linear, such as those found in biological systems [33]. A particular class of non-linear LVM has been demonstrated to work well for a variety of complex high-dimensional data, through a Gaussian Process embedding. Lawrence introduced the Gaussian Process Latent Variable Model (GPLVM) initially for the purpose of dimensionality reduction, finding a non-linear function that smoothly maps a low-dimensional latent space to a high-dimensional observation space [34].

GPLVM is closely related to Probabilistic PCA (PPCA) [35] and Dual Probabilistic PCA (DPPCA) [34]. Both PPCA methods assume that d-dimensional latent points $\mathbf{X} = [x_1, \ldots, x_N]$ are related to D-dimensional observation points $\mathbf{Y} = [y_1, \ldots, y_N]$ through a probabilistic mapping

$$y_i = \mathbf{W}x_i + \eta_i, \tag{2.2}$$

where mapping $\mathbf{W} \in \mathbb{R}^{D \times d}$ and Gaussian noise $\eta_i \in \mathbb{R}^{D \times 1}$. In PPCA, the prior distribution is placed over the latent variable $\mathbf{X}$ which yields the marginal likelihood

$$p(\mathbf{Y}|\mathbf{W}) = \prod_{i=1}^{N} \mathcal{N}(\mathbf{y}_i|\mathbf{0}, \mathbf{W}\mathbf{W}^T + \sigma^2\mathbf{I}). \tag{2.3}$$

In DPPCA, the prior distribution is instead placed over the mapping $\mathbf{W}$ which yields

$$p(\mathbf{Y}|\mathbf{X}) = \prod_{j=1}^{D} \mathcal{N}(\mathbf{y}_j|\mathbf{0}, \mathbf{X}\mathbf{X}^T + \sigma^2\mathbf{I}). \tag{2.4}$$

Lawrence shows that both PPCA and DPPCA are equivalent eigenvalue problems, noting that the latter has a linear kernel term $\mathbf{X}\mathbf{X}^T$. GPLVM is obtained by replacing the this term with a non-linear kernel [34].

A key advantage of GPLVM over existing non-linear LVMs is its flexibility. Its non-parametric formulation allows the model to grow and accommodate both the size and complexity of the data. This is advantageous in medical fields, where personalized data can be complex yet limited [36]. Furthermore, its Bayesian framework allows model assumptions to be automatically inferred, through hyperparameter optimization of the marginal likelihood. This flexibility allows it to be augmented and applied in a wide variety of machine learning applications [28].

GPLVM can be augmented in a variety of ways, depending on the type of data and problem. Three common strategies are described as follows: (1) *Constraint-based*: a specific prior can be imposed onto the latent variables; (2) *Generation Process-based*: multiple sources are explicitly assumed to model multi-view or multi-modal data; (3) *Kernel Method-based*: different types of kernels can be used [28]. However, this list is not exhaustive, and the strategies are not mutually exclusive in their implementation. Further, with the current range of extensions, the class of GPLVM models can tackle virtually any problem setting.

Of particular interest is a class of GPLVM that is used to model high-dimensional time series data, called Gaussian Process Dynamical Systems (GDPS). The general strategy is to add dynamical priors in the latent space, either auto-regressively or indexed by time. This strategy has been demonstrated to work well to model human motion [7] and video sequences [8] (Figure 2.5). Some fields that have found utility in the model are robot control, computer vision, and computational biology [10].

2.5 Sit-to-Stand Motion

Sit-to-Stand (STS) refers to rising up to an upright standing position from a seated position. It is one of the fundamental activities of daily living, necessary

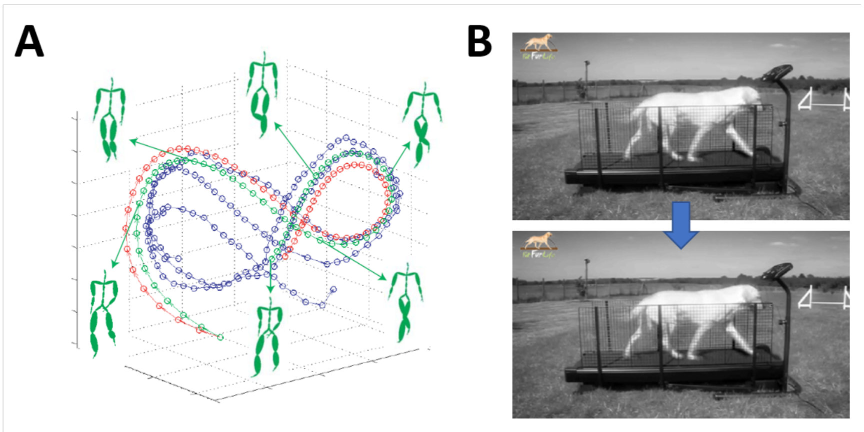

Figure 2.5: Applications of GPDS, **(A)** modeling human motion by mapping smooth three-dimensional latent space to observation space [7], **(B)** generating smooth test video frames from training frames [8].

for many routine activities and performed numerous times in a day (i.e. sixty times on average) [37]. The task is deceptively simple, but is actually a mechanically demanding task. It requires the rapid shift of many body segments to an unstable position over a small base of support. Individuals with physical and neural impairments require assistance or rehabilitation to live independently.

The importance of understanding STS movement is recognized in the medical field and is considered a good indicator of functional independence [38]. Its different roles, as intervention [1], exercise [39], and performance indicator [40], have been the focus of numerous studies. However, analysis of the STS movement can be challenging due to the inconsistency in methodology across studies. Unlike gait analysis, a commonly accepted methodology on sit-to-stand (STS) movements does not exist [41].

Early attempts at characterizing STS has been largely limited by the technology of the time. An often-cited work analyzing both kinetics and kinematics considered only seven segments to represent the whole body [9], with many other studies using similar models [42] (Figure 2.6). Environmental parameters that affect STS have also been analyzed, such as chair height, use of armrest, or changing foot positions [43]. Later studies included more abstract and more sensitive modalities such as synergies and muscle activation [2].

Of value in the varied STS studies is how each experiment was standardized. As with other studies on human movement, defining what is "normal" is a challenging task. Human motor behavior is intrinsically variable, causing successive performances of movement to vary even for simple movement [21] [44]. Individual differences in shape and size can lend to distinctive movement styles [41]. Healthy, normal movement patterns cannot simply be used as a norm, due to our propensity to adapt to task constraints [45]. In an effort to limit variability during motion, constraints such as fixed arm positions are often imposed. However, care must be taken that the subject should not be restricted too much, as such constraints could result in observations of the STS movement in ways not normally performed [41]. For our purposes, we opt to define only a few restrictions (i.e. the starting posture), allowing the subject to respond naturally to task constraint.

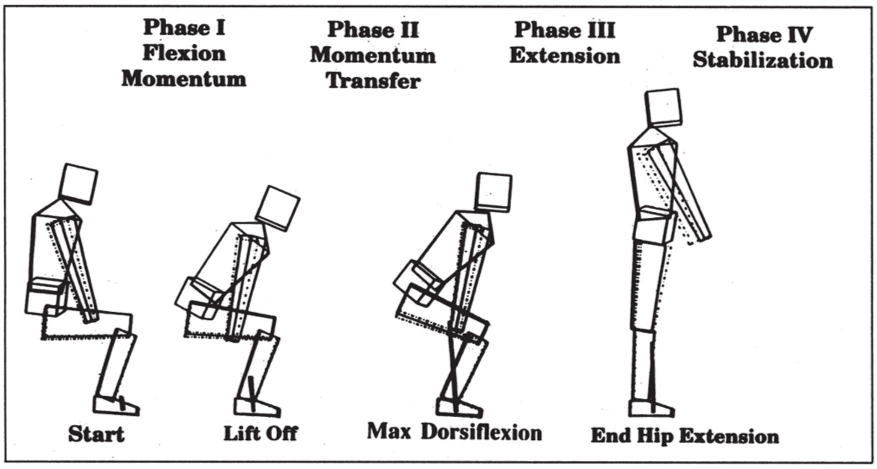

Figure 2.6: Four phases of STS marked by key kinematic and kinetic events [9].

3 Linear Model for Therapist Contribution

Defining good motor behavior is challenging, due to the complex organization of systems that make up the human motor system. A comprehensive analysis that includes many aspects of biomechanics might be ideal, but such large amounts of information is difficult to communicate effectively to both clinicians and patients. A possible approach is to use a dimensionality reduction algorithm with components having physiological interpretation.

The concept of muscle synergies suggests how the central nervous system is able to overcome the problem of controlling a large number of degrees of freedom using a low-dimensional controller [46]. The few, invariant synergy components correspond to the functional sub-tasks that comprise complicated motor tasks. This approach could provide a foundation to characterize normal motor behavior, as well as the deviations introduced by intervention.

In this chapter, we describe a muscle-synergistic approach in analyzing sit-to-stand motion in the context of physical therapy. Neuromuscular data is measured while healthy subjects perform STS, with and without assistance. Muscle synergies are extracted from the collected data, and properties of intervention are inferred from the results. In doing so, we are able to elucidate some neuromuscular properties of natural sit-to-stand behavior as well as the corrective changes induced during intervention.

In the following subsections, we first discuss the mathematical formulations of the muscle synergy model and comparison metrics, followed by a description of the data collection and processing procedures. Subsequently, we present the results from the analysis then discuss them in relation to other STS studies.

3.1 Methods

3.1.1 Non-negative Matrix Factorization

The muscle synergy model is used to represent the motor strategies performed for each STS condition. We used a non-negative matrix factorization algorithm to extract synergy components from the experiment data [47]. Each synergy vector represents a strategy sub-task, while its corresponding activation coefficients describe the magnitude and timing of synergy activation. Formally, the muscle synergy model used in this study is expressed as a linear combination of k non-negative vectors

$$\mathbf{a}(t) = \sum_{i=1}^{k} h_i(t)\mathbf{w}_i + \epsilon(t) \tag{3.1}$$

where $\mathbf{a}(t)$ is an N-dimensional vector representing the activation of N muscles at time t; $\mathbf{w}_i$ is the ith muscle synergy; $h_i(t)$ is a non-negative coefficient series, scaling the amplitude of the ith synergy; and $\epsilon(t)$ is the residual.

The number of synergies is conventionally chosen based on a cutoff of the total percent variability explained [31]. This can be calculated using the coefficient of determination (R^2) curve. The R^2 curve yields two common metrics: the total variation explained (TVE) and the mean squared error (MSE) of portions of the R^2 curve. We follow the MSE convention to stop adding synergies when a significant decrease in error is detected, i.e. MSE $< 10^{-4}$ [48].

In this analysis, we introduce the use of an aggregation approach for muscle synergy extraction. This approach allows the direct comparison of the temporal components $\mathbf{H}$, by matching the spatial components $\mathbf{W}$ of all samples. This differs from the conventional approach, where both spatial component $\mathbf{W}$ and temporal component $\mathbf{H}$ vary across samples. Our modification affects the succeeding analyses by making the synergy activation coefficients more relevant.

The approaches differ in the frequency of the extraction step and how multiple samples are combined. In the the conventional approach, individual extractions are performed for each subject [31] [48], where the resulting spatial components are then averaged (Fig. 3.1A). For the proposed approach, on the other hand, data from all trials are initially aggregated into one data matrix for a single simultaneous extraction of synergies (Fig. 3.1B). The latter approach has the effect where each condition or subject may or may not use all synergies in the group synergy set.

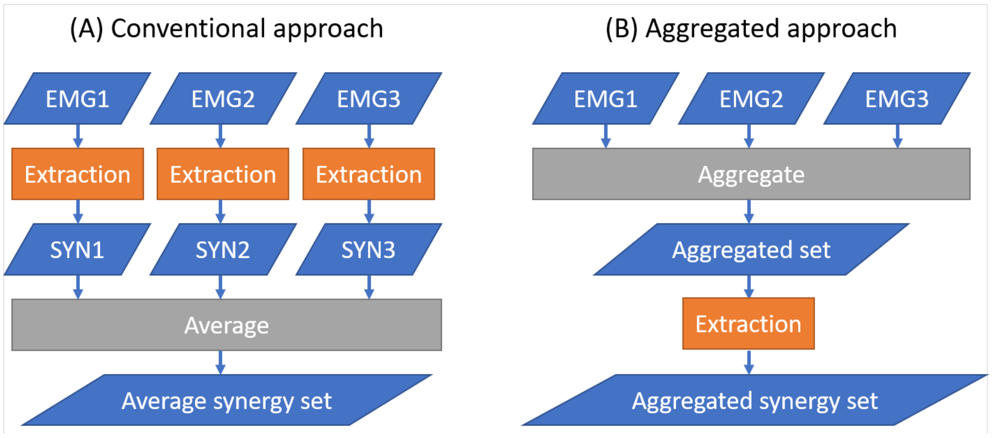

Figure 3.1: Muscle synergy extraction approaches, where EMGN represents the EMG data from subject N and SYNN represents the corresponding extracted synergy. **(A)** conventional approach, **(B)** aggregated approach.

3.1.2 Comparison Metrics

Due to the use of a group synergy set, each subject has access to the same synergies. Thus, analysis through the activation coefficients alone has become possible. The metrics employed in the succeeding analysis of activation coefficient series data is described in this subsection.

Synergy Inactivity

The terms *shared* and *specific* synergies have been used to describe the similarity and differences of synergies extracted from different subjects, respectively. Conventionally, these concepts are dependent on the *cosine similarity*. On the one hand, shared synergies have multiple pairs of similar synergy vectors, representing the common motor strategies. On the other hand, specific synergies have little or no similarity among subjects, representing subject- or condition-specific strategies [49]. However, our modified extraction procedure forces all subjects to recruit from the same synergy set, causing all subjects to have perfect cosine similarity. This renders the role of synergy vectors irrelevant. Instead, specific synergies can be determined based on the active synergies, defined by the amplitude of activation coefficients.

To determine the activity or inactivity of a synergy, we use the signal energy of the activation coefficient signal normalized to the subject's maximum signal energy output. The relative signal energy is expressed as

$$E_{rel} = \frac{E_i}{\max\{E_i\}_{i=1}^{k}}, \tag{3.2}$$

where

$$E_i = \sum_{n=0}^{100} |h_i(n)|^2, \tag{3.3}$$

h_i is the discrete coefficient activation signal, k is the size of the group synergy set, and n is normalized time. We consider a synergy to be inactive when the relative signal energy $E_{rel} \leq 0.05$.

Synergy Activation Similarity

Differences in synergy activation is typically assessed only through differences in peak activation time [2] [32] [50]. In the conventional synergy extraction scheme, any changes in recruitment amplitude might be reflected on the synergy vectors themselves. However, since the group synergy set stays the same across the subjects, any changes in activation amplitude will only reflect on the activation coefficients. To define the similarity between two discrete activation sequences, we use cross correlation

$$r = \frac{\sum_{n=0}^{100} h_i(n)h_j(n)}{\sqrt{\sum_{n=0}^{100} h_i(n)^2 \sum_{n=0}^{100} h_j(n)^2}} \tag{3.4}$$

where h_i and h_j are the discrete signal pairs. The denominator ensures a similarity of 1 for auto-correlation. A pair of synergies is considered similar when their normalized correlation $r \geq 0.95$. All active synergy coefficients were matched for checking similarity.

Synergy Modulation

Synergy activation coefficients can be modulated based on timing [2] and magnitude of the signal. On the one hand, timing difference can be measured using the the lag between the cross-correlation of two signals [50], defined as

$$r_{ij}(l) = \sum_{n=0}^{100} h_i(n)h_j(n-l), \tag{3.5}$$

$$l_{ij} = \arg\max_{l}(r_{ij}), \tag{3.6}$$

where h_i and h_j are the discrete signal pairs, and lag $l = 0, \pm1, \pm2, ...\pm100$. Signal h_i lags signal h_j if lag l_{ij} is positive. On the other hand, magnitude difference can be measured using the ratio of energy signals

$$E_{ij} = \frac{\sum_{n=0}^{100} |h_i(n)|^2}{\sum_{n=0}^{100} |h_j(n)|^2} \tag{3.7}$$

where h_i and h_j are the discrete signal pairs. Signal h_i is of a higher scale than signal h_j if $E_{ij} > 1$.

3.1.3 Data Collection

Participants

Seven healthy adult males (age: 27±2 years, weight: 65.8±8.7 kg) were recruited for the role of *subject* (person who stands up), while one PT (30 years experience) was recruited for the role of *expert* (person who assists). All participants gave informed consent to participate in the experiment.

Experiment Protocol

There are three unassisted conditions and one assisted condition. At the start of every trial, a subject is asked to sit comfortably on an armless, backless chair of fixed height (0.45 m). For unassisted conditions, each subject was asked to stand up at a self-selected pace. The conditions are dictated by placing the arms lightly: (1) FOLDED across the chest, (2) to the FRONT on the knees, and (3) to the body SIDES. For the assisted condition (i.e. EXPERT), the PT was asked to induce the desired sit-to-stand motion as he usually performs for his patients. This form of assistance is characterized by a light touch on the arms, requiring the subject to use his own strength to stand up. Each condition was repeated for 5 successful trials.

EMG Recording

The measurement system is composed of a wireless multi-channel surface electromyograph (EMG) with a sampling rate of 1000 Hz. Surface EMG signals from twelve mono- and biarticular muscles (Table 3.1) were recorded while the subjects performed each condition.

Data Pre-Processing

The measured EMG signals were converted to muscle activation before muscle synergies were extracted. The EMG data were rectified, low-pass filtered (second-order Butterworth; 5 Hz cutoff), then normalized to the maximum voluntary contraction (MVC) activity [51]. To allow comparison among trials with different

Table 3.1: EMG channels

Channel	Label	Target muscle	Alias
1	A	Latissimus Dorsi	LDOR
2	B	Erector Spinae	ESPI
3	C	Gluteus Maximus	GMAX
4	D	Biceps Femoris Long Head	BFLH
5	E	Rectus Abdominis	RABD
6	F	Vastus Medialis	VMED
7	G	Rectus Femoris	RFEM
8	H	Vastus Lateralis	VLAT
9	I	Tibialis Anterior	TANT
10	J	Biceps Femoris Short Head	BFSH
11	K	Gastrocnemius	GAST
12	L	Soleus	SOLE

sit-to-stand duration, the muscle activation data was time-normalized from 0% to 100%. Finally, data from all trials were aggregated into a single data matrix before muscle synergy extraction.

3.2 Results

In this section, we describe the resulting synergies and metrics calculated from the experiment data. Subsection 3.2.1 describes the extracted synergies and their associated functions. Subsection 3.2.2 and 3.2.3 highlight the specificity and persistence of natural synergies. Finally, Subsection 3.2.4 explains the strategy changes induced by the therapist.

3.2.1 Group Synergy Set

The group synergy set was first established by extracting an appropriate number of synergies. The fraction of total variation explained increases with each additional synergy. However, too many synergies will lead to a redundant representation, where some synergies only modulate a single muscle. We adapted the procedure of adding new synergies until no significant increase in variability is observed [48]. Fig. 3.2 shows the MSE and TVE curves for the 12 total possible synergies. Six synergies were chosen, corresponding to 86.15% total variation explained.

Fig. 3.3 shows the group synergy set. Each synergy (SYN) corresponds to a functional sub-task, based on the muscle groups that work together. SYN1 mainly activates VMED and VLAT, both contributing to knee extension. SYN2 mainly activates RFEM and BFSH, for the flexion of both knee and hip. SYN3 mainly activates GMAX and BFLH, corresponds to knee flexion and hip extension. SYN4 mainly activates GAST and SOLE, activating both ankle flexors and extensors. SYN5 mainly activates RABD, corresponding to abdominal flexion. SYN6 mainly activates LDOR and GMAX, corresponds to lumbar extension and hip extension.

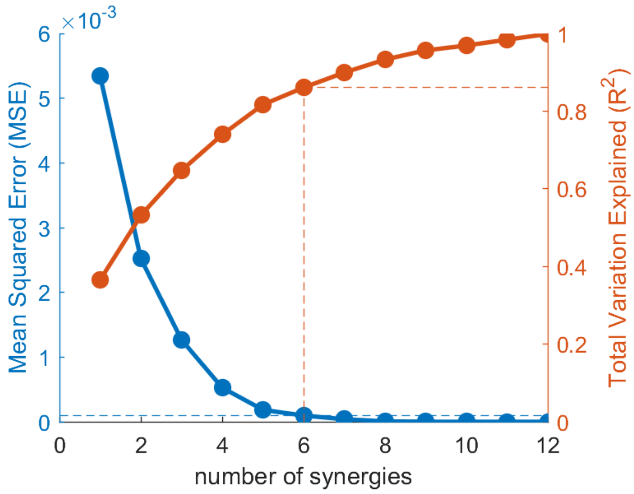

Figure 3.2: MSE and TVE curves for choosing the number of synergies.

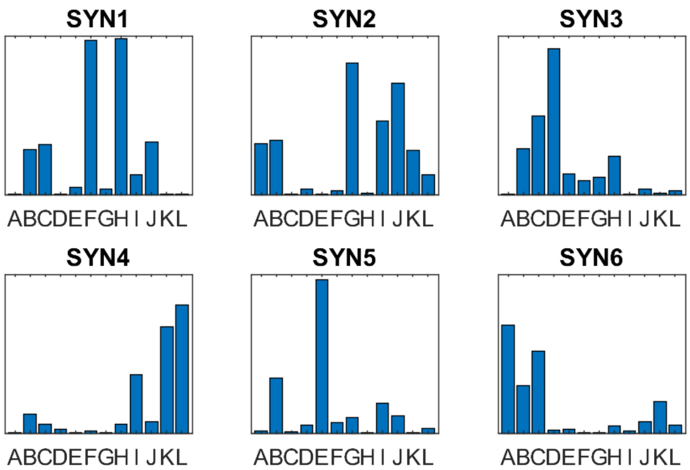

Figure 3.3: Extracted synergy vectors.

3.2.2 Subject-Specific Sit-to-Stand Strategies

The relative signal energy E_{rel} was computed for each subject. Fig. 3.4 shows the percentage of signal energy during unassisted conditions. The results indicate that individual subjects don't recruit the same set of synergies from the group synergy set. Some signals are even weak enough to be considered inactive, effectively excluding them from subjects' specific synergy sets. These suggest that natural STS strategies are not shared naturally.

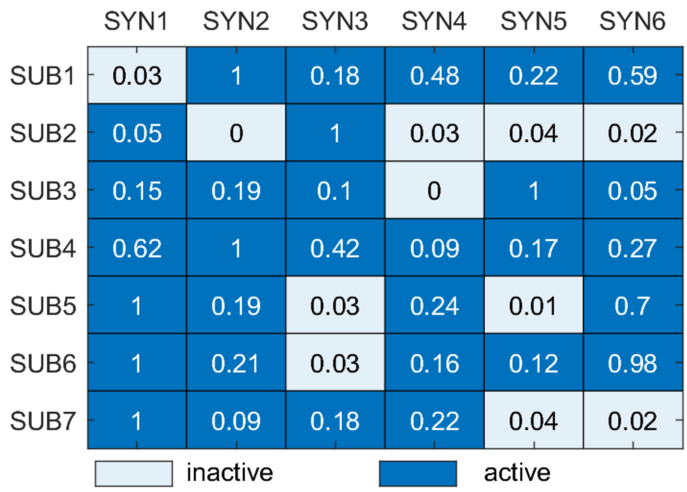

Figure 3.4: Relative signal energy of each subject and synergy.

3.2.3 Persistence of Natural Sit-to-Stand Strategies

The persistence of synergies across different conditions was evaluated using the mean synergy activation similarity r. Fig. 3.5 shows that each subject tends to use the same activation patterns during unassisted sit-to-stand, regardless of the arm position constraint. However, these natural activation patterns are generally dissimilar to the corresponding assisted conditions. Although the same synergies might be recruited, timing and magnitude can influence the identified differences.

3.2.4 Therapist Contribution

The contributions of the therapist to natural sit-to-stand pattern were identified through lag l_{xy} and ratio r between the assisted and unassisted conditions for each subject. The lag l_{xy} indicates delaying or advancing of the original activation signal by the therapist. A positive lag indicates a delayed signal, while a negative lag indicates a leading signal. Fig. 3.6 shows that the therapist tends to delay synergy activation if timing changes are induced. Considering all pairs of synergy activations: 52.38% have positive lag, 13.49% have negative lag, and 34.13% have

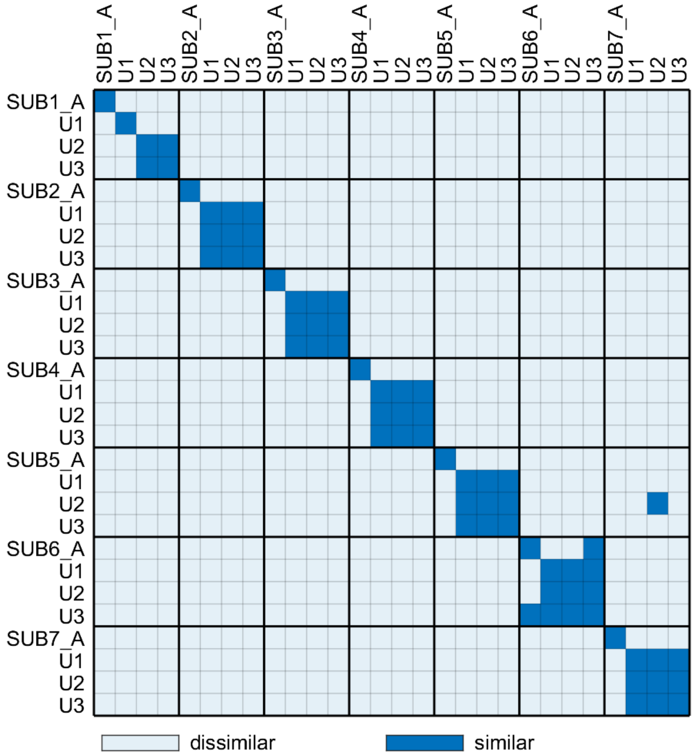

Figure 3.5: Synergy activation similarity between conditions. A: EXPERT, U1: FOLDED, U2: FRONT U3: SIDES.

no lag. Disregarding inactive synergies: 52.69% have positive lag, 9.68% have negative lag, and 37.63% have no lag.

A signal energy ratio $r > 1$ indicates amplification, while a ratio $r < 1$ indicates reduction. Fig. 3.6 shows that the therapist tends to reduce synergy activation strength. Considering all pairs of synergy activations: 76.98% are reduced while 23.02% are amplified. Disregarding inactive synergies: 73.12% are reduced while 26.88% are amplified.

3.3 Discussion

Neuromuscular activity was measured during unassisted and assisted sit-to-stand motion, and muscle synergies were extracted. The present study showed that six muscle synergies were sufficient to form a group synergy set, describing sit-to-stand strategies of a group of healthy people. On the one hand, naturally-

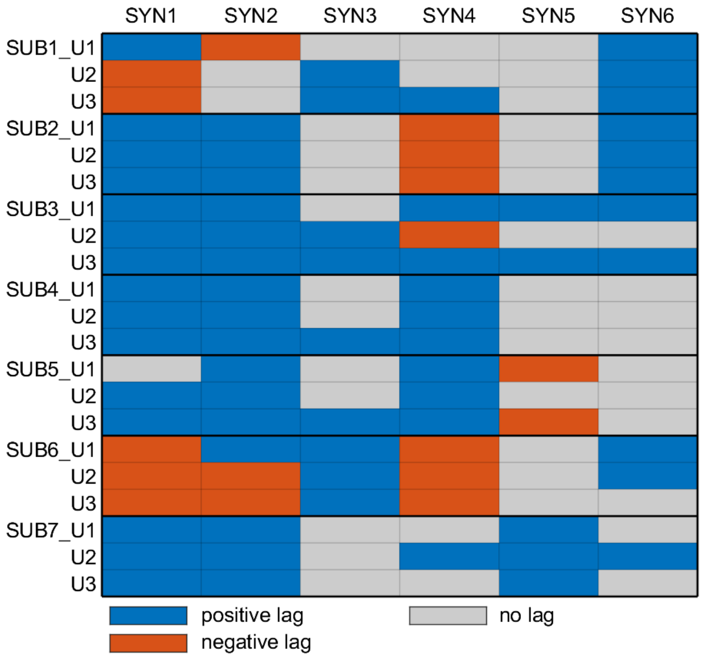

Figure 3.6: Lag between assisted and unassisted conditions, respectively. U1: FOLDED, U2: FRONT U3: SIDES. A signal is considered delayed if the lag $l_{xy} > 0$.

occurring sit-to-stand strategies were found to persist for different arm configurations, forming sets of subject-specific sit-to-stand strategies. On the other hand, the therapist was found to modulate existing strategies instead of introducing new ones. The modulations were often performed through the delaying of the unassisted synergies or reduction of the signal energy. In the discussion that follows, we compare our results with findings in related literature.

3.3.1 Synergy Interpretation

The selection of an appropriate number of muscle synergies aims to accurately explain a sufficient amount of data, while removing unwanted components [49]. The present study utilizes six synergies to describe both unassisted and assisted sit-to-stand conditions. Similar studies report about four or five synergies to sufficiently explain data variation, using conventional synergy extraction methods.

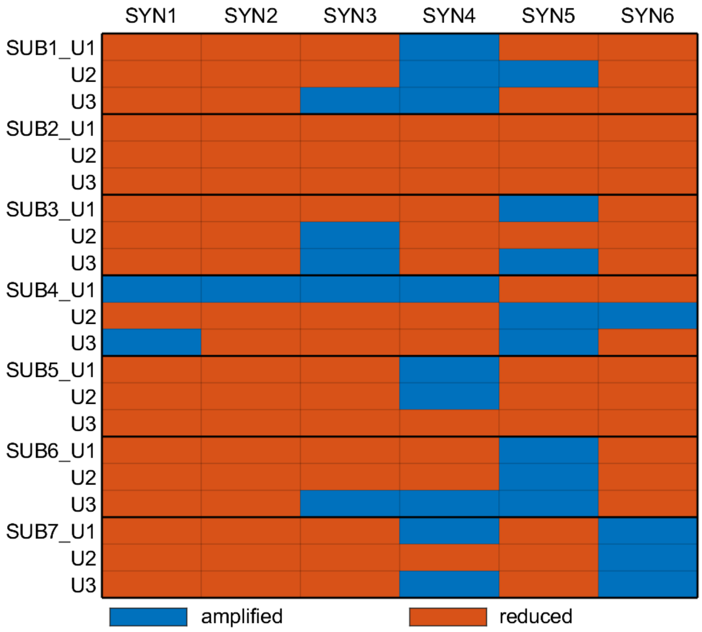

Figure 3.7: Ratio of signal energy between assisted and unassisted conditions, respectively. U1: FOLDED, U2: FRONT U3: SIDES. A signal is considered amplified if the ratio $r > 1$.

These synergies were reported to yield 90% for a postural control task [31], 73% to 82% for a fast-reaching movement task [48], and 89% to 93% for sit-to-stand tasks [2].

Among the six identified synergies, some key synergies are necessary to fulfill the four phases of standing up [52]: (1) the flexion momentum phase, (2) the momentum transfer phase (3) the vertical extension phase, and (4) the stabilization phase. SYN2 is responsible for generating forward momentum. SYN1 and SYN6 are responsible for lifting the body center of mass. SYN4 is responsible for supporting for supporting the ankle joint. SYN3 and SYN5 have supporting roles in the motion. Direct comparison of synergies across studies is inherently difficult, since factors such as choice of muscle groups and extraction criteria can affect the resulting synergies [53]. Although the identified synergies do not exactly match synergies from other studies, some semblance can be seen across these studies, as the overall functional task is the same.

3.3.2 Natural Synergies

Identifying invariant natural sit-to-stand strategies helps establish a baseline to identify changes when assistance is introduced. The *momentum transfer* and *stabilization* strategies are often used to described natural sit-to-stand strategies, based on the velocity of trunk movement and base of support rearrangement [54]. The momentum transfer is generally attributed to healthy persons, while the stabilization strategy is attributed to older adults with functional limitations [55].

In a study by Yang et al, the difference between these strategies have been demonstrated as a delay in the activation of the hip-rise synergy [32]. In their experiment, the subjects were asked to intentionally perform the two specific strategies, explicitly changing motor intent. In our current setup, the subjects were asked only to change their arm configurations. This is to test if the same synergies are recruited, in the presence of usual experimental constraints (arms FOLDED) or natural constraints (arms to the SIDES and arms to the FRONT). The subjects were found to activate the same synergies for performing the three unassisted conditions, suggesting that arm position has little or no effect on the motor intent during standing up.

3.3.3 Therapist Contributions

Changes on the unassisted synergy activation patterns can be measured to identify the contributions of intervention during standing up. A study by Kogami et al aimed to clarify the effects of intervention on the motion of hemiplegic patients [2]. Physical therapists were observed to induce a significantly *earlier* activation of a hip-rise synergy. However, our results show that intervention performed on healthy subjects caused delays for about half of all synergy pairs, mostly for synergies SYN1 (knee extension) and SYN2 (knee flexion and hip flexion). These findings don't necessarily disagree. Their scope included older healthy people and older hemiplegic patients. The hemiplegic patients' hip-rise synergy was found to be activated later than the healthy subjects'. Although the intervention induced an earlier activation, it was not to the same level as the healthy subjects. These suggest that the therapists' activation timing is after the healthy subjects', but before the hemiplegic patients'.

An important factor to consider among studies on STS assistance is the technique or principle employed by clinicians. Techniques such as the "Kamiya scheme" [1] and the "Bobath concept" [2] have been used in STS studies. However, it is not well-understood if these techniques affect sit-to-stand strategies in the same way or have totally different goals altogether. Two observable differences among these techniques is the positioning of the clinician and the point of contact with the subject standing up. In the Kamiya scheme and in our setup, the clinician stands in front of the subject, lightly touching the hands and arms [1]. On the other hand, the clinicians in Kogami et al's study sit to the side of the subjects, touching both the front and the rear of the subject [2]. Clarifying differences in assisting technique allows better comparison among studies of STS.

3.3.4 Effects of Aggregated Extraction

The analysis in this chapter utilizes the proposed "aggregated" extraction scheme. This change enables a straightforward approach to muscle synergy analysis. Certain factors such as motor task and physiology may have an impact on the individual synergy structures, and it would be impractical to track all these relationships. Therefore, we discuss only *general* differences between the proposed approach and the conventional approach. This subsection focuses on reconstruction performance and how our approach addresses the synergy-matching issues.

The R^2 curve can be used to determine the number of synergies necessary to perform the motor task. This curve is a function of the reconstruction error, which is an indicator of the model's performance. We obtain the R^2 for both extraction schemes and set the same model parameters, i.e. $MSE < 10^{-4}$ and $TVE > 85\%$. The results reveal that a similar R^2 curve can be observed for both conventional subject-wise extraction (Fig. 3.8) and aggregated extraction (Fig. 3.2). We find that while the aggregated approach requires six synergies, the conventional approach requires, on average, slightly less (Table 3.2). This result is, in fact, expected as modeling more data and variation in data is expected to need larger dimensionality. However, the parameter setting allows some degree of flexibility in the number of synergies, which further mitigates this difference. In general, the results suggest that the reconstruction performance are similar between the two approaches.

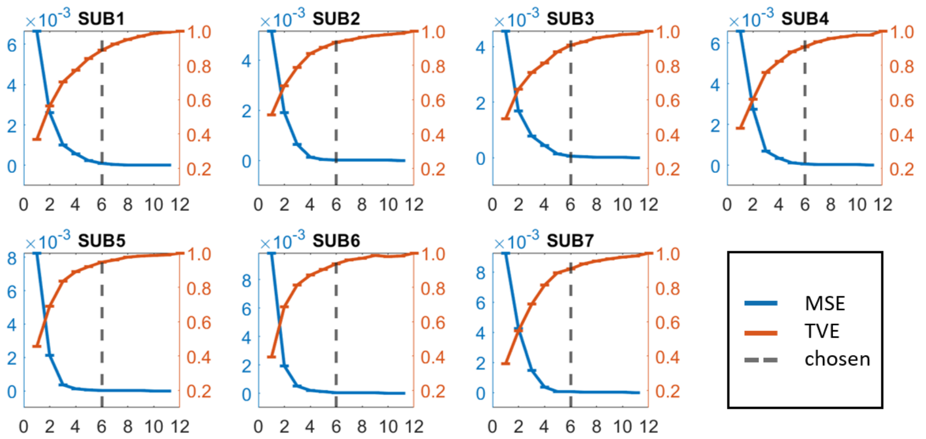

Figure 3.8: R^2 curves of conventional subject-wise synergy extraction. The dashed lines indicates the chosen number of synergies $k = 6$, necessary to model the motion.

The proposed aggregated approach solves three practical synergy-matching issues, which then enables temporal analysis. The first issue with extracting synergies from multiple subjects is choosing the appropriate number of synergies for each subject. As seen in Table 3.2, the appropriate number for each subject may be different despite all subjects performing the same motions. Subjects with different numbers of synergy k cannot match all their synergies. This issue is

Table 3.2: Selected number of synergies for conventional approach

Subject	MSE	TVE
SUB1	6	6
SUB2	5	4
SUB3	6	5
SUB4	6	5
SUB5	5	4
SUB6	6	4
SUB7	5	5
Mean	5.57	4.71

addressed by our proposed "inactive synergy" designation.

The second issue involves synergy mismatching, which we encountered even in the simple case where the same k for all subjects is assumed. Fig. 3.9 illustrates this problem, where six synergies were extracted from all seven subjects, from our experiment data. The orange dots indicate synergies where muscle E (RABD) is the dominant contributor, while the yellow dots indicate synergies where muscle A (LDOR) is the dominant contributor. The E-dominant synergy can be found across all subjects while the A-dominant synergy cannot be found in SUB4. Some studies propose matching schemes that solve the simple case of uniform number of synergies k. These solutions are based on the cosine similarity with ranking [49] and hierarchical clustering [48]. We avoid this issue by performing one extraction step for all subjects instead of performing the step for each subject.

The last issue involves the random ordering of the resulting synergies as a consequence of the factorization method NMF. Principal Component Analysis (PCA), another common factorization method, is able to order its components by component variance; the first in the order usually represents the most information-dense component. On the other hand, NMF yields no meaningful ordering like this, as each component tends to contribute a similar amount of variance [31]. While this problem can be partially addressed by introducing a ranking system [49], our approach avoids this issue by using only one shared pool. Our proposed "group synergy set" renders ranking and comparison of spatial components irrelevant.

3.3.5 Limitations

Although our study reveals some novel findings regarding the effects of intervention in an important motor task, some limitations should be noted. This study focused on the analysis on the parameters of muscle synergies, which is just one aspect of human biomechanics. Other features of human biomechanics need to be analyzed as well for our findings to be of practical clinical use. In contrast with other methods, our modified synergy extraction procedure guarantees exact-matching synergies. However, this approach may not work well if a functionally different task is introduced. The extraction procedure needs to be repeated in such a case. Also, the scope of this section is limited to young healthy subjects assisted by one therapist.

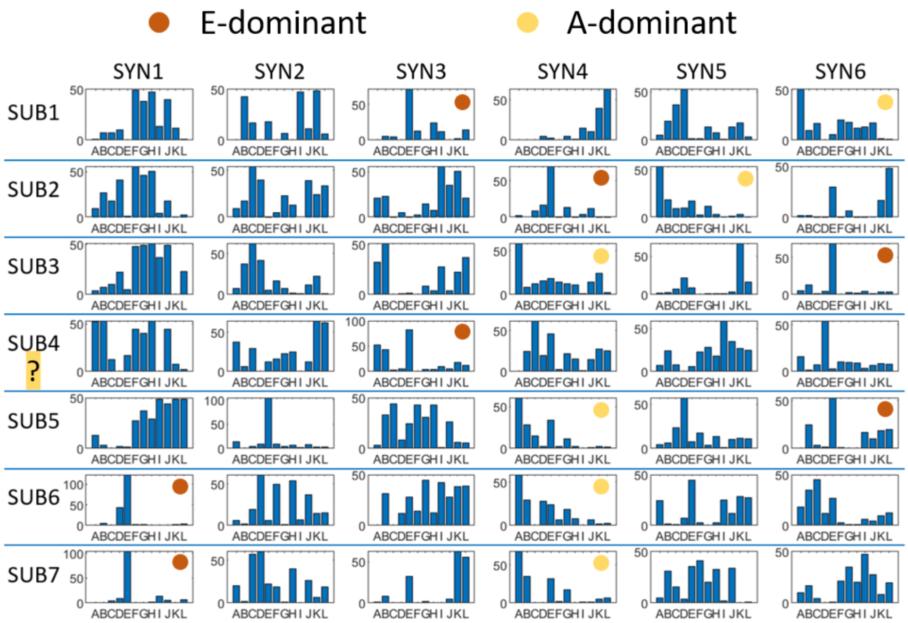

Figure 3.9: Simplifying case where the assumed number of synergies $k = 6$ for all subjects. Orange dots indicate muscle E-dominant synergies, while yellow dots indicate muscle A-dominant synergies. The A-dominant synergy cannot be found in SUB4.

4 Non-linear Model for Augmented Feedback

Telehealth has emerged as one solution to the healthcare service shortage. Many recent applications attempt to replicate the experience of conventional therapy through immersive technologies. However, these systems use single-purpose frameworks, which cannot readily provide personalized treatment, given the limited time and the variety of exercises necessary. A possible approach is to use a generative model which can learn expert-induced motion.

The Gaussian Process Dynamical model is part of a class of generative latent variable models which can represent high-dimensional observation data in a low-dimensional latent space [34]. GPDM, a dynamical variant, has been demonstrated to work well with just few a samples of human motion data. The key idea is to embed and organize meaningful task demonstrations in the same latent space. However, there is currently no principled way to compare characteristics between multiple demonstrations.

In this chapter, we describe modifications to the GPDM to reorganize the latent space for a more intuitive representation. From the modified latent space, we extract features that convey meaningful feedback mechanisms for self-correction. We then demonstrate our model's utility on the motion data of the sit-to-stand task, with conditions that reflect natural, induced, and learned strategies. In doing so, we show that using only a limited amount of personalized data, expert-level feedback can be obtained.

In the following subsections, we will start with mathematical formulations of the GPDM and proposed modifications, followed by a description of the data collection and processing procedures. Subsequently, we present the results of the modifications then discuss them in relation to similar studies.

4.1 Methods

4.1.1 Gaussian Process Dynamical Model

The Gaussian Process Dynamical Model (GPDM) is a non-linear dimensionality reduction technique proposed by Wang et al [10]. It is a dynamical extension of the Gaussian Process Latent Variable Model (GPLVM), a class of latent variable models that allows non-linear mapping from a low-dimensional latent space to a high-dimensional observation space [34]. The GPDM extends this model by introducing a dynamical prior in the latent space [10]. For human motion, data is a sequence of poses indexed by discrete time t. The observation space is defined by a sequence of vector-valued poses $\mathbf{y}_t \in \mathbb{R}^D$, while the latent space is defined by a corresponding lower-dimensional sequence $\mathbf{x}_t \in \mathbb{R}^d$. Either can be written in the form

$$\mathbf{x}_t = \sum_i \mathbf{a}_i \phi_i(\mathbf{x}_{t-1}) + \mathbf{n}_{x,t}, \tag{4.1}$$

$$\mathbf{y}_t = \sum_j \mathbf{b}_j \psi_j(\mathbf{x}_t) + \mathbf{n}_{y,t}, \tag{4.2}$$

for weights $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \ldots]$ and $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \ldots]$, basis functions ϕ_i and ψ_j, and zero-mean white Gaussian noise $\mathbf{n}_{x,t}$ and $\mathbf{n}_{y,t}$.

The GPDM is calculated by marginalizing over parameters of the mappings (i.e. $\mathbf{A}$ and $\mathbf{B}$) and optimizing the latent coordinates of the training data. To obtain the data likelihood over the observations $\mathbf{Y}$, we assume an isotropic Gaussian prior on each $\mathbf{b}_j$ and marginalize over $\mathbf{B}$ to obtain

$$p(\mathbf{Y}|\mathbf{X}, \overline{\beta}) = \frac{|\mathbf{W}|^N}{\sqrt{(2\pi)^{ND}|\mathbf{K}_Y|^D}} \exp\left(-\frac{1}{2}\mathrm{tr}(\mathbf{K}_Y^{-1}\mathbf{Y}\mathbf{W}^2\mathbf{Y}^T)\right), \tag{4.3}$$

where $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_N]^T$ is a design matrix of poses, $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_N]^T$ contains the corresponding latent coordinates, $\mathbf{W} \equiv diag(w_1, \ldots, w_D)$ is a scaling matrix, and $\mathbf{K}_Y$ is a kernel matrix. The elements of the kernel matrix are defined by a kernel function, $(\mathbf{K}_Y)_{i,j} = k_Y(\mathbf{x}_i, \mathbf{x}_j)$, chosen to be the default "RBF+bias+white" [34]. This kernel function for *mapping* is defined as:

$$k_Y(\mathbf{x}, \mathbf{x}') = \beta_1 \exp(-\frac{\beta_2}{2}||\mathbf{x} - \mathbf{x}'||^2) + \beta_3 + \beta_4^{-1}\delta_{\mathbf{x},\mathbf{x}'}, \tag{4.4}$$

where $\beta_1, \beta_2, \beta_3, and \beta_4$ are the kernel hyperparameters and δ is isotropic noise.

The density over the latent coordinates can be obtained in a similar manner. We assume an isotropic Gaussian prior on each $\mathbf{a}_i$ and marginalize over $\mathbf{A}$ to obtain

$$p(\mathbf{X}|\overline{\alpha}) = \frac{p(\mathbf{x}_1)}{\sqrt{(2\pi)^{(N-1)d}|\mathbf{K}_X|^d}} \exp\left(-\frac{1}{2}\mathrm{tr}(\mathbf{K}_X^{-1}\mathbf{X}_{2:N}\mathbf{X}_{2:N}^T)\right), \tag{4.5}$$

where $\mathbf{X}_{2:N} = [\mathbf{x}_2, \ldots, \mathbf{x}_N]^T$, $\mathbf{K}_X$ is the $(N-1)\times(N-1)$ kernel matrix constructed from $\mathbf{X}_{1:N-1} = [\mathbf{x}_1, \ldots, \mathbf{x}_{N-1}]^T$, and $\mathbf{x}_1$ is given an isotropic Gaussian prior. The dynamics are chosen to be the default "RBF+linear+white" [10]. This kernel function for *dynamics* is defined as:

$$k_X(\mathbf{x}, \mathbf{x}') = \alpha_1 \exp(-\frac{\alpha_2}{2}||\mathbf{x}-\mathbf{x}'||^2) + \alpha_3\mathbf{x}^T\mathbf{x}' + \alpha_4^{-1}\delta_{\mathbf{x},\mathbf{x}'}, \tag{4.6}$$

where $\alpha_1, \alpha_2, \alpha_3, and \alpha_4$ are the kernel hyperparameters and δ is isotropic noise. Finally, the latent mapping, priors, and dynamics define a generative model for time series of the form

$$p(\mathbf{X}, \overline{\alpha}, \overline{\beta}|\mathbf{Y}) \propto p(\mathbf{Y}|\mathbf{X}, \overline{\beta})p(\mathbf{X}|\overline{\alpha})p(\overline{\alpha})p(\overline{\beta}), \tag{4.7}$$

where simple uninformative priors $p(\overline{\alpha}) \propto \prod_i \alpha_i^{-1}$ and $p(\overline{\beta}) \propto \prod_i \beta_i^{-1}$ are assumed. The GPDM is learned by minimizing the joint negative log-posterior of the unknowns $-\ln p(\mathbf{X}, \overline{\alpha}, \overline{\beta}|\mathbf{Y})$.

The definition for GPDM in Equation (4.7) is trained using pose sequence $\mathbf{Y}$, implying a single instance of motor behavior. However, Wang et al also describe how the model can be extended to multiple sequences, explicitly modeling multiple instances simultaneously. To do so, the associated latent trajectories need to be embedded in a *shared* latent space. Now the observation space sequences $\{\mathbf{Y}^{(1)}, \ldots, \mathbf{Y}^{(P)}\}$ are still trained as a single data matrix, but each sequence is made independent by ignoring the temporal transitions between the last pose of sequence $i-1$ and the first pose of sequence i. Consequently, the associated latent trajectories $\{\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(P)}\}$ become disconnected [10].

4.1.2 Organizing Latent Trajectories

The original formulation of GPDM is insufficient for comparing multiple demonstrations. As is, we see a disorganized latent space where extraction of meaningful features becomes difficult (Figure 4.1A). We reorganize the space by introducing common reference points.

Common Pose

We hypothesize that the latent trajectories can be organized naturally by appending exact copies of common reference points to each motion sequence. We introduce the concept of a *common pose* (CP), which is appended to either end of each latent trajectory. The CP is calculated twice, based on the mean pose of trajectory end points. The first is done for the start of the motion sequence, and a second time for the end of the motion sequence. Formally, the common start pose is defined as

$$CP_{start} = \frac{1}{P}\sum_{p=1}^{P} \mathbf{y}_1^{(p)}, \tag{4.8}$$

where P is the number of sequences. The common end pose is similarly defined as

$$CP_{end} = \frac{1}{P}\sum_{p=1}^{P} \mathbf{y}_N^{(p)}, \tag{4.9}$$

where where N is the number of poses in a sequence. The CP_{start} is appended to the *start* of each motion sequence $\mathbf{Y}^{(p)}$, while the CP_{end} is appended to the *end*. In our tests, we found that appending fifteen instances to each trajectory end works well.

Zone of Intermediate Poses

By connecting all latent trajectories through the common poses, any pair of trajectories creates an enclosed zone bounding all intermediate poses between them (Figure 4.1B). This Zone of Intermediate Poses (ZIP) has two useful properties for simple user feedback.

The geometric area of the ZIP is a measure of performance similarity in the latent space. A large area indicates dissimilar movement patterns; a small area indicates similar movement patterns; and zero area (coincident lines) indicates exact-matching movement patterns. The area given by the coordinates (x_i, y_i) of two connected trajectories is defined by Gauss's Area Formula:

$$A = \frac{1}{2}\sum_{i=1}^{n}[x_i \cdot (y_{i-1} - y_{i+1})], \tag{4.10}$$

where A is the area of the polygon and n is the total number of vertices. The first and last points that create the polygon connect to each other, defined as $y_0 = y_n$ and $y_{n+1} = y_1$. Since our definition of similarity explicitly uses an area formula, the latent space is necessarily two-dimensional.

Poses sampled within the ZIP yield a smooth pose sequence, due to the proximity of the trajectories. Each latent point has an associated level of uncertainty in the pose space, with higher precision yielding better pose estimates. Precision is highest on the training points, but decreases rapidly as points are sampled farther away. In our implementation, the uncertainty of each point in the pose space is visualized by gray-scale coloring in the latent space. High-precision poses are indicated by a light color, while low-precision poses are indicated by darker colors (Figure 4.1C).

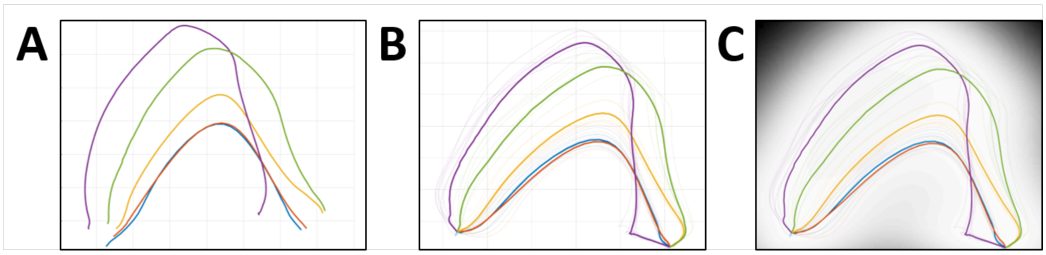

Figure 4.1: Latent space representation, **(A)** GPDM, **(B)** GPDM with CP, **(C)** with precision.

4.1.3 User Performance Feedback

We propose two ways to present both types of augmented feedback. Knowledge of Results is presented through a performance score, while Knowledge of Performance is presented by visualization of corrective poses. In other words, users can confirm if their performance is improving, and, if not, how to correct their mistakes.

Performance Score

The user is given a *Performance Score* to indicate the quality of performance, starting from base performance to the desired expert-induced motion. The Performance Score is defined as

$$PS = 1 - \frac{A_{curr}}{A_{ref}}, \tag{4.11}$$

where A_{ref} is the area between the baseline and desired trajectories; A_{curr} is the area between the current and desired trajectories; and both A_{ref} and A_{curr} are calculated using (4.10). This convention implies that the baseline condition is assigned a 0% score while a goal condition is assigned a 100% score. Generally, a user's progress starts from 0% and improves all the way to 100%. However, it is possible to perform worse than the baseline by misinterpreting the expert's instructions. In this case, the score can go below 0%. The possible scoring outcomes are visualized with their corresponding ZIP in Figure 4.2.

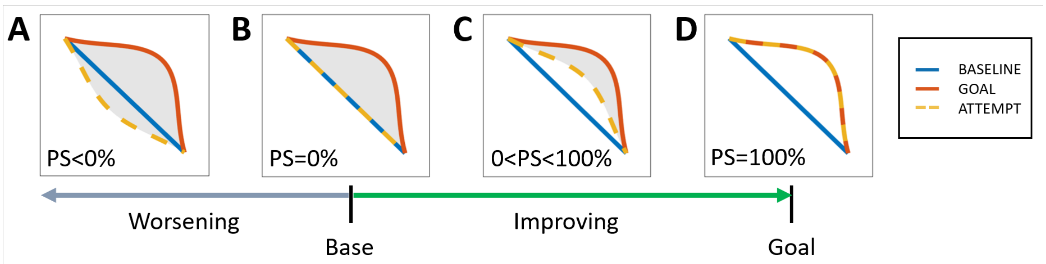

Figure 4.2: Area-based performance score system, possible scores **(A)** worsening: $PS{<}0\%$, **(B)** base: $PS = 0\%$, **(C)** improving: $0{<}PS{<}100\%$, **(D)** mastery: $PS = 100\%$

Corrective Action

Poses can be visualized by sampling points from the latent space. Self-correction is facilitated by sampling a trajectory from the ZIP. By tracing a line from a *current* trajectory to a *desired* trajectory, a smooth corrective pose sequence can be inferred. The pose vector is obtained by solving for the mean of the Gaussian Process while the prediction variance is visualized using the variance [10]. The mean and variance are defined as a function of the latent space position $x\mathbf{x}$ as follows:

$$\mu_Y(\mathbf{x}) = \mathbf{Y}^T \mathbf{K}_Y^{-1} \mathbf{k}_Y(\mathbf{x}), \tag{4.12}$$

$$\sigma_Y^2(\mathbf{x}) = k_Y(\mathbf{x}, \mathbf{x}) - \mathbf{k}_Y(\mathbf{x})^T \mathbf{K}_Y^{-1} \mathbf{k}_Y(\mathbf{x}), \tag{4.13}$$

where $\mathbf{K}_Y$ is the mapping kernel function.

4.1.4 Data Collection

Participants

Nine healthy adult males (age: 27.2±1.5 years, weight: 62.7±10.7 kg) were recruited for the role of *subject* (person who stands up), while one PT (30 years experience) was recruited for the role of *expert* (person who induces change). All participants gave informed consent to participate in the experiment.

Experiment Protocol

Each subject was asked to perform a number of conditions during their respective session. At the start of every condition, a subject is instructed to sit comfortably on an armless, backless chair of fixed height (0.45m), while the knee is flexed to 90-degrees. Before recording, a subject is given a few minutes to familiarize themselves with the movement of the current standing condition. There are three types of conditions, performed in the following order:

1. NATURAL: A subject is asked to stand up naturally, i.e. a self-selected pace and strategy. This condition is further subdivided into three typical sit-to-stand conditions. These conditions are distinguished by the arm position: (1) *N.folded*: folded across the chest, (2) *N.front*: to the front on the knees, and (3) *N.sides*: to the sides.

2. INDUCED: The PT is asked to induce the desired sit-to-stand motion as usually performed for his patients (Figure 1.1). This form of guidance is characterized by a light touch on the arms, requiring the subject to use his own strength to stand up.

3. LEARNED: A subject is asked to recall the new strategy that was learned from the INDUCED condition. The subject is then asked to replicate the motion as close as possible to the taught movement. No additional instructions on timing or strategy was given.

Each subject performed all five conditions during their respective session, where each condition was repeated for 6 successful trials. A trial was considered a failure if data capture was affected in any way, e.g. occlusion of markers. A total of 270 successful trials (5 conditions×6 trials×9 subjects) were collected for analysis. The experiment protocol is summarized in Figure 4.3.

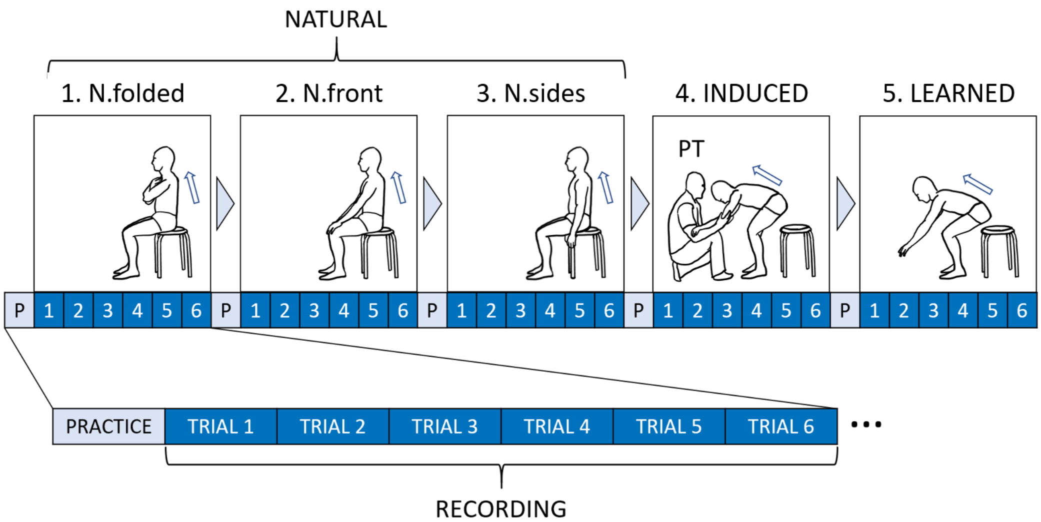

Figure 4.3: Experimental flow diagram for each subject.

Motion Capture Recording

While the sit-to-stand tasks were being performed, the subject's whole-body motion was being recorded. The setup was an indoor MAC3D motion capture system (Motion Analysis Corp.), with 16 cameras mounted around the capture space (Figure 4.4A). The Cortex software from the same company provides the control panel for all devices and the tools for processing raw motion capture data.

Before data recording, twenty-nine passive retroreflective markers were fitted to a subject's whole body, followed by a standard calibration procedure. The Helen Hayes marker set (Figure 4.4B) was used as a reference [56]. During recording, the marker trajectories were sampled at 200 Hz with measurement units in millimeters. The x, y, z positions of each marker were continuously recorded, for a total of 87 channels (29 markers$\times$3 dimensions). An audible beep signals the subject when a trial starts and ends. The trial is ended a few seconds after the subject is fully standing.

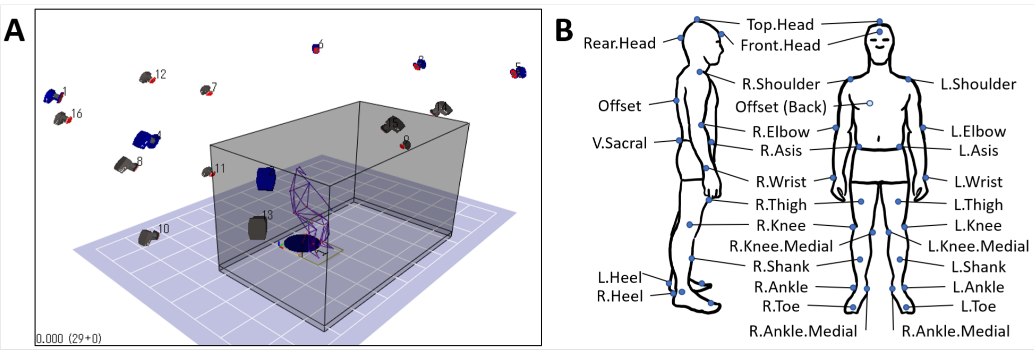

Figure 4.4: Motion capture setup, **(A)** capture space, **(B)** Helen Hayes markerset.

Data Pre-processing

The marker data were first pre-processed before analysis, using built-in tools in Cortex (Motion Analysis Corp.) and custom code in Matlab (The MathWorks, Inc.). The procedures were performed in the following order:

1. **Noise removal**: Each trial was visually examined and corrected for occlusions and noise. The markers were then labelled and smoothed using a fourth-order Butterworth (6 Hz low-pass) filter.

2. **Data translation**: The coordinate system was standardized across trials. A common origin point was obtained using the static point between the R.Heel and L.Heel markers. The average point between the two markers were calculated, and the coordinates of all other markers were subtracted by this value. This procedure was performed for all trials individually.

3. **Data normalization**: To reduce inter-individual differences, the length units are normalized to a unitless value based on height [57] [58]. For each trial, all coordinate values are divided by the vertical component of the Top.Head marker.

4. **Event standardization**: Each trial was truncated to retain only the relevant portion of the sit-to-stand motion. A *start* and an *end* event were defined based on a stable reference marker [59]. The start event is defined as the moment when the speed of the R.Shoulder marker is greater than zero in the sagittal plane, while the end event is defined as the moment when the R.Shoulder reaches its highest vertical position.

4.2 Results

A reorganized latent space was successfully extracted from the experiment data. Relevant properties of the new latent space are discussed in this section.

4.2.1 Latent Space Behavior

In conventional GPDM, multiple motor behaviors have no apparent relation in a shared latent space. This is true even when the *same* motor task is performed repeatedly. To organize the latent trajectories, we proposed to connect them according to known matching poses. Specifically, we appended reference common poses to both ends of each latent trajectory.

Results indicate that the latent trajectories have successfully connected at the common poses, CP_{start} and CP_{end}, found at either end. The trajectories of similar conditions stay close together, forming two subgroups. The NATURAL conditions stay close together, while the INDUCED and LEARNED conditions also stay close. However, the order within the subgroups vary among subjects. These results suggest that even small differences in the pose space can cause latent points to stay far apart. By connecting the trajectories through *exactly matching poses*, an organized latent space can be achieved. The extracted latent trajectories are shown in Figure 4.5.

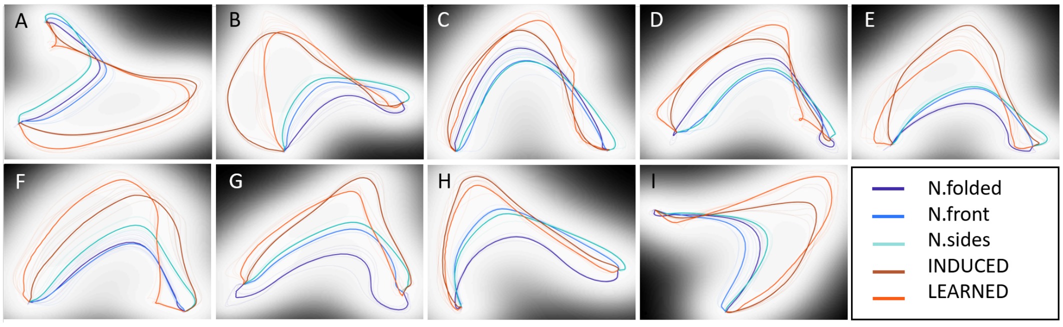

Figure 4.5: Latent trajectory representation for all subjects (subject A to I).

4.2.2 Performance Score

Knowledge of results (KR) is one of two types of augmented feedback shown to be positively linked to motor skill learning [25]. We proposed the Performance Score (PS) as a measure of performance success. The score uses the normalized geometric area between two connected latent trajectories as a measure of similarity between a baseline and a desired behavior. We assigned N.folded as the baseline condition and INDUCED as the desired condition.

Results indicate that the INDUCED scores are always higher than the LEARNED scores, while the LEARNED scores are always higher than the NATURAL scores. This consistent ordering suggests that all subjects were able to remember and perform part of the expert-induced movement. While all LEARNED scores are positive, the *N.front* and *N.sides* conditions report some negative values. The inconsistent negative scores of the other NATURAL conditions suggest that the expert-advised movement is not naturally achieved. These results indicate that the Performance Score can capture the expected performance improvements. A summary of all scores is shown in Table 4.1.

4.2.3 Corrective Pose Sequence

Knowledge of performance (KP) is the other type of augmented feedback for motor skill learning [25]. We proposed sampling from the Zone of Intermediate Poses (ZIP) as a simple yet robust solution for movement self-correction, since a smooth pose sequence can be visualized by simply sampling adjacent points from

Subjects	N.folded	N.front	N.sides	INDUCED	LEARNED
A	0	2.91	-14.69	100	75.52
B	0	23.06	27.44	100	36.95
C	0	-44.47	-56.51	100	52.14
D	0	-36.70	-38.41	100	65.77
E	0	15.34	11.46	100	38.97
F	0	-2.44	10.43	100	36.34
G	0	47.74	49.69	100	66.00
H	0	43.20	30.52	100	73.90
I	0	-19.33	4.08	100	62.48

Table 4.1: Performance Score summary.

the latent space.

Results show that by sampling along a trajectory, known pose sequences can be reconstructed. For example, tracing a NATURAL trajectory and the INDUCED trajectory shows the prototypical pose sequences. We can see some distinction between the two conditions as the INDUCED poses are lower and more forward-leaning (Figure 4.6). Furthermore, by sampling the ZIP between the LEARNED trajectory and the INDUCED trajectory, one can visualize the corrective pose sequence (Figure 4.7). These results indicate that generative portion of GPDM is unaffected by our modification, while the formed ZIPs can be used to identify the erratic portions of movement.

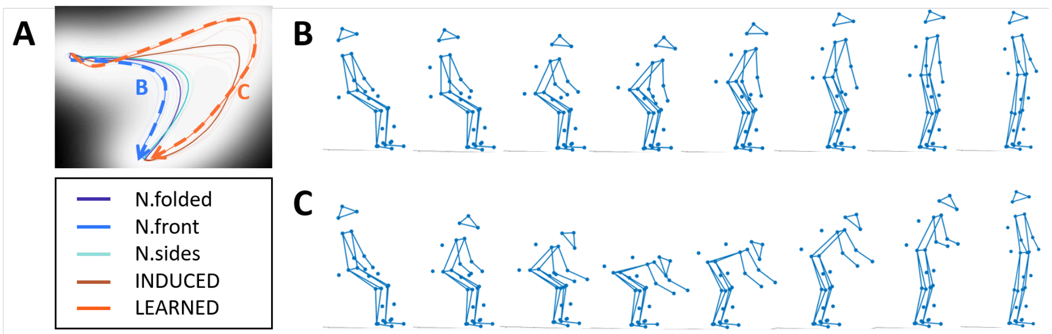

Figure 4.6: Reconstructing demonstrated poses, **(A)** latent space, **(B)** N.folded sequence, **(C)** INDUCED sequence.

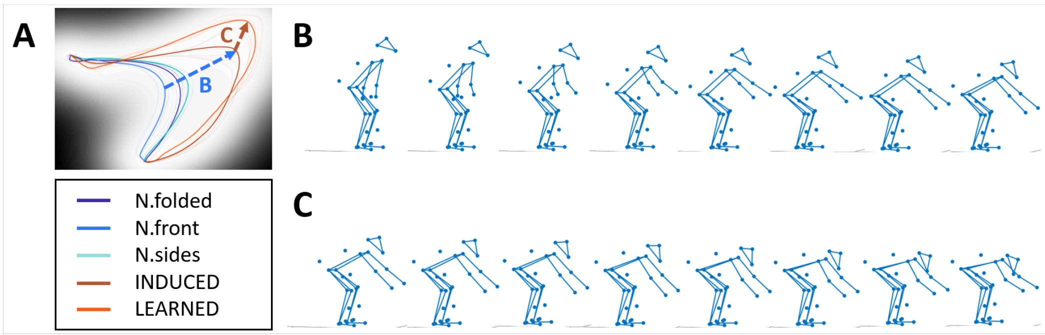

Figure 4.7: Inferring corrective poses, **(A)** latent space, **(B)** N.folded to LEARNED sequence, **(C)** LEARNED to INDUCED sequence.

4.3 Discussion

Our goal is to develop a data-efficient framework for personalized physiotherapy exercises. Our modified GPDM approach solves the common problems of existing telehealth applications, by providing personalized feedback for whole-body exercises, based on few expert-induced demonstrations. Specifically, two types of augmented feedback were extracted from the reorganized latent embedding, conveying both performance quality and the corrective action. By analyzing the feedback outcomes from sit-to-stand experiments, we confirmed the utility of our proposed method for an important physiotherapy exercise. In the discussion that follows, we compare our findings with those in related literature.

4.3.1 Latent Space Behavior

The introduction of the *common pose* solves the problem of relating multiple latent trajectories. The desired connecting effect is achieved because the common pose points “gather” nearby similar points. Effectively, the appended common poses reintroduce the *still* poses to both ends of the movement, i.e. subject is sitting still and standing still. The key differences with original *still* poses are that they now match exactly and are shared by all sequences. In this section, we discuss how these changes compare to related studies and what latent space properties prove most useful.

Multiple GPLVM studies demonstrate that multiple trajectories lie separately in the shared latent space, despite sharing common poses. One study modeled four golf swings from the same golfer, using conventional GPDM (Figure 4.8A) [10]. Another study modeled sitting motion on surfaces of different heights, using Observation Driven GPLVM (Figure 4.8B) [11]. In both types of motion, the start poses are known to be the same pose, yet the latent starting points are represented by different points. One key difference with our target movement, i.e. sit-to-stand, is that the end pose is also known to be the same. This condition appears to be unique to our study.

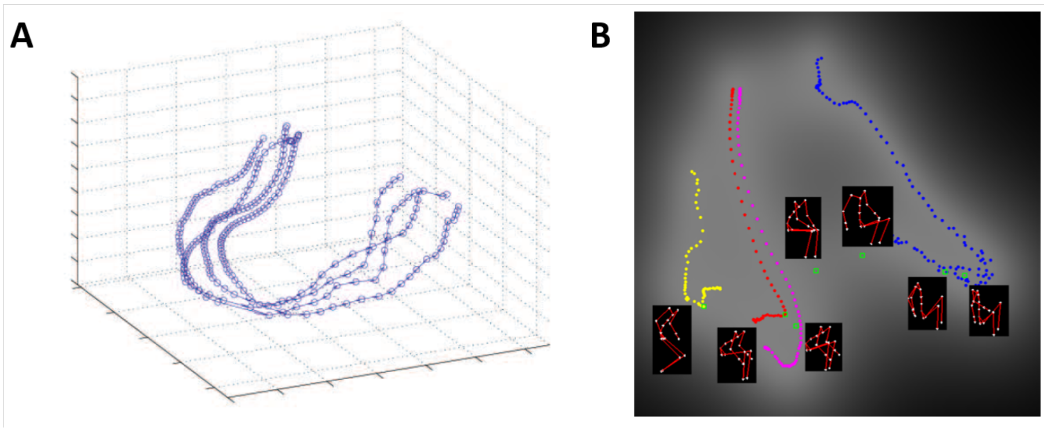

Figure 4.8: Disjoint trajectories found in motion with shared poses, **(A)** four golf swings [10], **(B)** sitting down onto different heights [11].

The spatial pattern of the latent trajectory is often associated with the behavior of the data itself. Several authors try to match their intuition of the latent trajectory shape as a qualitative confirmation of the goodness of the embedding. In the study by Wang et al, acyclic motion was demonstrated to form U-shapes (Figure 4.8A) while cyclic motion formed circular loops, signifying that the start and end of cyclic motion must occupy the same latent point. In a study using Variational Autoencoders to encode human movements, the latent trajectories were found to convey spatial information about the type of movement performed (i.e. balancing, kicking, punching, taichi, and walking). For example, walking is represented by a loop, kicking occupies a large range, and balancing creates sideways branches (Figure 4.9A). Finally, the spatial pattern can also convey literal spatial information as in the case of simultaneous localization and mapping (SLAM), where WiFi signal strength was used to obtain the topography of a floor building in the latent space (Figure 4.9B) [12].

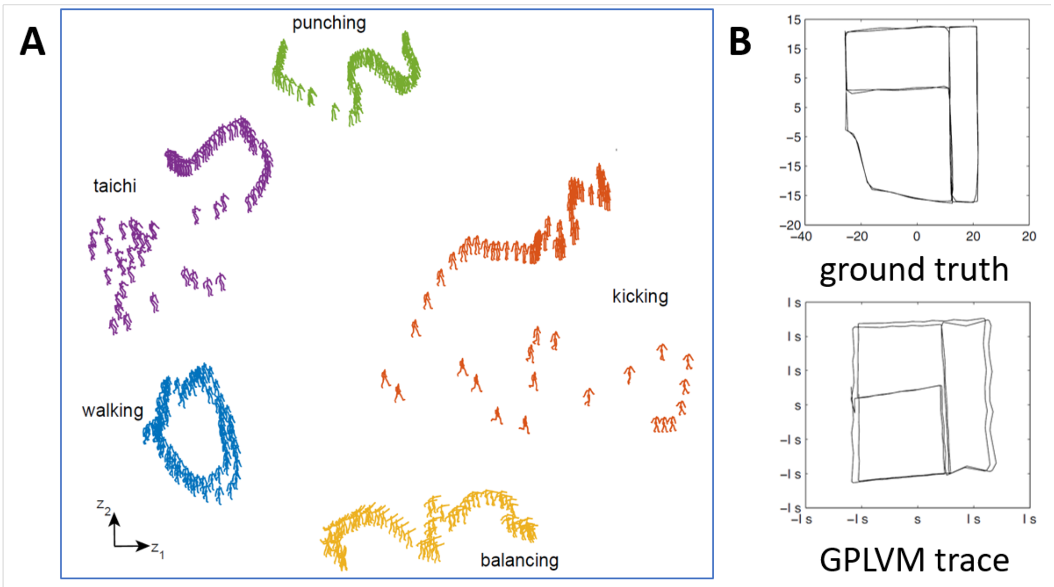

Figure 4.9: The spatial patterns in the latent space are able to embed different types of information, **(A)** five human movements [12], **(B)** topography of a building floor, where top is the ground truth and bottom shows the GPLVM-generated trace [13].

It was suggested that models with visually smooth latent trajectories also achieve better quantitative results. These can be partially quantified by a smoothness and an inverse-variance criteria. Wang et al demonstrates this visually for walking motion. In Figure 4.10A, an un-smooth trajectory shows clumpy and disconnected points, which is distinctly different from the walking trajectory found in Figure 4.9A. The inverse-variance plot of the same trajectory is characterized by "clouds" of clumpy high-confidence regions. On the other hand, Figure 4.10B highlights the desired smooth trajectory, which shows high-confidence poses wrapped in a smooth "tube" across the whole length of the trajectory.

In addition to the above criteria, our application relies heavily on low-variance areas *between* trajectories as well. The inferred poses from these regions are

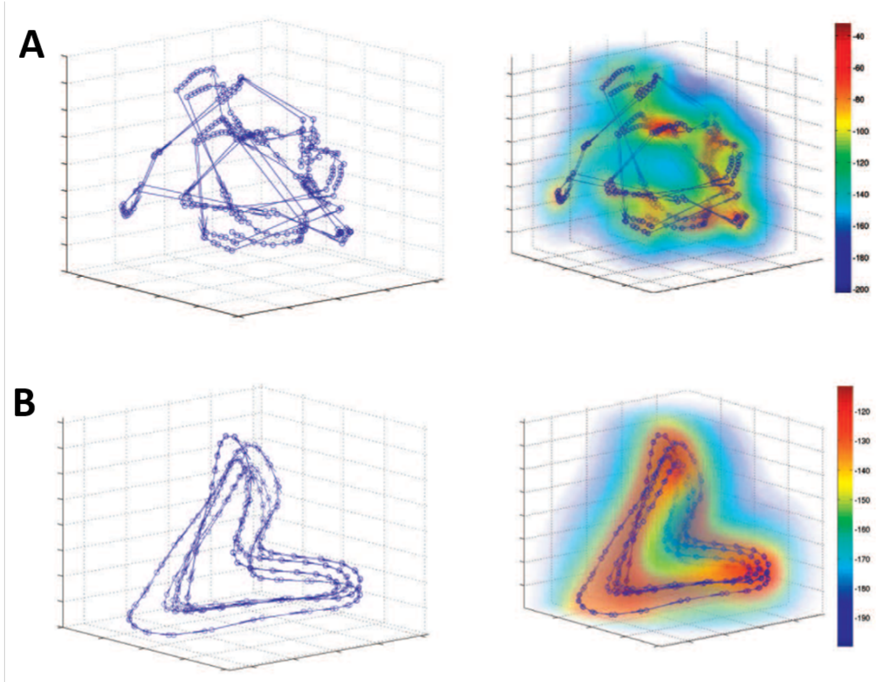

Figure 4.10: Criteria for goodness by Wang et al [10], **(A)** un-smooth trajectory with patches of high-confidence regions (red), **(B)** smooth trajectory with tube of high-confidence regions.

important in the presentation of the corrective poses, which has not been explicitly examined in previous studies. We found that both the smoothness and the amount of high-confidence areas are functions of the number of training iterations T. Figure 4.11A illustrates that as T increases, smoothness tends to increase slowly while the high-confidence regions decreases rapidly (i.e. more dark areas). In lower values of T, we can observe degenerate poses only outside the outermost loop created by the trajectories. However, degenerate poses are also inferred in the regions between latent trajectories for higher values of T. Figure 4.11B illustrates both a sensible pose and a degenerate pose inferred from the same model when $T = 100$. Our current setting of $T = 15$ offers a nice balance for both smoothness and pose inference.

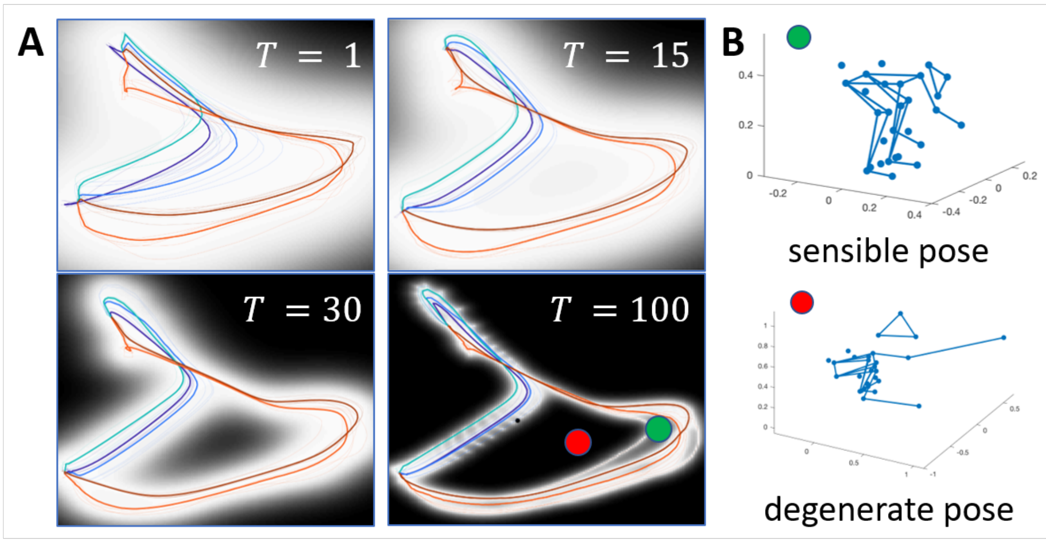

Figure 4.11: Criteria for goodness of the latent space (dark colors are low-confidence areas), **(A)** effects of increasing the number of iterations, **(B)** inferred poses, where the top pose is sampled from a high-confidence point and the bottom pose is sampled from a low-confidence point.

4.3.2 Augmented Feedback

Our approach provides Knowledge of Results through the Performance Score, calculated as a function of the geometric area between two connected latent trajectories. Since the mapping from latent space to pose space is non-linear, the score does not necessarily convey the *scale* of progress, i.e. flexing a joint 50% more does not necessarily mean a subject improved by 50%. Other single-valued quantities for progress have been previously proposed, yet the scale of progress remains difficult to quantify exactly. Some examples include time to accomplish the task [4] [60], joint flexion/extension angle [61], and root mean square error between poses [62]. Our proposed Performance Score has two advantages over the aforementioned metrics. First, it incorporates both spatial and sequential information. Second, calibration is not necessary when generalizing to other motions.

One note regarding the Performance Score is that the area term in Equation (4.11) is a valid metric, i.e. distance function. Given any pair of connected latent trajectories, the four conditions of non-negativity, identity of indiscernibles, symmetry, and triangle inequality are all satisfied.

Our approach provides Knowledge of Performance through estimating the intermediate poses between two motor behaviors, allowing the user to visualize a corrective pose sequence. Due to the complexity of human motion, presenting the optimal amount and type of information is challenging. To reduce cognitive load, popular approaches use immersive technologies or only limb-specific movements. For example, some existing systems ask the user to move a target bone vertically and horizontally [63] or move the shoulder laterally to a target number of degrees [3]. Virtual reality applications, in particular, tend to be limited to upper-body movement for safety reasons. Our approach balances the amount of information captured and presented, by focusing only on the corrective poses of a whole-body model. Thus, removing the need to arbitrarily isolate body parts.

An important note regarding the corrective pose sequence is that it involves the *visualization* of the latent space, limiting the dimensionality to a maximum of three. Nonetheless, it would be of interest to discuss how different choices of dimensionality may affect our application. In particular, we highlight a three-dimensional example by Wang et al [10] and a nine-dimensional example by Damianou et al [64]. In the first example, the original GPDM model was used to describe golf swing motion, and the latent space was set to three dimensions. The latent trajectories resulted in U-like shapes. In the second example, an extension of GPDM (dynamical variational GPLVM) was used to describe walking and running motions. In their model, the latent space was initially set to nine dimensions, but the model selects three "true" dimensions. Within the visualized three-dimensional space, each motion resulted in circular shapes with some distance between them.

We can see that relatively simple motions, such as golf swing, walking, and running, form "flat" trajectories. Despite being assigned three dimensions, the flat shapes of each motion suggest that these simple motions can be embedded in a two-dimensional space with some small loss. Although we can instead decide on a three-dimensional space for visualization, we argue that "navigating" a two-dimensional space can offer a more familiar experience. Since many commercial devices, such as smartphones and computer screens, offer a two-dimensional user interface, a user can more readily interact with the proposed system without the need to learn new forms of interaction. Thus, from the practical considerations discussed above, the simplest and most direct approach remains to be the two-dimensional representation.

4.3.3 Interpretation of Motor Knowledge

Subjects were found to interpret intervention differently after being guided by the PT. This difference can be observed by looking at the LEARNED condition, which is the subject's attempt to repeat the INDUCED movement. The PT's general strategy was observed to be guiding the subject lower and more forward than natural. However, some subjects undershoot while some overshoot the target motion. This behavior can be observed by plotting the body center of mass (CoM) in the sagittal plane (Figure 4.12). Furthermore, the expert-INDUCED CoM trajectories were observed to be different across subjects. These results suggest that both treatment and subject response are personalized in practice.

We observed that the subjects with the lowest LEARNED performance scores (i.e. subjects B, C, E, F in Table 4.1) were performing a posture called "augmented arm", where arms are extended forward at shoulder height. We speculate that these subjects have the same interpretation of the therapist's intended change in motor behavior; thus, the same posture. Since this posture has the tendency to produce lower scores, such postures should be identified and avoided. This finding suggests that "low-score" postures may exist in other exercises as well, and identifying particular bad postures may help in self-correction. This posture may also be a reaction to counteract the slower motion induced by the PT, as it has been demonstrated to reduce standing up time [65].

4.3.4 Alternative Models

Human motion can be modeled reasonably well in a few other ways. One can look at variants of GPDM, GPLVM, or deep learning techniques, which have the capacity to model a variety of high-dimensional dynamical data. We first discuss some extensions of GPDM and their application to different types of human motion. We also discuss two models which use conceptually different techniques

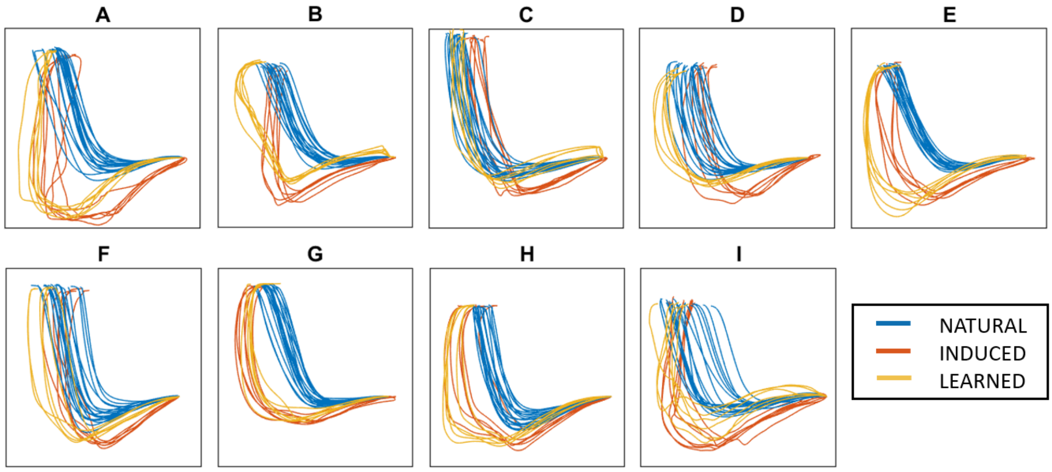

Figure 4.12: Body center of mass in the sagittal plane for each subject.

to model dynamics.

GPDM has been demonstrated to work well with different types of whole-body motion. Wang et al uses the GPDM to describe walking and golf swing movements, demonstrating the model's capability to model simple cyclic and acyclic motions [10]. Chen et al extends this concept by introducing a switching mechanism to account for motion sequences that involve switching dynamics, such as in salsa dancing [66]. To our knowledge, a GPDM-based approach has not yet been applied to sit-to-stand motion, but we should note that Gupta et al have applied it to the related stand-to-sit task [11]. Together, these studies demonstrate the applicability of GPDM as a model for describing full-body motion, which can reasonably include sit-to-stand.

Hierarchical GPLVM (HGPLVM) can be considered an alternative implementation of dynamics for GPLVM. The main difference is that GPDM is autoregressive while HGPLVM is not. Instead, HGPLVM takes timestamps as inputs [7]. This is advantageous if uniform sampling is difficult to achieve. But in the controlled environment of telehealth, such a precaution is generally unnecessary. HGPLVM also offers the option to "decompose" the subject model into component parts, allowing an isolated view of selected parts. However, this increases the number of visualized subspaces and requires a decision on the appropriate decomposition for each target activity. While still considering whole-body information, the single latent space representation of GPDM is a more straightforward approach to visualization.

Deep learning models are widely considered as universal approximators, which can work well with a large variety of data [67]. Given enough data and resources, deep learning models can exceed the performance of whatever specific-purpose model. In fact, a single network can be demonstrated to generalize well to multiple types of human actions. A trained model can simultaneously perform classification and prediction of novel poses with very little computational cost [68]. The main downside with such a model is the amount of resources necessary to perform training. In the context of single subjects with limited sessions, such large amount of resources is simply unavailable.

4.3.5 Motion Data Format

One practical note regarding the model is the data format used. In several GPLVM-based works on human motion data, no motion experiments were actually performed. Instead, the popular CMU Graphics Lab Motion Capture Database (mocap.cs.cmu.edu) was often used. The format used in this database contains *joint angle* information instead of the *marker coordinate* information we used in our study.

Although both formats are functionally similar, formats that store joint angles typically need to define a skeleton. The main advantage to this is that bone segments can be calibrated to each user, and limb lengths can be fixed. On the other hand, we can also argue that coordinate-based formats are more *accessible* as sensors and algorithms themselves measure anatomical *points*. Currently, in-home telehealth applications often use the ubiquitous Kinect sensor (Microsoft Corp.) which tracks 3D skeletal landmarks of the users. As computer vision algorithms become more advanced, ordinary images and videos are increasingly used to extract similar coordinate-based anatomical key points as well [69].

4.3.6 Latest Motor Skills Training Systems

Various training systems have been developed for motor skills learning. We surveyed systems developed in the past year to see how our system stacks against the latest innovations in the field. We found that many of the same fundamental challenges from the past decade still exist today, but interesting innovations provide the groundwork for better solutions. In this section, we discuss the merits and demerits of three studies, spanning three different applications: rehabilitation, sports, and exergaming.

Many studies regarding rehabilitation systems are difficult to compare because of the differences in target movement, system design, and lack of standardization. Da Gama et al recognize that many systems, typically VR and AR implementations, do not follow standard motion descriptions set by the International Society of Biomechanics (ISB). Their paper provides a framework which incorporates clinical conventions and jargon familiar to the rehabilitation therapists. In lieu of comparing human joint positions to a pre-recorded movement, they set allowable thresholds called Movement Tolerance Margin (MTM). MTM allows the standardization and flexibility to scale exercises for a variety of users for standard exercises. This threshold is set by the physiotherapist, customized to the therapy needs of the individual patient [70]. The study provides a nice alternative to template-driven systems by allowing the physician to specify familiar medically-relevant parameters, reducing friction in providing flexible yet personalized programs. The problem remains, however, that the therapist still needs to set the parameter values for each user and exercise. This prevents the system from scaling efficiently to novel exercises or a large number of users.

A common difficulty in automated self-training systems is the presentation of the appropriate amount of feedback. Oshita et al posit that it is impractical to show all of the motion features at the same time. Thus, they propose a system that highlights the aspect of the movement that will provide the most improvement when changed. The study focuses on the self-training of tennis shots, comparing trainee motion with an expert-demonstrated motion. Their algorithm compares the two movements by drawing an arrow on the component considered most relevant for correction (Figure 4.13) [14]. The study provides a simple solution to the feedback visualization problem by highlighting only a single component, i.e. the one that provides the greatest improvement. In many ways, sports science and rehabilitation share many of the same concepts for motor skill learning, and many innovations like these can inspire solutions in other fields.

Perhaps a point of contention is the use of pre-recorded templates performed by a person besides the trainee. Whenever motions performed by different individuals are compared, physical characteristics of each individual introduce multiple factors that may be difficult to take into account. Our system offers the alternative where the template itself is performed by the trainee as well. This effectively eliminates physical differences between trainee and template.

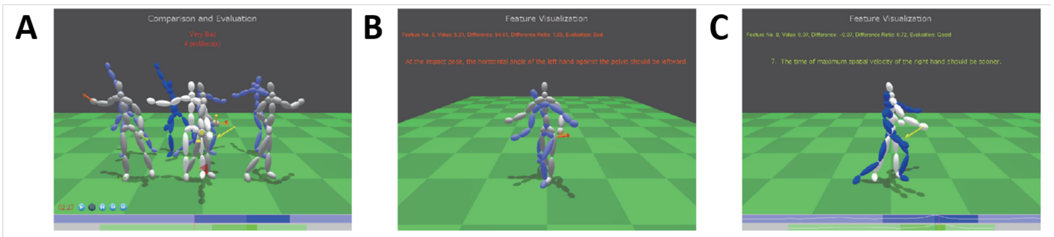

Figure 4.13: Self-training system for tennis shots (adapted from [14]), **(A)** comparison and evaluation mode for trainee and target motions, **(B)** isolated key poses indicated by gray and light blue, **(C)** isolated trainee pose in white and target pose in blue.

"Exergames" are digital games designed to encourage physical activity. They share similar goals as the self-exercise systems discussed thus far, by providing interactivity and feedback to the user. However, exergames mainly focus on the problem of adherence (i.e. compliance in doing the assigned exercises); it is considered to be one of the greatest challenges to rehabilitative self-exercise. Uzor et al designed an exergame system for reducing fall risk in seniors. The games were designed to improve muscle strength and balance, using movements from existing rehabilitation programs. The user's performance is evaluated by a game score, based on the quality of the movement, quantified by range, pace, and repetition count [15]. The study provides the most simple and interactive elements among the systems we have discussed thus far. Further, quantifiable measures were used to provide a simple score as feedback. However, it must be noted that the system does not put a large importance on performing movements accurately, which may be a problem for certain exercises requiring fine-grained movements. Nonetheless, a gamified approach can certainly be coupled with a fine-tuned scoring system to provide an overall better system.

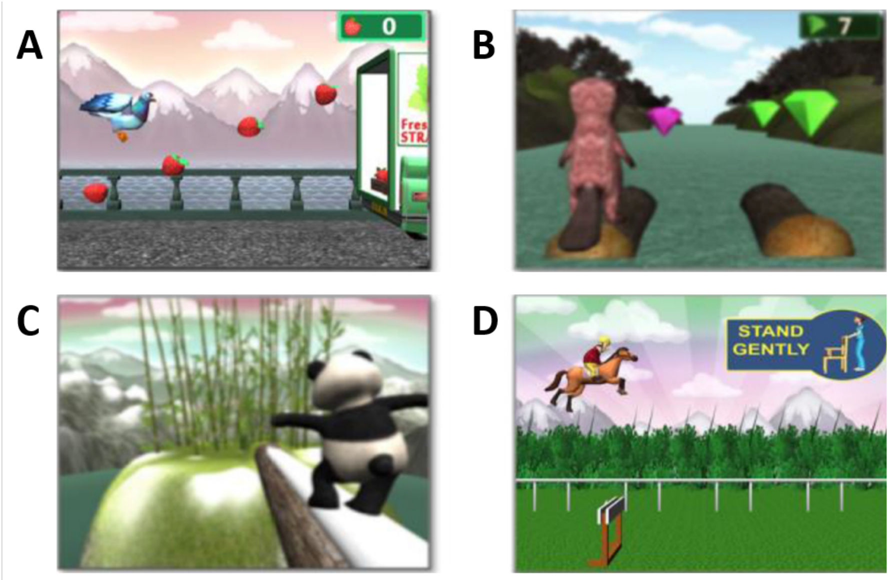

Figure 4.14: Exergames to reduce fall risk in seniors (adapted from [15]), **(A)** Pigeon Express Game for sit-to-stand exercise, **(B)** River Gems Game for side steps exercise, **(C)** Panda Peak Game for marching exercise, **(D)** Horse Hurdles Game for knee bends exercise.

4.3.7 Limitations

Our approach was only demonstrated to work well in a controlled environment. Tests on a larger variety of users and motor tasks would be necessary to confirm its clinical utility. We discuss some of the methodological limitations here.

The proposed framework was tested on sessions between one therapist and multiple healthy subjects. Realistically, different PTs may have different individual preferences, and it would be interesting to investigate intervention strategies across multiple PTs. Currently, we made the simplifying assumption that a patient's attending PT can best prescribe the personalized exercises. Furthermore, none of the subjects had any motor impairments. Identifying the appropriate impairments and conditions for our system would be challenging and is outside the scope of the current study.

Learning the GPDM involves numerical optimization in estimating the model unknowns $\{\mathbf{X}, \overline{\alpha}, \overline{\beta}\}$. In our study, we needed to set both the number of learning iterations and the number of appended CPs. We found that setting a low number for both quantities saves on computational costs. In studies using GPLVM-based methods, the best working settings are generally reported without explanation. Some examples include iteration $T = 15$ for GPLVM [34], outer loop iteration $I = 100$ for GPDM [10], and no mention for Hierarchical-GPLVM [7]. Notably, these methods have been demonstrated to generalize well despite few iterations and training samples.

We considered only motion capture data as a modality since it is naturally learned and can be measured remotely. However, incorporating other modalities can be useful for characterizing different motions, as we saw with body CoM for sit-to-stand (Figure 4.12).

5 Conclusion

5.1 Summary

Physiotherapy is a labor-intensive process that has become increasingly inaccessible. Telehealth has emerged as an alternative to conventional therapy, but with various limitations. Chief among them is the model choice for the human subject. Due to the complexity of the human motor system, matching the appropriate model to a particular motor task is non-trivial. This problem is addressed using latent variable models, due to its desirable properties in the context of physiotherapy. In this dissertation, we presented the application of two latent variable models in the physical intervention of the sit-to-stand task, an important physiotherapy exercise.

In Chapter 3, a linear latent variable model was used to quantify neuromuscular changes during physiotherapy. We introduced an aggregated approach to muscle synergy extraction, which resulted in a common group synergy set. This allowed the extraction of three new metrics for comparing motor behavior. It was demonstrated that half of the original dimensionality was sufficient to form a group synergy set, and that the coefficient-based metrics are able to elucidate the therapists' contribution to the timing and magnitude of natural motor strategies.

In Chapter 4, a non-linear latent variable model was used to encode expert-induced movements in a data-efficient manner. We introduced a reorganization approach to a Gaussian Process embedding of multiple motor task performances. This allowed the extraction of simple yet meaningful augmented feedback mechanisms. It was demonstrated that the proposed framework is able to preserve the predictive ability of the Gaussian Process Dynamical Model, provide simple feedback mechanisms, and train personalized models with reasonably few demonstrations.

Overall, our current methods are able to capture some potential factors that can cause motor strategies to deviate from normal behavior, and what kinds of changes are induced. By comparing the extracted metrics through various natural conditions and expert-induced conditions, we were able to gain some insight on the expert strategy as well as build tools for providing expert-level feedback. The proposed approaches present simple low-dimensional features that can serve as feedback mechanisms for both patient and therapist. Our hope is that models and metrics such as those proposed may be helpful to systematically evaluate the current intervention methods and empirically evaluate patient recovery progress.

5.2 Future Directions

A comprehensive solution for solving the various issues in physiotherapy is an open-ended problem. A wide range of factors need to be considered, such as the type of clinician, impairment, and exercise. Thus, future work first involves the consideration of a wider range of conditions.

A patient often needs to consult with different types of healthcare providers. It would be natural to assume that intervention strategies are not necessarily shared among the different healthcare professionals. Similar to how natural motor action is highly individualized, so can the intervention or assistance performed. In particular, it would be useful to characterize the assistance provided by nurses, caregivers, as well as other physical therapists.

Testing the system for generalizability for different types of movements is another possible research direction. There are many types of impairments and each with their corresponding sets of exercises. Since our tests focus on a movement which inherently involves the whole body, it would be interesting to test the system for exercises with naturally isolated movements, such as shoulder abduction and adduction [3].

In the context of telehealth, the type of modality plays a big role in the data collection and system implementation. We currently considered only electromyography and joint position data due to practical reasons. Both types of data can be logged wirelessly and the raw sensor data need not be transformed. However, other biomechanical quantities, such as kinematic (e.g. spatial direction) and kinetic (e.g. force) parameters, are also usual modalities to consider. Moreover, with the ubiquity of wearables, modalities such as heart rate, muscle activation, and acceleration have becoming increasingly accessible. A multi-modal model can be an interesting direction to consider.

Finally, we have demonstrated a "patient-centric" framework where only the patient's body is modeled. The alternative "therapist-centric" approach is an interesting possibility where therapist's body is modeled instead. In this framework, the learner is not the patient, but the novice therapist. This turns from learning expert-induced movements to learning the expert motor skill itself. This approach need not be limited to physiotherapy application, but it can be reasonably applied to arbitrary motor tasks where expert demonstrations are available, such as sports science.

Acknowledgements

The path towards a PhD degree has been difficult, where challenges and failures were plentiful. Overcoming it is perhaps impossible without the support of those whose paths have fortunately crossed with mine. For your kindness, patience, and time, I thank you.

Foremost, I wish to express my deepest gratitude to my advisor Professor Kazushi Ikeda, for accepting me as a student and allowing me to work under his guidance. He has provided me with the support and the flexibility to pursue my academic endeavors. I would also like to thank my thesis committee members and journal referees, for their encouragement, insightful comments, and thought-provoking questions. I am particularly grateful to Associate Professor Tomoya Tamei, who has been working closely with me from the beginning. Through him, I have learned many practical skills in research, as in life.

I thank my fellow lab members in the Mathematical Informatics Laboratory. It's always a joy to see and interact with them in the lab. I am grateful for the constant support and encouragement of the professors and secretaries, who are always looking out for us students. I want to give a shoutout to Christopher Bishop and Angelika Hofmann for molding the budding scientists in the MI Lab. I want to extend this gratitude to all the nice people of NAIST, that made my PhD experience a meaningful one. Life wouldn't be as nice if it was *all* research.

I want to express my gratitude to the MEXT Scholarship Program of Japan for the opportunity of a lifetime. Their unwavering support for education has created a comfortable research environment for me and my fellow scholars.

My biggest thanks to my family, for all the love they have showered me with!

References

[1] D. Chugo, K. Kawabata, H. Kaetsu, N. Miyake, E. Okada, H. Asama, and K. Kosuge, "Force assistance control for standing-up motion," in *Biomedical Robotics and Biomechatronics. The First IEEE/RAS-EMBS International Conference on*, pp. 135–140, IEEE, 2006.

[2] H. Kogami, Q. An, N. Yang, H. Yamakawa, Y. Tamura, A. Yamashita, H. Asama, S. Shimoda, H. Yamasaki, M. Itkonen, *et al.*, "Effect of physical therapy on muscle synergy structure during standing-up motion of hemiplegic patients," *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 2229–2236, 2018.

[3] R. Tang, X.-D. Yang, S. Bateman, J. Jorge, and A. Tang, "Physio@ home: Exploring visual guidance and feedback techniques for physiotherapy exercises," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 4123–4132, ACM, 2015.

[4] N. Kato, T. Tanaka, S. Sugihara, and K. Shimizu, "Development and evaluation of a new telerehabilitation system based on vr technology using multisensory feedback for patients with stroke," *Journal of physical therapy science*, vol. 27, no. 10, pp. 3185–3190, 2015.

[5] N. D. Lawrence, "Dimensionality reduction the probabilistic way," *ICML2008 Tutorial*, 2008.

[6] R. E. Singh, K. Iqbal, G. White, and T. E. Hutchinson, "A systematic review on muscle synergies: From building blocks of motor behavior to a neurorehabilitation tool," *Applied bionics and biomechanics*, vol. 2018, 2018.

[7] N. D. Lawrence and A. J. Moore, "Hierarchical gaussian process latent variable models," in *Proceedings of the 24th international conference on Machine learning*, pp. 481–488, ACM, 2007.

[8] A. Damianou, M. K. Titsias, and N. D. Lawrence, "Variational gaussian process dynamical systems," in *Advances in Neural Information Processing Systems*, pp. 2510–2518, 2011.

[9] M. Schenkman, R. A. Berger, P. O. Riley, R. W. Mann, and W. A. Hodge, "Whole-body movements during rising to standing from sitting," *Physical Therapy*, vol. 70, no. 10, pp. 638–648, 1990.

[10] J. Wang, D. Fleet, and A. Hertzmann, "Gaussian process dynamical models for human motion.," *IEEE transactions on pattern analysis and machine intelligence*, vol. 30, no. 2, p. 283, 2008.

[11] A. Gupta, T. Chen, F. Chen, D. Kimber, and L. S. Davis, "Context and observation driven latent variable model for human pose estimation," in *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8, IEEE, 2008.

[12] N. Chen, M. Karl, and P. Van Der Smagt, "Dynamic movement primitives in latent space of time-dependent variational autoencoders," in *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*, pp. 629–636, IEEE, 2016.

[13] B. Ferris, D. Fox, and N. D. Lawrence, "Wifi-slam using gaussian process latent variable models.," in *IJCAI*, vol. 7, pp. 2480–2485, 2007.

[14] M. Oshita, T. Inao, S. Ineno, T. Mukai, and S. Kuriyama, "Development and evaluation of a self-training system for tennis shots with motion feature assessment and visualization," *The Visual Computer*, vol. 35, no. 11, pp. 1517–1529, 2019.

[15] S. Uzor and L. Baillie, "Recov-r: Evaluation of a home-based tailored exergame system to reduce fall risk in seniors," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 26, no. 4, pp. 1–38, 2019.

[16] L. Schopp, B. Johnstone, and D. Merrell, "Telehealth and neuropsychological assessment: New opportunities for psychologists.," *Professional Psychology: Research and Practice*, vol. 31, no. 2, p. 179, 2000.

[17] United Nations, "World population ageing 2019: Highlights," 2019.

[18] M. Higo and H. T. Khan, "Global population aging: Unequal distribution of risks in later life between developed and developing countries," *Global Social Policy*, vol. 15, no. 2, pp. 146–166, 2015.

[19] A. C. Lee, T. E. Davenport, and K. Randall, "Telehealth physical therapy in musculoskeletal practice," *journal of orthopaedic & sports physical therapy*, vol. 48, no. 10, pp. 736–739, 2018.

[20] C. E. Lang, J. R. MacDonald, D. S. Reisman, L. Boyd, T. J. Kimberley, S. M. Schindler-Ivens, T. G. Hornby, S. A. Ross, and P. L. Scheets, "Observation of amounts of movement practice provided during stroke rehabilitation," *Archives of physical medicine and rehabilitation*, vol. 90, no. 10, pp. 1692–1698, 2009.

[21] N. Bernstein, "The co-ordination and regulation of movements," *The co-ordination and regulation of movements*, 1966.

[22] L. H. Ting and J. L. McKay, "Neuromechanics of muscle synergies for posture and movement," *Current opinion in neurobiology*, vol. 17, no. 6, pp. 622–628, 2007.

[23] A. Tang and K. Dillman, "Towards next-generation remote physiotherapy with videoconferencing tools," tech. rep., University of Calgary, 2013.

[24] C. E. Lang, J. R. MacDonald, and C. Gnip, "Counting repetitions: an observational study of outpatient therapy for people with hemiparesis post-stroke," *Journal of Neurologic Physical Therapy*, vol. 31, no. 1, pp. 3–10, 2007.

[25] Y. Sunaryadi, "The role of augmented feedback on motor skill learning," in *6th International Conference on Educational, Management, Administration and Leadership*, Atlantis Press, 2016.

[26] B. Parmanto and A. Saptono, "Telerehabilitation: State-of-the-art from an informatics perspective," *International journal of telerehabilitation*, vol. 1, no. 1, p. 73, 2009.

[27] M. Ayoade and L. Baillie, "A novel knee rehabilitation system for the home," in *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 2521–2530, ACM, 2014.

[28] P. Li and S. Chen, "A review on gaussian process latent variable models," *CAAI Transactions on Intelligence Technology*, vol. 1, no. 4, pp. 366–376, 2016.

[29] S. Giszter, V. Patil, and C. Hart, "Primitives, premotor drives, and pattern generation: a combined computational and neuroethological perspective," *Progress in brain research*, vol. 165, pp. 323–346, 2007.

[30] D. B. Lockhart and L. H. Ting, "Optimal sensorimotor transformations for balance," *Nature neuroscience*, vol. 10, no. 10, p. 1329, 2007.

[31] L. H. Ting and S. A. Chvatal, "Decomposing muscle activity in motor tasks," *Motor Control Theories, Experiments and Applications. Oxf. Univ. Press, New York*, pp. 102v–138, 2010.

[32] N. Yang, Q. An, H. Yamakawa, Y. Tamura, A. Yamashita, and H. Asama, "Muscle synergy structure using different strategies in human standing-up motion," *Advanced Robotics*, vol. 31, no. 1-2, pp. 40–54, 2017.

[33] J. Walleczek, *Self-organized biological dynamics and nonlinear control: toward understanding complexity, chaos and emergent function in living systems*. Cambridge University Press, 2006.

[34] N. Lawrence, "Probabilistic non-linear principal component analysis with gaussian process latent variable models," *Journal of machine learning research*, vol. 6, no. Nov, pp. 1783–1816, 2005.

[35] M. E. Tipping and C. M. Bishop, "Probabilistic principal component analysis," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 61, no. 3, pp. 611–622, 1999.

[36] R. T. Harbourne and N. Stergiou, "Movement variability and the use of nonlinear tools: principles to guide physical therapist practice," *Physical therapy*, vol. 89, no. 3, pp. 267–282, 2009.

[37] P. M. Dall and A. Kerr, "Frequency of the sit to stand task: an observational study of free-living adults," *Applied ergonomics*, vol. 41, no. 1, pp. 58–61, 2010.

[38] K. Kerr, J. White, D. Barr, and R. Mollan, "Standardization and definitions of the sit-stand-sit movement cycle," *Gait & Posture*, vol. 2, no. 3, pp. 182–190, 1994.

[39] H.-F. Liao, Y.-C. Liu, W.-Y. Liu, and Y.-T. Lin, "Effectiveness of loaded sit-to-stand resistance exercise for children with mild spastic diplegia: a randomized clinical trial," *Archives of physical medicine and rehabilitation*, vol. 88, no. 1, pp. 25–31, 2007.

[40] P. K. Schot, K. M. Knutzen, S. M. Poole, and L. A. Mrotek, "Sit-to-stand performance of older adults following strength training," *Research Quarterly for exercise and Sport*, vol. 74, no. 1, pp. 1–8, 2003.

[41] B. Etnyre and D. Q. Thomas, "Event standardization of sit-to-stand movements," *Physical therapy*, vol. 87, no. 12, pp. 1651–1666, 2007.

[42] P. J. Millington, B. M. Myklebust, and G. M. Shambes, "Biomechanical analysis of the sit-to-stand motion in elderly persons," *Archives of Physical Medicine and Rehabilitation*, vol. 73, no. 7, pp. 609–617, 1992.

[43] W. G. Janssen, H. B. Bussmann, and H. J. Stam, "Determinants of the sit-to-stand movement: a review," *Physical therapy*, vol. 82, no. 9, pp. 866–879, 2002.

[44] K. M. Newell and D. M. Corcos, *Variability and motor control.* Human Kinetics Publishers Champaign, IL, 1993.

[45] A. Daffertshofer, C. J. Lamoth, O. G. Meijer, and P. J. Beek, "Pca in studying coordination and variability: a tutorial," *Clinical biomechanics*, vol. 19, no. 4, pp. 415–428, 2004.

[46] A. d'Avella, P. Saltiel, and E. Bizzi, "Combinations of muscle synergies in the construction of a natural motor behavior," *Nature neuroscience*, vol. 6, no. 3, p. 300, 2003.

[47] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.

[48] A. d'Avella, A. Portone, L. Fernandez, and F. Lacquaniti, "Control of fast-reaching movements by muscle synergy combinations," *The Journal of neuroscience*, vol. 26, no. 30, pp. 7791–7810, 2006.

[49] A. d'Avella and E. Bizzi, "Shared and specific muscle synergies in natural motor behaviors," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 8, pp. 3076–3081, 2005.

[50] A. Saito, A. Tomita, R. Ando, K. Watanabe, and H. Akima, "Similarity of muscle synergies extracted from the lower limb including the deep muscles between level and uphill treadmill walking," *Gait & posture*, vol. 59, pp. 134–139, 2018.

[51] D. J. Berger, R. Gentner, T. Edmunds, D. K. Pai, and A. d'Avella, "Differences in adaptation rates after virtual surgeries provide direct evidence for modularity," *The Journal of Neuroscience*, vol. 33, no. 30, pp. 12384–12394, 2013.

[52] P. O. Riley, M. L. Schenkman, R. W. Mann, and W. A. Hodge, "Mechanics of a constrained chair-rise," *Journal of biomechanics*, vol. 24, no. 1, pp. 77–85, 1991.

[53] K. M. Steele, M. C. Tresch, and E. J. Perreault, "The number and choice of muscles impact the results of muscle synergy analyses," *Frontiers in computational neuroscience*, vol. 7, p. 105, 2013.

[54] M. A. Hughes and M. L. Schenkman, "Chair rise strategy in the functionally impaired elderly," *Journal of rehabilitation research and development*, vol. 33, no. 4, p. 409, 1996.

[55] D. M. Scarborough, C. A. McGibbon, and D. E. Krebs, "Chair rise strategies in older adults with functional limitations.," *Journal of Rehabilitation Research & Development*, vol. 44, no. 1, 2007.

[56] Motion Analysis Corporation, *EVaRT 5.0 User's Manual.* Motion Analysis Corporation, 2006.

[57] A. L. Hof, "Scaling gait data to body size," *Gait & posture*, vol. 3, no. 4, pp. 222–223, 1996.

[58] F. Bahrami, R. Riener, P. Jabedar-Maralani, and G. Schmidt, "Biomechanical analysis of sit-to-stand transfer in healthy and paraplegic subjects," *Clinical Biomechanics*, vol. 15, no. 2, pp. 123–133, 2000.

[59] E. A. Tully, M. R. Fotoohabadi, and M. P. Galea, "Sagittal spine and lower limb movement during sit-to-stand in healthy young subjects," *Gait & posture*, vol. 22, no. 4, pp. 338–345, 2005.

[60] U. Yang and G. J. Kim, "Implementation and evaluation of "just follow me": an immersive, vr-based, motion-training system," *Presence: Teleoperators & Virtual Environments*, vol. 11, no. 3, pp. 304–323, 2002.

[61] M. Piqueras, E. Marco, M. Coll, F. Escalada, A. Ballester, C. Cinca, R. Belmonte, and J. M. Muniesa, "Effectiveness of an interactive virtual telerehabilitation system in patients after total knee arthroplasty: a randomized controlled trial," *Journal of rehabilitation medicine*, vol. 45, no. 4, pp. 392–396, 2013.

[62] F. Anderson, T. Grossman, J. Matejka, and G. Fitzmaurice, "Youmove: enhancing movement training with an augmented reality mirror," in *Proceedings of the 26th annual ACM symposium on User interface software and technology*, pp. 311–320, ACM, 2013.

[63] E. Velloso, A. Bulling, and H. Gellersen, "Motionma: motion modelling and analysis by demonstration," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1309–1318, ACM, 2013.

[64] A. C. Damianou, M. K. Titsias, and N. D. Lawrence, "Variational inference for latent variables and uncertain inputs in gaussian processes," *The Journal of Machine Learning Research*, vol. 17, no. 1, pp. 1425–1486, 2016.

[65] P. W. Kwong, S. S. Ng, R. C. Chung, and G. Y. Ng, "Foot placement and arm position affect the five times sit-to-stand test time of individuals with chronic stroke," *BioMed research international*, vol. 2014, 2014.

[66] J. Chen, M. Kim, Y. Wang, and Q. Ji, "Switching gaussian process dynamic models for simultaneous composite motion tracking and recognition," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2655–2662, IEEE, 2009.

[67] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural networks*, vol. 2, no. 5, pp. 359–366, 1989.

[68] J. Butepage, M. J. Black, D. Kragic, and H. Kjellstrom, "Deep representation learning for human motion prediction and classification," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 6158–6166, 2017.

[69] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime multi-person 2d pose estimation using part affinity fields," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7291–7299, 2017.

[70] A. E. F. Da Gama, T. de Menezes Chaves, P. Fallavollita, L. S. Figueiredo, and V. Teichrieb, "Rehabilitation motion recognition based on the international biomechanical standards," *Expert Systems with Applications*, vol. 116, pp. 396–409, 2019.

Publication List

Journals

[1] **Bryan Lao**, Tomoya Tamei, Kazushi Ikeda. Data-Efficient Framework for Personalized Physiotherapy Feedback. Frontiers in Computer Science. 2020.

[2] Clark Kendrick Go, Monamie Ringhofer, **Bryan Lao**, Takatomi Kubo, Shinya Yamamoto, Kazushi Ikeda. A mathematical model of herding in a horse-harem group. Behavioural Processes. Submitted.

International Conferences

[1] **Bryan Lao**, Tomoya Tamei, Kazushi Ikeda. Characterizing Strategic Contributions of Physical Therapy to Natural Standing Motion in the Muscle Synergy Space. In 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2019). pp. 2311-2315. 23-27 July 2019.

[2] **Bryan Lao**, Tomoya Tamei, Kazushi Ikeda. Muscle Synergies as Index for Expert Sit-to-Stand Therapy. In 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2018). 17-21 July 2018.