A Formal Investigation of Human Spatial Control Skills: Mathematical Formalization, Skill Development, and Skill Assessment

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Bin Li

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor of Philosophy

Advisor: Bérénice Mettler

July, 2016

Acknowledgements

There are a multitude of people without whom this dissertation would not have been completed and to whom I am greatly indebted.

First and foremost, to my mentor and advisor, Prof. Bérénice Mettler, who brought to me her novel insights and passion to this academic world of human skills, guided me to pursue ideas that were inspiring and challenging, patiently taught me about research approaches, presentation techniques, and written skills, I cannot express enough thanks for your continuous help and encouragement during all these years. I truly enjoyed working with you, and I hope our collaboration lasts for a lifetime.

To my other dissertation committee members, Prof. Demoz Gebre-Egziabher, Prof. Timothy Kowalewski and Prof. Yohannes Ketema, I would like to express my sincere gratitude for your availability, for your times reading my dissertation, and for the valuable advice you provided in your areas of expertise.

Great thanks to my colleagues and friends, Jon Andersh, Zhaodan Kong, Abhishek Verma, Andrew Feit, and Kuo-Shih Tseng, at the Interactive Guidance and Control Lab, who shared with me a lot of insights and from whom I learned a lot of skills.

I would also like to thank my friends, Xiaochuan Chai, Yintao Song, Xian Chen, Guanda Wu, Bingzhe Li, Qiannan Li, Peng Liu, Han Zhang, Yin He, Bin Hu, Shu Wang, and Bo Peng, for your encouragement and moral support which made my stay and study in Minnesota enjoyable.

Special thanks to the administrative staff of the Department of Aerospace Engineering and Mechanics, the Graduate School, the International Student and Scholar Services and the Center for Writing. They are always supportive and helpful.

Last but not least, I would like to thank my family: my parents: Yong Li and Lianzi Yang, my elder sister Yun Li, and my fiancée Yan Yang. You have always stood by my

side, encouraged me, unconditionally loved me, taught me values of hard work, of being humble, of being patient, and of integrity. I owe everything that I have achieved to you.

Dedication

To my dearest grandmother up above.

Abstract

Spatial control behaviors account for a large proportion of human everyday activities from normal daily tasks, such as reaching for objects, to specialized tasks, such as driving, surgery, or operating equipment. These behaviors involve intensive interactions within internal processes (i.e. cognitive, perceptual, and motor control) and with the physical world. This dissertation builds on a concept of interaction pattern and a hierarchical functional model. Interaction pattern represents a type of behavior synergy that humans coordinates cognitive, perceptual, and motor control processes. It contributes to the construction of the hierarchical functional model that delineates humans spatial control behaviors as the coordination of three functional subsystems: planning, guidance, and tracking/pursuit. This dissertation formalizes and validates these two theories and extends them for the investigation of human spatial control skills encompassing development and assessment.

Specifically, this dissertation first presents an overview of studies in human spatial control skills encompassing definition, characteristic, development, and assessment, to provide theoretical evidence for the concept of interaction pattern and the hierarchical functional model. The following, the human experiments for collecting motion and gaze data and techniques to register and classify gaze data, are described. This dissertation then elaborates and mathematically formalizes the hierarchical functional model and the concept of interaction pattern. These theories then enables the construction of a succinct simulation model that can reproduce a variety of human performance with a minimal set of hypotheses. This validates the hierarchical functional model as a normative framework for interpreting human spatial control behaviors. The dissertation then investigates human skill development and captures the emergence of interaction pattern. The final part of the dissertation applies the hierarchical functional model for skill assessment and introduces techniques to capture interaction patterns both from the top down using their geometric features and from the bottom up using their dynamical characteristics. The validity and generality of the skill assessment is illustrated using two the remote-control flight and laparoscopic surgical training experiments.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Human spatial control behavior accounts for a large portion of everyday activities. Figure 1.1 illustrates a number of example activities, which include normal tasks, such as handwriting and reaching for objects, as well as specialized tasks, such as driving, surgery, or operating equipment. Human spatial control behavior typically involves the manual control of an end-effector (e.g. a tool or a dynamical system) with relation to the immediate surrounding environment.

The examples in Figure 1.1 show humans being confronted with increasing interactions associated with the agility of end-effector dynamics (e.g. fast response and high sensitivity to human control) and the complexity of task elements and environments. However, given adequate practice, humans can acquire spatial control skills to conquer these difficulties and conduct spatial control behavior successfully with high efficiency and consistency.

This dissertation focuses on establishing a framework that can delineate skill components and their functionalities in coordinating spatial control behavior, illuminate the mechanism underlying skill development progress, and enable a detailed and robust skill assessment. We start the chapter with our motivations for investigating human spatial control behaviors. Following, in Section 1.2, we illustrate what is currently missing from the existing literature and provide the main thesis of this dissertation. We then present an outline of this dissertation in Section 1.3 and conclude this chapter in Section 1.4 with a list of publications corresponding to each of the following chapters.

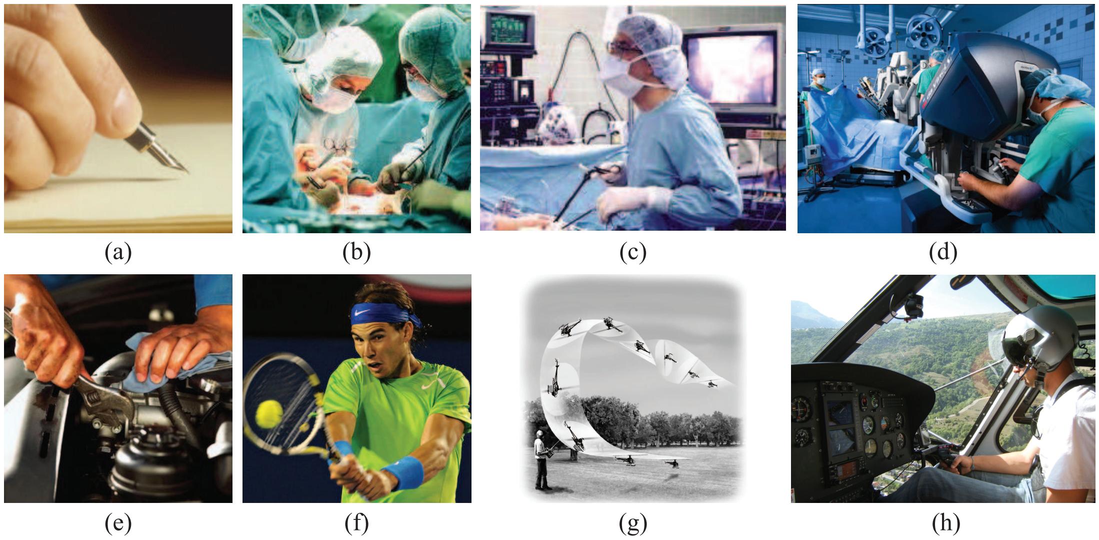


Figure 1.1: Illustrations of spatial control behavior. (a) Handwriting (image from jobs.guardian.co.uk). (b) Manual surgery (image from animalpetdoctor.homestead.com). (c) Augmented surgery [1]. (d) Surgery with da Vinci surgical system [1]. (e) Vehicle maintenance (image from www.gilbertautoford.com). (f) Tennis (image from www.nytimes.com). (g) Remote-control flight. (h) Airplane operating [1].

1.1 Motivation

1.1.1 Machine Systems and Human Roles

During the last century, human roles in human-machine systems have been migrating from manual, to augmented, and ultimately to supervisory functions [1]. However, autonomous machines are still far from replacing all manual work. In fact, the "irony of automation," proposed by Bainbridge [7], claims that complete automation is impossible due to human limitations. First, humans are always subject to mistakes, so automated systems, being designed by humans, can never be free of error. Second, automated systems are unable to deal with conditions that human designers have failed to consider. Moreover, Bibby and colleagues [8] stressed that, "even highly automated systems, such as electric power networks, need human beings for supervision, adjustment, maintenance, expansion and improvement."

Increasingly advanced automation systems greatly alleviate human workload and reduce human intervention in that they can detect various abnormalities and address them automatically. However, at the same time, these systems increases demands for higher human skills when, once the systems break down, human operators need to deal with conditions that occur infrequently and necessitate faster and more precise responses. For instance, in the accident of US Airway Flight 1549 on January 15, 2009, the plane lost engine power after a collision with a flock of Canadian geese during its initial ascent. The auto-pilot system and auto-landing system were unable to respond to that condition. Human pilots took over emergency control and landed the plane in the Hudson River with no loss of passenger lives.

The above examples underscore the need to guarantee the proficiency of human operators in tackling normal conditions accurately and consistently, and being well-prepared for unpredicted conditions. This necessitates standardized training and credentialing [9, 10, 11]. Achieving this objective relies on the advancement in the understanding of human behaviors.

Human behaviors are constrained by biological and psychophysical limitations in nearly all functionalities. From the motor control perspective, human procedural memory can only record approximate limb poses and rough force scales, such that only limited precision can be reached even with extensive practice. The simplest human reactions are measured in tens and hundreds of milliseconds [12]. From the perceptual perspective, human ears are sensitive to sound in the frequency range from 20 to 20000 Hz, and can make precise distinction between 1000 and 5000 Hz; human eyes are sensitive to light in the wavelength range from 400 to 700 nm [12]. From the cognitive perspective, most human adults can retain no more than about seven distinct elements in their working memory for a short duration around 10-15 seconds [13, 14]. The firing frequency of neurons varies from 250 to 2000 Hz, which implies a minimum of 0.5 ms inter-signal intervals. An act of recognition takes nearly a second [15, 12]. Moreover, human behaviors are subject to other factors such as fatigue, emotions, and experience.

Although the above limitations prevent humans from achieving complete optimality, accuracy, and consistency, thousands of years of evolution have endowed humans with a variety of survival capacities. One survival capacity is adaptability. As Parasuraman and Riley [16] indicated, humans are excellent in "responding to changing or unforeseen conditions." More importantly, humans do not passively adapt to surrounding environments. Humans can perform self-directed learning to acquire skills that improve task performance. Another survival capacity is versatility. It enables humans to use their skills on different tasks. This versatility builds upon the fact that humans can adeptly detect inherent similarities or coherence among tasks.

However, human performance encompassing adaptability and versatility is still not fully understood by researchers. Over the last decades, human skills have been an active focus of research covering a broad range of disciplines, including movement science [17, 18], neuroscience[19], cognitive science [20], and psychology [21]. These studies have illuminated that all human skills involve the coordination of cognitive, perceptual, and motor control components [22]. Different types of skills distinguish themselves in which of these three components is emphasized [23]. Existing surveys on human skills primarily focused on motor control or movement skills, and some of them considered the roles of perception and information processing in performing and developing these motor control skills [24, 23, 25, 26, 27, 28].

1.1.2 Human Spatial Control Skills

This dissertation is primarily concerned with spatial control skills that involve intensive interactions encompassing all cognitive, perceptual, and control components. In typical spatial control tasks, the environment may not be fully known, the task space may not be well structured, and the task elements may change over time. How do humans achieve successful performance of versatile and adaptive spatial control behaviors under these difficulties? This question has spurred an active research on human spatial control skills for more than a century. Early studies focused on the motor control outcomes of spatial control skills, for instance, precision and reaction time of repetitive tasks [29] and the effect of knowledge of results (KR) [30]. During the mid-twentieth century there was a surge of skills research due to wartime demands for tracking tasks, such as flying, driving, or aiming [23]. Researchers increasingly analyzed tracking skills, using the theory of servo control and, later, modern control. These models treated a human operator as an element in a “closed-loop” system [31, 32]. Ecological psychology, which flourished in the 1960s and 1970s, provided a more holistic view, focusing on the interactions between humans and environments [33, 34]. In opposition to closed-loop theories, Keele [35] proposed the concept of motor programs and Schmidt [36] extends it to a schema theory, suggesting that humans could develop motor programs as generalized muscle patterns that function largely as open-loop processes. In the 1980s, skills research migrated to an information processing perspective, emphasizing the role of the central nervous system as an information-processing system. The research also looked at the effects of humans' limited-capacity information channel [13] and intermittent servo-action pattern [23]. Finally, the advent of brain recording and imaging enabled the neuro-physiological perspective and the related motor control models.

The primary difficulty in understanding spatial control skills is that spatial behavior relies on a combination of perceptual, cognitive, and sensorimotor control processes. Humans must develop spatial and temporal movement patterns but also representations of both the physical world and body configuration, for instance, through joint angles and muscle strength. Furthermore, they need to acquire efficient mechanisms for extracting cues that provide relevant information for task performance. These cues are used by cognitive processes to generate executable control plans. At the same time, achieving a satisfactory performance relies on monitoring outcomes.

More importantly, perceptual, cognitive, and sensorimotor control processes are coupled and operate as a system, with each component combining with the others to enable versatile and adaptive capabilities [37]. For instance, a mental model can predict task-relevant cues and guide attention to areas where the relevant cues have a high probability of occurring. Andersh and colleagues [38] investigated human perceptual function in the remote control of miniature rotorcrafts, suggesting that gaze plays a role in estimating the rotorcraft states necessary for tracking, as well as updating information of anticipatory goal interception for planning. Although the planning process generates action plans for the motor control process, motor control skill can in turn affect the selection of action plans. For instance, Pratt and colleagues [39] found that children with significant impairments in motor skill had difficulties with planning.

The investigation of human spatial control skills also requires capturing the closed-loop and active interactions with the environment. According to the ecological psychology theory [34], while performing spatial control behaviors, humans are immersed in the environment and become an active component in a closed-loop system, rather than a feedback servo element described in [32]. Human perception actively interacts with the environment and identifies actions that are affordable for human musculoskeletal system.

During the last decades, heightened requirements have changed the way humans interact with the environment and the task [1]. For example, automobiles are increasingly equipped with features to enhance safety and reduce workload. As well in the surgical domain, Da Vinci and other robotic surgery systems have enabled procedures in a confined space, with an extended view through magnification, high resolution, and high degrees of freedom. All of these technologies alter the way humans process information and perform tasks.

1.1.3 Skill Development and Assessment

All the above situations exaggerate the complexity of human spatial control behaviors and emphasize the heightened requirement for human expertise. Especially in domains, such as surgery, where safety is the highest priority, the performance relies on the operator's rapid perceptual cue extraction, comprehensive information processing, efficient planning, and dexterous motor control. Training and evaluation of specialized operators in these domains is therefore crucial, yet challenging.

Human skill development generally conforms to the "power law of practice" [40, 41]. Subjects show rapid performance improvement in the early phase, followed by decreasing improvements with further practice. Control performance is traditionally measured by quantities such as reaction time and variability [42]. Although the power law indicates that high levels of performance can generally be achieved, approaching the upper limit requires geometrically increasing effort. Therefore, the certifications of professionals such as airline pilots, surgeons, or production line workers have minimum training time requirements. According to the regulations of the Federal Aviation Administration (FAA), an airline transport pilot should have at least 1,500 hours of total flight time, including 500 hours of cross-country, 100 hours of night flight, and 75 hours under instrument conditions [43]. Similarly, according to the American Board of Surgery, the minimum requirement for general surgery certification involves five years of progressive residency and at least 48 weeks of full-time clinical activity in each residency year [44].

However, skill development requires more than simply reducing reaction time and variability. Gentile [45] suggested that motor skill development involves two stages: getting the idea of movements and refining characteristics of goal attainment. McRuer and

Krendel [46, 47] studied the skill development in tracking tasks and idealized skill development in tracking tasks as a progression from a compensatory response to a pursuit-type response, finally culminating in an open-loop response. The open-loop response relies on fully planed and precisely implemented control profiles. Dreyfus [48] also studied skill development across diverse domains including driving and chess playing, and proposed a five-stage model for skill development. This model indicates that humans enhance their skills by identifying domain-specific features, distinguishing situations, and extracting relationships between situations and responses.

Another question concerning human spatial control skills is the measurement of human expertise. This is important for providing feedback to training and credentialing human operators. Existing rating systems in aviation are heavily based on models that date back to crewed operations. These models focus on vehicle-handling characteristics. As well in the surgical field, traditional metrics, such as procedure time, precision, and economy of hand movement, still dominate the assessment and measurement of surgeon skills. Even when combined into a multi-metric scoring system (MMSS) [49], these metrics do not provide an understanding of the heterogeneous and multi-dimensional aspects of skill development and training. For instance, human motor control is subject to a speed-accuracy trade-off [50], so it follows that speed and accuracy should not be considered independently. Finally, the power law of practice characterizing the learning curve suggests that it can be difficult to differentiate between expert and intermediate subjects using these simplistic metrics. Therefore, the development of a more accurate and robust skill assessment system necessitates a more detailed and complete understanding of human spatial control skills.

1.2 Research Statement

Section 1.1 provides a brief background of human spatial control behavior that underscores the increasingly heightened requirements for the training of human operators and assessment of human expertise. These requirements motivate the construction of a comprehensive framework that can delineate human spatial control skills from a system's perspective. The framework should build on the understanding of interactions across internal processes and embedding the subject in the task environment. More importantly, the framework should delineate the basic components of skills as well as their functionalities in coordinating spatial control behaviors and their evolution through skill development progress.

This dissertation build skill modeling framework on interaction patterns (which will be described in Chapter 4) as primary units of spatial control behavior that handle coordination of cognition, perception, and sensorimotor control. These interaction patterns are derived from experimental data with minimal human interventions, and the investigation of skill development verifies their emergence from extensive training. This type of functional element was proposed by Mettler et al. [51] as the key organizational principle that the central nervous system employs for reducing structural and computational complexities. Similar to motor primitives or muscle synergies, interaction patterns bridge the gap between high-level planning and low-level motor control. This understanding was used to construct a hierarchical functional model that integrates three subsystems: planning, guidance, and tracking/pursuit. The interaction patterns and the hierarchical functional model are used in this dissertation to create a skill assessment framework. The specific focus of this work is to achieve a systematic and formal description of the contributions of each subsystem to the evaluation of spatial control performance. This dissertation also investigates the following research questions:

- Do humans employ interaction patterns as behavior units to represent and store their knowledge of spatial control skills? what are the contributions of interaction patterns to human organization of spatial control behaviors?

- How can we formalize human spatial control behavior using mathematical languages?

- Based on the formalization, can we construct a rigorous framework of human spatial control skills that delineates the coordination of skill components to tackle with physical, biological, and neurological constraints in human-machine systems?

- How do human spatial control skills progress through the development process and prompt the emergence of interaction patterns and the subsequent refinement?

- Can we construct a skill assessment system based on the human spatial control skill framework that emphasizes the dimension of behavior units to provides detailed and robust descriptions of human expertise?

- Finally, what conclusions can we reach from empirical analysis to advance our understanding of human spatial control skills and guide the design of human-machine systems?

1.3 Dissertation Outline and Contributions Per Chapter

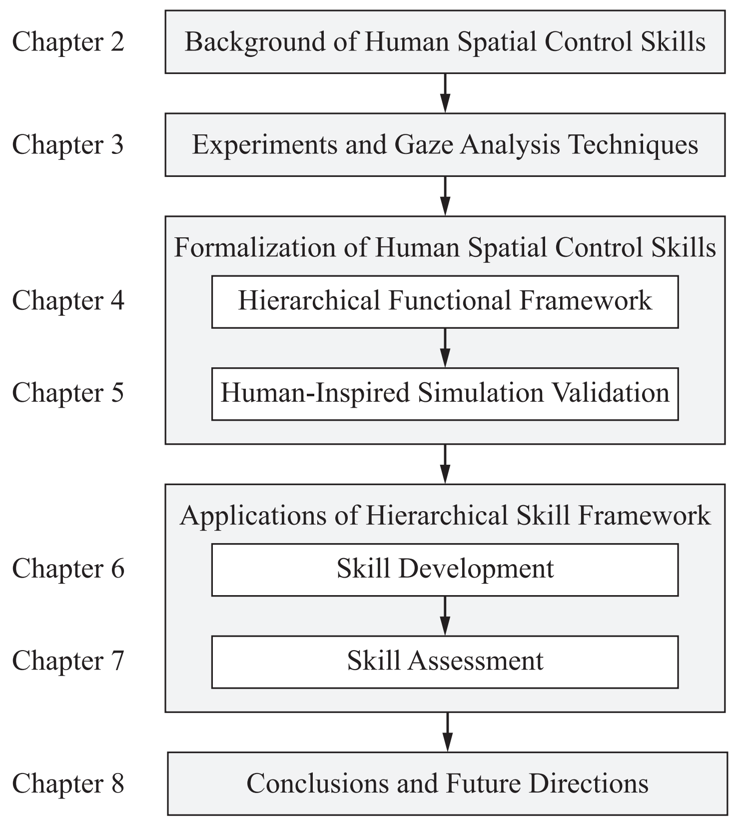


Figure 1.2: Dissertation outline.

Figure 1.2 illustrates the organization of the dissertation. The contributions of each chapter are outlined as follows.

- **Chapter 2**

This chapter provides a detailed review of human spatial control skills, encompassing the definition, characteristics, taxonomy, development, and assessment. The review underscores an inherent structure of human spatial control skills across different behavior domains. Moreover, from the skill development perspective, the review highlights the emergence of a skill hierarchy and emphasizes that human central nervous systems may employ behavior units as the primary dimension for representing knowledge and organizing behavior.

- **Chapter 3**

  This chapter describes the experimental part of the dissertation. It first illustrates two experiments as representative examples of today's human-machine systems: remote control of a miniature helicopter and laparoscopic surgery training. These two experiments demand human skill level beyond that is required by everyday tasks. The chapter then introduces the lab facilities and experimental procedures for the two experiments. The lab facility allows us to collect data not only from human subjects' control data and agent dynamics, but also human subjects' gaze locations. This chapter introduces gaze pre-processing techniques for investigating gaze functions in spatial control behaviors in the following chapters.

- **Chapter 4**

  This chapter and the next one cover the core theoretical part of the dissertation. In these two chapters, we mathematically formalize and validate the hierarchical functional model proposed in previous works as the framework of human spatial control skills. These two chapters mainly serves two purposes: one provides theoretical description of the framework and fundamental analysis for the framework to accommodate with humans' limitations; the other one validates the framework through the construction of a simulation model.

  In this chapter, human spatial control behavior is mathematically formalized as a motion planning problem subject to physical, biological, and neurological constraints in human-machine systems. Solving this problem is NP-hard. This chapter extends the components of the hierarchical functional model to model spatial control skills. Specifically, this chapter elucidates and formalize the concept of interaction patterns and uses them as the essential form of behavior units that humans use to coordinate all of human motor control, perceptual, and cognitive functions. Following that, this chapter describes the hierarchical functional model as a procedural strategy that decomposes a global human guidance problem into multiple levels of lower-dimension local optimization problems, and theoretically proves that the model is compatible with human limitations.

- **Chapter 5**

  In this chapter, the understanding from Chapter 4 is used to construct a human-inspired simulation model. Using the hierarchical functional model as the framework, this simulation model specifies two types of interaction patterns and uses them as units for both planning and motor control. This enables a parsimonious form of the simulation model that can reproduce human performance by tuning only two parameters. It validates the hierarchical functional model as a platform for the interpretation of human spatial control behavior.

- **Chapter 6**

  This chapter investigates human spatial control skill development using a circle task in the remote-control flight configuration. This task involves agile system dynamics, multiple control inputs coordination, and long-duration repetitions. This chapter's primary focus is the mechanisms of human behavior organization leading to the emergence of interaction patterns. This chapter also discusses the specificity and generality of practice for developing skills with respect to unfamiliar/familiar task setups.

- **Chapter 7**

  This chapter focuses on the application and validation of the hierarchical functional model for skill assessment. The hierarchical functional model enables systematic description of human operator's skill that encompasses the three primary levels of planning, guidance, and tracking/pursuit, and also accounts for visual perception. The key feature of the skill assessment is the capturing of interaction patterns. It enables the evaluation of human performance on the dimension of atomic behavior units. Interaction patterns represent a type of sensori-motor synergy that is consistent across task configurations and therefore ensure a robust evaluation of human expertise. The performance of the proposed skill assessment system is illustrated in both the remote-control flight and laparoscopic surgery experiments.

- **Chapter 8**

  This chapter gives key conclusions to the work presented in this dissertation and lays out some possible future directions for the investigation of human spatial control skills.

1.4 List of Publications by Chapters

Below list the publications corresponding to each of the chapters[1] :

- **Chapter 2**

  B. Li and B. Mettler, "Hierarchical Functional Model for Human Spatial Control Skill Assessment, Part I: Review and Theory", *IEEE Transactions on Human-Machine Systems*, final preparation.

- **Chapter 3**

  B. Li, A. Jonathan, and B. Mettler, "Classification of Human Gaze in Spatial Guidance and Control", *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, HongKong, 2015.

  B. Li and B. Mettler, "Hierarchical Functional Model for Human Spatial Control Skill Assessment, Part II: Approach and Applications", *IEEE Transactions on Human-Machine Systems*, final preparation.

  B. Li and B. Mettler, "An Experimental Investigation on Human Spatial Control Skill Development", *1st IFAC Conference on Cyber-Physical & Human-Systems*, submitted.

- **Chapter 4**

[1] Our other publications that are not included in this dissertation are [2, 1, 51]

B. Li and B. Mettler, "Hierarchical Functional Model for Human Spatial Control Skill Assessment, Part I: Review and Theory", *IEEE Transactions on Human-Machine Systems*, final preparation.

B. Li and B. Mettler, "Human Inspired Hierarchical Model for Spatial Control Behavior", *Frontiers in Robotics and AI*, final preparation.

- **Chapter 5**

  B. Li and B. Mettler, "Investigation of Hierarchical Architecture of Human Guidance Behavior for Skill Analysis", *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, HongKong, 2015.

  B. Li and B. Mettler, "Human Inspired Hierarchical Model for Spatial Control Behavior", *Frontiers in Robotics and AI*, final preparation.

- **Chapter 6**

  B. Li and B. Mettler, "An Experimental Investigation on Human Spatial Control Skill Development", *1st IFAC Conference on Cyber-Physical & Human-Systems*, submitted.

- **Chapter 7**

  B. Li and B. Mettler, "Application of trajectory segmentation techniques for operator skill evaluation", *American Control Conference (ACC)*, Washington DC, 2013.

  B. Li and B. Mettler, "Hierarchical Functional Model for Human Spatial Control Skill Assessment, Part II: Approach and Applications", *IEEE Transactions on Human-Machine Systems*, final preparation.

Chapter 2

Background and Related Works

This chapter provides the context necessary to understand the significance of the problem formulated in this dissertation, and it will highlight the main differences between the approaches taken here and those of previously established work in the field of human spatial control skills. The chapter first provides background on the definition and characteristics of both general human skills and spatial control skills in Section 2.1, followed by the description of a human skill taxonomy in Section 2.2. Section 2.3 discusses skill development in terms of knowledge representation and emergent hierarchy. Finally, Section 2.4 reviews and categorizes existing skill assessment methods.

2.1 Definition and Characteristics of Human Skills

This section presents background on general human skills by discussing existing definitions and summarizing characteristics. Next, the section will discuss the definition of spatial control skills and specific challenges for this category of skills.

2.1.1 Nature of Human Skills

The concept of *skill* is synonymous to that of *competence*, *expertise*, and *capacity* [52]. However, expertise and competence concentrate on assessment, referring to the level of performance required for a particular task; capacity implies inherent characteristics of a human individual. Skill is used to describe a function or mechanism underlying behaviors. As Burton and Miller [26] defined in *Movement Skill Assessment*, behavior is "the observable act of moving," and skill is "a qualitative expression of behavior performance." In other words, the relationship between behavior and skill can be described as an observation and its associated distribution function. To characterize a human skill requires sufficient observations of associated behavior.

Researchers have been working on deriving a formal definition of skill for a long time. Leplat [52] defined *skill* as the capacity to execute a class of tasks, focusing on the versatility of skills. Suzuki [53] proposed that "a skill is a motion, or a sequence of motions, which can achieve a certain result," indicating that skills are developed with specific purposes. Welford [54] argued that "one does not observe a skill, only its manifestations," indicating that skills are the mechanisms underlying motor behaviors. Bainbridge [22] defined skill as the efficient use of appropriate behaviors, and argued that efficiency develops with experience.

However, consensus on the definition of skill is difficult to achieve among researchers due to their diverse research objectives. Moreover, it is difficult to provide concrete descriptions of human skills with a simple definition, because human skills have a variety of characteristic features. These features are summarized as follows:

- Skills are "learned" [52, 22]. They are not innate and cannot be inherited from parents. They are developed through direct interactions with tasks and surrounding environments.

- Skills are "goal-directed" [52]. Although for domain-general skills the goal can be vague, skills usually require specific goals to drive the acquisition process.

- Skills are "adaptive" [52]. This feature enables humans to modulate behavior for a specific task configuration.

- Skills are developed in a hierarchical way. Lower-level skills serve as "coordinated units" [55] that can be integrated into a more complex skill. For instance, tennis skill requires coordination of footwork, posture, and stroke, and each of these components also requires coordination of multiple muscles.

- Skills can be measured with workload [22, 52]. This comes from an established fact that execution error will increase if task demands are aggregated. Moreover, the implementation of the second task will degrade the performance of the primary task. The reason for this is that humans have limited information processing channels and computational resources.

Another thing to be noted about skills is that a skill or expertise is typically domain-specific [56]. An expert in one domain may not be professional in other areas. However, skills usually share common elements, structures, and processes. For example, singing and dancing, though different in their associated motor actuators and sensing functions, are both tools to express ideas and emotions. The dancing's physical movement performs a similar function as the singing voice. This similarity suggests that the same basic processes may be employed in these two activities to articulate and organize single units. The commonalities or similarities between skills can also be interpreted as prior knowledge when learning a new skill. As indicated by Bainbridge [22], "no person coming to a task for the first time is a novice in all ways. The prior knowledge can lead to interference or facilitation in learning a new task." Therefore, the existence of a common framework for skills is important to explain human versatility and adaptability.

2.1.2 Human Spatial Control Skills

The spatial control skills considered in this paper concentrate on activities that "involve changing or maintaining the position of the body or the positions of objects in space" [45]. These activities emphasize the coordination of cognitive, perceptual, and motor control components. Activities that primarily involve one or two components (e.g. weight-lifting, reading, speaking, and puzzle solving) are not considered.

Spatial control skills are an essential aspect of many human-machine systems. In these systems, humans are connected with machines and serve as an active component in a closed-loop agent-environment system [32]. Humans must develop spatial and temporal movement patterns but also representations of both the physical world and body configuration. Furthermore, they need to acquire efficient mechanisms for extracting cues that provide relevant information for task performance. These cues are used by cognitive processes to generate executable control plans. At the same time, achieving a satisfactory performance depends on monitoring outcomes.

Therefore, spatial control skill relies on a combination of perceptual, cognitive, and sensorimotor control processes, and these processes must run concurrently, continuously, and in real-time [37]. These processes are coupled and operate as a system, with each component combining with the others to enable versatile and adaptive capabilities. For instance, Pratt and colleagues [39] found that children with significant impairments in motor skill had difficulties with planning.

Moreover, human spatial control skills emphasize human interaction with the task and environment. They belong to the category of "open" skills proposed by Poulton [**?**], as against "closed" skills which may be run off without reference to the environment. More importantly, the interaction with the environment is bi-directional, involving human affecting environmental states via motor control and extracting information from environment via perception [57]. Ecological psychology [34] also emphasizes human perception. It indicates that humans are immersed in the environment acting as an active component in a closed-loop integrated system, rather than a feedback servo element described in [32].

2.2 Taxonomy of Human Skills

A taxonomy of human skills delineates the overall structure of human skills that are common among various skill domains. The construction of a taxonomy relies on integrating knowledge from multiple disciplines, such as motor control science, neuroscience, and psychology. This section first surveys existing taxonomies of human skills. Next, the section presents understandings of each modality of human skills, including motor control, perceptual, and cognitive skills.

2.2.1 Overview

Taxonomy of human skill have been studied broadly. Avermaete [58] suggested classifying human skills into categories of technical and non-technical skills. Technical skills are defined as bodily movements that have external physical manifestations and, therefore, can be easily measured and analyzed using indexes such as speed, fluidity of performance, and error rate. Non-technical skills are associated with internal activities. As an example, Van Avermaete [58] described a framework of non-technical skills for a multi-pilot aircrew. This framework consists of cooperation, leadership and management skills, situation awareness, and decision-making. Similar investigations of non-technical skills were also conducted in the surgical field [59].

However, this type of classification makes it difficult to categorize skills associated with visual perception. Visual skill involves gaze movement, which is the coordination of eye, head, and sometimes trunk movement. These movements can be ostensibly observed, but they have no direct relevance with the movement of task elements.

Considering the above discussion, we constructed a skill taxonomy (as illustrated in Figure 2.1). This taxonomy classifies skills according to the biological and psychophysical modalities (i.e. motor, sensory, and neural systems) associated with them. In particular, motor control skills require efficient coordination of muscles and joints. Perceptual skills deal with the constraints imposed by sensory system. Cognitive skills are enabled by the neural systems. The following sections discuss the details of these three groups of skills.

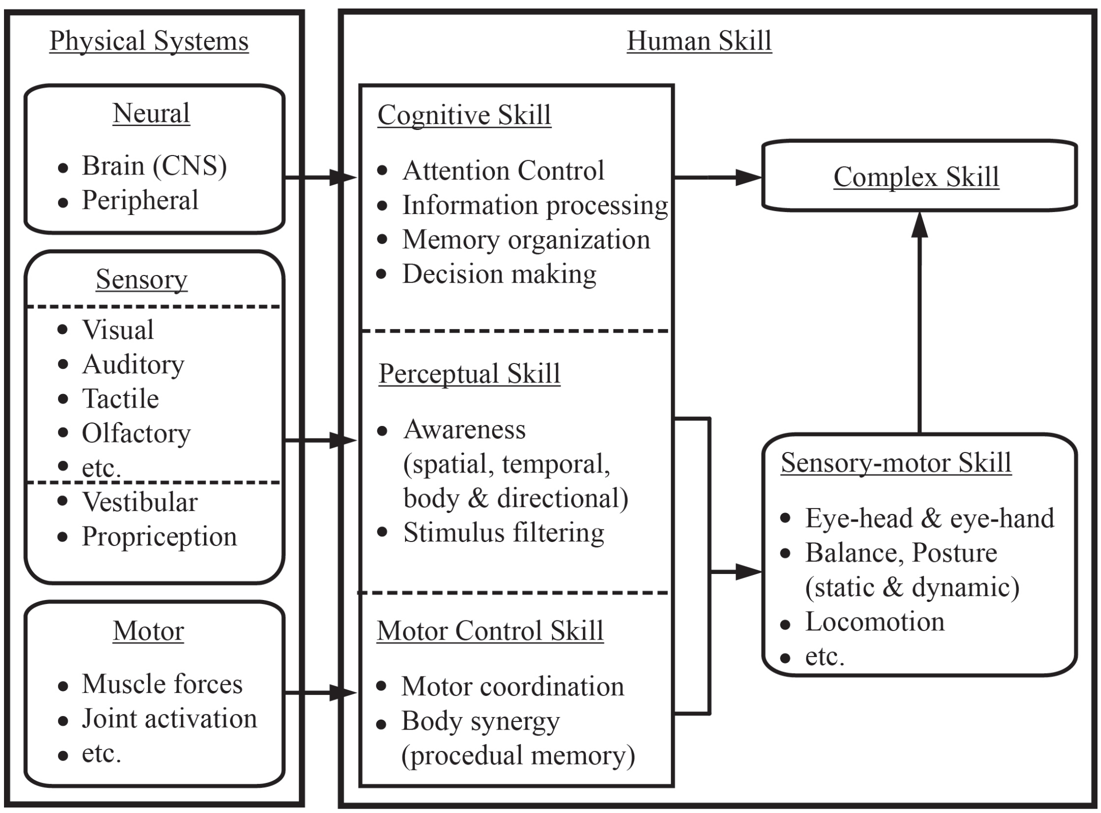


Figure 2.1: Taxonomy of human skills based on human biophysical modalities

2.2.2 Motor Control Skills

Motor control skills involve motion generation through the coordination of muscles to satisfy performance requirements in terms of accuracy and efficiency. They emphasize both movement and the outcome of motions [56].

The acquisition of motor control skills is through directly implementing actions. Proper instructions can facilitate the acquisition and development of motor control skills but cannot replace on-site training. First, humans need to calibrate sensory systems through training by relating actuator movements (e.g. joint angle, muscle strength) with sensory consequences. Second, replicating movements from demonstrations is usually difficult due to the difference in individuals' physical strength, task understanding, perceptual perspectives (i.e. first person versus third person), etc. Moreover, humans are ultimately embedded in a task environment, which involves the combination of a large number of interactions.

An important question associated with motor control behaviors arises from the so-called motor equivalence [60]. Bernstein [61] later rephrased the motor equivalence concept as the degree-of-freedom problem. This concept indicates that the same outcome can be achieved through various motor actions. Humans always have more elements than necessary to conduct motor tasks. For instance, the body has more joints than needed to position the center of mass, the arm has more joints than needed to orient the hand to an object, and the hand has more fingers than needed to generate a grasp [62]. This redundancy gives humans flexibility in achieving efficient behaviors in their daily lives but also makes it challenging to choose an effective combination of muscles [63].

To address the degree-of-freedom problem, Bernstein proposed the concept of muscle synergy [61]. In this concept, the central nervous system (CNS) does not search for a unique solution by simply eliminating redundant degrees of freedom but rather uses the apparently redundant set of joints to ensure more accurate (less variable) performance of the task [62]. This is achieved by learning the appropriate region and size of the synergy space [63].

2.2.3 Perceptual Skills

Perceptual skills involve the acquisition and interpretation of sensory stimuli of environmental elements and internal physical states. They play a significant role in spatial control behaviors by providing adequate information sources [64].

Researchers usually mix up perceptual skills with sensory modalities. However, sensory modalities are the elementary physical functions. They establish the foundation for perception and can be categorized into two groups. Exteroceptive sensory functions, such as visual, auditory, tactile, olfactory, and taste senses, allow humans to interact with external environments. Interoceptive sensory functions, such as vestibular system and proprioception, enable the awareness of internal states (e.g. body position and orientation).

Perceptual skills rely on sensory functions, but, more importantly, perceptual skills integrate information collected from a multitude of sensory functions. A high level of perceptual skills can help compensate the low acuity of a single sensory function. Therefore, perceptual skills are better discriminators of human performance than sensory functions [64, 65].

Perceptual skills consist of three basic functions: awareness, sensorimotor alertness, and filtering.

Awareness

Awareness is the direct representation of current states. It involves the following aspects [66]:

- Spatial awareness: This type of awareness represents the understanding of external space surrounding an individual. It also specifies the spatial constraints that restrain an individual's ability to function motorically in and through space.
- Temporal awareness: This type of awareness provides understanding of the relationship between instances on the dimension of time. The experiment of Grosjean and Terrier [67] indicated that the presence of temporal awareness is a good indicator of performance.

- Body awareness: This type of awareness represents the understanding of the positions and orientations of various body parts. It also involves the understanding of constraints with respect to body parts (e.g. angle of limbs and force strength).
- Directional awareness: This type of awareness is the understanding of laterality and directionality.

Note that perceptual skills do not process the signals taken in. Therefore, the precision of awareness depends on the accuracy of information sources. Moreover, the extent of awareness is heavily limited by the time available [68].

Sensorimotor Alertness (Bottom-Up Filtering for Salient Stimuli)

Sensorimotor alertness is a prerequisite for the more complex attention selectivity [69]. Alertness refers to the rapid switching of focus to stimuli with salient features such as flashing, vibration, or contrasted colors. Humans are very sensitive to these types of stimuli, which sometimes indicates some danger. The instinct of sensorimotor alertness provides spontaneous protection from sudden circumstances, such as collision avoidance. Moreover, sensorimotor alertness is also considered in the design of alarm system in human-machine systems.

Top-Down Filtering by Sensitivity

Top-down sensitivity control modality serves as a gate valve to control the information streaming from all sensory functions. This modality can filter out useless information in the immediate task domain and improve the signal-to-noise ratio [70]. More importantly, the filtering relaxes the limitations of human information processing and memory allocation.

2.2.4 Cognitive Skills

Cognitive skills relies on functions of central nervous system. They encompass processes such as attention direction, information processing, memory organization, learning, and decision making. Cognitive skills allow humans to deal with complicated tasks and environments.

Attention Selection

Attention selection is different from sensorimotor alertness [71] because the former represents the active control of attention. With extensive practice, experts can shift their attention to precursors or cues that are the most information-rich. This skill can also help reject distractions. It enables experts to perceive task-relevant information more efficiently, and frees time to perform more complex cognitive processes, such as planning and decision making [72]. To understand the attention control mechanism, researchers have recently taken a rising interest in the use of eye movements as the external manifestations of visual attention [64, 72].

Attention is mediated by working memory. Working memory refers to structures and processes for temporarily storing and manipulating information. The capacity of working memory is limited. As suggested by Miller [13], working memory can only accommodate seven plus-or-minus two elements. The concepts of working memory and short-term memory, though generally synonymously used, are distinct because short-term memory only refers to the transient storage of information [73].

Bainbridge [74] suggested that working memory is a type of "goal structure." Due to limited capacity, information is evaluated before being imported to working memory, based on their relevance to current tasks, or more specifically, their importance in determining actions toward a goal. Knudsen [70] supported this idea and presented a framework of attention, where top-down knowledge in memory and the bottom-up salient stimuli need to undergo a competitive selection process to gain access to working memory.

Therefore, the contents in working memory are the result of deliberate thinking about an immediate task, including the internal model of human-machine systems, action plans and their evaluation [75, 76, 74].

Information Processing

Information processing refers to the broader mechanism used to encode or register information. It is usually referred to as a knowledge representation process, which involves transformations (analyzing, segmentation, and blending) of the information in working memory into long term knowledge needed for decision making.

Because the environment can present overwhelming information sources beyond human information processing capacity, strategies to simplify the structure of knowledge are required. Moreover, a well-organized structure of knowledge not only accelerates information extraction, but also facilitates manipulation of memory storage and improves recall [72].

One strategy of information processing is to map the problem space into a lower dimension so as to conquer "the curse of dimensionality" [77]. Perceptual invariants and other invariants in the structure of the agent-environment interaction are believed to play an important role in information processing. The information is tailored to the immediate task and represented as domain-specific features. For instance, while driving on the road, drivers have no need to record absolute locations of all traffics but only the distance to the car in front of them [78].

Another strategy is to assimilate the information in working memory according to previously encountered conditions such that the details of the information can be ignored. This strategy relies on the recall paradigm. The more experienced with current conditions, the higher the information processing speed and accuracy that can be achieved.

Memory Organization

Memory organization refers to the ability to store knowledge in memory and recall knowledge from memory when needed. A classic paradigm is the modal model, which proposed that memory can be classified into three levels of storage, including sensory stores, short-term memory, and long-term memory [79, 80, 81]. These three levels have different retention characteristics. Sensory stores enables a registration and retention of the presented stimuli across several hundred milliseconds [79]. Short-term memory has a limited capacity and holds for about 18 seconds without rehearsal [82]. Long-term memory has no known limit in capacity, and the decay is either very slow or non-existent [80]. However, there are ongoing debates about multi-store theory [83, 84].

No matter the storage, human memory is maintained by building connections between elements of knowledge. Horn [71] suggested two types of groupings in memory organization. The first one is aids, or chunking. In aids, the grouping of elements of the same type (e.g. digits, shapes, colors) gives no actual meaning. For instance, a person can use an algebraic aid to memorize a telephone number by coding it into sets of three or four digits, yet these sets have no semantic interpretation.

Chunking has also been supported in the studies of chess players [85]. In these experiments, experts using *perceptual chunks*. Each perceptual chunk consists of a familiar sub-configuration of pieces.

Another grouping type is concepts. Each concept is established as a category that humans use to classify objects or phenomena (e.g. plants, animals, and sports) [71]. The hypothesis is that concepts are organized following a hierarchical structure by clustering elements with the same features as a higher-level element. This structure will improve the recall paradigm by strengthening the connections between elements of similar characteristics, and it will also facilitate knowledge updates. However, because the hierarchical structure enhances the storage by removing the relationships between less relevant elements, the missing details will inevitably lead to some imperfect knowledge.

Internal Model

Internal models are the mental representations of the world to simulate and predict responses. Due to humans' limited computational capacity, internal models cannot exactly describe operating laws of the system (e.g. the dynamics of a vehicle). They only represent each person's understanding of these laws, and individuals' internal models about the same law can be quite different [16].

The primary functionality of internal models is anticipating a future event based on current information. This functionality is often regarded as one of the most crucial cognitive functions enabling effective motor performance. Experts have precise internal models in their domains that allow them to relate information to structured prior knowledge such that they can successfully anticipate future demands [65]. In the Kawato model proposed in [86], a well constructed internal model serves as a feed-forward element used to decrease error in the feedback loop of human motion control.

Humans can develop internal models at multiple complexity levels. This capacity gives humans flexibility to choose the one based on available time duration and performance requirements. In time-relaxed conditions, humans can use a complicated internal model that considers more features and enable a higher prediction accuracy. A model with lower complexity can be applied in time-constrained situations to allow quick predictions.

As to the structure of internal models, Bainbridge [87] pointed out that "the operator's knowledge of process dynamics may not be in the form of control equations, but may be the result of simple correlational learning, which could lead to, or be related to, conditional propositions about general aspects of process behavior." Suzuki [88] supported this idea and claimed that humans do not always interact with machines continuously but may change discontinuously their own controllers corresponding to the subdivided operations in a sequential task.

Decision Making

Decision making refers to the processes used to select an action or a plan among several alternative possibilities available in working memory. Human decision making is limited by available time, available information. Simon claimed that humans can only achieve "bounded rationality" [15].

Recognition is the major and also the most efficient decision making process to handle everyday tasks. Task-specific cues in working memory can activate a recognition process if associated knowledge can be retrieved from memory. This knowledge includes a solution plan that has been successfully implemented before for the situations associated the cues. For instance, in Schmidt's schema theory for motor control [36], a motor program stores an abstract representation for a class of movements. A specific response can be determined with a recall schema based on the initial conditions and the outcome of the movement. Through this process, many problems can be solved in a few seconds without conscious analysis [15].

Exploration and exploitation are two decision making processes when no immediate recognition can be retrieved [89]. Exploration is employed when task environment is unknown or humans attempt to search for new opportunities. During exploration, humans begin by making a random or a planned decision, either of which allows humans to take steps to the unknown and acquire new knowledge. The acquired knowledge can then be used to learn the task environment.

In exploitation, humans will apply heuristic search to determine an "optimal" action plan based on available information. If the task domain is highly structured, a schematic algorithm may be applied to guide the search with task-specific heuristics (e.g. time duration and energy consumption) [15]. If the task domain has no obvious structure, "weak methods" can reduce the complexity of the search and still arrive at satisfactory solutions [15]. Satisficing is one of the weak methods. This method sets an expectation based on relevant experience and halts search immediately when a solution meets the expectation. Another weak method is means-ends analysis. This method attempts to reduce unknowns by setting subgoals and then solves the problem as a sequence of sub-problems that have been encountered before [20].

Exploration and exploitation have their own merit and disadvantages. For instance, exploration can incur a large cost with minimal knowledge gain. Conversely, exploitation is likely to be trapped into suboptimal solutions. Therefore, in situations when recognition is not available, rational decision making needs to maintain an adequate balance between exploration and exploitation [89].

2.2.5 Integration of Human Spatial Control Skills

Fig. 2.2 illustrates an overall structure of human spatial control behavior by integrating all components in the perceptual, cognitive, and motor control processes. Specifically, perceptual system filters the information from sensory functions to achieve adequate awareness of the entire human-machine system including both the human subject, agent, and task environment. These information then feeds into working memory in the cognition system. The contents in the working memory, if no exact match can be recalled from memory, activate the retrieval of internal models in specific task domains so as to generate candidate control profiles. Decision making process then determines an "optimal" or satisfactory plan that drives human both motor control and perception systems.

The fact that motor control, perceptual, and cognitive skills rely on different biophysical elements seems to imply that they are working independently. However, dynamical interactions exist among these skill components, which make the decomposition of a skill difficult, or even impossible. Moreover, these skills can hardly be developed independently. Even though a large proportion of human processes are running unconsciously, at the early stage of development they require the coordination of all of motor control, perceptual and cognitive skills [23]. Impairments in any components of these skills may degrade the skill level of a human individual.

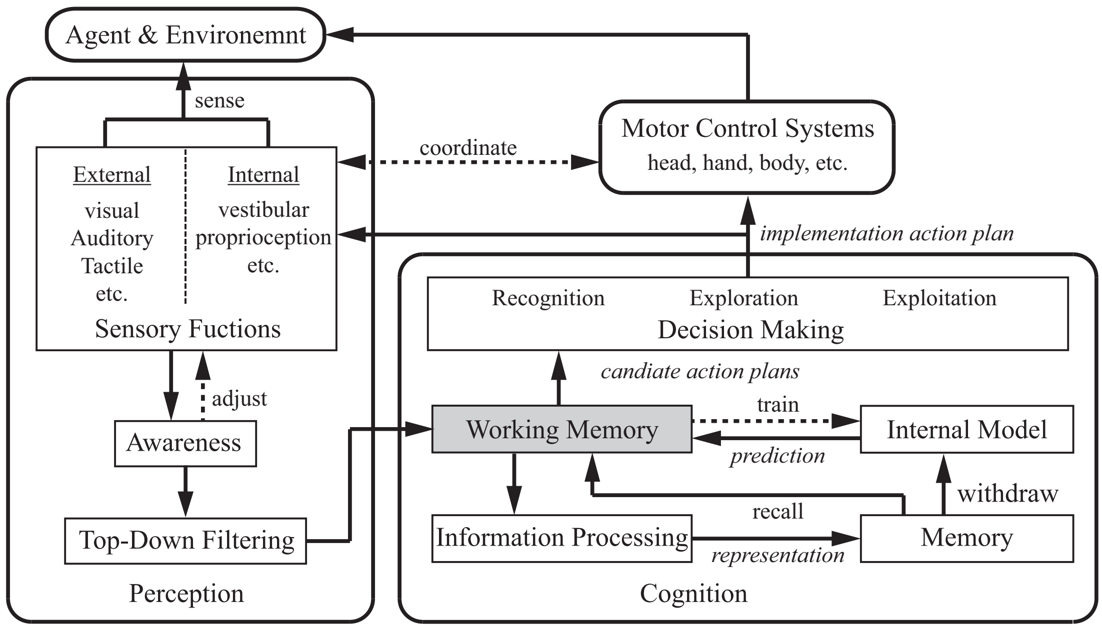


Figure 2.2: Illustration of human spatial control behavior structure. Humans conduct spatial control behavior primarily through three systems: perception, cognition, motor control systems. Perception system prepares useful information obtained with sensory functions for the immediate tasks and feeds to the cognition system. Working memory plays the role of information buffer and decision making generates action plans that drive motor control and perception systems.

It is therefore believed that another architecture of skills exists that can help explain humans' capacity in spatial control behavior organization. The investigation of skill development may provide additional insights into constructing such a skill architecture.

2.3 Skill Development and Emergent Hierarchy

This section presents existing theories on human skill acquisition and development and discusses how these theories lead to an understanding of the emergence of skill hierarchy.

2.3.1 Motor Skill History

Human skill development has been an active topic for more than a century. Early studies focused on the performance of spatial control skills, for instance, precision and reaction time of repetitive tasks [29] and the effect of Knowledge of Results [30]. Researchers at that time primarily focused on speed-skills with short cycle-time and one degree-of-freedom signal-response. There was a surge of skills research in 1940s, due to wartime demands for tracking tasks, such as flying, driving, or aiming [23]. Researchers increasingly analyzed tracking skills, using the theory of servo control and later, modern control. These models treated a human operator as an element in a “closed-loop” system [31, 32]. Ecological psychology was then developed that emphasizes the interactions between humans and environments [33, 34]. Later based in part on Miller’s information channel capacity theory [13], skills research migrated to information processing perspective with the idea that humans have a limited-capacity information channel and an intermittent sensorimotor pattern [23].

2.3.2 Skill Development Theories

The best-known empirical theory in the skill development is the power law of practice. This theory indicates that performance time, when considered as the task criterion in perceptual-motor skills, tends to decrease with practice as a power function [56]. Depicting performance time against amount of practice in a log-log plot exhibits a linear decrease followed by an asymptotic approach to an “incompressible” performance time [40]. A strong evidence is Crossman’s cigar-making example [40]. This example shows the reduction in cycle-time across different subjects with more than 10 million cigar-making machine operations over a 7-year period.

Newell [41, 56] proposed that the power law function is a general law of learning but also pointed out several potential limitations. First, the power law function is typically demonstrated in tasks where task completion time is the dependent variable but lack of the evidence for other motor performance variables such as the ones associated with perception and cognition. Second, the power law cannot accommodate tasks depending on an ordinal scale (e.g. producing relative motions). Third, the power law cannot account for changes occurred on multiple performance dimensions [56].

The power law theory indicates that skill development proceeds as a continuous and linear process. This suggests that the skill development process is cumulative, e.g. the effects of experience can be carried over to aid later performance. However, Shaw and Alley [90] argued that skill development can be discontinuous or nonlinear. In typical spatial control tasks, humans are embedded in a rich perceptual-motor environment with multi-dimension interactions. Therefore, learning is a functional rather than a function. It is difficult to assure that learning takes continuous change across all dimensions [90]. Linear development can hold as a special case.

Moreover, in a task depending on an ordinal scale, the qualitative properties of the coordination mode may change from trial to trial, leading to discontinuous changes in performance measures over practice time [56]. Resolving the continuity-discontinuity controversy requires detailing what evolves as skill development progresses.

McRuer and Krendel [46, 47] and Salmoni [25] all proposed that skill development can be idealized as a progression of three phases:

- *Early phase (compensatory mode)*: In this phase, the operator has no prior knowledge and relies on a trial-and-error strategy. All control depends on external sensory information. The operator constructs the internal model based on the discrepancy between action and its result. Therefore, this phase involves a high level of attention and cognitive processing.

- *Intermediate phase (pursuit mode)*: In this phase, the learned internal model enables the operator to make short time prediction of actions. The operator can therefore implement control actions to track a reference signal with a few seconds anticipation, resulting in more stable performance.

- *Final phase (precognitive mode)*: In this phase, the internal model is well developed and can generate accurate action plan. The operator needs to make few corrections, and therefore the control strategy is shifted from closed-loop to open-loop.

Another well-established skill development model was proposed by Dreyfus [48]. He investigated human skill development in both spatial control tasks (i.e. vehicle driving) and tasks depending primarily on cognitive processes (i.e. chess). He suggested that human skill development follows five stages:

- Stage 1 (Novice): At this stage, the beginner determines the action on the basis of merely domain-independent features.

- Stage 2: (Advanced Beginner): At this stage, humans begin to develop situation- or domain-specific understandings from examples.

- Stage 3: (Competence): With more experiences, humans achieve competence by filtering overwhelming relevant elements and developing rules and reasoning procedures. Humans are also emotionally involved at this stage by taking responsibility for one's successful and unsuccessful choices.

- Stage 4: (Proficiency): At this stage, the performer's skill theory, as represented by rules and principles, will gradually be replaced by situational discriminations. However, deficiency still exist for mapping discriminations to reactions.

- Stage 5: (Expertise): The experts can make more subtle and refined discriminations and see immediately how to achieve the goal.

Both the three-phase model and the five-stage model assume the task and environment to be stable over the entire practice profile. However, this assumption can hardly hold for tasks with rich environment elements and complicated agent dynamics. Therefore, Bainbridge [22] proposed a skill schema theory. This theory suggested that skills are developed in different situations rather than in different phases, and each situation is associated with a specific skill category.

- Perceptual-motor skills are developed in familiar specific situations that are frequently recurring and sufficiently stable. In these situations, a skilled person can react appropriately to the environment without using conscious attention and working memory.

- Familiar cognitive skills are developed in familiar general tasks. In these situations, methods and knowledge are already available but the environment is unfamiliar. A skilled person can accomplish the task without first having to work out a plan.

- Problem-solving skills are developed in unfamiliar tasks. However, people may recognize familiar properties of tasks allowing them to rely on general problem-solving methods. confidence.

One conclusion of these development theories is that spatial control skills are developed through the construction and perfection of internal models. McRuer and Krendel [46, 47] suggested that the construction of internal model is a goal-directed process aimed at minimizing prediction errors and prolonging prediction range. Dreyfus [48] suggested that internal models are enhanced through the identification of domain-specific features, discrimination of situations, and construction and strengthening of situation-response correlations. As Crossman [40] claimed, "the expert's ability seems to lie rather in knowing exactly the right method to use in each situation that arises in the task, than in having superior coordination, acuity or timing."

Another conclusion is that skilled human operators can proceed to an automatic phase. Shiffrin and Schneider [91] investigated the properties of automatic processing in information processing. Automatic processing is fast and non-conscious, and in the automatic phase performers change from continuous to intermittent reliance on feedback [92, 93]. These properties can greatly alleviate human attention workload in this phase.

In brief, skills development progresses through the construction and perfection of internal models and culminates into an automatic processing mode.

2.3.3 Knowledge Representation

The preceding conclusions raises the questions about what is the mechanism enabling automatic processing and in what form the knowledge of internal models is stored.

A fundamental assumption of skill development is that learning is a consequence of acquiring more appropriate representations of actions, in other words, the policies between actions and task demands [56]. Early motor control research also suggested that practice strengthens connections between stimuli and responses. That is, acquired knowledge can be represented as Stimulus-Response (SR) pairs. To elucidate how these stimulus-response pairs are stored and implemented, Keele [94] proposed the concept of motor program. A motor program is defined as a set of structured muscle commands that can be enacted without being influenced by peripheral feedback.

Adams [95] proposed a closed-loop theory of motor learning that extends the motor program concept. In Adams' theory, a motor program consists of two memory states: memory trace and perceptual trace. The memory trace stores initial conditions of the movement. It works like a recall function that activates the motion given an initial condition. The perceptual trace stores the sensory consequences of earlier responses. It works like a recognition function that compares sensory feedback with stored knowledge and determines corrections.

Adams' closed-loop theory has two limitations. First, the theory cannot explain continued skill development without Knowledge of Results. Moreover, in this theory, a motor program is a one-to-one relation between the movement and its outcome. Therefore this theory cannot explain the acquisition of general skills. The motor program theory also has difficulty in explaining the specificity of practice, i.e. skill development can be surprisingly specific [93]. For example, Keetch and colleagues [17] found that basketball players achieved a much higher success rate at foul line than expected by generalizing their success rates at neighboring distances. In other words, the practice at foul line has less benefit to general shooting skill.

As a response to the limitations of Adams' theory, Schmidt [36] proposed a schema theory which emphasizes open-loop control. The schema theory shifts the emphasis from recognition to recall. In this theory, a motor program is generalized as the invariants across a class of movements. It can be described with a set of parameters including initial conditions, response specification, sensory consequence, and outcome. The generalized motor program enables a one-to-many relation. Therefore, it alleviates memory demands and provides necessary principles to accommodate new aspects of movement dynamics [56]. The primary concerns of Schmidt's schema theory are its use of rule-based action representation and the difficulties in mapping dynamics [56].

The knowledge represented by motor programs is essentially a form of declarative knowledge that can be expressed as propositions. Anderson [96, 97] suggested that skill development is realized through a shift from declarative knowledge to procedural knowledge. He proposed an ACT (Adaptive Control of Thought) model that integrates a memory network with a production system. The memory network stores declarative facts such as past memories, current goals, and current stimuli. The production system connects facts to a set of control procedures that can be executed automatically in a given situation.

2.3.4 Elements of Skill Hierarchy

General motor programs in the schema theory, and procedures in the ACT model, exemplify a human strategy of *chunking*, or modularity. These functional elements might be the key organizational principle that the central nervous system employs for reducing both structural and computational complexities [98]. Considering Miller's information capacity limitation [13], humans can only retain seven plus-or-minus two distinct elements in working memory. The chunking strategy can address this limitation by combining elements into larger units.

The chunking strategy is considered as the general learning mechanism of humans [99]. Newell [41] indicated that human skill development based on chunking could model the power law of practice. Moreover, chunking can improve recall performance. Through the experiments with chess players, Chase and Simon [85] suggested that experts can develop perceptual chunks that consist of a familiar sub-configuration of pieces to improve the recall performance. More recently, Fonollosa and colleagues [100] simulated the learning of sequences using a representation in the form of chunking, and the results demonstrated the robustness of chunking in recall.

A specific form of chunking in motor control skill is a muscle synergy. Bernstein [61] proposed the existence of muscle synergies as a neural strategy to address the degree of freedom problem. The degree of freedom problem states that due to the biomechanical redundancy of human musculoskeletal structure, a given movement can be realized in an infinite number of musculoskeletal coordinations. Muscle synergies are defined as motor patterns that activate multiple muscles with one neural command. Bernstein [61] suggested that unskilled performers initially reduce the degrees of freedom by locking joints and operate the remaining joints as a unit through an equation of constraints. With experience, the performer gradually unlock the joints to coordinate an increasing number of degrees of freedom.

The transition from the locking to the unlocking of joints was supported by the limb link analysis of Newell and van Emerick [101] which revealed that naturally right-handed subjects using their nondominant limb locked their distal limb joints compared to the organization displayed by the dominant limb. However, Newell and van Emerick also suggested that coordination is strongly influenced by a variety of constraints (e.g. past experience) because left-handed subjects did not demonstrate this transition.

Tresch and Jarc [102] argued that introducing muscle synergies does not reduce neural complexity because the number of coded muscle combinations could be many more than that of muscles. Rosenbaum [93] suggested that this issue can be addressed by considering functional linkages between different parts of the body. He presented an example showing some coordinations of wrist and elbow are easier to be done than the others. However, synergies are not mandatory linkage, such as blinking while sneezing, but soft constraints that they can still be controlled at will [93]. This property allows synergies to be learned or changed.

Another question in movement skills is which of the many possible degrees of freedom do the central nervous system control [103]. The operational definition of a variable being "controlled" is that the variable is stabilized against perturbations. The stability of the variable can be measured using the variability across trials [102].

Based on the definition of controlled variable, Scholz and Schöner [103] analyzed an experiment on the sit-to-stand transition. They found that the central nervous system preferentially stabilizes those degrees of freedom that are task-relevant [104]. This result in some way stands in contrast to the muscle synergy hypothesis that the central nervous system freezes task-irrelevant degrees of freedom. Tresch and Jarct [102] indicated that the results can be explained from the control efficiency perspective, i.e. attempts to correct task irrelevant variability would be an unnecessary waste of effort. Therefore, they suggested to extend the concept of muscle synergies by considering them as the regulation of task-relevant variables rather than the grouping of muscle activations. This idea is also referred to as the "uncontrolled manifold" [103] or "minimum intervention" hypothesis [105].

2.3.5 Hierarchical skill development

The preceding theories characterized basic units of motor control skills. The following question is whether these units can be integrated into larger skills with extensive practice, or in other words, whether skill development can be represented as hierarchical structure. The hypothesis of hierarchical skill development has the potential to prove the skill generality that at a high level the development of all skills follows the same principle [93].

Hierarchical skill development was first investigated by Bryan and Harter [106] in the late 19th century. They analyzed the practice curves of telegraph operators in receiving (1) letters not making words, (2) letters making words but words not making sentences, and (3) letters making words and words making sentences. The features in the practice curves (i.e. ascends and plateaus) indicates that in the early phase of learning, telegraph operators process letters as a unit, then words, and eventually groups of words or sentences. Mackay and Bowman [107] confirmed this hierarchical skill development hypothesis. In their experiment with bilingual subjects, the recitation practice of a sentence in one language greatly improved the recitation speed of a new sentence in the other language that has the same meaning. The results suggested a high-level unit corresponding to the meaning of the sentence had emerged.

Leplat [52] and Bainbridge [22] also believed that human skills are developed following a hierarchical tree structure by continuously integrating groups of smaller units into larger units when the smaller units become automatic. Therefore, both control and conscious attention can be allocated at the highest level of operation. This hierarchical development hypothesis has been investigated in many activities. Schack and Mechsner [19] analyzed the tennis-playing performance among individuals and determined that proficient players had structurally more elaborate mental representations of the observed serves than beginners. Reiley and colleagues [108, 109] investigated surgical operation skills and constructed a four-level description of human performance, which they call: dexeme, surgeme, task, and procedure level.

Some researchers suggested that hierarchical skill development is not homogeneous across all levels. That is, moving to the next level does not merely involve removal of details of information but more importantly involves a shift in representation structure. For instance, Fischer [110] proposed a hierarchical structure of cognitive development theory in attempt to provide a common framework for skill development analysis. The theory claims that across the development profile, skills are gradually transformed from sensory-motor actions to representations and eventually to abstractions. Similarity, Rasmussen [111] developed a model representing human performance as a hierarchical structure consisting of skill-, rule-, and knowledge-based levels. These levels process different visual information gains from signals, then signs, to symbols, which involve increasingly more abstract meaning.

These theories of hierarchical skill development explains how humans adapt to more complicated tasks with practice following an aggregation process, i.e. iteratively combining smaller elements into groups. However, currently there is no biophysical evidence supporting this type of aggregated structures.

2.4 Skill Assessment

Setting minimum time requirement is not an efficient way for skill certification and cannot ensure qualification. Here are also specific maneuvers and procedures that trainees have to successfully demonstrate. This section presents a brief background of existing skill assessment methods. The section first introduces methods based on human grading, followed by statistical analysis methods and methods emphasizing behavior units or movement primitives. Finally, the section completes with a discussion of constructing a skill assessment framework based on the hierarchical functional model.

2.4.1 Human Grading

Human grading has been applied in a variety of domains, such as education, sports competition, and professional certifications. This approach relies on human raters' observations and its reliability is limited by the subjectivity.

Several issues lead to subjectivity in human grading. First, human raters grade performance of operators based on their experience. Raters may differ in their understanding of task difficulty and their standard of optimal performance. The quality of observation can also bias human grading. Obtaining sufficient information for evaluation requires proper positioning and attention. Recorded video has become a preferred medium since. All observers share the same information and replay can be used to compensate for distraction. The development of data visualization tools can further assist skill evaluation [112].

The design of qualification tasks is limited by the trade-off between simplicity and expressiveness. It is worthwhile to consider Bernstein's degree-of-freedom problem [61] during the design. The degree-of-freedom problem states that, because of the redundancy of human muscle coordination, a motor task has more than one solution under normal conditions. Therefore, evaluation of a specific skill requires carefully constraining the task domain while not reduce subjects' flexibility [113].

Structured human grading was proposed to further help overcome limitations due to subjectivity. In this method, raters are given a checklist of carefully selected performance criteria with guidelines. Hence grading lowers the requirement for rater expertise. More importantly, the multiple metrics in the checklist enables a detailed evaluation. Structured human grading has been widely used in athletic competitions, such as gymnastics and platform diving. Similar checklists have also been devised in the surgery field, e.g. objective structured clinical examinations (OSCE), objective structured assessment of technical skills (OSATS), and global rating system (GRS) [109].

Another way to reduce subjectivity is to increase the number of expert raters. The ratings from experts are more convincing because experts have a better understanding of task difficulty and of the components that constitute optimal performance and merit attention. However, increasing the number of experts is usually impractical due to insufficient number of experts in a specific field and the associated costs. These limitations are a strong motivator for data-driven skill assessment methods.

2.4.2 Statistical Analysis

Traditional metrics still dominate the assessment of spatial control skills, such as procedure time, precision, and economy of hand movement [114, 115, 116]. Statistical analysis is a simple method to obtain these metrics. It assumes that data drawn from a sample is subject to a random variation and then captures the characteristics of variations using mean or variance. Statistical analysis can highlight consistent features while eliminating variations associated with human sensory-motor processes, as well as emotions, fatigue, and motivation, averaging over multiple trials is required [117].

The moving average method extends statistical analysis by capturing more dynamic characteristics using a time window [118]. The window size determines the cutoff frequency, the dynamics below which are discarded. However, the decrease of window size does not guarantee the capturing of transient and dynamical features of spatial control behavior.

2.4.3 Assessment Methods based on Behavior Units

Human behavior and natural languages share many commonalities as both exhibit structure and satisfy a form of "grammar" [119]. The general idea behind the methods in this category is to identify atomic behavior units from motion data. These behavior units can illuminate inherent mechanisms that human operators use to organize spatial control behavior. Therefore, they provide a better measure of human expertise.

The methods of behavior units identification fall primarily into two categories. One proceeds in a top-down manner, using unsupervised segmentation methods, and the other in a bottom-up manner, using supervised clustering methods. Unsupervised segmentation methods assume no prior knowledge of behavior units. They use kinematic (i.e. spatial-temporal) features (e.g. velocity minimal, maximum heading change, and minimum jerk) to determine candidate transition points that segment a sequence [120, 121, **?**]. The resulting subsequences describe behavior units. These units, however, usually exhibit highly diverse characteristics and are less useful for explaining behavior organization.

Supervised clustering methods require models of behavior units. Exact models are either learned from demonstrations [122] or chosen by an expert [123, 124]. Example models include piecewise polynomials [125], Bézier curves [126] or piecewise autoregressive exogenous (PWARX) model [127, 78]. The model used to describe behavior units determines the insights of the method. Simple models such as polynomials can describe agent dynamics but cannot capture the interactions with the environment and task elements. Therefore, to describe interaction patterns, the model structure should be carefully considered. PWARX is used in [128] to formalize interaction patterns in the state space form, considering the agent, environment, and task elements as a closed-loop system [**?**]. With specified models, the sequence of behavior units can then be recognized by using techniques such as dynamic time wrapping (DTW) [129, 130] or hidden Markov model (HMM) [131, 132, 133, 108, 117, 134].

By capturing behavior units, these models enable a variety of performance metrics, such as number of pauses, motions, or operational phases, and consistency of operational phases. These metrics provide detailed descriptions of human behavioral structure and organization.

Chapter 3

Experiments and Gaze Analysis Techniques

This dissertation uses two example applications to illustrate the human spatial control skill framework that will be described in Chapter 4. A remote-control flight task exercises human skills in remote-control of miniature rotorcrafts, and a laparoscopic surgery training task tests surgeons with basic clinical maneuvers. Both experiments demand a level of human skills beyond that is required by everyday tasks to tackle with task complexities such as nonlinear agent dynamics or confined movement region. They distinguish themselves in that the flight task is structured with specified task elements while the surgical training task is unstructured with contingent start and target. These two applications are representative examples of today's human-machine systems. Section 3.1 introduces the lab facilities and experimental procedures of these two experiments. One advantageous feature of our lab facilities is the capturing of gaze movement, and Section 3.2 describes gaze pre-processing techniques to provide sources for the analysis in the following chapters.

3.1 Experiment Facilities and Configurations

This section presents the lab facility, experimental setup, and task configurations of the two experiments: remote-control of miniature rotorcrafts and laparoscopic surgery training. The lab facilities allow us to collect data not only from human subjects' control data and agent dynamics, but also human subjects' gaze locations. Human gaze data provides sources of psycho-physical information for investigating human spatial control skills.

3.1.1 Remote-Control Flight Task

The flight experiments are conducted in the Interactive Guidance and Control Lab (IGCL) at the University of Minnesota. (See [135, 136] for details.) Figure 3.1 provides an overview of the experimental facility [2].

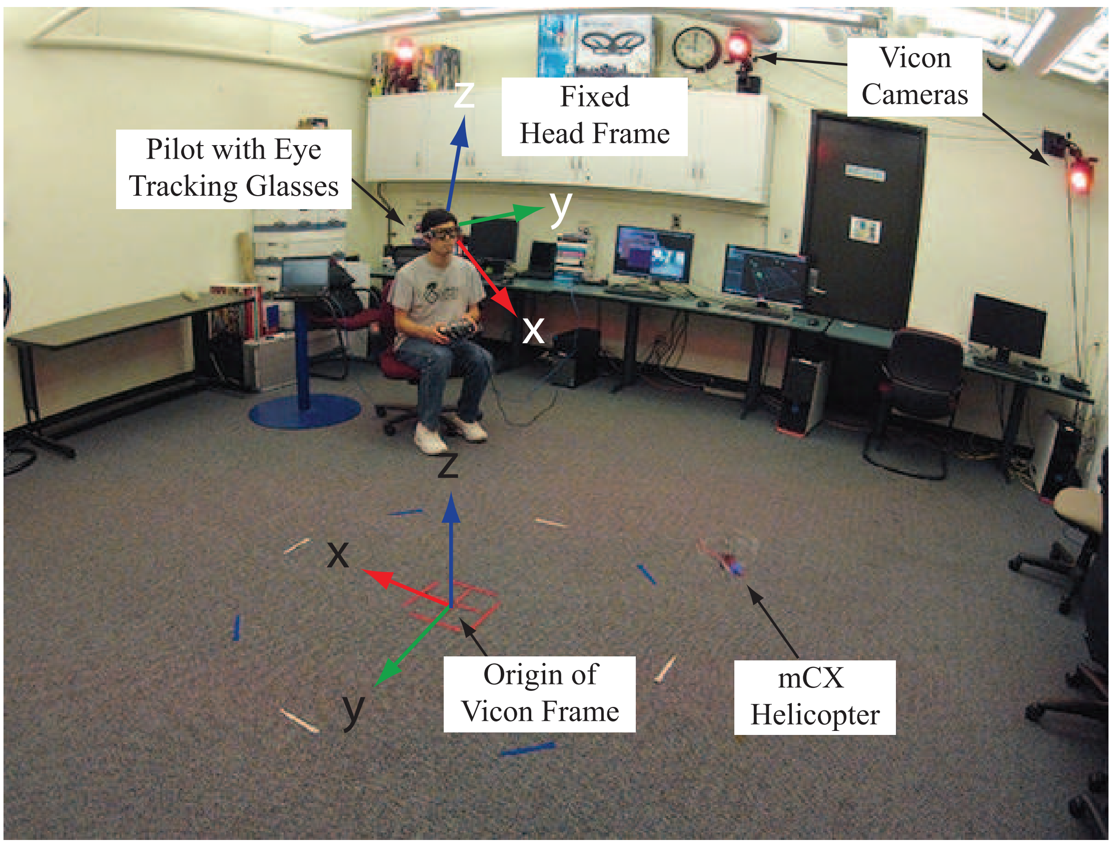
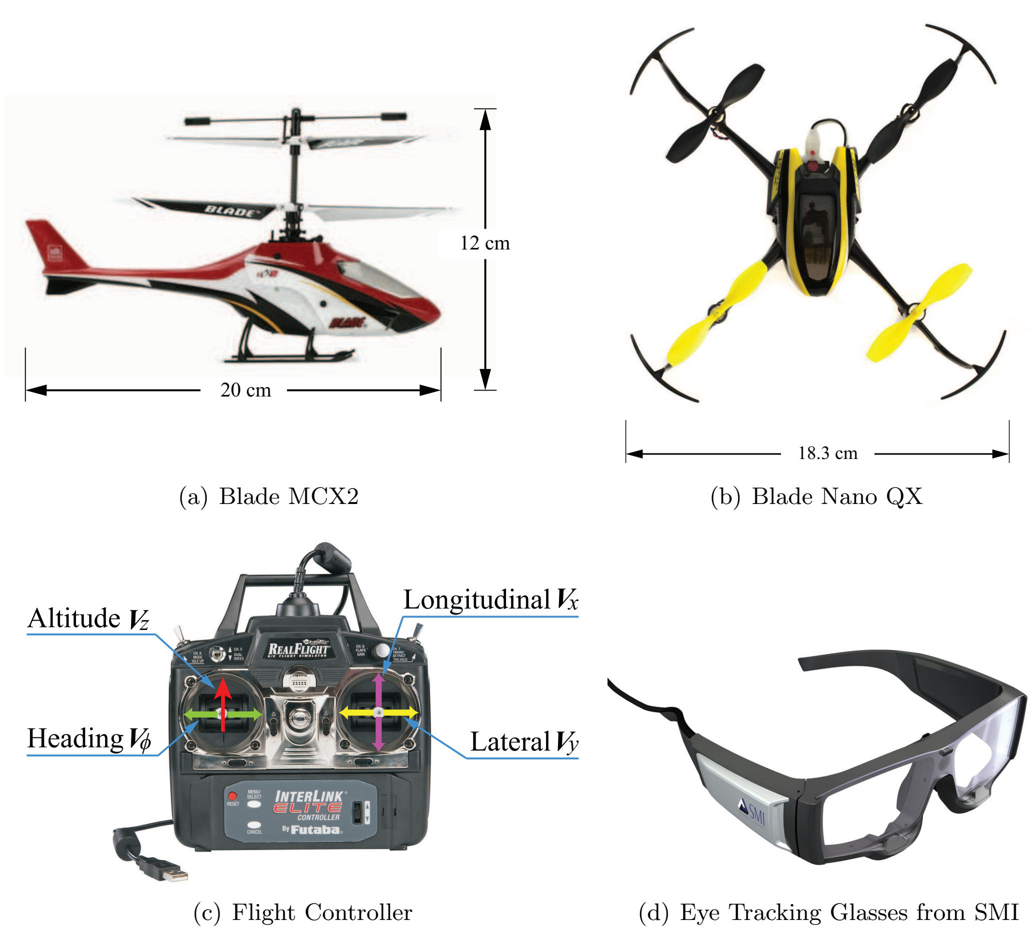


Figure 3.1: Experiment in the lab environment showing rotorcraft, Vicon motion tracking system, and gaze tracking [2].

The experiments use miniature rotorcrafts (e.g. Blade MCX2 in Figure 3.2(a) and Blade Nano QX in Figure 3.2(b)) as test vehicles to exercise human pilots. Their operational capacities and coupled nonlinear dynamics allow humans to exercise their skills across the complete hierarchy including sensory-control, perceptual, and cognitive mechanisms. These vehicles provide a unique platform to collect trajectory data and investigate fundamental questions pertaining to human guidance and control skills, and more generally for the development of analysis and modeling tools.

Human pilots maneuver these rotorcrafts via a remote controller with four primary control inputs: longitudinal and lateral cyclic to the lower rotor (denoted as control v_x and v_y), and collective and differential rotor rpm (denoted as altitude control v_z and heading control v_ϕ). Figure 3.2(c) illustrates these four control inputs.



(a) Blade MCX2

(b) Blade Nano QX

(c) Flight Controller

(d) Eye Tracking Glasses from SMI

Figure 3.2: Experimental facilities.

The lab is equipped with six high-speed Vicon^{TM} tracking cameras to capture real-time pose (including position and orientation) of the configured objects (e.g. rotorcraft motion and pilots' head movements). The captured motion data is sampled at 100Hz.

Pilots' eye movements are recorded using a pair of eye tracking glasses (ETG) from Senso-Motoric Instruments (SMI). (See Figure 3.2(d).) Eye movement is stored as a time history of 2D locations on the image plane with 1280-by-960 resolution and at a frequency of 30 frames per second.

In this dissertation, we collect data from two tasks: a circle task and a guidance task. The data are smoothed using a Butterworth filter and a zero-phase filter with well-tuned cutoff frequency.

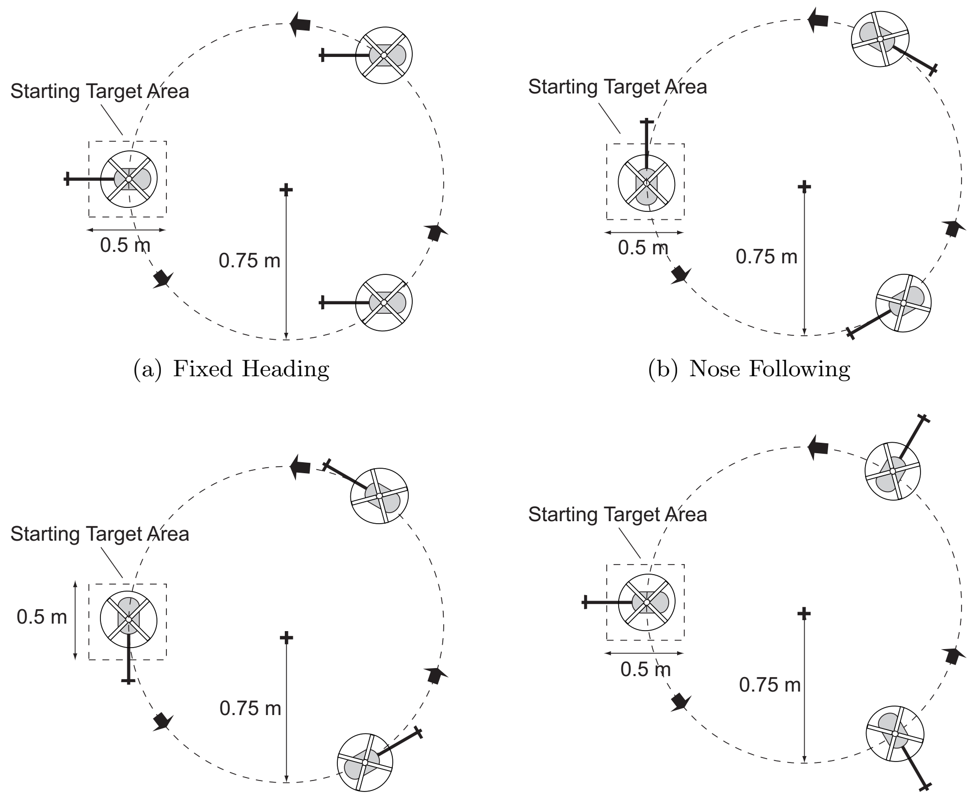
Circle Task

In the circle task, we instruct human pilots to track a reference circle on the ground with maximum precision and minimum operation time, following four heading guidelines in Figure 3.3. The radius of tracking reference is 0.75m, marked on the ground with eight stripe stickers. For each configuration, a pilot needs to complete three trials. Each trial consists of 10 to 15 circles. Successful circles satisfy a radius constraint, $r \in [0.5\text{m}, 1\text{m}]$, and a time constraint, $t \leq 15\text{s}$.

Guidance Task

The guidance task consists of intercepting a target from an array of starting positions scattered across the lab area, while avoiding obstacles at pre-specified locations. Four different task configurations are defined by specifying the magnitude and heading of the target velocity and by introducing extra obstacles, as shown in Figure 3.4 [3]. Human pilots start the helicopter from a stationary hover, above a starting location with an open initial heading. Multiple trials are made for each starting location to capture variability of a human pilot.

For successful trials, the trajectory should have no collision with obstacles and not exceed the maximum travel time of 10 seconds. Successful trials are also subject to terminal constraints defined by the positional tolerance, $W_y = \pm 0.3$ meters, and the course angle tolerance, $W_\Xi = \pm 30$ degrees.



(a) Fixed Heading

(b) Nose Following

(c) Tail Following

(d) Pirouette

Figure 3.3: Circle task configurations. Pilots are instructed to maneuver the micro-helicopter to track the circle following these four heading guidelines. The illustrations assume pilots track the circle in a counter-clockwise direction.

3.1.2 Laparoscopic Surgery Training Task

The surgical dataset in [112] are collected using the electronic data generation and evaluation (EDGE) platform (Simulab Corp. Seattle, WA) in Figure 3.5(a). The dataset are from clinicians of various skill levels at three different clinical teaching hospitals in the United States.

This dissertation uses the dataset from peg transfer tasks (shown in Figure 3.5(b)). In this task the clinicians use Maryland graspers to transfer blocks in the shortest period of time and with minimal drops. The blocks must be picked up by one hand with a laparoscopic tool and then transferred mid-air to the other hand tool. Successful

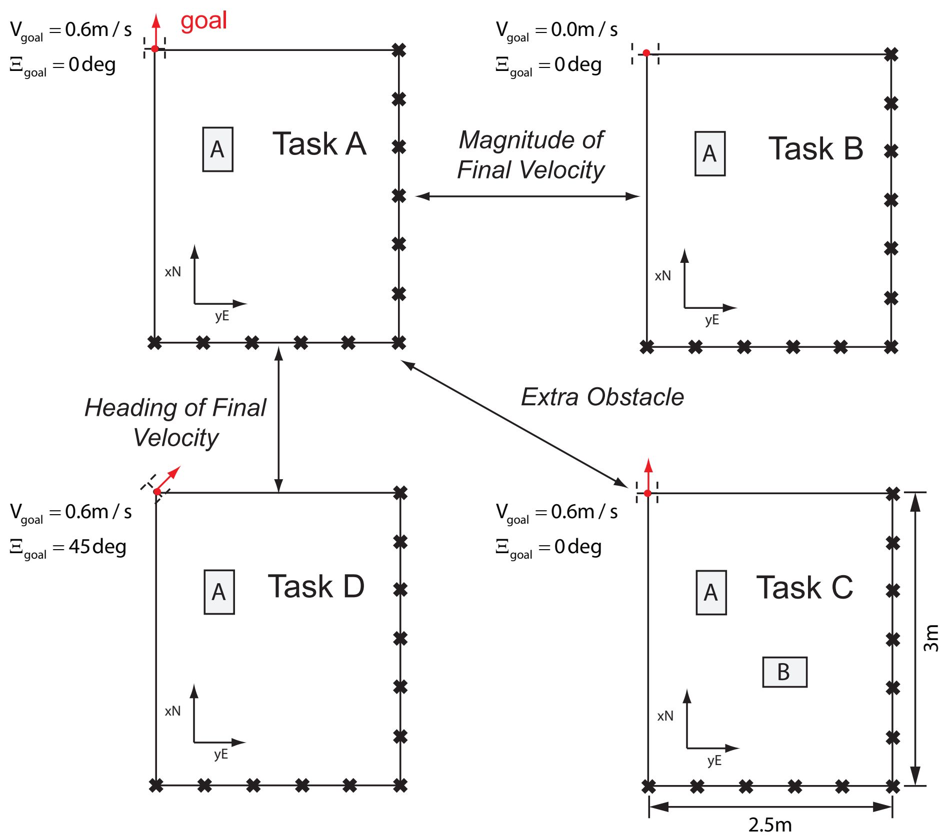
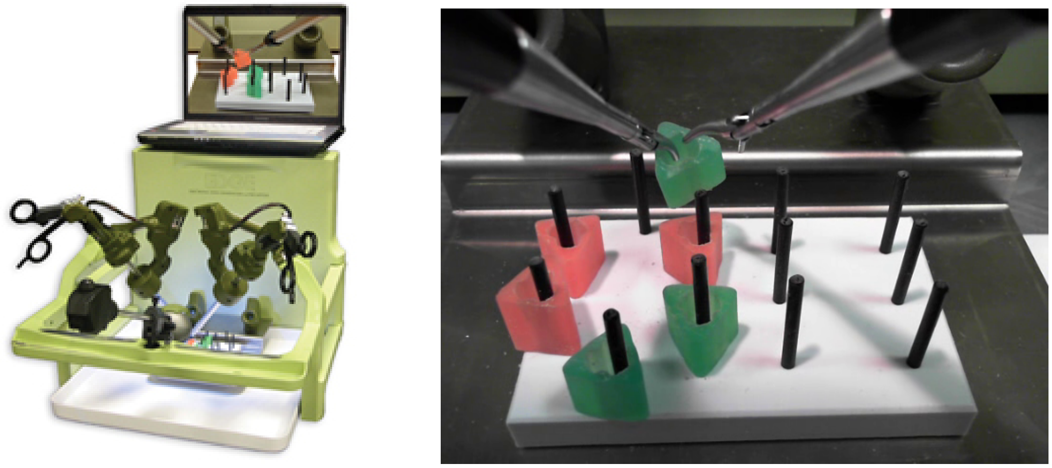


Figure 3.4: Illustration of the four different task settings (A, B, C, and D). Three environmental factors were manipulated: the i) magnitude of the final velocity, ii) direction of the terminal velocity and iii) presence of an extra obstacle [3]. Starting locations are indicated using crosses.

accomplishment of the task requires moving all six blocks from poles on the one side of the board to poles on the other side and then returning the blocks to the initial side.

The motion data includes tool tip positions, orientations, grasp angles, and grasp forces for both hands, with a sampling frequency of 30Hz. A camera is mounted above the working platform and streams real-time images to a laptop. A Tobii EyeX eye tracker is mounted under the screen and records surgeon eye movements at a frequency of 60Hz.



(a) EDGE (b) Peg Transfer Task

Figure 3.5: The EDGE platform (a) and screen shot of the FLS peg transfer task (b) used for this work [4].

Ratings by faculty clinicians, via blinded video review, establish three reliable categories of expertise based on a combination of criteria in Table 3.1. Complete details are available in [137].

Table 3.1: Summary of rating criteria used to establish expertise groups.

Group	**Criteria**
Expert	Practicing laparoscopists (over 100 lapr. procedures): surgeons' and fellows' best FLS-scoring logs with 3/5 or greater average OSATS video review scores.
Intermediate	15th percentile of FLS scores about midpoint FLS score determined between lowest Expert FLS score and highest Novice FLS score $[.59, .73]$.
Novice	All logs below 15^{th}-percentile FLS score.

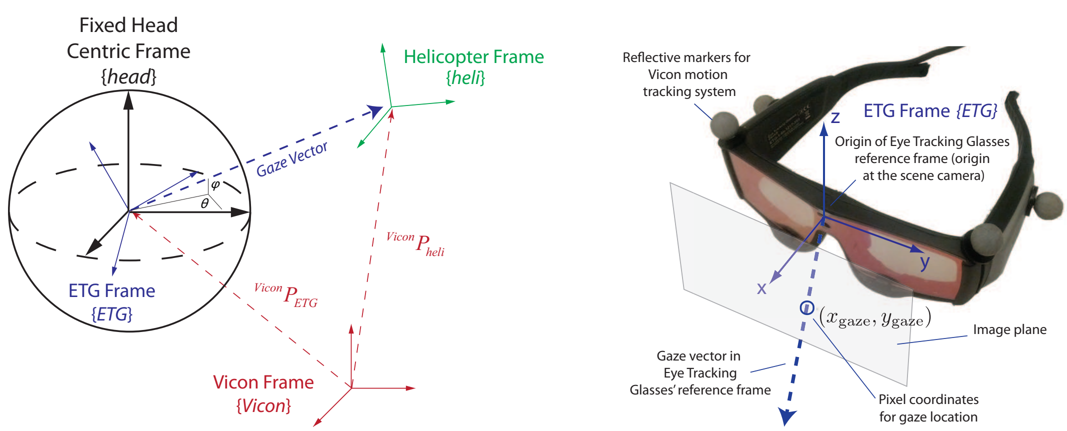
3.2 Gaze Analysis Techniques

Visual perception provides the primary source of information for humans' everyday activities. To achieve visual perception, humans need to actively orient a foveated visual system through the coordination of eye movement and head movement. Gaze movement describes this coordinated motion.

Gaze movements consist of three basic patterns: fixation, smooth pursuit, and saccade. These patterns have distinct kinematic characteristics and functions. Fixations are the intervals during which the gaze is stabilized typically on stationary points [138]. Smooth pursuits are similar to fixations but differ in that the gaze follows moving visual stimuli [139]. Saccades are fast gaze movements used to redirect the eyes to a new location [140].

This section presents methods to register hand-eye coordination and to identify gaze patterns.

3.2.1 Gaze Registration



(a) Spherical reference frame.

(b) Eye Tracking Glasses reference frame.

Figure 3.6: Reference frames for generating 3D gaze location [2].

Eye movements are recorded by the ETG and head movements are captured by the Vicon cameras. It requires a registration process that integrates eye and head movements to obtain gaze positions in an appropriate reference frame. This reference frame should ensure the capturing of gaze movement magnitude and the use by human decision making and motor control systems. Huston and Krapp [141] proposed a spherical head centric coordinate frame to describe the visual receptive field of flies. This dissertation applies it to represent human gaze as azimuth (θ) and elevation (ϕ) angles, as illustrated in Figure 3.6(a).

The gaze registration is achieved with the following transformations:

$$\begin{aligned}\text{Image plane (IP)} &\Rightarrow \text{ETG frame (ETG)}\\ &\Rightarrow \text{Vicon frame (Vicon)}\\ &\Rightarrow \text{Fixed head centric frame (Head)}\end{aligned}$$

IP $\Rightarrow$ ETG

The reference frame defined by the ETG has its origin centered on the scene camera of the ETG as shown in Figure 3.6(b). Gaze location is represented by the pixel location in the scene image (1280x960 resolution). Based on the field of view (FOV) of the scene camera (60° horizontal and 46° vertical), the gaze direction is converted to an angular vector in the ETG frame, $[\theta_{\text{ETG}}, \phi_{\text{ETG}}]$:

$$\theta_{\text{ETG}} = \frac{x_{\text{gaze}} - x_{\text{size}}/2}{x_{\text{size}}}\ \text{FOV}_{\text{hori}} \tag{3.1}$$

$$\phi_{\text{ETG}} = \frac{y_{\text{gaze}} - y_{\text{size}}/2}{y_{\text{size}}}\ \text{FOV}_{\text{vert}} \tag{3.2}$$

ETG $\Rightarrow$ Vicon

Given the head pose (position ${}^{\text{Vicon}}H$ and Euler angles ${}^{\text{Vicon}}A$) captured by the Vicon system, a rotation matrix from the Vicon frame to the ETG Cartesian frame, ${}^{\text{ETG}}_{\text{Vicon}}R$, can be constructed. The unit vector of gaze in the ETG frame, ${}^{\text{ETG}}V = [\cos(\theta_{\text{ETG}})\cos(\phi_{\text{ETG}}), \sin(\theta_{\text{ETG}})\cos(\phi_{\text{ETG}}), \sin(\phi_{\text{ETG}})]$, can be projected to the 3D Vicon frame by

$${}^{\text{Vicon}}V = {}^{\text{ETG}}_{\text{Vicon}}R^T \cdot {}^{\text{ETG}}V \tag{3.3}$$

The depth information d cannot be captured by the ETG. It is recovered by intersecting the gaze vector with the horizontal plane at the helicopter altitude. This is formulated as

$$d = \frac{(O - {}^{\text{Vicon}}H) \cdot n}{{}^{\text{Vicon}}V \cdot n} \tag{3.4}$$

$${}^{\text{Vicon}}P = {}^{\text{Vicon}}H + {}^{\text{Vicon}}V \cdot d \tag{3.5}$$

where O is the helicopter position in the Vicon frame and n is the vertical vector.

Vicon ⇒ Head

The fixed head frame is determined as the average head pose during an experimental run. Gaze registration is accomplished through the mapping from the Vicon frame to the fixed head frame:

$$ {}^{\text{Head}}P = {}^{\text{Head}}_{\text{Vicon}}R \cdot ({}^{\text{Vicon}}P - {}^{\text{Head}}H), \tag{3.6} $$

representing gaze as

$$ \theta_{\text{Head}} = \tan^{-1}(x_{\text{Head}}/y_{\text{Head}}) \tag{3.7} $$

$$ \phi_{\text{Head}} = \sin^{-1}(z_{\text{Head}}/d) \tag{3.8} $$

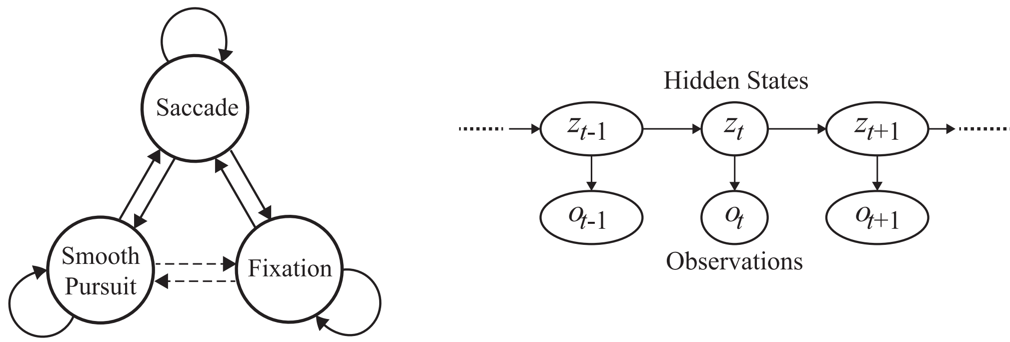
With the fixed head frame serving as the natural frame for humans' perceptive and cognitive processes, the signal of visual stimuli (including the helicopter, task configuration and environment cues) can also be transformed to this frame with Eqs (3.6), (3.7) and (3.8).

3.2.2 Gaze Classification

This dissertation uses Hidden Markov Model (HMM) to classify fixations, smooth pursuits, and saccades. HMM is a statistical technique which includes a doubly stochastic process. A Markov chain governs the characteristic change of the system, or the transition of finite hidden/latent states. The latent state at a time instant, denoted as z_t, is one of the three gaze patterns. Each state of the Markov chain is then associated with an observation function, $o_t = f(z_t)$, which is a stochastic process or a distribution governing the generation of high-dimension observations [142]. The observation functions is described as gaze models in this section. Figure 3.7 illustrates a HMM in both temporal evolution form and state transition from.

Gaze Models

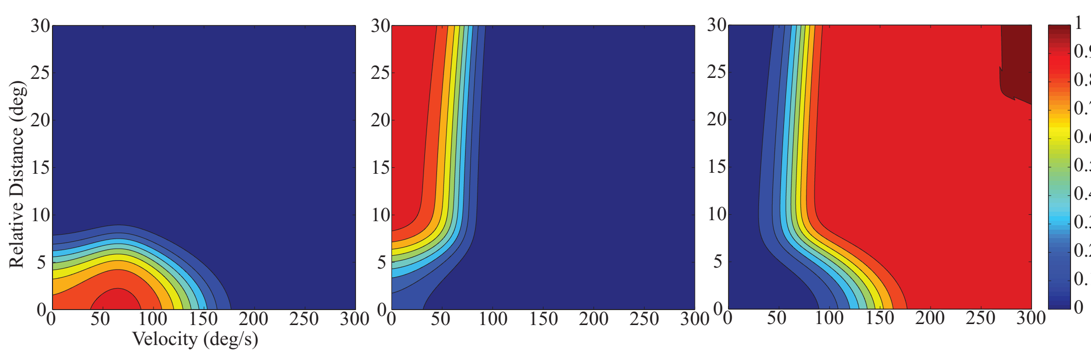
This dissertation construct gaze models using both gaze kinematic characteristics (velocity) and task feature (distance between gaze and agent derived from the registration process) as features. The three gaze patterns can then be described as multi-variate Gaussian distributions with different means and variances.



(a) State transition (b) Temporal evolution

Figure 3.7: Hidden Markov Model (HMM) algorithm. The dash lines in state transition plot indicate no direct transition between fixations and smooth pursuits.

However, the two features of gaze model are positive, such that the distributions are not symmetric and the integration over the positive range is no equal to one when directly using the Gaussian distributions assumption. We solve this issue by applying a saturation range for each feature (specifically, [0, 300] deg/sec for velocity and [0, 30] deg for distance between gaze and helicopter) and standardizing the probability distribution within that range. As an example, Figure 3.8 illustrates the three gaze models in a flight experiment. Gaze models are presented as the probabilities of an observation o_t given patterns k normalized by their sum, that is, $\mathrm{P}(o_t|k)/\sum_k \mathrm{P}(o_t|k)$. This normalization does not influence gaze classification.



(a) Smooth Pursuit (b) Fixation (c) Saccade

Figure 3.8: Gaze models as probability distributions in feature space

In the feature space integrating gaze velocity and distance between gaze and agent, the three gaze patterns exhibit as probability distributions centered at distinct regions. The boundaries of the gaze patterns shown as gradient are the primary regions that the HMM functions to determine gaze patterns and to reduce identification error. The steep gradient indicates that the couplings among the gaze patterns are within a limited range.

Emission, Transition and Initial Probability

The emission probability of each observation $\mathrm{P}(o_t|k)$ can then be computed with respect to the three gaze patterns and be incorporated into an emission probability matrix B. The entry in tth row and kth column represents the probability of the observation at t sample given the pattern k. The three gaze patterns are numbered, 1 for smooth pursuit, 2 for fixation and 3 for saccade.

Another essential component of HMM is the transition probability matrix A. The entry in ith row and jth column represents the probability that the next observation is in cluster j given that the current point is in cluster i, that is, $A_{ij} = \mathrm{P}(z_{t+1} = j|z_t = i)$. It is noted that there is no immediate transition between smooth pursuits and fixations, emphasized with dash lines in Figure 3.7(a). Therefore, both $A_{12} = \mathrm{P}(\text{fixation}|\text{pursuit})$ and $A_{21} = \mathrm{P}(\text{pursuit}|\text{fixation})$ are zero.

Initial probability π is the probabilities of the three gaze patterns being at the beginning of a gaze trajectory.

The final step is to determine the gaze pattern sequence by applying the Viterbi algorithm. The Viterbi algorithm searches both forward and backward for a sequence of maximum likelihood given the observations.

Chapter 4

Formalization of Human Spatial Control Behavior

In this and the next chapter, we formalize and validate a hierarchical functional model as the framework of human spatial control skills. This chapter mainly provides theoretical descriptions of the framework. First, Section 4.1 presents a brief review of motion planning. This provides theoretical foundations to formalize human spatial control behavior in Section 4.2 as a motion planning problem subject to physical, biological, and neurological constraints in human-machine systems. Retrieving the solution to this problem is NP-hard. Section 4.3 then presents a hierarchical functional model that delineates human spatial control skills as the coordination of three functional subsystems: planning, guidance, and tracking/pursuit. This model describes a procedural strategy that decomposes a global human guidance problem into multiple levels of lower-dimension local optimization problems. With this strategy, humans can retrieve satisfactory solutions with humans' limited information processing and computational resources.

4.1 Related Work on Motion Planning

Motion planning is one of the fundamental problems in robotics [143, 144, 145, 146, 147]. The general form of a motion planning problem consists of finding a solution in a configuration space with a start as the initial constraint, a goal as the terminal constraint, and obstacles as the environmental constraints. The configuration space describes the set of all possible poses and actions that can be afforded by the robot. The solution is a continuous sequence of legal states in the configuration space. The solution is usually subject to a variety of other constraints, including the agent's dynamics, sensing range, and obstacle environment. It has been proven that the time to find an optimal solution is exponential to the number of a robot's degrees-of-freedom [148].

Primary approaches to solving this problem can be roughly categorized into four groups: grid-based, geometric, potential-field, and sampling-based. Grid-based approaches discretize the configuration space and apply heuristic search algorithms (e.g. Dijkstra, A*) to connect grid points. Geometric algorithms (e.g. visibility graph [149], Voronoi diagrams [150]) represent the configuration space as a set of nodes including a start, a goal, and vertices of obstacles, and search for the shortest path using a connectivity map of the nodes. Potential-field algorithms are mathematically elegant by treating robots as point-mass vehicles driven by attraction to goals and repulsion from obstacles. These algorithms require choosing adequate potential functions for attraction and repulsion to avoid local minima [151]. Sampling-based algorithms (e.g. rapidly exploring random tree (RRT) [152] and probabilistic roadmaps (PRMs) [144]) are currently the most widely used approach, because they can consider agent dynamics and incorporate agent states into configuration space. (See [153, 154, 155] for detailed reviews.)

Online motion planning is a more challenging problem in which the task is performed in a partially or entirely unknown environment. The planner can only rely on local sensory information to plan the trajectory. Moreover, most online motion planning problems involve uncertainties in sensing and control, which further increase the difficulty of the problems. Existing online motion planning algorithms include Laplacian potential field [156], partially observable Markov decision processes (POMDPs) [157], and receding horizon control [158, 159].

4.2 Human Spatial Control Behavior Formalization

Human spatial control behavior involves more complexities than robotic motion planning. In robotic motion planning, the dynamic model of the agent is generally known, such that the planner can compute an optimal control profile. However, for human spatial control behaviors, according to ecological psychology, engaged humans become an active component in a closed-loop agent-environment system [46, 47, 160]. They are embedded in the system while closely interacting with the vehicle, surrounding environment, and task elements. Humans need to undergo processes including sensing necessary information cues, integrating information for deriving control plans, and implementing control actions that best match current conditions. All these processes need to be accomplished simultaneously, in parallel, and in real time. Moreover, these processes are constrained by a variety of biological and neurophysical limitations.

Human spatial control behavior can be defined as a collection of state trajectories resulting from the closed-loop agent-environment system [161, 51]:

$$\dot{\boldsymbol{x}} = f(\boldsymbol{x}, \boldsymbol{u}). \tag{4.1}$$

The agent states $\boldsymbol{x}$ augment states of both the vehicle and the human operator:

$$\boldsymbol{x} = [\boldsymbol{x}_v; \boldsymbol{x}_h], \boldsymbol{x}_v \in X^V \text{and } \boldsymbol{x}_h \in X^H \tag{4.2}$$

The coupling in the holistic dynamics is intensified in situations when the human operator is embedded in the vehicle, e.g. bicycle riding.

The fundamental process of human spatial control behaviors is the determination of control commands $\boldsymbol{u}(t)$. This process can be formalized as a trajectory optimization problem that minimizes a cost function (4.3) within constraints imposed by the agent dynamics (4.1), environment (4.4), human biophysical system (4.5), and information processing capacity (4.6).

$$\boldsymbol{u} = \mathrm{argmin}_{\boldsymbol{u}} \boldsymbol{J}(l(\boldsymbol{u}), t_{\mathrm{f}}, w) \tag{4.3}$$

s.t.:

$$G(\boldsymbol{x}(t), \boldsymbol{e}(t)) \leq 0, \quad \forall t \in [0, t_{\mathrm{f}}] \tag{4.4}$$

$$F_1(\boldsymbol{x}(t), \boldsymbol{u}(t)) \leq 0, \quad \forall t \in [0, t_{\mathrm{f}}] \tag{4.5}$$

$$F_2(\boldsymbol{x}(t), \boldsymbol{i}(t)) \leq 0, \quad \forall t \in [0, t_{\mathrm{f}}] \tag{4.6}$$

The objective function (4.3) considers performance indexes such as path length $l(\boldsymbol{u})$, time duration t_{f}, and workload w. Workload w is imposed by control intensity and information intensity:

$$w(t) = w(\boldsymbol{u}(t), \boldsymbol{i}(t)). \tag{4.7}$$

Control intensity can be measured as cumulative jerks or variation change in control commands. Economic control can reduce the possibility of making mistakes and save energy for information processing. Information intensity reflects the task complexity in information processing. It is influenced by the channel size, precision, and time available to capture required information.

Figure 4.1 shows the result of an advanced human pilot performing a guidance task with the Blade Nano QX. This result provides evidence for the influence of workload on trajectory planning. The pilot initially chose the path in between the two obstacles but after two trials changed to the path circumventing the obstacles. The first path, though shorter in length and time duration, requires a high volume of measurement feedback and fast and precise responses to avoid collision.

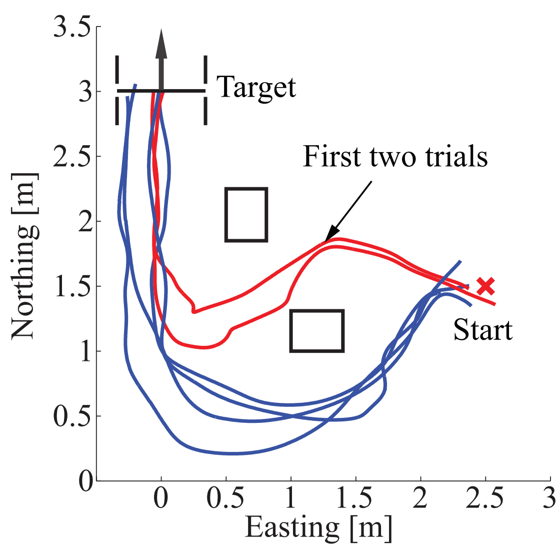


Figure 4.1: Influence of workload in a human guidance task.

4.3 Hierarchical Functional Model

The complexity of human spatial control behavior demands that humans rely on a structure that can harmoniously integrate lower-level sensorimotor functions with high-level cognitive processes. Based on the idea of behavior invariants (e.g. motor program and muscle synergy) and skill hierarchy presented in Chapter 2, Mettler and colleagues [161, 51] proposed the concept of interaction patterns as primary units that human central nervous system uses for behavior organization. The authors suggested that interaction patterns capture the behavior invariants associated with a domain of performance. Interaction patterns also contribute to the construction of a hierarchical functional model of human spatial control behaviors [5].

This section first provides an overview of the hierarchical functional model. Next the section describes the concept of interaction patterns. We then present the description of each subsystem in the hierarchical model, followed by a discussion of perceptual function across the hierarchy. Finally, the section summarizes the benefits of humans employing the hierarchical functional model.

4.3.1 Overview

Figure 4.2 illustrates the structure of the hierarchical functional model. It consists of three subsystems: planning, guidance, and tracking/pursuit. Each subsystem is responsible for a different function. The planning subsystem at the top transforms a task into a sequence of subtasks. These subtasks are characterized by interaction patterns and linked by subgoals. The guidance subsystem refines interaction patterns into a fully specified control profile by using local information. The local information includes the gap between the current states and the immediate subgoal and environment structure. Finally the tracking and pursuit subsystem modulates and regulates motor control processes in order to implement control profiles specified by the guidance subsystem. These three subsystems are coupled and operate as a system. Moreover, perceptual functions participate at all levels of the hierarchical framework, where they provide different types of sensory information for each subsystem.

4.3.2 Interaction Patterns as Behavior Units

The concept of interaction patterns (IPs) is derived from the concept of muscle synergy as a type of neural coordinative structure that activates multiple muscles with one neural command. IPs extend the muscle synergy concept by emphasizing human interactions with the agent and environment. As suggested in McRuer's crossover model [162], the interactions result in the integration of a human subject and agent as a holistic system.

IPs represent a type of behavioral synergy that encapsulates invariant coordination of human motor control, perceptual, and cognitive processes involved in the human-agent system. From the control perspective, IPs can be described as control profiles, $D_j = \boldsymbol{u}(0 : t_j^*) = D(\bar{s}, \alpha)$, where $\bar{s}$ is the segment to employ the IP and α are the associated parameters. The parameters include muscle type, action (e.g. translation, rotation), intensity (e.g. joint angles, force strength), and timing. They specify the coordination of human muscles and joints.

IPs allow spatial control behaviors to be characterized as searching within a lower-dimension motor problem space. The development of a library or repertoire of IPs further enables humans to perform more complicated tasks. At the same time, humans acquire IPs with their independent hands-on experience. They do not compute movements of muscles or joints, or models of agents or tasks due to the high degrees of freedom associated with their musculoskeletal system and limited information processing and computational capacity. Therefore, the library of IPs, $\mathscr{D} \equiv \{D_j\}_{j=1:m}$, can also exhibit diversity across human pilots.

Most importantly, IPs result in a hierarchical representation of spatial control behavior. In this hierarchical representation, planning and motor modulations are connected through IPs, but they can be performed largely independently. This functional property of IPs enables humans to achieve versatility and adaptability regardless of changing task configurations or evolving machine systems.

4.3.3 Planning Function

The planning function is responsible for reducing a global planning task into a sequence of subtasks. Subtasks are connected by subgoals, and each subtask is accomplished using an interaction pattern. Humans arrange subtasks in order to minimize a cost function with respect to metrics such as travel time, control energy, and/or path length.

Subgoals, $\{g_k\}$, are defined spatial points in the specific environment that satisfies the equality environmental constraints:

$$G(g_k, \boldsymbol{e}) = 0 \tag{4.8}$$

For instance, given a rectangular obstacle as shown in Figure 4.3, the environment variable becomes a vector including the length and width of the obstacle and the size of safe margin. That is, $e = [a, b, \delta]$. Then the environmental constraint can be defined as:

$$G(p, \boldsymbol{e}) = \left(|x| - \frac{a}{2} - \delta\right)^2 + \left(|y| - \frac{b}{2} - \delta\right)^2. \tag{4.9}$$

Four subgoal candidates satisfy the above constraint, but the determination of the subgoal for actual implementation relies on global path planning.

The identification of subgoal sequence involves both environmental information and knowledge of behavior patterns. First, the set of candidate subgoals $\mathscr{G}_0 \equiv \{g_k; G(g_k, \boldsymbol{e}) = 0, k = 1 : n\}$ are identified as environmental critical points (i.e. start, target, and obstacle corners). A subset $\mathscr{G} \subseteq \mathscr{G}_0$ of size n' will be chosen to construct the sequence, with x_0 and x_f as the start and end,

$$s = x_0 g^1 \cdots g^{k'} \cdots g^{n'} x_\mathrm{f}, \quad g^{k'} \in \mathscr{G}. \tag{4.10}$$

The segments of the trajectory are encoded as:

$$\begin{aligned} s^0 &= x_0 g^1 \\ s^{k'} &= g^{k'} g^{k'+1} \\ s^{n'} &= g^{n'} x_\mathrm{f} \end{aligned}$$

The cost of the sequence is a collective sum of objective function values for all segments. The value for a single segment $s^{k'}$ is approximated using the default information of IPs:

$$\min \boldsymbol{J}_{s^{k'}}(\boldsymbol{u}) = \min_{D_j} \boldsymbol{J}(l_{s^{k'}}(D_j), t_j^*, w(D_j, \hat{\boldsymbol{\imath}})) \tag{4.11}$$

where $D_j = D_j(s^{k'}, \hat{\alpha})$. $\hat{\alpha}$ are the default parameters stored in memory, and $\hat{\boldsymbol{\imath}}$ is an approximation of information intensity corresponding to the interaction pattern.

The central optimization problem at the planning level is finding a sequence of IPs connecting at subgoals

$$\boldsymbol{u} = D^1 \cdots D^{k'} \cdots D^{n'} \tag{4.12}$$

which minimizes the objective function:

$$\min_s \boldsymbol{J}(\boldsymbol{u}) = \min_s \left\{ \sum_{s^{k'}} \left[\min_{D_j} \boldsymbol{J}(l_{s^{k'}}(D_j), t_j^*, w(D_j, \hat{\boldsymbol{\imath}})) \right] \right\} \tag{4.13}$$

By using IPs as units of behavior instead of dealing with activations of multiple muscles, planning becomes a low-dimensional search problem within a small-size library. Moreover, the planning function only uses the default information of IPs in calculating path cost and does not have to compute specifications of IPs. These characteristics dramatically reduce the planning complexity. At the same time, path segments become environmentally unconstrained, reducing the task complexity in the following subsystems.

4.3.4 Guidance Function

The guidance level is responsible for fully specifying an interaction pattern D_j to close the gap s^k from g_{k-1} to g_k. The type of the IP for the segment s^k has been preselected when estimating the path cost at the planning level. Now, with detailed local information associated with the immediate subgoal, humans can refine the IP with complete specifications, so as to adapt to changes in task elements or reject effects of disturbance.

The central optimization problem for guidance is to adjust the associated parameters of the determined IP (or change the type of IP when needed), to minimize the objective function as follows:

$$\min \boldsymbol{J}_{s^{k'}}(\boldsymbol{u}) = \min_{j,\alpha} \boldsymbol{J}(l_{s^{k'}}(D_j(s^{k'}, \alpha)), t_j^*, w(D_j(s^{k'}, \alpha), \hat{\boldsymbol{\imath}}) \tag{4.14}$$

Skilled operators possess a library of stable IPs. These IPs have already been optimized with extensive practice through refining parameters and reducing variability. Therefore, IPs are ready to be implemented consistently and accurately with few computation. More importantly, IPs represent feasible or legal sets of parameters in the solution space. That is, their implementation is within humans' physical, information processing, and computational constraints, as well the constraints associated with agent dynamics and task environment. They dramatically reduce the attention workload. These properties enable IPs as efficient and at the same time robust solutions to the unconstrained gap closure problem.

4.3.5 Tracking/Pursuit Function

The tracking and pursuit function is responsible for modulating and regulating motor control processes. It is also responsible for rejecting disturbances and mitigating effects of uncertainties. McRuer [46] has defined tracking as "manual processes for minimizing visually perceived errors by exercising essentially continuous control so as to match visually presented input and output signal." He has also pointed out that tracking skill develops from high-bandwidth closed-loop to open-loop control. The emergence of IPs can explain this development process.

When IPs are not available, human operators need to continuously modulate control to reduce implementation error in system states (e.g. position and speed). In this mode, the central objective can be formalized as:

$$\min \boldsymbol{J}_t(\boldsymbol{u}) = \min_{\boldsymbol{u}} \left| \boldsymbol{x}(t) - \boldsymbol{x}_{s^k,D_j}(t) \right| \tag{4.15}$$

where $\boldsymbol{x}_{s^k,D_j}(t)$ are the projected states given the current segment and the desired IP. In real-world tasks, the projected state trajectory can be captured by averaging multiple repeated trials.

IPs emerge as human operators improve their understanding of system dynamics and their ability to predict outcomes of control responses [46]. In this mode, IPs serve as feedforward control profiles with high tracking precision, and humans employ IPs as the reference to regulate inputs, with no need for tracking error feedback. The optimization problem at this level is minimizing the discrepancy between parameters of implemented and optimal IP:

$$\min \boldsymbol{J}_t(\boldsymbol{u}) = \min \left| \alpha(\boldsymbol{u}) - \alpha(D_j) \right|. \tag{4.16}$$

4.3.6 Perceptual Function Encompassing the Hierarchy

Visual perception provides the primary source of information and participates throughout the spatial behavior hierarchy. The hierarchical functional model in Figure 4.2 details the roles of gaze in each subsystem. During planning, visual perception provides the general layout of the task elements and environment. During guidance, visual perception identifies the immediate motion gap associated with the currently active subgoal. During tracking and pursuit, gaze movements update the dynamical state information necessary for motor control modulation [2].

4.3.7 Benefits

The hierarchical functional model delineates human spatial control skills as a sequence of local optimization problems within each functional subsystem. Fig. 4.4 illustrates how the model and the interaction pattern concept help reconcile the complexity of spatial control behavior within humans' bounded rationality.

Human spatial control behavior is analogous to an optimization problem searching within a solution space. This problem is extremely difficult considering coupled system dynamics and a variety of constraints. Humans alleviate the difficulty by decomposing the problem to be solved at three levels. During planning, humans employ IPs as units for global path planning. IPs reduce the trajectory dimension and simplify the global optimization problem as a set of subtasks. Moreover, subgoal allocation removes environmental constraints from these subtasks and narrows the search space. During guidance, because IPs represent feasible/legal solutions of subtasks within human-agent dynamics and neurophysical constraints, the use of IPs further narrows the solution space and facilitates the search for a satisfactory solution that approaches the optimal one. During tracking/pursuit, humans only have to follow the derived solution and reduce implementation error.

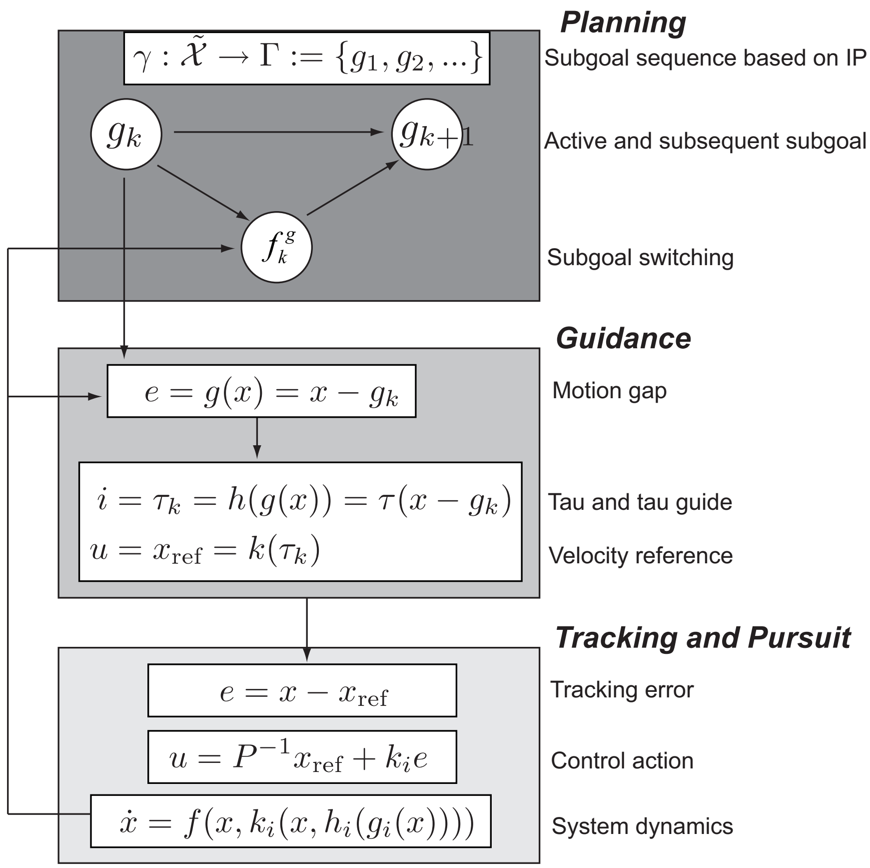


Figure 4.2: Hierarchical functional model of human spatial control behaviors [5]. Spatial control behavior is achieved through interactions among three subsystems. The top subsystem performs the planning based on the deconstruction of the task and environment, in terms of interaction patterns (IPs). The plan is codified as a subgoal sequence. The guidance function closes the motion gap between the current state and the reference defined by the currently active subgoal. The third subsystem implements desired motions for the agent.

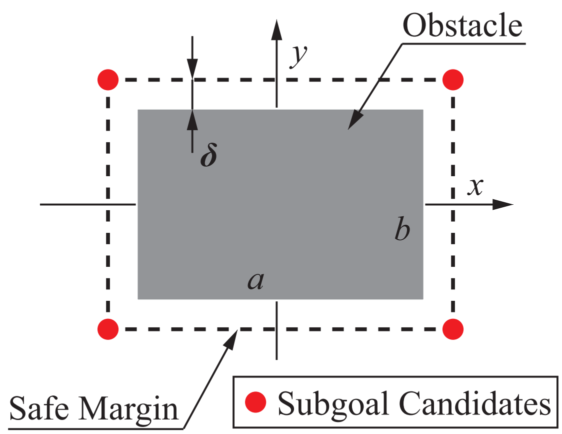


Figure 4.3: Illustration of subgoal candidates.

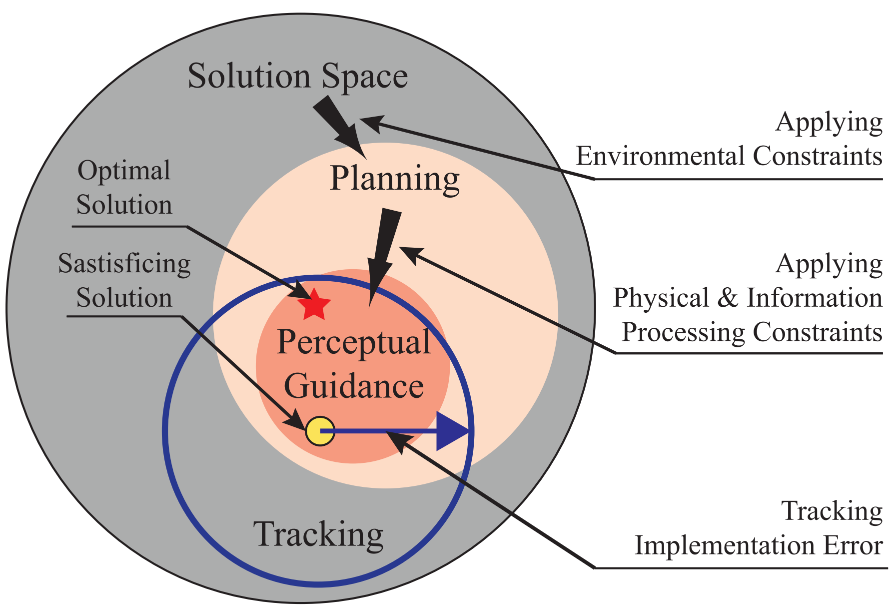


Figure 4.4: Description of the hierarchical functional model as a search strategy in the solution space [6]. The hierarchical framework reconciles behavior complexity with human limitations. The planning function considers environmental constraints in subgoal sequence generation. The guidance function exploits interaction patterns to consider physical and information processing constraints. These two processes narrow the feasible solution space and facilitate the search for a satisfactory solution that approaches an optimal solution.

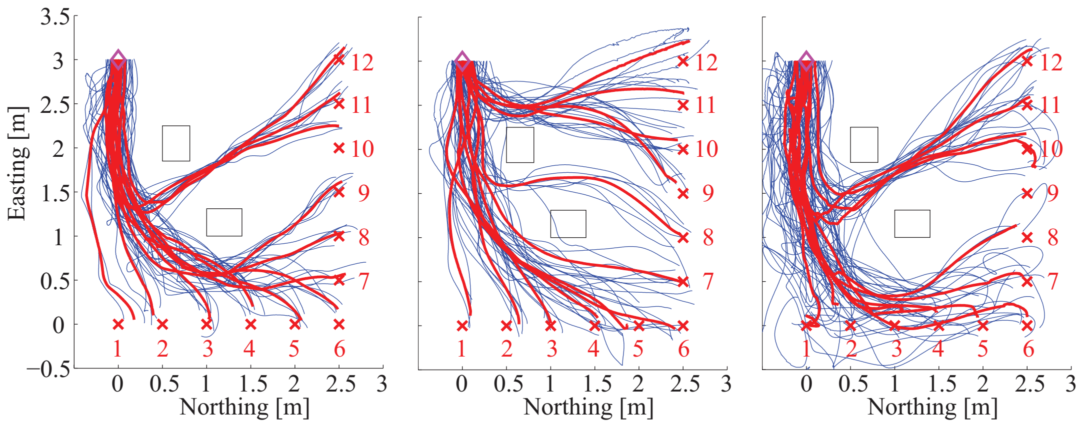
Chapter 5

Human-Inspired Simulation Model

This chapter constructs a human-inspired simulation model in the remote-control flight configuration as a validation of the skill framework described in Chapter 4. The central guideline follows the theory of bounded rationality [15] that in complicated real-time tasks, humans resort to simple strategies to approach optimal solution. Therefore, the simulation model needs to be succinct, which means that the model involves a minimal number of parameters and builds on a minimal set of hypotheses on the functional subsystems of the hierarchical model. By reproducing human results, this model provides a so-called normative framework for interpreting human spatial control behavior. Section 5.1 describes human data in a remote-control guidance task and a technique to capture representative trajectories from a trajectory manifold. The next three sections construct the human-inspired simulation model corresponding to each functional subsystem by introducing a set of hypotheses. Finally, Section 5.5 presents the simulation results and the understanding on skill development enabled by the simulation model.

5.1 Human Experiment Data

This chapter uses the data collected from the remote-control flight experiment presented in Section 3.1. The dataset includes three pilots performing a guidance task in configuration C (with two obstacles) using Blade MCX2. Figure 5.1 shows successful trajectory manifolds of the three pilots.



(a) Pilot 1 (b) Pilot 2 (c) Pilot 3

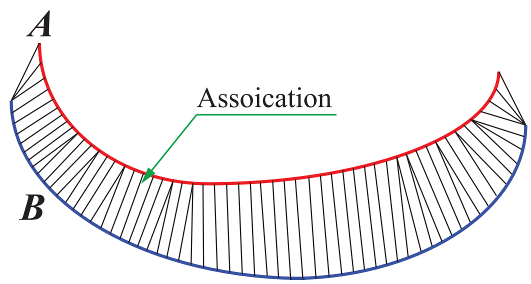
Figure 5.1: Trajectory manifolds of the three human pilots. Red lines indicate representative trajectories obtained with Dynamic Time Warping Barycenter Averaging.

High-level analysis of human spatial control behavior requires excluding control variability so as to highlight underlying behavior organization mechanism. This chapter implements a dynamic time warping barycenter averaging (DBA) method to capture representative trajectories for each run (from each starting location) given multiple trials. DBA method is proposed in [163] for averaging a set of sequences based on the Dynamic Time Warping (DTW) algorithm.

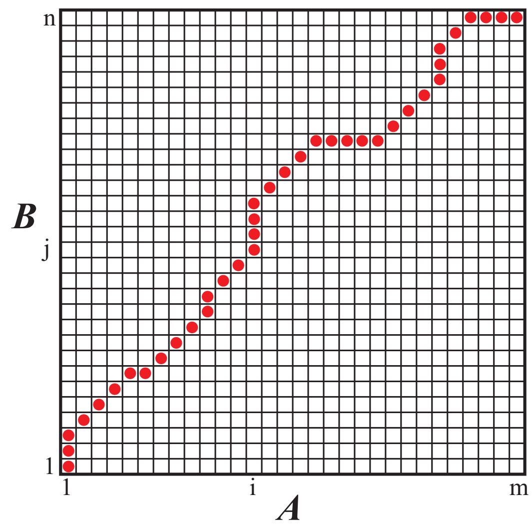
DTW is an algorithm for measuring the similarity or distance between two time series, which is originally developed for speech recognition [164] and have been applied to areas such as human motion recognition [129, 130]. The idea of DTW is to find an optimal alignment between two time series that may differ in time length or speed. Let $A =< a_1, a_2, ..., a_m >$ and $B =< b_1, b_2, ..., b_n >$ be the two time series, and a_i and b_j are the respective coordinates of the two time series. The similarity $D(A_m, B_n)$ between A and B can be obtained by recursively computing

$$D(A_i, B_j) = \delta(a_i, b_j) + \min \begin{Bmatrix} D(A_{i-1}, B_{j-1}) \\ D(A_{i-1}, B_j) \\ D(A_i, B_{j-1}) \end{Bmatrix} \tag{5.1}$$

where $\delta(a_i, b_j)$ is the distance between a_i and b_j. The optimal alignment associates points in two time series, as shown in Figure 5.2(a). The associations can be described with an association table as illustrated in Fig. 5.2(b). They forms a complete path from (a_1, b_1) to (a_m, b_n).



(a) Time series alignment



(b) Example of association table

Figure 5.2: Illustration of two time series aligned with Dynamic Time Warping.

DBA proceeds by iteratively refining the representative sequence. The aim is to minimize the sum of squared DTW distances from the representative sequence to all time series in the set [163]. DBA works in the following steps:

1. Choose an initial representative sequence (e.g. pick the time series in the set that is of the median time length).

2. Associate the representative sequence with each of the time series in the set.

3. Update each coordinate of the representative sequence by averaging all associated coordinates.

4. Iterate step 2 and 3 for refinement.

The red lines in Figure 5.1 show the representative trajectories obtained using DBA.

5.2 Planning Function

The planning problem formalized in (4.13) is solved through determination of a subgoal sequence and using approximate information of interaction patterns to estimate path cost. De Camp [165] indicated that decrease in path length is a characteristic of learning and can be used as a measure of optimality. Therefore, this simulation model uses path length as the cost.

If there is no dynamic constraints, the planning problem becomes a minimum-path-length trajectory optimization problem that connects subgoals with straight lines. The planning function can then be modeled using visibility graph algorithm (VGA).

VGA is designed for point-mass vehicles to find the shortest path. It is based on the knowledge that the shortest path grazes polygonal obstacles at their vertices, and builds a roadmap connecting pairs of visible vertices. The solution is complete and optimal, with a time complexity of $O(N^2)$ [166]. The algorithm is codified following the pseudo-code described in [167], and applied to the task configuration.

Figure 5.4(a) shows the result of directly implementing VGA. The discrepancy between the VGA result and the human data in Figure 5.1 indicates that the point-mass assumption is insufficient and humans consider system dynamics during path planning.

Humans cannot construct the internal model that exactly replicates vehicle dynamics but can develop a systematic mapping to approach the true dynamics for path cost estimation [48]. Therefore, we can construct the mapping of path cost by adding most relevant factors.

Turning angle at subgoals can be an important factor in considering the dynamics. Human pilots tend to avoid making sharp turns with an agile helicopter because it requires high control gain and precision. This may also be one reason that the pilot dismissed the path of the first two trials in the example of Figure 4.1.

The simulation model penalizes for turning angles at subgoals to obtain human estimation of path length. Figure 5.3 illustrates the penalization. A candidate path that smoothly intercepts two subgoals with course angle change is the curvilinear path. Let the length of the straight path to be l. The increased proportion, Δl, by using the curvilinear path l' can be derived:

$$\frac{l'}{l} = \frac{R \cdot \theta}{2R\sin(\theta/2)} \tag{5.2}$$

$$\Rightarrow \quad \Delta l = l' - l = l \cdot \left(\frac{\theta}{2\sin(\theta/2)} - 1 \right). \tag{5.3}$$

where θ is the turning angle.

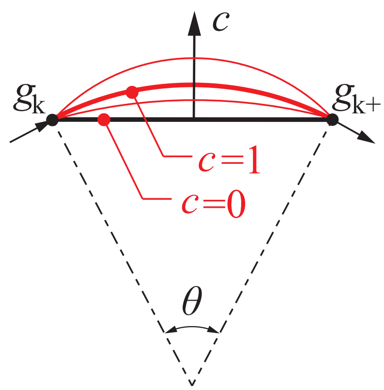


Figure 5.3: Penalization of path length for turning angles.

The penalized path length, l^* can be defined by placing a penalty weight c on the increased proportion:

$$l^* = l + c \cdot \Delta l = l \cdot \left(\frac{c \cdot \theta}{2\sin(\theta/2)} - c + 1 \right). \tag{5.4}$$

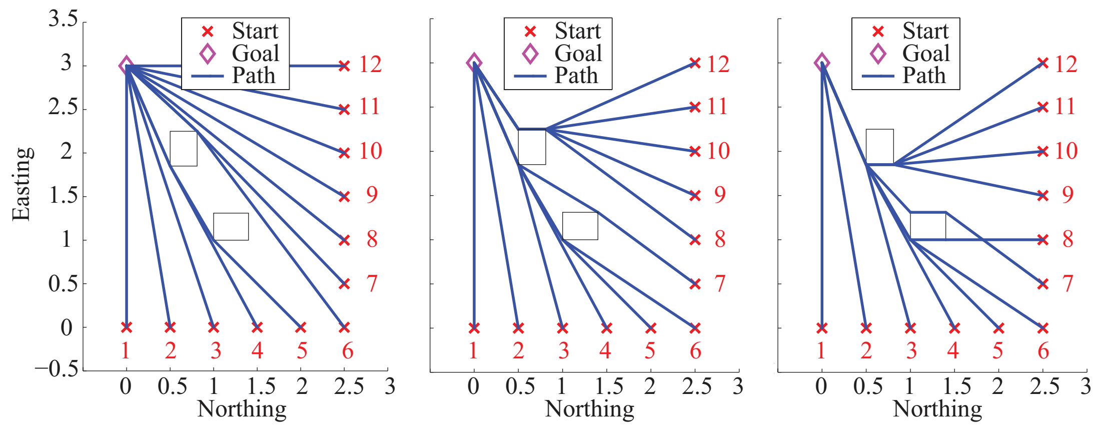
This equation exaggerates the penalty for large turning angles.

Furthermore, terminal constraints require more precise implementation because they are used to qualify successful trials. Therefore, this factor is considered by multiplying c at the target.

The increase of the penalty value c leads to discrete changes in subgoal sequences. Figures 5.4(b) and 5.4(c) show the results considering the turning angle penalty, with $c \in [4.7, 6.2)$ and $c \in [10.2, \infty)$ respectively.

5.3 Guidance Function

Given the subgoal sequence, the guidance function is responsible for implementing interaction patterns to close the gap between consecutive subgoals. This section defines two



(a) No penalty $c = 0$ (b) $c \in [4.7, 6.2)$ (c) $c \in [10.2, \infty)$

Figure 5.4: Visibility graph with turning penalty

types of interaction patterns.

5.3.1 Point Pursuit Behavior

Point pursuit behavior is the behavior to close the gap to a specified target. This type of gap closure behavior can be described using tau guidance theory [168]. Tau of a gap is the time it would take to close the gap x at current closing rate $\dot{x}$:

$$\tau = \frac{x}{\dot{x}}. \tag{5.5}$$

The attractive proposition of tau guidance theory is that the tau of a motion variable can be sensed directly by human pilots without measuring motion gap size and closing velocity. This ensures human visual perception system to be robust and efficient [169, 170].

If distance is the only feature of a motion gap, point pursuit behavior takes the form of a straight path towards the subgoal. In practice, the gap may involves multiple features (e.g. position, velocity, and orientation) to be closed simultaneously. These motions need to be tau-coupled:

$$\tau_y = k\tau_x. \tag{5.6}$$

τ_x is the intrinsic motion. Its form determines the motion types for point tracking behavior. For instance, $\tau_x = t - T$ indicates constant velocity guidance, and $\tau_x = \frac{1}{2}(t-T)$

indicates constant deceleration guidance. (See [170] for more details.)

5.3.2 Boundary Tracking Behavior

Boundary tracking behavior maintains the motion relative to a boundary, that is, $\tau = 0$, in an attempt to minimize or prevent an excursion. This boundary can be a physical barrier (e.g. an obstacle), a vehicle limit (e.g. velocity, turning rate), or the operator's physiological constraints (e.g. attention workload) [171].

Boundary tracking behavior takes two forms in the simulation assuming humans operate the vehicle in constant velocity. The first form is maintaining a constant distance to an obstacle, where the tau to obstacle is zero. The second form is holding maximum turning rate, where the tau to maximum turning rate is zero.

5.4 Tracking/Pursuit Function

The tracking/pursuit function implements the chosen interaction pattern (point pursuit or boundary tracking behavior) in real time.

During the implementation, human pilots can sense the time to contact obstacles with saccade eye movement [172]. If the vehicle approaches too close to obstacles, it will trigger a transient behavior called boundary avoidance behavior. The timing is govern by the tau to obstacle [170].

In this simulation, we assume the helicopter is flying in constant velocity, thus turning rate is the only control input. When the boundary avoidance behavior is activated, human pilots will increase the turning rate to deviate from obstacles.

The implementation of tracking/pursuit function also involves errors induced by imperfect actuators of human pilots. This error can reflect human expertise and can be captured using DBA described in section 5.1. The simulation assumes no implementation error.

5.5 Results and Discussions

5.5.1 Simulation Results

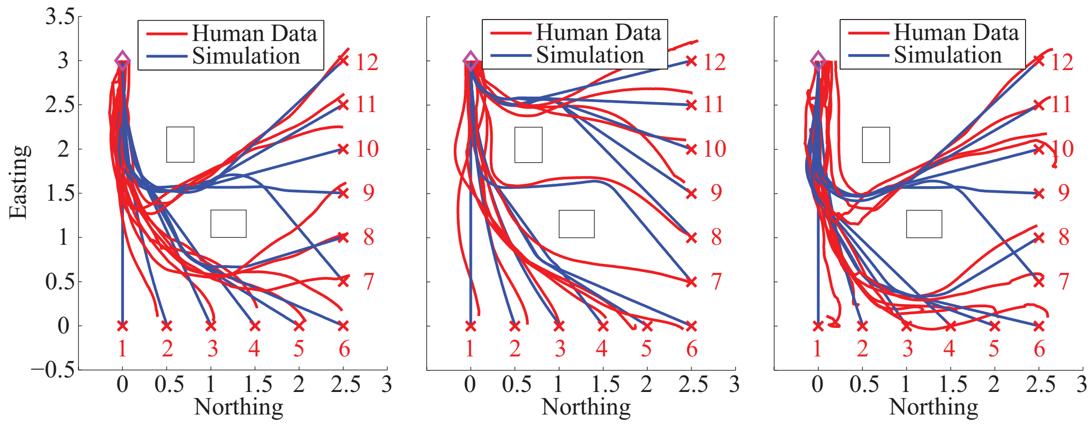
The above-described hierarchical simulation model is applied to the guidance task configuration. This simulation model only involves two parameters: penalty on turning angle for planning and tau to obstacle governing boundary avoidance. The simulation is run by setting these two parameters.

For each pilot, the parameters can be identified by comparing the difference between the VGA results and human data. Specifically, the identification of the penalty value c involves manually discretizing human representative trajectories into subgoal sequences and counting the total number of mismatched subgoals between VGA results (in Fig. 5.3) and human data. c is within the range of $[4.7, 6.2)$ for the expert and novice, and $[10.2, \infty)$ for the intermediate pilot.

The discrepancy at starting points 7, 8, and 9 indicates human pilots may consider additional criteria in path planning. For instance, pilots may resort to strategies that reduce task complexity or attention load, such as minimizing the number of decision points, sticking on an exploited path instead of exploring a new one, or utilizing a consistent turning direction (clockwise or counter-clockwise). Including these factors can increase the similarity between simulation and human data but will also increase the model complexity.

Similarly, the identification of the tau to obstacle is through running simulation by setting tau values from 0 to 4 with 0.1 increment and choosing the one with the minimum total distance between human representative trajectories and the simulation result. The tau values are identified to be 1 sec, 1.2 sec and 2.3 sec respectively for the three pilots.

Figure 5.5 shows the simulation results in blue lines, along with red lines showing the representative human trial data. The trajectories derived from the simulation resemble with human trials by only tuning two parameters. This result strongly support the hierarchical functional model as a normative framework for interpreting human spatial behavior.



(a) Pilot 1 (b) Pilot 2 (c) Pilot 3

Figure 5.5: Trajectories of hierarchical framework simulation

5.5.2 Insights for Skill Development

The human-inspired simulation model can be extended to describe skill development at all three functional subsystems.

Planning Function

The result in Figure 5.4 shows that the subgoal sequence is sensitive to planning criteria. Planning skill can be developed by identifying more principal criteria and their influence on path cost. This can be formulated as a weighting system that places different bias on candidate criteria such as path length (or time), turning angles, terminal constraints, and exploration risk.

The development of planning skill also relies on increased knowledge of lower level capabilities, because the planning function needs to ensure the plan be successfully realized. Possible strategies include detaching subgoals from obstacle corners with a margin to reduce possibilities of triggering boundary avoidance behavior, or enriching subgoal information by augmenting position with velocity vector.

Guidance Function

The guidance skill accounts for the mastering of interaction patterns. The development process of IPs can have two directions.

One direction of development is to have more subtle and refined situational discrimination and map situations to reactions [48]. For instance, if subgoals are augmented with velocity vector, point pursuit behavior can be a set of refined control profiles corresponding to different course angle changes. This process requires a large size of memory to record the corresponding parameters. However, it can significantly shorten the time and improve the accuracy when adapting to practical situations.

The other direction is to generalize interaction patterns in a coherent structure. Our earlier study [51] showed that experts can abstract interaction patterns with a time-to-go (TTG) function. By propagating the learned TTG function back from a subgoal, a TTG map can be obtained that contains sufficient information to conduct quasi-optimal control. The generalization eliminates mode transition processes which may require much more attention workload than parameter adaptation within a mode. This can be analogous to the problem of searching in a series of books as opposed to searching through sections within a book.

Tracking/Pursuit Function

Motor control skill is the capacity of implementing specified interaction patterns and attending to contingencies.

The implementation error is inevitably due to humans' imperfect sensing and actuators. At the initial phase of skill development, the tracking/pursuit function is achieved with a compensatory control strategy. However, humans can decrease this error by proceeding to a precognitive mode when they have complete knowledge of the inputs' future [46]. In this mode, skilled operators can directly implement an optimized IP with no need for instantaneous visual feedback.

On the other hand, the primary contingency in the guidance task is collisions with obstacles. Padfield [173, 170] conducted terrain-following flight experiment and showed that the tau to obstacle governs the timing of boundary avoidance behavior. More importantly, the timing is almost constant regardless of the initial forward speed and height. As to the flight experiment setup in this paper, the shortest path always grazes obstacles at edges or corners. In pursuit of high performance, skilled pilots have the tendency to improve their confidence of approaching obstacles. This phenomenon can be observed in the simulation results that the tau to obstacle threshold decreases with respect to human expertise.

Chapter 6

Skill Development and Interaction Pattern Emergence

Human skill development generally conforms to the "power law of practice", indicating that subjects show rapid performance improvement in the early phase, followed by decreasing improvements with further practice. Therefore, the development of professional human skills is usually time-consuming and demands strategies to accelerate the process. This chapter investigates the development of human spatial control skills through a remote-control flight task. The research primarily focuses on validating the emergence of interaction patterns during skill development. Section 6.1 describes experimental data of human pilots performing a circle task in the remote-control flight experiment. Section 6.2 presents theoretical description of skill development progress. The chapter ends in Section 6.3 by presenting experiment results of human pilots and their implications of skill development.

6.1 Human Experiment Data

This chapter uses the data collected from four test pilots performing the circle task in the remote-control flight experiment using Blade MCX2. The pilots are different in their years of experience and holding of a pilot license, as listed in Table 6.1. Due to the difficulty of tracking the development profile of a single pilot, this chapter uses these four pilots to represent different stages in the development process. As an example,

Figure 6.1 presents the helicopter trajectories under the maneuver of the four pilots in the nose following configuration.

Table 6.1: Test pilots

Pilot	Pilot License	Years of Experience
1	Yes	2
2	No	1.5
3	Yes	0.5
4	No	0.2

The circle tasks in this dissertation provide a realistic setup for studying human spatial control skills development. First, repetitions enable stabilizing human performance. Early studies have also investigated repeated tasks, but they primarily focused on transient motions within 3 seconds. The control patterns in these motions, as Schmidt [36] has indicated in the schema theory, are essentially open-loop and ballistic, or using knowledge of results (KR) to correct following trials. The circle tasks here involve longer repetitive sessions and agile system dynamics (e.g. fast response and high sensitivity to control commands), demanding real-time modulations.

Moreover, accomplishing the circle tasks requires coordinations of multiple control commands, compared with single-degree-of-freedom tasks in early studies. The optimal control profiles of these circle tasks can be easily derived, as listed in Table 6.2. They ensure all human pilots perform the planning function at the same level and emphasize behavior pattern emergence.

Table 6.2: Optimal control profiles of circle tasks

Task Setup	Optimal Control Profile
Fixed Heading	(collective) sinusoid x and y control
Nose Following	constant x and heading control
Tail Following	constant x and heading control
Pirouette	constant y and heading control

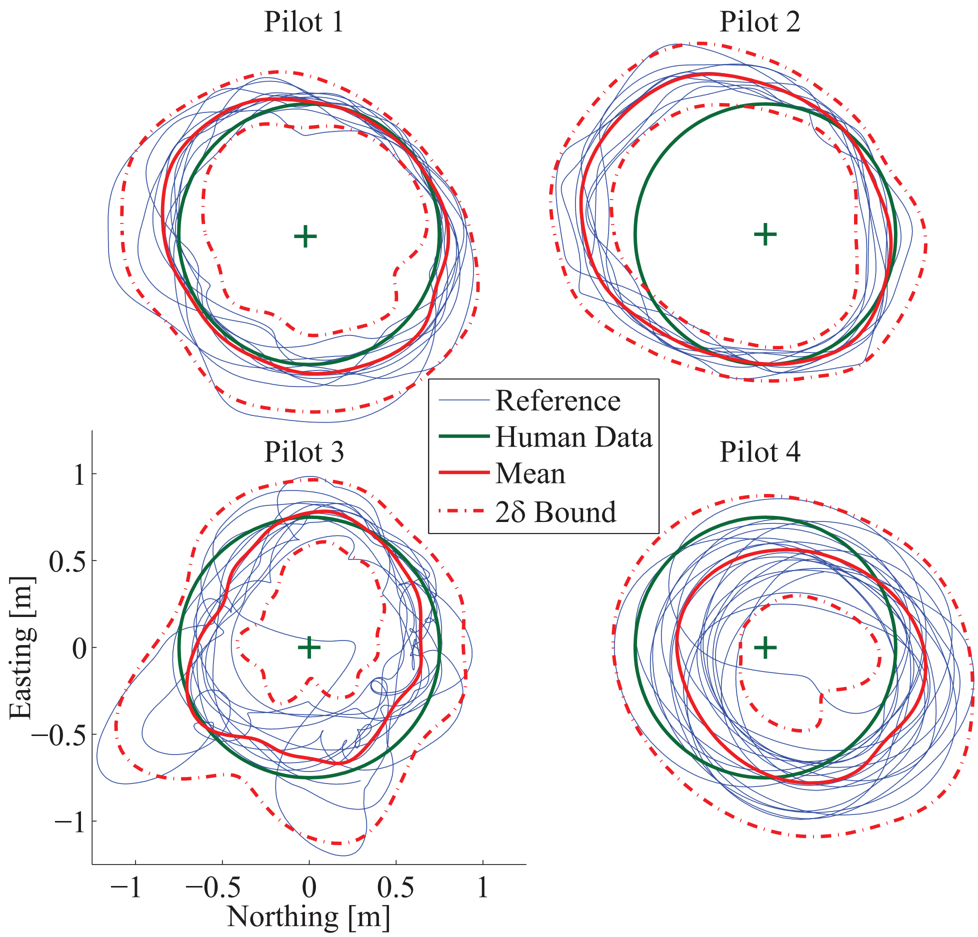


Figure 6.1: Circle task trajectories in nose following configuration for the four pilots. Red solid lines represents the mean circles obtained using exponential moving averaging.

6.2 Skill Development Progress

The circle task can be formalized from the computational perspective as actuating remote controller to maneuver the helicopter following the tracking reference. This research focuses on the engineering perspective on how human operators address this problem through the coordination of skill components. Figure 6.2 delineates the coordination in the circle task. However, this progress model can also be generalized to other spatial control tasks.

The skill development primarily involve four phases, as shown in Fig. 6.3.

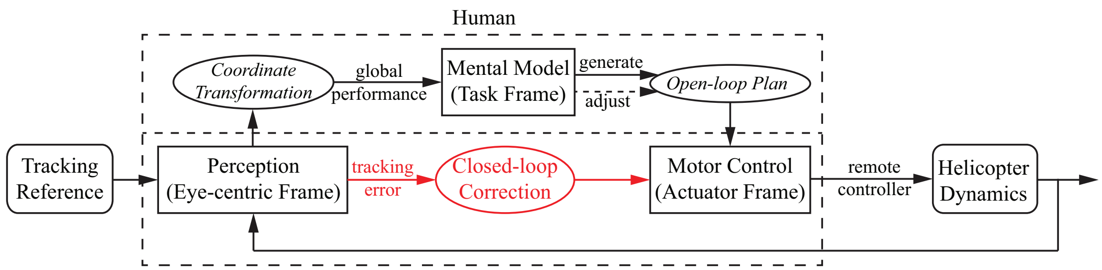


Figure 6.2: The coordination and evolution of spatial control skill components. The black components describe the coordination at the early phase of skill development.

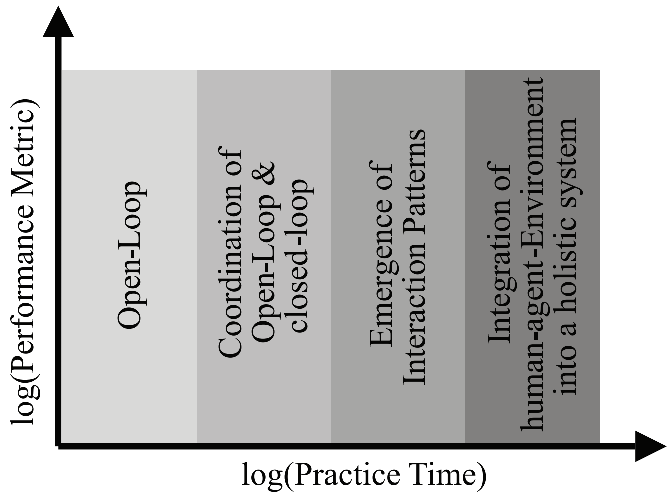


Figure 6.3: The four phases of skill development.

At the early phase of skill development, the human vision system captures the information of tracking reference and feeds to the mental model after coordinate transformation. At this time, human operators can only generate an open-loop plan by assuming the vehicle in control to be a point-mass system. This open-loop plan is then transferred to the actuator frame associated with remote controller inputs. The vehicles usually involves complicated dynamics. For instance, the helicopter in the experiment exhibits diverse dynamics in the longitudinal and lateral directions. Through practice, humans can adjust the open-loop plan by observing the discrepancy between their global performance and the task reference.

Due to limited computational resources and biophysical constraints, human operators are unable to construct exact mental model for complex machine systems. This results in the tracking error between human expectation and actual performance. To reduce tracking error, human operators initiate closed-loop corrections. However, early corrections, usually resulting in tracking error increase and/or overshoot, have low success rate. Continuous practice enables human operators to improve their closed-loop correction mechanisms.

The coordination of a well-tuned open-loop plan and closed-loop correction mechanisms incur the emergence of interaction patterns (IPs). IPs represent a type of sensorimotor synergy who characteristics are stable across task trials. Moreover, as described in Chapter 4, each IP can be activated with a neural command, which reduces human computational and behavior organization complexity.

At the final phase, as indicated in crossover model theory [162], human operators can integrate human adaptive motor control and agent dynamics into a holistic system that can be represented using a single mental model. This integrated mental model generalizes human interactions with the agent, environment, and task elements, which enables human versatility encompassing different configurations.

6.3 Experiment Results and Analysis

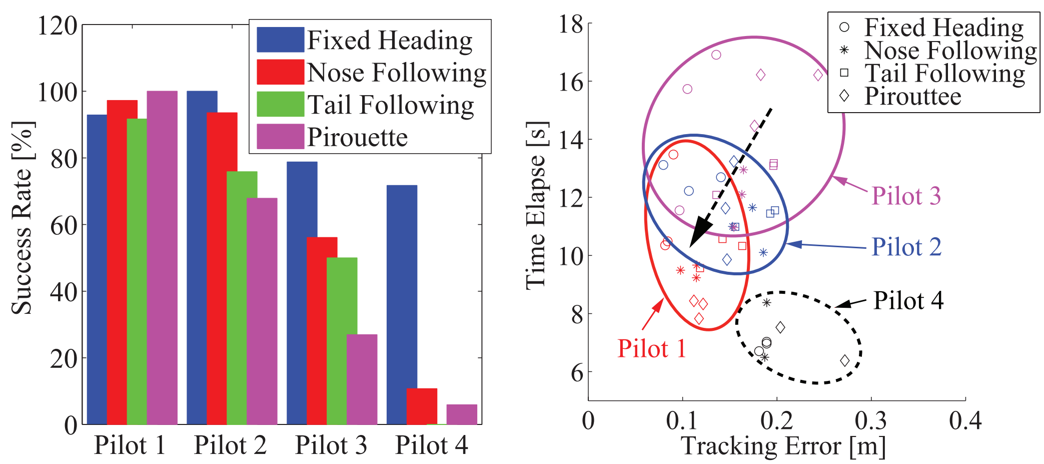
This section analyzes human experiment results to provide evidence for the theory of skill development progress. The section first discusses the overall performance, followed by discussions of multi-input coordination and interaction pattern emergence.

6.3.1 Overall Performance

The overall performance is evaluated using three metrics: success rate, tracking error, and time elapse per circle.

Success rate is the ratio between the numbers of successful and total circles. Successful circles satisfy a radius constraint, $r \in [0.5\text{m}, 1\text{m}]$, and a time constraint, $t \leq 15\text{s}$. Fig 6.4(a) shows the success rates of the four pilots, which are highly correlated with their skill levels. Pilot 1 completed all four heading setups with over 90% success rates, while pilot 4 failed more than 90% of circle tracking in setups except fixed heading.

Tracking error and time elapse are calculated based on successful circles. Tracking error is defined as the standard deviation of the differences between tracking radius and the reference (0.75m). Time elapse is the averaged time per successful circle. Figure 6.4(b) shows a plot of tracking error versus time elapse per circle of each trial for the four pilots. The trend from pilot 3 to 1 implies a continuous and linear trend of skill development reflected in the reduction of operation time and behavioral variance. However, the dramatic difference in pilot 4's performance suggests a transition takes place at the early phase. This transition will be discussed below.



(a) Success Rate

(b) Tracking Err. vs. Time Elapse

Figure 6.4: Evaluation result of pilots. Note that the metrics are based on successful circles. Pilot 4 failed one trial of nose following, one trial of pirouette, and all trials of tail following. The short travel time that pilot 4 achieved was by sacrificing success rate.

6.3.2 Coordination of Control Inputs

Accomplishing the circle tasks requires coordination of multiple control inputs. Lee [168] proposed a tau guidance theory to explain human control coordination strategy. Tau of a motion gap is the time needed to close the gap at the current closing rate,

$$\tau = \frac{x}{\dot{x}}. \tag{6.1}$$

A motion gap refers to a perceived difference between an observer's current and desired goal state. The gap can take various forms, such as spatial distance or heading difference.

An advantageous proposition of the tau guidance theory is that humans can sense the tau of a motion variable directly without measuring gap size and closing velocity. This property enables humans to achieve efficient and robust guidance.

Further, Lee [168] developed the tau coupling theory to consider the closing of two or more motion gaps:

$$\tau_y = k\tau_x. \tag{6.2}$$

Among all gap closures, one (i.e. τ_x) serves as the intrinsic control and the reference for synchronizing the others.

In the circle tasks, except for the fixed heading configuration, heading control plays the role of the intrinsic control. Figure 6.5 shows the heading data for the four pilots in the nose following configuration. The linearity indicates constant heading control. The tau coupling strategy dramatically alleviates control complexity and benefits task accomplishment such that pilots can concentrate on modulations in positional control.

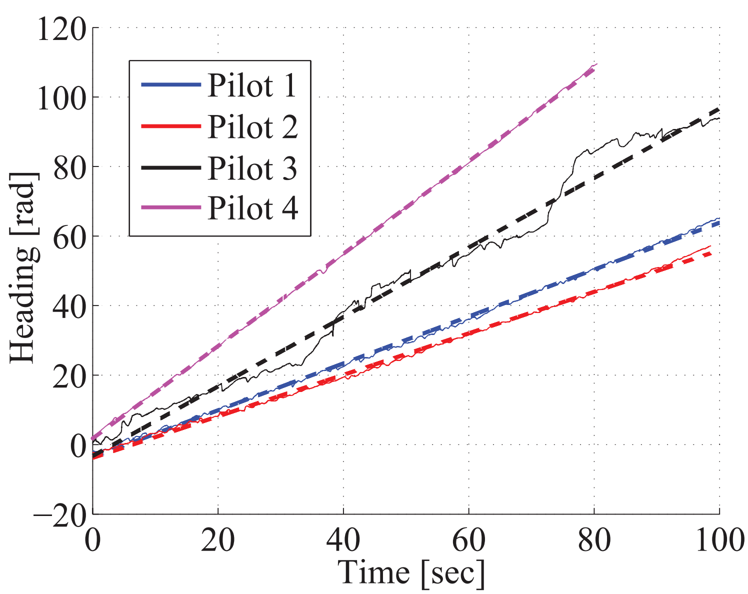


Figure 6.5: Time histories of heading angles in the nose following configuration for the four pilots. Dashed lines indicate the linear fitting.

6.3.3 Motor Control Modulations

Figures 6.1 and 6.4(b) show the convergence of tracking error with respect to the increase of human expertise. It is more essential to understand how human pilots achieve this progress.

In configurations except for fixed heading, human pilots implement heading control as the intrinsic control and simplify the tasks to an approximately single-degree-of-freedom problem. For instance, in nose following, pilots only need to modulate the longitudinal control input. Figure 6.6 presents the longitudinal control commands of the four pilots. It can be observed that human control is intermittent. That is, human control consists of plateaus and intermittent ballistic curves [174]. Intermittent control leads to the difficulty of determining modulation frequency using frequency identification methods such as power spectral analysis. For instance, Table 6.3 lists the crossover frequency of longitudinal control for the four pilots. The difference between pilot 4 and the other pilots is much smaller in crossover frequency than can be observed in Figure 6.6.

Table 6.3: Crossover Freq. of Long. Control

Pilot	Pilot 1	Pilot 2	Pilot 3	Pilot 4
Crossover Freq. (Hz)	0.24	0.23	0.28	0.22

This dissertation resorts to an empirical way to capture human modulation frequency. It uses peaks of helicopter trajectory curvature as the beginning of modulations. This is based on the fact that high curvature is associated with high control gain. Table 6.4 lists the identified modulation frequency for the four pilots, which shows a clear gap between pilot 4 and the others.

Table 6.4: Modulation frequency of circle tasks (Unit in Hz)

Task Setup	Pilot 1	Pilot 2	Pilot 3	Pilot 4
Fixed Heading	0.66	1.43	0.88	0.07
Nose Following	0.31	0.29	1.00	0.04
Tail Following	0.56	0.63	0.83	0.01
Pirouette	0.22	0.74	1.01	0.01

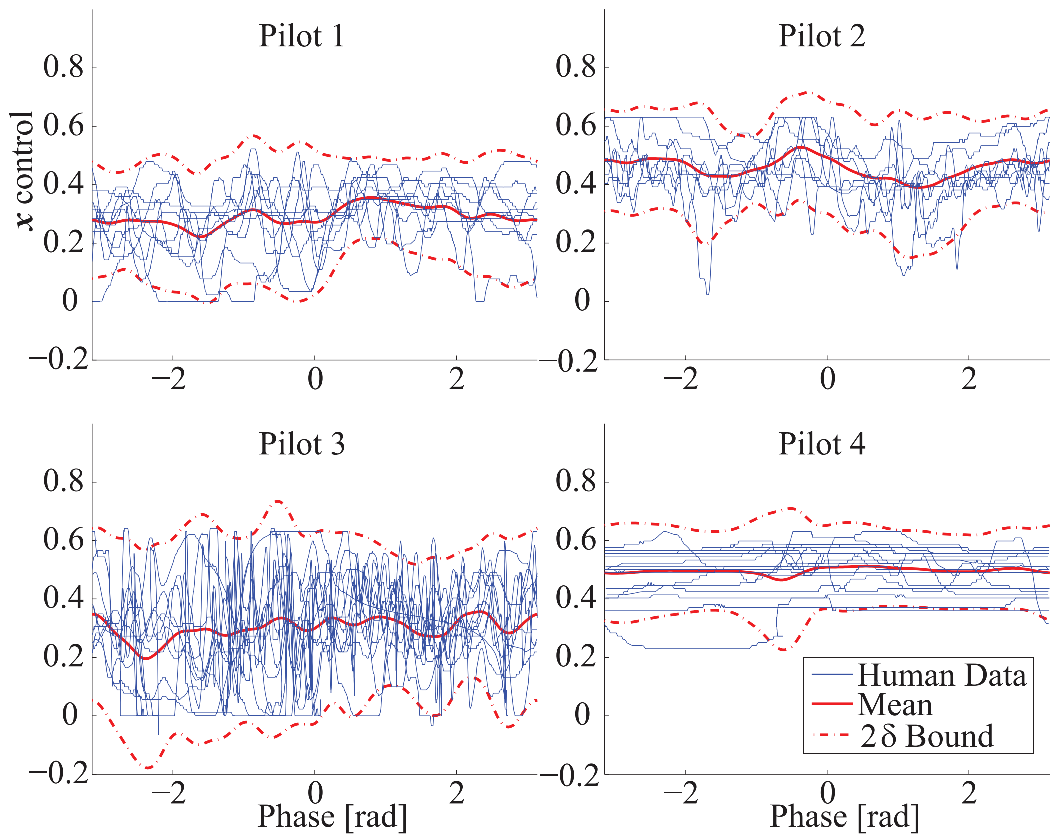


Figure 6.6: Longitudinal control commands in the nose following configuration for the four pilots. These data exhibit human control intermittency.

Specifically, Figure 6.6 and Table 6.4 both show that pilot 4 implemented few modulations. This phenomenon is further investigated on fixed heading data. Assuming the current phase of the helicopter in the circle to be ϕ and the phase lag of human control to be θ, the optimal control profile can be derived that the longitudinal control $v_x = -V\sin(\phi+\theta)$ and the lateral control $v_y = V\cos(\phi+\theta)$. Figure 6.7 presents curve fitting results of pilot 4 in fixed heading, and Table 6.5 lists the correlation coefficients of the four pilots. Both indicated that the control commands of pilot 4 is highly correlated with the optimal control profile. These results strongly suggest that the novice essentially implements open-loop control.

Moreover, the progress from pilot 4 to plot 1 experiences a surge and a subsequent decline in modulation frequency. The frequency surge from pilot 4 to pilot 3 is due to

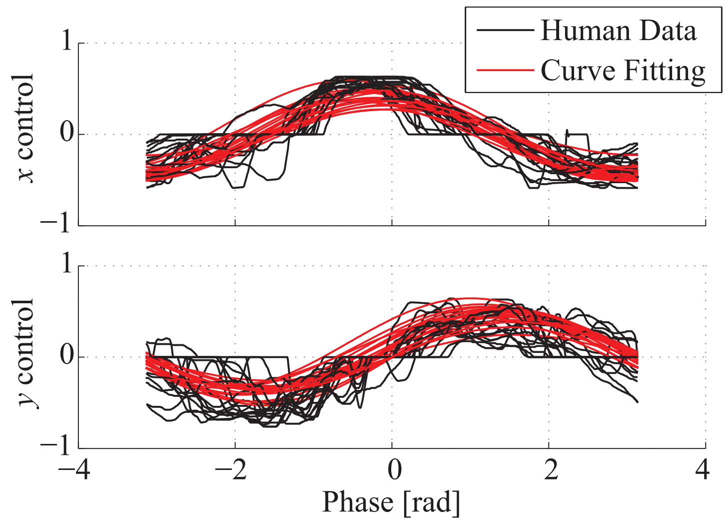


Figure 6.7: Control commands of pilot 4 in the fixed heading configuration. Curve fitting data represent optimal control profile.

Table 6.5: Correlation Coef. in Fixed Heading

Pilot	Pilot 1	Pilot 2	Pilot 3	Pilot 4
Correlation Coef.	0.76	0.71	0.61	0.87

the introduction of closed-loop modulations.

We hypothesize that the subsequent decline resulted from increased modulation accuracy. To prove this hypothesis, we first define an empirical modulation accuracy as the proportion of correct modulations that reduce tracking error in an immediate time window (i.e. 0.5s). Modulations within 0.05m around the reference are removed to consider humans' limited perception accuracy. Table 6.6 lists the results for pilots 1, 2, and 3. The results confirm the reduction in modulation accuracy with respect to human expertise increase.

6.3.4 Interaction Pattern Emergence

Mettler and colleagues [161, 51] suggested that extensive practice of human spatial control behavior enables the acquisition of interaction patterns. Figure 6.1 shows that the tracking of pilots 1 and 2 becomes consistent across different circles. For pilot 2,

Table 6.6: Modulation accuracy of circle tasks

Task Setup	Pilot 1	Pilot 2	Pilot 3
Fixed Heading	0.76	0.75	0.66
Nose Following	0.92	0.83	0.65
Tail Following	0.69	0.67	0.63
Pirouette	0.79	0.67	0.61

there is skewing toward the southeast (upper-left) corner. Figure 6.8 shows that this skewing does not exist in fixed heading and tail following, indicating that it did not result from pilot 2's perception error. Moreover, unlike in fixed heading and tail following, the trajectory in nose following is quite smooth and does not swing around the reference trajectory. This evidence suggests that the consistent performance of pilot 2 in nose following is not simply due to the improved closed-loop modulation capacity but may result from the emergence of a circle-wise interaction pattern.

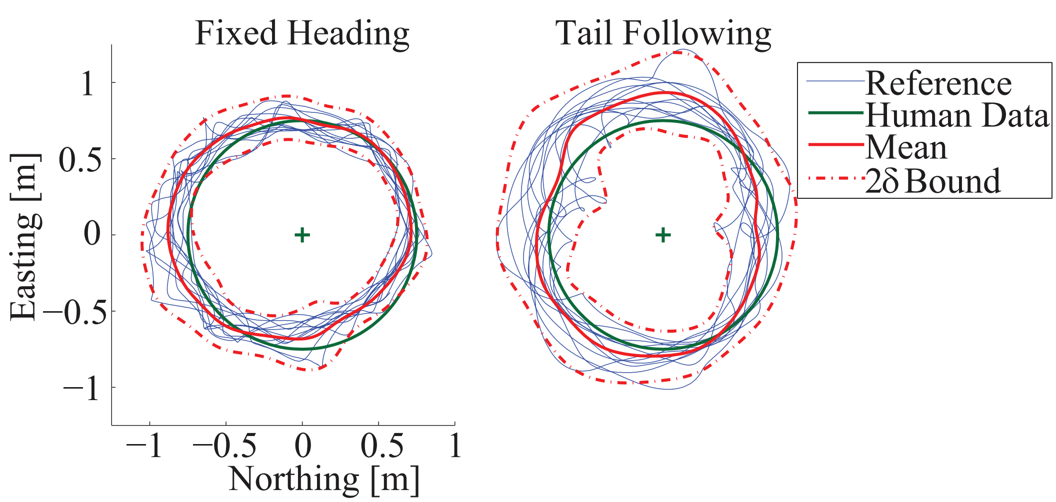


Figure 6.8: Circle task trajectories of pilot 2 in the fixed heading and tail following configurations.

6.3.5 Skill Specificity and Generality

The optimal control profiles of nose following, tail following, and pirouette all consist of a constant heading control and a constant positional (i.e. longitudinal and lateral) control. Therefore, the control complexities associated with them are almost the same.

Among these three configurations, nose following involves a more natural control that pilots have practiced during earlier experiments. This can explain the results in Figure 6.4 in that the performance of pilots 2, 3, and 4 is better in nose following than in the other two configurations. This result suggests that skill development initially takes place in the specific training task setup and does not allow transfer to other configurations. This phenomenon is defined in [17, 93] as “the specificity of practice.”

However, the performance of pilot 1 shows only minor distinctions encompassing different configurations. This suggests that advanced pilots can generalize their skill for achieving versatility.

Chapter 7

Hierarchical Skill Assessment

The assessment of human skills in performing complex spatial control tasks remains challenging. This chapter proposes a skill assessment framework building on the concept of interaction patterns and the associated hierarchical functional model presented in Chapter 4. The model enables systematic descriptions of a human subject's skill that encompass the three primary levels of planning, guidance, and tracking/pursuit, and also accounts for visual attention. The central feature of the model is the identification of interaction patterns. This chapter illustrates the model using two representative examples presented in Chapter 3: remote control of a miniature helicopter and laparoscopic surgery training. The experimental evaluation demonstrates the general validity of the proposed framework.

The chapter is organized as follows. Section 7.1 introduces the proposed skill assessment framework. Section 7.2 briefly describes the human data collected from the remote-control flight and laparoscopic surgery training tasks. Section 7.3 describes the results, and Section 7.4 presents a discussion of the results.

7.1 Hierarchical Skill Assessment Framework

The skill assessment framework is developed from the hierarchical functional model shown in Figure 4.2. The framework evaluates a full range of human functional capabilities in spatial control, from high-level behavior, like planning, to low-level behavior, like perception and control. Figure 7.1 illustrates an overview of the assessment framework by delineating analytical approaches at each functional level for characterizing behavior mechanisms and deriving skill metrics. The following subsections provide detailed descriptions of these approaches. Specifically, skill assessment for planning, guidance, and tracking/pursuit are analyzed in Sections 7.1.1, 7.1.2, and 7.1.3, respectively. Finally, this section presents a skill assessment of gaze movement.

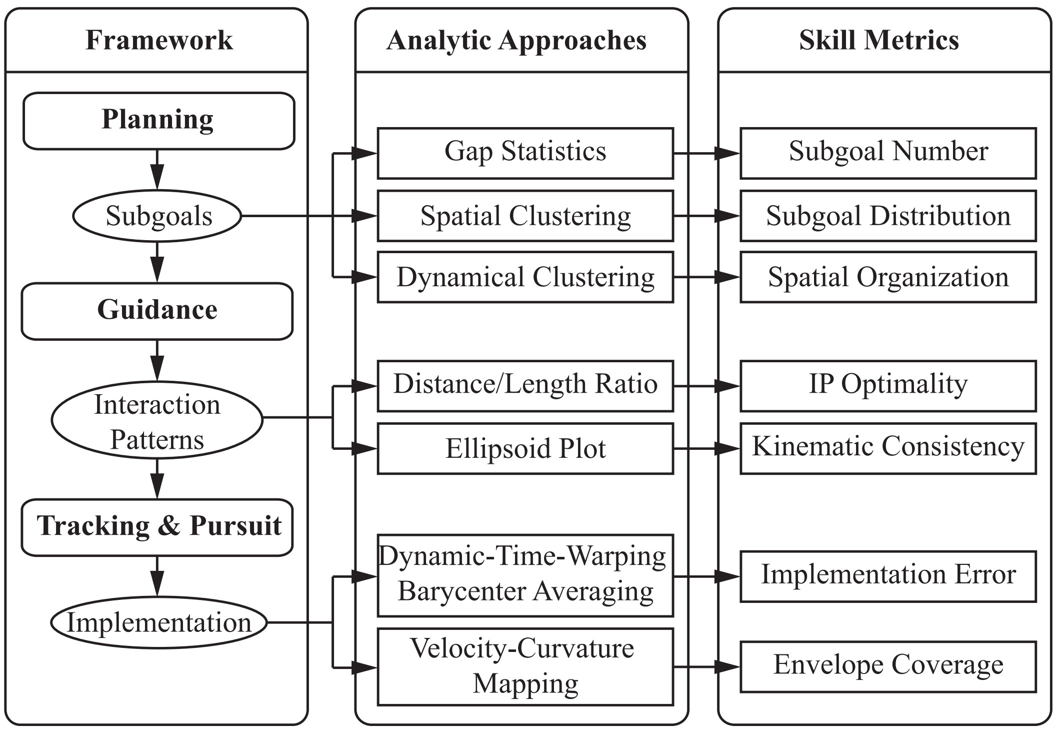


Figure 7.1: Overview of the assessment system for human spatial control skill.

7.1.1 Skill Assessment for Planning

The planning function reduces a global planning task to a sequence of subtasks. Each subtask is characterized by an interaction pattern and its associated subgoals. By capturing these features, subtasks are identified using techniques including spatial clustering, dynamical clustering, and gap statistics.

Spatial Clustering for Subgoal Identification

The spatial clustering approach was originally proposed in [161]. It determines subtasks from the top down using geometric information. It proceeds by identifying subgoals from trajectory ensembles. Kong and Mettler [175] defined subgoals as locations where two trajectories, starting from different initial states, converge and eventually follow the same path to a destination. Specifically, for two trajectories S_1 and S_2, the subgoal, if it exists, is the starting point of the sub-trajectory $\bar{S}$ that satisfies

$$\bar{S} = S_1 \cap S_2 \neq \emptyset. \tag{7.1}$$

However, the locations of subgoals may vary across trials due to the dynamical interactions with the agent and environment during real-time implementations [161]. Although this prevents the direct implementation of Eqn (7.1) to determine subgoals, it still allows us to represent subgoals as stochastic distributions. This fact suggests a two-step method of subgoal identification, first identifying subgoal candidates, and second, clustering subgoal candidates.

By relaxing Eqn (7.1), subgoal candidates can be extracted when two trajectories come within close vicinity for a sufficiently long duration. Specifically, two segments $\bar{S}_1 \in S_1$ and $\bar{S}_2 \in S_2$ of a duration T are assumed to be in vicinity, if they satisfy the following constraints:

$$\|\bar{S}_1 - \bar{S}_2\|_2 \leqslant \delta_2 \quad \text{and} \quad \|\bar{S}_1 - \bar{S}_2\|_\infty \leqslant \delta_\infty. \tag{7.2}$$

$\|\cdot\|_2$ and $\|\cdot\|_\infty$ are Euclidean and infinity norm, and δ_2 and δ_∞ are corresponding deviation thresholds, respectively.

The second step is to cluster subgoal candidates. The subgoal candidates are assumed to be observations of some *hidden* subgoals [161]. The distribution of the subgoal candidates can be modeled using Gaussian mixture models, representing the hidden subgoals as the center of clusters. A clustering method (e.g. K-means in the Euclidean space) is then performed to capture hidden subgoals.

The identified subgoals divide trajectories into segments. The segments or sub-trajectories can be grouped according to their associated subgoals. The segmentation provides information about partitioning the spatial task environment, because sub-trajectories in each partition close the gap to the same subgoal.

Dynamical Clustering for Interaction Pattern Identification

The spatial clustering approach described above works for static environments but is not suited to dynamic environments involving moving targets or moving obstacles. A dynamical clustering approach can address this shortcoming. This approach determines subtasks from the bottom up using dynamics information. It proceeds by by identifying interaction patterns (IPs) as clusters of trajectory points that are similar in their dynamics.

This apporach uses a piecewise autoregressive exogenous (PWARX) model to describe the system dynamics. The PWARX model represents human spatial control behavior as a hybrid system, wherein each ARX model defines an IP as a continuous dynamical mode in a polyhedral partition of the regressor space [127]. In this dissertation, ARX models are expressed in a linear state-space form:

$$\begin{bmatrix} x \\ y \\ v_x \\ v_y \end{bmatrix}_{k+1} = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ a_{31} & 0 & a_{33} & 0 \\ 0 & a_{42} & 0 & a_{44} \end{bmatrix} \begin{bmatrix} x \\ y \\ v_x \\ v_y \end{bmatrix}_{k} + \begin{bmatrix} 0 \\ 0 \\ b_3 \\ b_4 \end{bmatrix}, \tag{7.3}$$

where Δt is the sampling time.

The dynamical clustering process introduced in [127] can be summarized by the following steps:

- Given state $x(k)$, construct a local data set ℓ_k of size c by combining $x(k)$ and $c-1$ neighboring points $\{\tilde{x}\}$ that satisfy

$$\|\tilde{x} - x(k)\| \leq \|\hat{x} - x(k)\| \quad \forall \hat{x} \in \varphi \backslash \ell_k, \tag{7.4}$$

  where φ is the entire data set, $\tilde{x}$ is the point in the local data set ℓ_k, and $\hat{x}$ is the points outside ℓ_k.

- Identify the feature vector ξ_k for ℓ_k. The feature vector ξ_k is a combination of model coefficients a and states centroid m. The states centroid is the mean of the states in a local data set. Model coefficients are estimated using linear regression.

- Partition the collection of feature vectors $\{\xi_k\}$ into clusters, using a hierarchical clustering method in [176]. The clustering proceeds by successively merging two clusters, C_r and C_t of minimum dissimilarity. The dissimilarity is defined as

$$D_{r,t} = \frac{n_r n_t}{n_r + n_t} \left\| \frac{1}{n_r} \sum_{\xi_i \in C_r} \xi_i - \frac{1}{n_t} \sum_{\xi_j \in C_t} \xi_j \right\|^2_{R^{-1}_{r,t}}, \tag{7.5}$$

where R is the covariance matrix of feature vectors. To reduce outliers, this section uses a weighted dissimilarity measure, by including a weight vector w that biases towards state centroid in feature vectors, i.e. $\xi = [a; w \cdot m]$.

Determination of Number of Subtasks

The implementation of both spatial clustering and dynamical clustering requires a technique to determine the number, k, of subtasks. A gap statistic approach is used to estimate k. (See [177] for details of the approach.)

The clustering error decreases monotonically as the number of clusters increases, as shown in Figure 7.2. The gap statistic compares the clustering error curves of real data and a reference distribution (e.g. normal distribution). The optimal number is the point at which there is the largest gap between the two error curves.

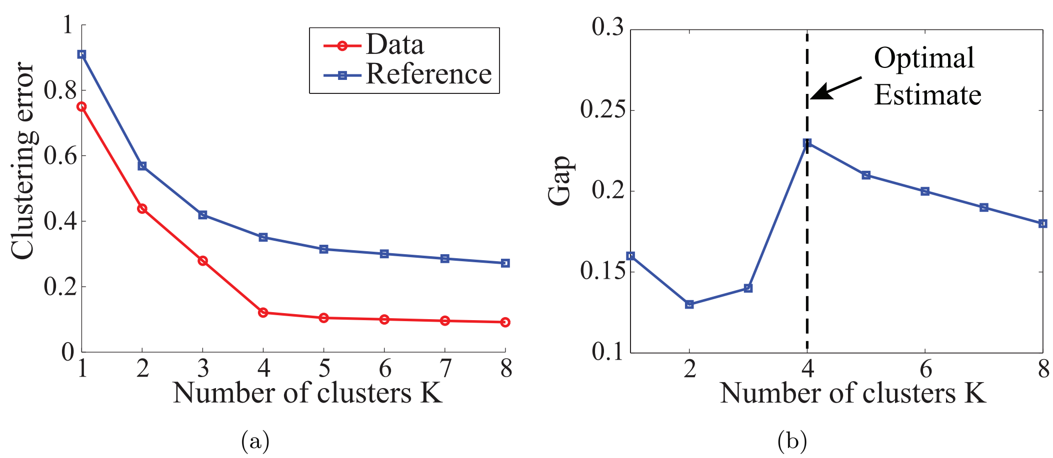


Figure 7.2: Illustration of gap statistic analysis. (a) Clustering errors of the real data and reference data. (b) Gap between the two clustering errors

Derivation of Planning Skill Metrics

Human spatial control skills at the planning level manifest through the optimality of subtasks arrangement.

The spatial clustering approach assumes that humans use subgoals to facilitate motion planning. Experts consider more factors when planning the path, and their trajectories are less variant across trials. This variance is also exhibited in the distribution of subgoal candidates. Therefore, the variance in subgoal candidates distribution can provide a performance metric for human planning skills. Specifically, the metric is defined as a collective variance of the distances between each subgoal candidate and its corresponding subgoal, that is,

$$\sigma_s = \sum_k \sum_{sc_i \in c_k} \|sc_i - s_k\|, \tag{7.6}$$

where sc_i and s_k are subgoal candidates and subgoals, respectively.

The dynamical clustering approach assumes that humans accomplish a spatial task using interaction pattern (IP) as behavior units. Each IP is associated with a procedure or phase of the task. For instance, as will be discussed in Section 7.3.3, surgeons decomposed a peg transfer task into three phases: starting, maneuvering, and interception. Phase margins represent IP transitions. A clear margin indicates a stable strategy in planning subtasks. The margins can be captured by applying a Fisher classifier to the Cartesian coordinates of clustered trajectory points. Planning skills can then be evaluated using misclassification ratio as a performance measure.

7.1.2 Skill Assessment for Guidance

The guidance function implements a coordinated control profile to close the motion gap between current state and the active subgoal. Human skill at this level can be assessed based on the optimality and consistency of implemented guidance behavior.

Optimality of Guidance Behavior

The optimality of guidance behavior is usually described as the difference between the implemented and optimal guidance behavior. However, in practice, the definition of optimal behavior for humans is not trivial. De Camp [165] has indicated that decrease in path length is a characteristic of learning and suggested using relative distance as a measure of optimality. Here the relative distance is calculated by comparing path length (L) and displacement (d).

The L-d difference can be described in various forms. One is the ratio of d/L. Another option is the integrated absolute difference between L and d across time history. This dissertation chooses the first one, because it is normalized in the range of [0 1].

Kinematic Consistency

The dynamical clustering approach describes interaction patterns (IPs) as a set of dynamical models with different coefficients. The coefficients can be identified using regression. As a result, the variances of these regressed coefficients for an IP contain information about its consistency. However, these coefficients have no physical meaning of human behaviors and their variances cannot directly explain the IP consistency. The solution is to transform these coefficients into kinematic coordinates using an ellipsoid plot.

The ellipsoid plot uses three kinematic characteristics: velocity, normal acceleration, and tangential acceleration. Since spatial control behaviors are recorded as trajectories, the kinematic characteristics can be easily calculated for each trajectory point. The ellipsoid plot can then depict IPs by their means and variances in the three-dimensional kinematic space.

Overlap of Interaction Patterns

IPs provide a frugal way of realizing movement behaviors. First, the invariant characteristics enable a small set of interaction patterns to cover a broad range of conditions and situations. More importantly, the overlap between interaction patterns needs to be small, so as to reduce the number of necessary IPs.

The ellipsoid plot visualizes IPs. It depicts the correlation among interaction patterns as the overlap in the ellipsoids. This dissertation defines a performance metric called connectivity to capture this correlation. Connectivity is defined as the ratio of the overlapping volume relative to the volume occupied by all ellipsoids.

7.1.3 Skill Assessment for Tracking and Pursuit

The tracking and pursuit function is responsible for minimizing tracking errors. Tracking is defined by McRuer as "manual processes for minimizing visually perceived errors by exercising essentially continuous control so as to match visually presented input and output signal" [46].

It is difficult to implement McRuer's definition for skill assessment. This definition builds on an experiment in which the operator follows an input signal presented on a display. In real world applications, such as remote-control flight and surgery, system inputs are not explicitly shown. Moreover, when performing in the precognitive mode, humans may not project a specific flight trajectory, because an open-loop control plan is implemented with no need for visual feedback. Even if projected trajectories exist, they are products of human internal processes and cannot be visually observed. Therefore, tracking error cannot be directly measured in these conditions. However, the tracking and pursuit performance can be assessed based on the variability in repeated tasks.

Variability in Trajectory Ensembles

In tasks with static environments and task elements, such as start, goal, and obstacles, tracking consistency refers to the repeatability of spatial control trajectories. It can be measured as the variability in trajectory ensembles. Specifically, if a representative trajectory can be determined, this variability can be calculated as the distances between the representative trajectory and all trials in the ensemble. This dissertation determines representative trajectories and calculates trajectory distances using the DTW barycenter averaging (DBA) method presented in Section 5.1.

Maneuver Envelope

According to the uncontrolled manifold hypothesis [103], the variability in states can measure the stability of control behavior. A "controlled" behavior is more stable against perturbations and exhibits less variability.

To capture the variability in states, we develop a method based on a maneuver envelope. A maneuver envelope is depicted by a mapping of velocity and curvature, because these two features are sufficient to describe agent states. The mapping is constructed using a 2D histogram to represent the frequency distribution of states [128]. Specifically, the velocity-curvature space is discretized into grids, and then the frequency is counted as the number of time samples when the agent states are within the range of a specific grid.

Two types of maneuver envelopes are defined: the implementation envelope of an operator and feasible envelope of a human operator. The former is a maneuver envelope that only considers the data of the interested human operator. The later is approximated as a collected maneuver envelope of all operators. Then a performance metric called coverage can be defined as the range of the implementation envelope relative to the feasible envelope. A low coverage value indicates high performance.

7.1.4 Skill Assessment on Gaze Movement

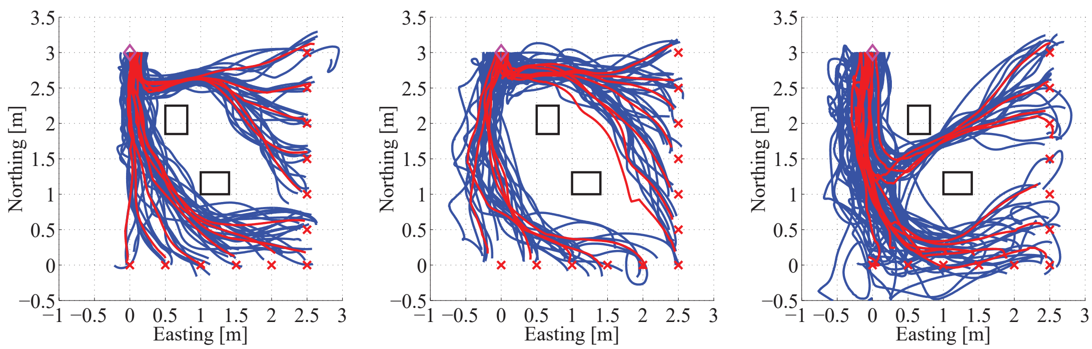
Gaze movements are used throughout the spatial behavior hierarchy., providing different types of information for each functional subsystems. Recent studies have used gaze movements as a new source of assessing spatial control skills [178, 64, 179].

Gaze movements primarily consist of three basic patterns: fixation, smooth pursuit, and saccade. These three gaze patterns have been widely accepted as the means for attention control. Most of these gaze movements are unconscious. Fixations and smooth pursuits are performed to acquire high-quality visual information. Fixations are tightly connected to critical task elements associated with the progress of the task [180]. Smooth pursuits estimate the dynamical state information of the agent under control [2]. Although high velocity and short duration render the visual system essentially blind during saccades [179], saccades may provide a measurement of motion gap in the form of time to closure [2]. This chapter performs gaze classfication that decomposes gaze movements into these patterns using the hidden Markov model (HMM) method presented in Section 3.2.2.

7.2 Human Experiment Data

Two applications are used to exercise the proposed skill assessment framework: a remote-control guidance task, and a laparoscopic surgery training task. The task configuration of these two experiments are described in Section 3.1.

The guidance task data analyzed in this chapter are collected using Blade MCX2 with three test subjects, including an expert, an intermediate, and a novice pilot. The expert holds a pilot license and has two years of prior experience with the guidance experiment. The intermediate subject also holds a pilot license but has only three months of experience. The novice holds no pilot license and has one year of experience. Figure 7.3 presents the trajectory ensembles of these three pilots.



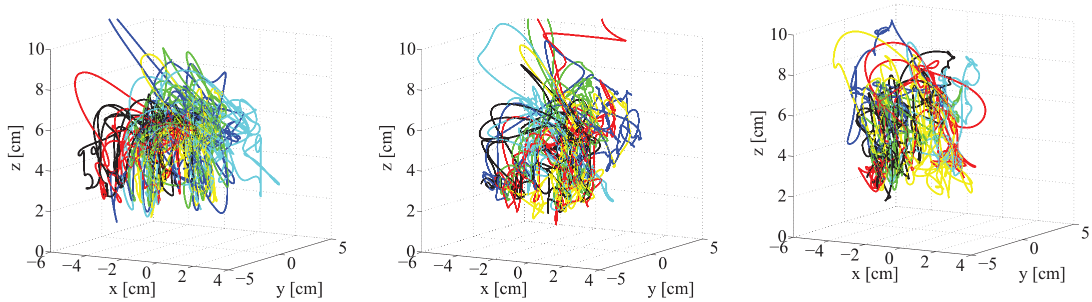
(a) Expert (b) Intermediate (c) Novice

Figure 7.3: Trajectory ensembles of human subjects in the guidance task. We use red lines to show representative trajectories obtained with DBA. The implementation error of the novice is larger than that of the intermediate and the expert.

The surgery simulation data are collected from a peg transfer task. We arbitrarily select a set of six complete peg transfer task instances to represent each of the three expertise groups: expert, intermediate, and novice. The gaze dataset is limited to four subjects: one expert, one intermediate, and two novices, due to the recent setup of the eye tracking system.

7.3 Results

This section presents results of applying the assessment framework to the flight and surgery datasets, and discusses first the overall task performance, followed by performance at each functional level, and finally, gaze movement.



(a) Expert (b) Intermediate (c) Novice

Figure 7.4: Trajectory ensembles of human subjects in the surgery training task. Colors indicate different trials. The trajectories of experts are more smooth and organized.

7.3.1 Overall Performance

We use success rate and travel time to describe the overall performance of the human pilots in the flight task. Successful trials are those satisfying the time and terminal constraints (described in Section 3.1.1) and having no collision with obstacles. The success rate of a pilot is defined as the proportion of successful trials performed by the pilot. The travel time is defined as the averaged time duration for a pilot to accomplish one trial. The results listed in Table 7.1 show that the expert is superior in both performance metrics. On the other hand, the novice has a higher success rate than the intermediate pilot but requires a much longer travel time. This indicates that the overall outcome metrics are insufficient for differentiating skill levels among multiple individuals.

Table 7.1: Overall performance in the flight task.

Performance Metrics	Expert	Intermediate	Novice
Success Rate	94%	84%	85%
Average Travel Time [sec]	5.69 (0.58)	6.14 (0.87)	8.03 (1.36)

* The numbers in the parentheses are standard deviations.

Overall performance in the surgery training task is described by task duration and the number of clamp trials. Task duration is defined as the time interval from the instant when a surgeon makes contact with the first block to the instant when he or she returns the last block. Perfect performance of the task involves a minimum of 12 clamps of each hand tool to pick up the blocks from the board or from the other hand tool. Therefore, clamps exceeding 12 indicate failed trials. Similar results in Table 7.2 show that the experts accomplished the surgery training task much faster and with far fewer failed trials. The standard deviations of these metrics indicate that the differences among the three groups are statistically significant.

Table 7.2: Overall performance in the surgery training task.

Performance Metrics	Expert	Intermediate	Novice
Average Time [sec]	72.64 (12.86)	143.39 (10.09)	207.71 (12.66)
Average Clamps	13.83 (0.75)	19.00 (4.34)	20.50 (4.37)

* The numbers in the parentheses are standard deviation.

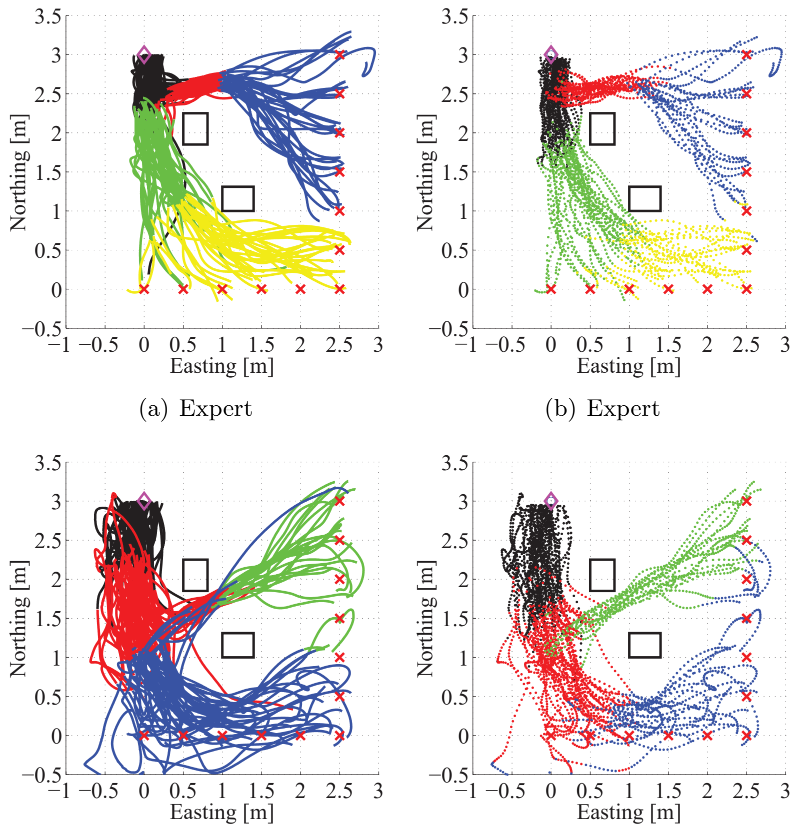
7.3.2 Planning Performance

In the flight task, we identify subgoals following the spatial clustering and the gap statistic and partition trajectory into segments. Interaction patterns are represented as manifolds of the segments associated with the same subgoal. Column 1 in Figure 7.5 shows the segmented trajectories of the flight task obtained using spatial clustering, and Column 1 in Figure 7.6 shows the associated aggregated interaction patterns. (Here only the results of the expert and the novice are presented.)

For comparison, we then apply the dynamical clustering approach. Column 2 in Figure 7.5 shows the trajectories with identified clusters, Column 2 in Figure 7.6 shows the aggregated interaction patterns, and Figure 7.7 shows the ellipsoid plots. First, two parameters are specified in order to implement the dynamical clustering. The first parameter, the size of a local dataset, is set to 35. It is a trade-off between noise (for smoothing) and outliers (local datasets with members in mixed categories). The second parameter, the bias on state centroids, is tuned to 4. This parameter enables the dynamic clustering to consider spatial information of trajectory points, thereby reducing clustering noises.

The comparison of the aggregated interaction patterns obtained with both methods in Figure 7.6 indicates that the dynamical clustering is more accurate in capturing coherent interaction patterns.

The above results also provide the following insights. First, the subtasks derived from both techniques are quite similar. This strongly supports the idea that humans organize spatial control behaviors with interaction patterns. Second, the margins between the



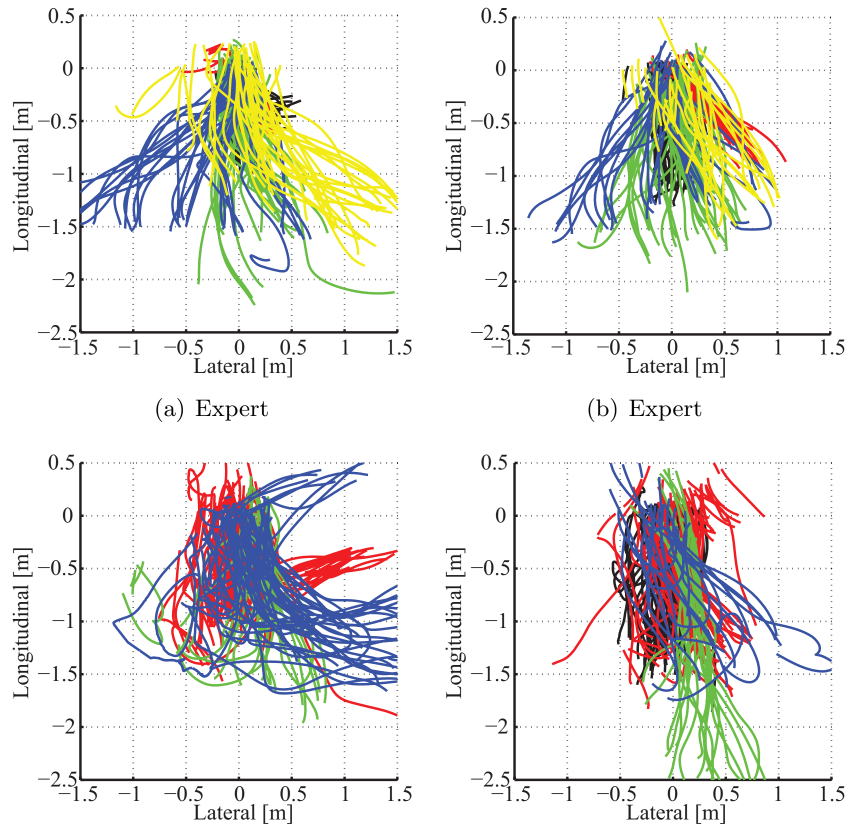
(a) Expert (b) Expert

(c) Novice (d) Novice

Figure 7.5: Trajectory segmentation results in the flight task. Column 1 is obtained using spatial clustering, where subgoals are first identified as the symbol for segmentation. Column 2 is obtained using dynamical clustering, where trajectory points within a cluster are identified to have almost the same dynamics.

subtasks, as delineated by the subgoals, are almost always located around the corners of obstacles. This phenomenon supports the hierarchical functional model in that humans account for environmental features when allocating subgoals during planning.

The results obtained with the planning performance metrics are listed in Table 7.3. The lowest subgoal distribution variance and the lowest misclassification ratio of spatial organization indicate that the expert has the most consistent arrangement of subgoals and interaction patterns. This can also be observed by comparing Columns 3 and 4 in Fig. 7.5, where the expert exhibits the least difference in the interaction patterns



(a) Expert (b) Expert

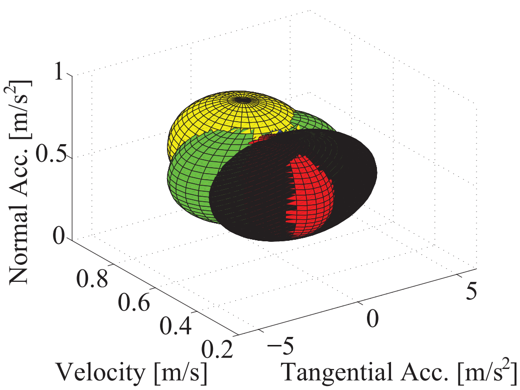
(c) Novice (d) Novice

Figure 7.6: Aggregated interaction patterns in the flight task. Column 1 is obtained using spatial clustering, and column 2 is obtained using dynamical clustering. The trajectory segments in each aggregated interaction pattern are transformed to align with the terminal velocity.

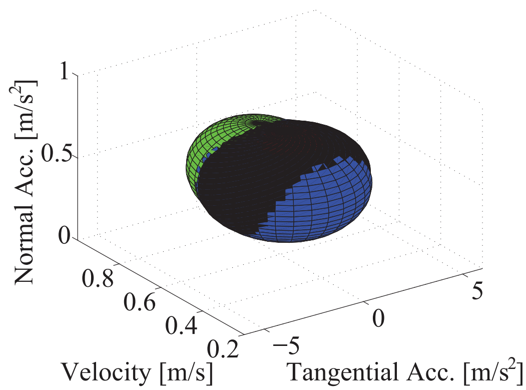
obtained with spatial clustering and dynamic clustering.

In the surgery training task, the starting point (location of a block) and the target (the location where a block is passed from one tool to the other) are not specified but change with every trial. Therefore, applying the spatial clustering approach to the surgery training task is not possible.

The dynamical clustering approach can solve this problem. The trajectories with identified clusters and the ellipsoid plots are shown in Figures 7.8 and 7.9, respectively. (Here only the results of the expert group and the novice group are presented.) Three



(a) Expert



(b) Novice

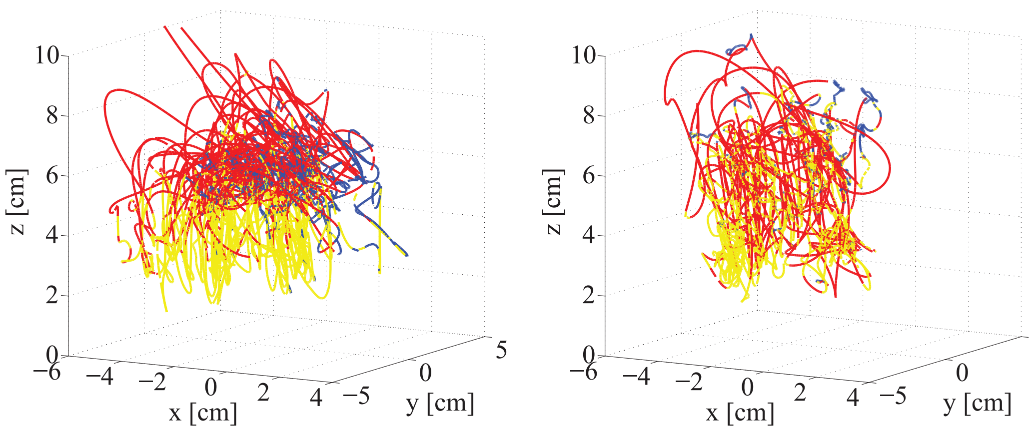
Figure 7.7: Ellipsoid plots of clusters obtained through dynamical clustering in the flight task. Clusters, representing interaction patterns, are depicted as colored ellipsoids with three kinematic characteristics: velocity, tangential acceleration, and centripetal acceleration. The size of ellipsoids reflects the consistency of interaction patterns, and the overlap among ellipsoids describes the similarity of interaction patterns.

Table 7.3: Planning skill metrics in the flight task.

Performance Metrics	Expert	Intermediate	Novice
Subgoal distribution variance [m^2]	0.146	0.256	0.335
Spatial organization [%]	9.6	10.3	11.3

clusters are identified for each group. These clusters correspond to three major phases: starting (cluster mode 2), maneuvering (cluster mode 3), and interception (cluster mode 1). They will be discussed in Section 7.3.3, when the performance at the guidance level is presented.

The framework uses the metric of spatial organization to analyze the margins between the task phases. This analysis is only performed between the starting phase and the maneuvering phase, because the trajectory points of the interception phase scatter widely, and there is no clear margin between the interception phase and the other two phases. This is due to the fact that for each trial, there is a different position at which the block is in transition. Table 7.4 lists the misclassification ratios for the three groups. The significant difference in the misclassification ratio among the three groups indicates that, as skill level increases, the movement behavior of a surgeon will display a better defined spatial organization. Improved task phase planning plays a key role



(a) Experts

(b) Novices

Figure 7.8: Trajectory segmentation obtained with the dynamical clustering in the surgery training task.

in overall performance (time and success rate). This is the hallmark of highly skilled surgeons, since it builds on all the other skill components.

Table 7.4: Spatial organization of subtasks in the surgery training task.

Spatial Organization [%]	Expert	Intermediate	Novice
Complete Groups	17.9	29.6	38.6
Leave-One-Out Mean	13.3 (4.4)	27.1 (5.6)	35.0 (6.5)

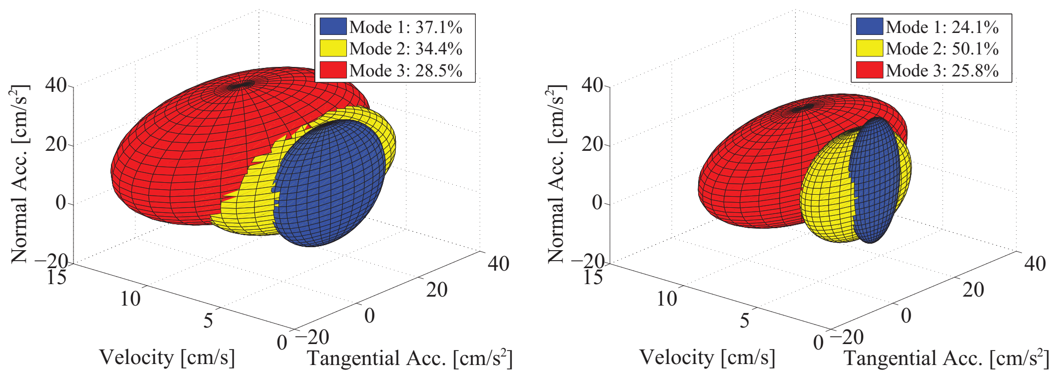
* The numbers in the parentheses are standard deviation.

7.3.3 Guidance Performance

Interaction patterns in the flight task lead to an arrangement of subgoals around the obstacle corners. (See Figure 7.5.)

Table 7.5 lists the averaged d/L ratio for each subtask obtained using spatial clustering. These results indicate that, although d/L ratios are relatively low for an entire trial, the d/L ratios are much higher for subtasks.

When comparing the three pilots, the expert achieved a higher d/L ratio for both the entire trajectory and individual subtasks. In the flight task setup (Figure 3.4), terminal constraints on position, heading, and speed are used to establish successful trials. When closing the gap to the final target (mode black), the expert and the intermediate had a



(a) Experts (b) Novices

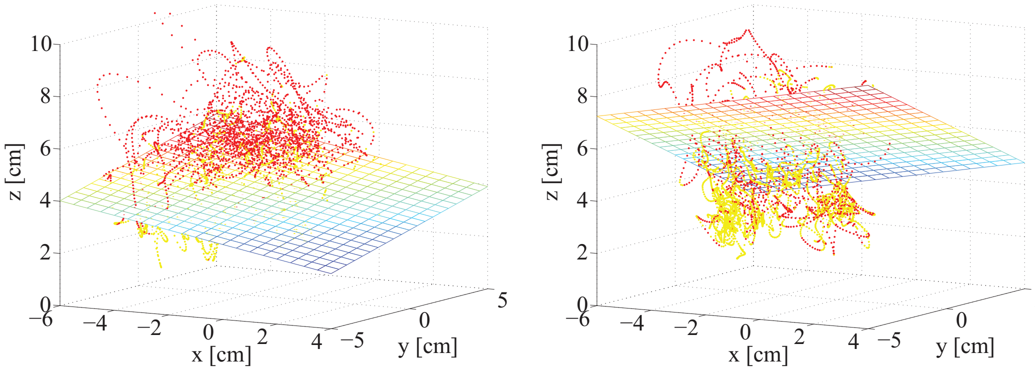
Figure 7.9: Ellipsoid plots of clusters obtained through the dynamical clustering in the surgery training task. Clusters, representing interaction patterns, are depicted as colored ellipsoids.

lower d/L ratio, suggesting that they performed finer control corrections in an attempt to satisfy the terminal constraints. On the other hand, the novice slowed the helicopter to a complete hover before the final stage and then engaged in a “straight” flight with a high d/L ratio, in order to increase the success rate.

The measure of connectivity describes the degree of differentiation of implemented interaction patterns. Table 7.6 lists the connectivity values of the three pilots. A certain degree of overlap is required to enable smooth transition among interaction patterns. However, the high value of the novice indicates that the implemented interaction patterns are essentially the same. This information can also be observed in the plot of the aggregated interaction patterns of the novice in Figure 7.6(d). In other words, the novice does not possess a differentiated library of interaction patterns. In contrast, the connectivity values of the expert and the intermediate subject are much lower. (Note that the novice allocated one less subgoal than the other pilots, which can be one factor of high connectivity value.)

Table 7.5: d/L ratio for subtasks obtained with spatial clustering in the flight task.

Mode	Yellow	Blue	Green	Red	Black	Trajectory
Exp	0.94	0.94	0.95	0.98	0.89	0.93
Int	0.90	0.96	0.94	0.94	0.80	0.89
Nov		0.79	0.91	0.89	0.92	0.88



(a) Experts (b) Novices

Figure 7.10: Spatial organization of the starting phase and the maneuvering phase in the surgery training task. Misclassifications in the partitioning indicate inconsistent organization of subtasks.

Table 7.6: Connectivity of interaction patters in the flight task.

Performance Metrics	Expert	Intermediate	Novice
Num. of subgoals	4	4	3
Connectivity [%]	61.4	52.6	77.5

For the surgery training tasks, the interaction patterns obtained using the dynamical clustering correlate with the three task phases. (See Figure 7.8.)

The starting phase (cluster mode 2) coincides with the surgeons picking up the blocks from the bottom of the poles to above the tips. The movement during this phase follows a medium velocity range. The poles constrain the trajectory in this phase. Therefore, the pick-up path is the same for all trials. The experts distinguish themselves through the use of high acceleration, without loss of accuracy.

The maneuvering phase (cluster mode 3) coincides with the surgeons moving the gripped blocks to the central area of the board. The movement is not restricted during the maneuvering phase, and the objective of the phase is to be as fast as possible. The surgeons try to reach a high velocity. The normal acceleration spans a large range to allow sharp curves to change the moving direction.

The interception phase (cluster mode 1) coincides with the period in which the blocks are transferred in the air between the laparoscopic tools. This phase is critical in that it requires coordination between both hands. Poor coordination will drop the blocks, leading to failure of trials and an increase in task duration. For these reasons, the interception phase is performed at a low velocity. Table 7.7 gives the proportions of time duration in each phase for the three groups. The results show that the experts spend the largest amount of time in this phase, while the novices spend the least.

Table 7.7: Phase durations for the three groups in the surgery training task.

Phase Duration [%]	Expert	Intermediate	Novice
Starting	34.4	54.7	50.1
Maneuvering	28.5	14.9	25.8
Interception	37.1	30.4	14.9

Table 7.8 lists the connectivity values of the three groups. These values are lower than those in the flight task because the three interaction patterns in the surgery training task differ significantly in their task elements and control objectives. However, it is the same as in the flight task in that the novice group had a larger connectivity value than the expert group, indicating a lower performance in switching between the three phases.

Table 7.8: Connectivity of interaction patters in the surgery training task.

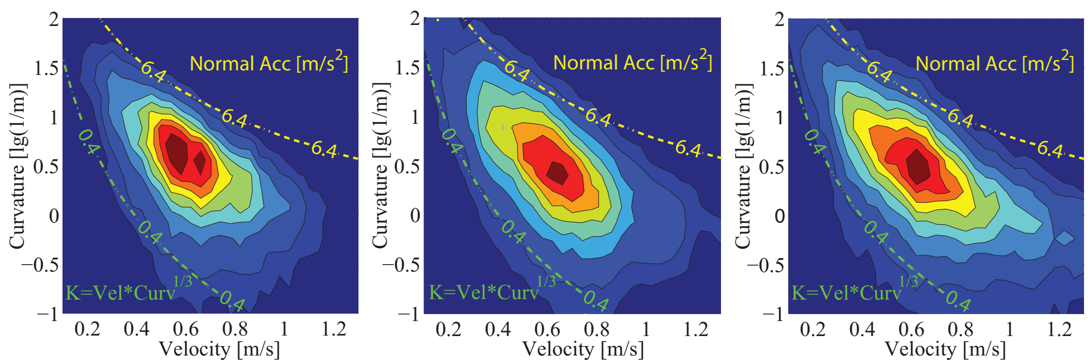
Performance Metrics	Expert	Intermediate	Novice
Num. of subgoals	2	2	2
Connectivity [%]	37.2	17.9	45.2

7.3.4 Tracking and Pursuit Performance

For the flight task, representative trajectories are obtained using the DTW barycenter averaging (DBA) for each ensemble of trials from the same starting location. DBA provides a within-group-sum-of-squares as a measure of implementation error. The averaged values are 0.40×10^4, 1.52×10^4, and 2.87×10^4 for the expert, intermediate, and novice, respectively. This result indicates that the expert's motor control behavior is significantly more consistent than that of the other two pilots.

Figure 7.11 presents the velocity-curvature mappings for the three pilots. The pilot's *implementation envelope* is empirically determined as the range of states with a frequency higher than 1% of the entire task duration. The implementation envelopes approximately form a triangular region. The top edge represents the helicopter's normal acceleration constraint. The value should be the same for all pilots. The left-bottom edge represents a 2/3 power law constraint relating the geometric and kinematic characteristics of human movements [181].

We then generate the *feasible maneuvering envelope* of the helicopter using the data of all pilots and following the same process as in determining the implementation envelope. The coverage can then be calculated. The respective values for the expert, intermediate, and novice pilots are 0.57, 0.59, and 0.68. This result indicates that the pilots tend to use a smaller range of states as their skill levels increase. The reduced state range can facilitate prediction of control outcome and improve maneuvering consistency.



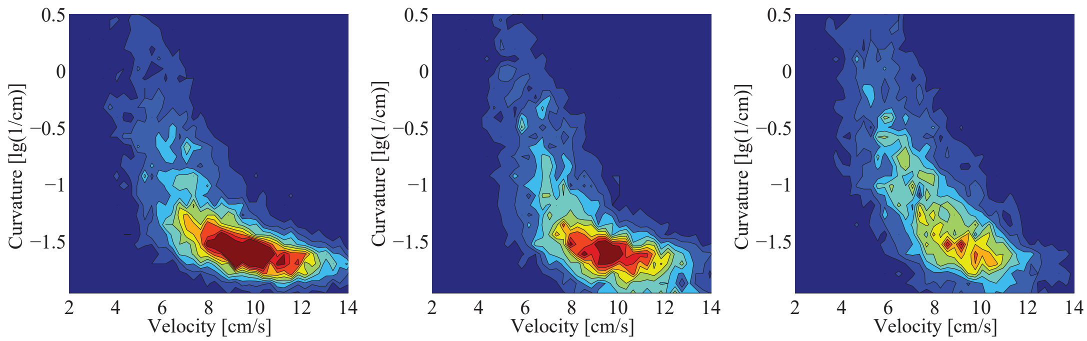
(a) Expert (b) Intermediate (c) Novice

Figure 7.11: Velocity-curvature mapping in guidance tasks. Yellow line indicates the normal acceleration constraint and green line indicates the 2/3 power law constraint. The coverage of the maneuver envelope for the expert is smaller than that of the novice.

Figure 7.12 shows the velocity-curvature mappings of the surgery training tasks. In this task, surgeons are controlling mechanical tools, and their manipulations are not restricted by the normal acceleration limit. However, the same phenomenon as in the flight task appears when comparing the three groups. The experts still exhibit a more focused implementation envelope with a smaller velocity range. The coverage values are 0.68, 0.71, and 0.78 for the expert, intermediate, and novice groups, respectively. This result supports the hypothesis that experts have a more focused motor control range.

7.3.5 Performance Associated with Gaze Behavior

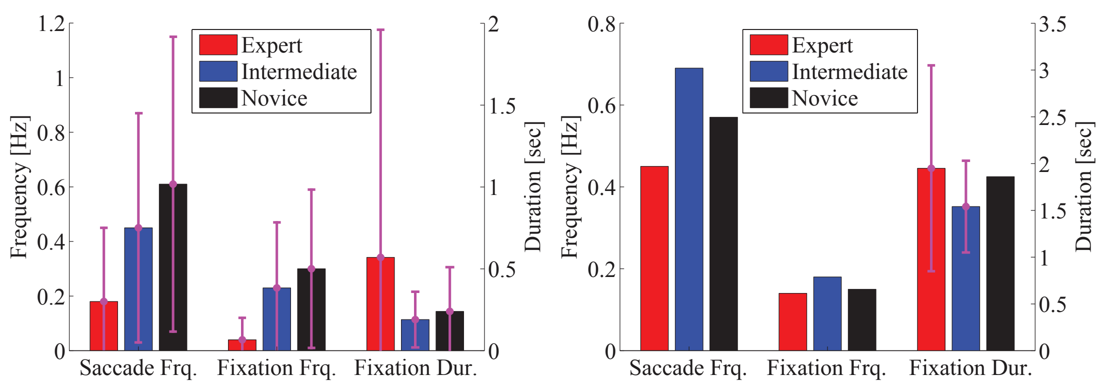
This chapter investigates the performance associated with gaze behavior using the three metrics: saccade frequency, fixation frequency, and fixation duration. Figure 7.13 shows



(a) Expert (b) Intermediate (c) Novice

Figure 7.12: Velocity-curvature mapping in the surgery training task.

the results for both the flight and surgery training tasks. It is the same in both tasks in that the experts exhibited fewer saccades, fewer fixations, and longer fixation durations. The results suggest that an expert has a better internal model and a better visual search strategy.



(a) Flight task (b) Surgery task

Figure 7.13: Gaze mode statistics.

7.4 Discussion

The objective of this chapter is to construct a skill assessment framework that can detail the performance of human operators in coordinating functional components of spatial control behavior. The section first integrates performance metrics derived in the previous sections to provide a detailed description of human performance. Following that, the section discusses couplings between the hierarchical functional levels and extends the concept of interaction pattern. Finally, the section ends with a discussion of the use of gaze behavior for skill assessment.

7.4.1 Integration of the Skill Metrics

Figure 7.14 shows a plot of overall performance combining the skill metrics for the functional skill components in the flight task. All metrics are normalized to the range of [0, 1], with 1 referring to the best performance among the three pilots. Notice that the expert's performance results in a more circular profile in the plot, indicating no obvious weakness for any performance functional components.

When compared with the intermediate pilot, the novice has achieved a higher success rate but lower performance in other areas such as travel time. This trade-off highlights the fact that humans may emphasize different aspects of performance. This further underscores that skill assessment based on a single criterion is insufficient.

The proposed assessment model combines skill metrics across the hierarchical functional levels and also includes performance on gaze behavior. The combination of these metrics enables a more detailed assessment and a more robust evaluation of spatial control skills, in particular since there exists coupling between the components which can confound the assessment.

The skill profile (Figure 7.14) enabled by the skill assessment framework highlights the weaknesses and strengths of operators. This information can help tailor the training by addressing specific areas of weakness. This knowledge can greatly save training time and improve training effectiveness.

7.4.2 Coupling between Functional Levels

Human spatial control skills rely on the coordination of all functional levels. The most prominent interaction exists between planning and motor control skills.

A spatial control behavior is complicated due to the large number of degrees of freedom associated with the agent's dynamics, the environment, and task elements, and as well is limited by humans' sensory-motor and computational constraints. However, as proposed by Mettler et al. [51], humans can alleviate planning complexity by dividing

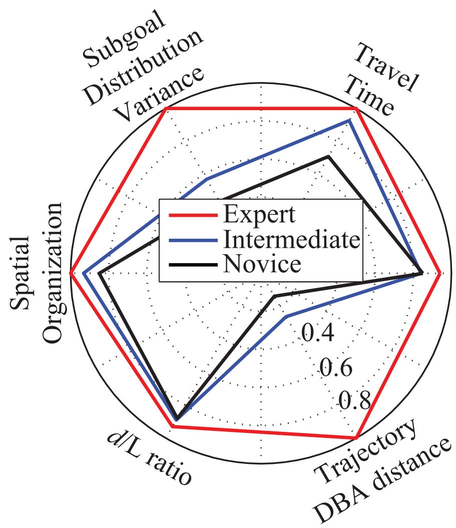


Figure 7.14: Skill assessment of human spatial control behavior in the flight task. A large value in the axis indicates a good performance for each metric.

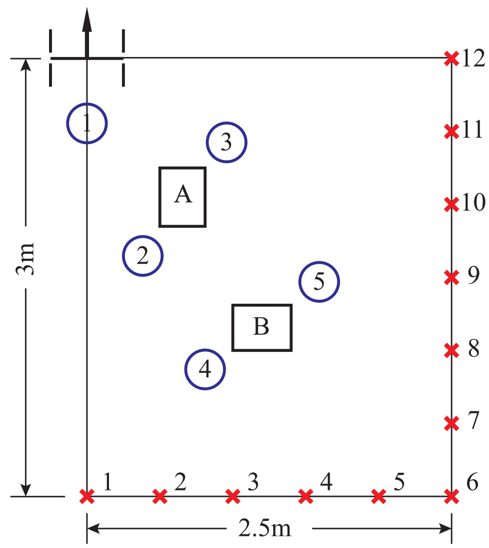
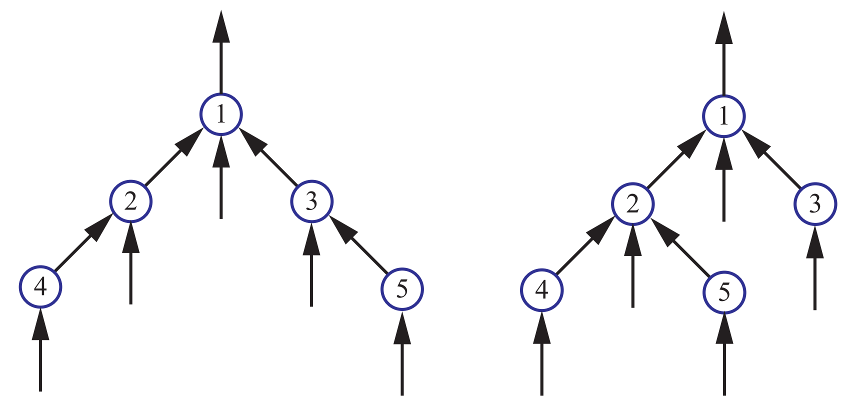


Figure 7.15: Subgoal candidates for trajectory planning.

a spatial control task into subtasks. We define subgoals as locations that connect two subtasks. This strategy is supported by the results in both the remote-control flight and laparoscopic surgery simulation.

The arrangement of subgoals is the result of the coordination between planning and motor control. From the planning perspective, pilots use environmental or task-relevant features to allocate subgoals. Take the flight task for example. Figure 7.15 illustrates the five possible subgoals derived from the clustering results for configuration C, and four of them are located around obstacle corners.

From the motor control perspective, pilots will allocate subgoals based on the knowledge of their own motor control expertise. For instance, Subgoal 1 is located just before reaching the target. It provides a buffer that allows human pilots to adjust the vehicle state in order to satisfy terminal constraints. Moreover, Figure 7.16 shows two candidate plans to arrange subgoals. The expert chose plan 1 that avoids the path from subgoal 5 to subgoal 2. Although this plan increases the path length, it greatly reduces the probability of collisions and increases the success rate.



(a) Plan 1 (b) Plan 2

Figure 7.16: Possible plans for subgoal allocation.

These subgoals are actually the combinational results of both environmental constraints and human sensori-motor constraints. For subgoals 2 to 5, they are not just locations that satisfy equality environmental constraints, by also locations where the obstacles minimizes human vision range. Moreover, at these locations, the agent is the closest to obstacles, which necessitates the activation of boundary avoidance behavior and the increase in the control bandwidth.

This chapter uses the spatial clustering to capture subgoals from the top down, by analyzing the spatial organization of trajectories, and the dynamical clustering to capture subgoals from the bottom up, using motor control characteristics. The equivalence between the results of these two methods underscores that spatial control skills are realized by coordinated planing and motor control processes.

7.4.3 Interaction Patterns

The above evidence motivates the interaction pattern concept as the strategy humans employ to bridge the gap between planning and motor control. They connect planning and motor control, but at the same time, allow these two functions to process largely independently. Specifically, interaction patterns represent the invariants or equivalences in human spatial control behaviors that regulates human motor control processes consistently and active these processes with a small number of neural commands. On the other hand, planning function can use interaction patterns as building blocks to constitute more complex behaviors. Therefore, interaction patterns contribute to a system architecture that clearly delineates planning, guidance, and tracking/pursuit functions.

The development of interaction patterns explains the gain associated with organizing spatial control behaviors as someone improves his or her skills. In both the flight and surgery training tasks, the novices did not demonstrate a consistent library of interaction patterns. On the other hand, the experts distinguished themselves, not only through their superior performance outcomes, but through their differentiated library of interaction patterns. This library is associated with tight and coherent performance across all levels, from the higher-level organization and planning to the lower-level control functions. Therefore, the analysis based on interaction patterns enables a systematic and detailed skill differentiation, making it possible to identify skill deficiencies. They can also help us understand the skill development phases.

7.4.4 Gaze Behavior for Skill Assessment

Gaze behavior participates in the movement behavior throughout the multiple levels of the functional hierarchy. The functions rely on different visual information (e.g. obstacle locations and vehicle velocity) obtained through a combination of gaze patterns [2, 51]. Interpreting these gaze patterns through statistical analysis is limited.

The combination of gaze with the functional components can provide additional insights into mechanisms of attention control, for instance, how attention is allocated for different functions. The comprehensive analysis of gaze and attention will help achieve a more complete understanding of human expertise, as well as provide new opportunities for training, for example through visual cueing.

Chapter 8

Conclusion and Future Directions

8.1 Conclusion

This dissertation extends the concept of interaction patterns and the hierarchical functional model to the investigation of human spatial control skills. By representing human spatial control behavior as a motion planning problem, this dissertation illuminates that humans can follow the hierarchical functional model to decompose the global problem into a sequence of subproblems at each functional level, and the complexity of these subproblems is compatible with constraints in human-machine systems. This dissertation then construct a simulation model that can reproduce human experiment data, validating interaction patterns as behavior units and the hierarchical functional model as a normative framework for interpreting human skills.

Through the investigation of skill development, this dissertation validates the emergence of interaction pattern as the result of a combination of a well-tuned open-loop plan and close-loop modulations. Moreover, the dissertation implies the refinement of interaction pattern as human internal model integrates the human subject, environment, and task elements into a holistic system. The refinement enables the generality of skills across multiple task configurations.

This dissertation then presents a skill assessment framework based on the hierarchical functional model. The framework covers the three hierarchical levels of planning, guidance, and tracking/pursuit, and also gaze behavior, which can be integrated into an overall skill profile. The techniques in the assessment framework capture interaction patterns both from the top down using geometric features and from the bottom up using dynamical characteristics. This result confirms the existence of interaction patterns. Moreover, their stable and invariant characteristics ensure the robustness of the skill assessment framework. The skill assessment framework presented in this dissertation has shown the capacity of identifying an individual's specific weaknesses, creating a basis for customized training. The validity and generality of the assessment framework is illustrated in both the remote-control flight and laparoscopic surgery applications.

8.2 Future Directions

As discussed in the introduction, human spatial control behaviors are complex due to the involvement of all internal processes (i.e. cognitive, perceptual, and sensori-motor control) and interactions with the agent and the environment. Using the hierarchical functional model as the foundation, this dissertation has constructed a simulation model that provides a good interpretation of human behaviors. It would be interesting to elaborate this framework by incorporating more factors and identifying their influence on human planning, guidance, and tracking, to reduce the discrepancy between the model and human performance.

This dissertation has also presented a new way to investigate human spatial control behavior from visual perception perspective. Patterns exist in human gaze movement, and they exhibit different functionalities in spatial control behaviors. In this dissertation, we have confirmed that the characteristics of these gaze patterns are associated with human expertise. It would further extend our understanding of human spatial control skills by capturing the relationship between interaction patterns and these gaze patterns, such as how the gaze patterns aid in the emergence of interaction patterns.

This dissertation highlights that interaction patterns enable a dimension of atomic behavior unit and a systematic perspective for the investigation of spatial control behaviors. The gained understanding of in the work following this direction can help improve training techniques by capturing the necessary interaction patterns for the tasks or machine systems and set up sessions for human operators to acquire them. On the other hand, knowing human interaction patterns and their emergence can assist the design of human-machine systems that facilitate human engagement.

References

[1] Bérénice Mettler and Bin Li. Human Skills and the Evolving Human-Machine System. In *H-CPS-I: Human Cyber Physical System Interaction - Control for the Human Welfare*, Paris (France), 2014.

[2] Jonathan Andersh, Bin Li, and Bérénice Mettler. Modeling Visuo-Motor Control and Guidance Functions in Remote-Control Operation. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2014)*, pages 4368 – 4374, Chicago, IL, 2014. IEEE.

[3] Zhaodan Kong and Bérénice Mettler. An investigation of spatial behavior in agile guidance tasks. In *IEEE International Conference on Systems, Man, and Cybernetics*, pages 2473–2480. Ieee, oct 2011.

[4] Bin Li, Bérénice Mettler, and Timothy Kowalewski. Towards Data-Driven Hierarchical Surgical Skill Analysis. In *M2CAI*, 2014.

[5] Bérénice Mettler and Zhaodan Kong. Hierarchical model of human guidance performance based on interaction patterns in behavior. In *2nd International Conference on Application and Theory of Automation in Command and Control Systems*, volume 1, London, UK, 2012.

[6] Bin Li and Bérénice Mettler. Investigation of Hierarchical Architecture of Human Guidance Behavior for Skill Analysis. In *IEEE International Conference on Systems, Man, and Cybernetics*, 2015.

[7] Lisanne Bainbridge. Ironies of automation. *Automatica*, 19(6):775–779, nov 1983.

[8] K. S. Bibby, F. Margulies, J. E. Rijnsdorp, R. M. J. Withers, and I. M. Makarow. Man's role in control systems. In *IFAC 6th Triennial World Congress*, Boston/Cambridge, Mass., 1975.

[9] L.P. Sturm, J.A. Windsor, P.H. Cosman, P. Cregan, P.J. Hewett, and G.J. Maddern. A systematic review of skills transfer after surgical simulation training. *Annals of surgery*, 248(2):166, 2008.

[10] Samer G Mattar, Adnan A Alseidi, Daniel B Jones, D Rohan Jeyarajah, Lee L Swanstrom, Ralph W Aye, Stephen D Wexner, José M Martinez, Sharona B Ross, Michael M Awad, et al. General surgery residency inadequately prepares trainees for fellowship: Results of a survey of fellowship program directors. *Annals of surgery*, 258(3):440–449, 2013.

[11] Jamie J. Coleman, Thomas J. Esposito, Grace S. Rozycki, and David V. Feliciano. Early subspecialization and perceived competence in surgical training: Are residents ready? *Journal of the American College of Surgeons*, 216(4):764 – 771, 2013.

[12] Dennis McBride and Dylan Schmorrow. *Quantifying human information processing*. Lexington Books, 2005.

[13] George A. Miller. The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *The Psychological Review*, 63:81–97, 1956.

[14] E Bruce Goldstein. *Cognitive psychology: Connecting mind, research and everyday experience*. Nelson Education, 2014.

[15] HA Simon. Invariants of human behavior. *Annual review of psychology*, 41:1–19, 1990.

[16] Raja Parasuraman and Victor Riley. Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors*, 39(2):230–253, jun 1997.

[17] Katherine M Keetch, Richard a Schmidt, Timothy D Lee, and Douglas E Young. Especial skills: their emergence with massive amounts of practice. *Journal*

of experimental psychology. Human perception and performance, 31(5):970–978, 2005.

[18] Rajesh Kumar, Amod Jog, Anand Malpani, Balazs Vagvolgyi, David Yuh, Hiep Nguyen, Gregory Hager, and Chi Chiung Grace Chen. Assessing system operation skills in robotic surgery trainees. *The international journal of medical robotics and computer assisted surgery-MRCAS*, 8(1):118–24, mar 2012.

[19] Thomas Schack and Franz Mechsner. Representation of motor skills in human long-term memory. *Neuroscience letters*, 391(3):77–81, jan 2006.

[20] J Sweller. Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2):257–285, jun 1988.

[21] Paul M. Fitts. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology: General*, 121(3):262–269, 1992.

[22] Lisanne Bainbridge. Development of skill, reduction of workload. In *Developing Skills with Information Technology*. 1989.

[23] Dennis Harry Holding. *Human Skills*. John Wiley & Sons Incorporated, 1989.

[24] Richard A. Schmidt and Tim Lee. *Motro control and learning*. 1988.

[25] A. W. Salmoni. Motor Skill Learning. In Dennis Harry Holding, editor, *Human Skills*, chapter 8, pages 197—-227. Wiley, New York, 2 edition, 1989.

[26] Allen William Burton and Daryl E. Miller. *Movement Skill Assessment*. Human Kinematics, 1998.

[27] David A. Rosenbaum. *Human Motor Control*. 2010.

[28] George E. Stelmach. *Information processing in motor control and learning*. Academic Press, 2014.

[29] R. S. Woodworth. The accuracy of voluntary movement. *Psychological Review*, 23:1–114, 1899.

[30] E. L. Thorndike. The law of effect. *American Journal of Psychology*, 39:212–222, 1927.

[31] K. J. W. Craik. Theory of the human operator in control systems. I: The operator as an engineering system. *British Journal of Psychology*, 38:56–61, 1947.

[32] a. Tustin. The Nature of the Operator's Response in Manual Control, and its Implications for Controller Design. *Journal of the Institution of Electrical Engineers*, pages 190–206, 1947.

[33] JJ Gibson. The concept of affordances. In *Perceiving, Acting, and Knowing: Toward an Ecological Psychology*, number December 2007, pages 67–82. 1977.

[34] James J. Gibson. *The Ecological Approach to Visual Perception*. Number ISBN:0-89859-959-8. Lawrence Erlbaum, 1979.

[35] SW Keele and JJ Summers. The structure of motor programs. In *Motor control: Issues and trends*, chapter 5, pages 109–139. Press, Academic, 1976.

[36] Richard A. Schmidt. A Schema Theory of Discrete Motor Skill Learning. *Psychological review*, 82(4):225–260, 1975.

[37] B. Mettler. Structure and Organizational Principles of Agile Behavior: Challenges and Opportunities in Cognitive Engineering. *Cognitive Critique*, 3, 2011.

[38] Jonathan Andersh, Bin Li, and Berenice Mettler. Experimental investigation of the visuo-motor in remote-control operation. In *submitted to the American Control Conference*, 2014.

[39] Michelle L. Pratt, Hayley C. Leonard, Hanna Adeyinka, and Elisabeth L. Hill. The effect of motor load on planning and inhibition in developmental coordination disorder. *Research in Developmental Disabilities*, 35(7):1579–1587, 2014.

[40] E. R. F. W. Crossman. A theohy of the acquisition of speed-skill. *Ergonomics*, 2(2):153–166, 1959.

[41] Allen Newell and Paul S Rosenbloom. Mechanisms of skill acquisition and the law of practice. *Cognitive skills and their acquisition*, 6:1–55, 1981.

[42] Frank E Ritter and Lael J Schooler. The learning curve. *International encyclopedia of the social and behavioral sciences*, 13:8602–8605, 2001.

[43] Federal Aviation Administration. *Airplane Flying Handbook*. Flight Standards Service, 2004.

[44] The American Board of Surgery. Booklet of Information—Surgery. 2015.

[45] Ann M Gentile. A working model of skill acquisition with application to teaching. *Quest*, 17(1):3–23, 1972.

[46] DT McRuer and ES Krendel. The human operator as a servo system element. *Journal of the Franklin Institute*, 267(5):381–403, 1959.

[47] ES Krendel and DT McRuer. A servomechanisms approach to skill development. *Journal of the Franklin Institute*, 1(269):24–42, 1960.

[48] Stuart E. Dreyfus. The Five-Stage Model of Adult Skill Acquisition. *Bulletin of Science, Technology and Society*, 24(3):177–181, jun 2004.

[49] Michael Bajka, Stefan Tuchschmid, Daniel Fink, Gábor Székely, and Matthias Harders. Establishing construct validity of a virtual-reality training simulator for hysteroscopy via a multimetric scoring system. *Surgical endoscopy*, 24(1):79–88, jan 2010.

[50] Paul M. Fitts. Cognitive Aspects of Information Processing: III. Set for Speed versus Accuracy. *Journal of experimental psychology*, 71(6):849–857, 1966.

[51] Bérénice Mettler, Zhaodan Kong, Bin Li, and Jonathan Andersh. Systems view on spatial planning and perception based on invariants in agent-environment dynamics. *Frontiers in neuroscience*, 8(439):1–14, 2015.

[52] Jacques Leplat. Cognitive skills at work. In *Developing Skills with Information Technology*. 1989.

[53] Daiki Suzuki and Kouhei Ohnishi. A skill proficiency evaluation method of individual based on accuracy and repeatability. In *IECON 2012 - 38th Annual Conference on IEEE Industrial Electronics Society*, pages 2144–2149, 2012.

[54] A. T. Welford. On the nature of skill. *Skill*, pages 21–32, 1970.

[55] William Thomas Singleton. The analysis of practical skills. *University Park Press*, 1:10, 1978.

[56] K. M. Newell. Motor skill acquisition. *Annual review of psychology*, 42:213–37, jan 1991.

[57] Zhaodan Kong. *A Formal Investigation of the Organization of Guidance Behavior: Implications for Humans and Autonomous Guidance*. Thesis or dissertation, University of Minnesota, 2012.

[58] J.A.G. Van Avermaete. *NOTECHS: Non-technical skill evaluation in JAR-FCL*. 1998.

[59] Steven Yule, Rhona Flin, Nicola Maran, David Rowley, George Youngson, and Simon Paterson-Brown. Surgeons' non-technical skills in the operating room: reliability testing of the NOTSS behavior rating system. *World journal of surgery*, 32(4):548–56, apr 2008.

[60] D. Hebb. *The Organization Of Behaviour: A Neuropsychological Theory*. Wiley, New York, New York, USA, 1949.

[61] N. A. Bernstein. *The co-ordination and regulation of movements*. Pergamon Press, New York, New York, USA, 1967.

[62] Mark L Latash, John P Scholz, and Gregor Schöner. Toward a new theory of motor synergies. *Motor control*, 11(3):276–308, 2007.

[63] Fady Alnajjar, Tytus Wojtara, Hidenori Kimura, and Shingo Shimoda. Muscle synergy space: learning model to create an optimal muscle synergy. *Frontiers in computational neuroscience*, 7(October):136, 2013.

[64] a. Mark Williams, Paul Ward, John M. Knowles, and Nicholas J. Smeeton. Anticipation skill in a real-world task: Measurement, training, and transfer in tennis. *Journal of Experimental Psychology: Applied*, 8(4):259–270, 2002.

[65] Paul Ward and AM Williams. Perceptual and cognitive skill development in soccer: The multidimensional nature of expert performance. *Journal of Sport and Exercise Psychology*, 25:93–111, 2003.

[66] Duane Millslagle. *Motor learning power point presentations in Magill's Motor Learning*. Macgraw-Hill Publishing, 7 edition, 2000.

[67] Vincent Grosjean and Patrice Terrier. Temporal awareness: pivotal in performance? *Ergonomics*, 42(11):1443–1456, 1999.

[68] Willem B Verwey, Christian Peham, and Wolfgang I Schöllhorn. Preface to the EWOMS special issue of Human Movement Science. *Human movement science*, 26(2):175–9, apr 2007.

[69] W. Sturm, A. de Simone, B.J. Krause, K. Specht, I. Hesselmann, V. Radermacher, H. Herzog, L. Tellmann, H.W. Müller-Gärtner, and K. Willmes. Functional anatomy of intrinsic alertness: evidence for a fronto-parietal-thalamic-brainstem network in the right hemisphere. *Neuropsychologia*, 37(7):797–805, 1999.

[70] Eric I. Knudsen. Fundamental Components of Attention. *Annual Review of Neuroscience*, 30(1):57–78, 2007.

[71] John L. Horn. Organization of Abilities and the Development of Intelligence. *Psychological Review*, 75(3):242–259, 1968.

[72] Derek T Y Mann, a Mark Williams, Paul Ward, and Christopher M Janelle. Perceptual-cognitive expertise in sport: a meta-analysis. *Journal of sport & exercise psychology*, 29(4):457–78, aug 2007.

[73] Nelson Cowan. What are the differences between long-term, short-term, and working memory? *Progress in Brain Research*, 169(07):323–338, 2008.

[74] Lisanne Bainbridge. Types of representation. In *Tasks, errors, and mental models*, pages 70–91. Taylor & Francis, Inc., 1988.

[75] Lisanne Bainbridge. Analysis of verbal protocols from a process control task. In *The Human Operator in Process Control*, pages 1–58. 1974.

[76] Lisanne Bainbridge. The representation of working storage, and its use in the organisation of behaviour. *Measurement of human resources*, pages 165–183, 1975.

[77] R. E. Bellman. Curse of dimensionality. In *Adaptive control processes: a guided tour.* Princeton, NJ, 1961.

[78] Toshikazu Akita, Tatsuya Suzuki, and Soichiro Hayakawa. Analysis and synthesis of driving behavior based on mode segmentation. In *International Conference on Control, Automation and Systems*, pages 2884–2889, 2008.

[79] Bennet B. Murdock. Recent Developments in Short-Term Memory. *British Journal of Psychology*, 58(3-4):421—-433, 1967.

[80] Richard C. Atkinson and Richard M. Shiffrin. Human memory: A proposed system and its control processes. *The psychology of learning and motivation*, 2:89–195, 1968.

[81] N Cowan. Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological bulletin*, 104(2):163–91, sep 1988.

[82] Russell Revlin. *Cognition: Theory and Practice*. Palgrave Macmillan, 2012.

[83] Gordon D a Brown, I a N Neath, and Nick Chater. A ratio model of scale invariant memory and identification. *Memory Lab Technical Report*, pages 1–94, 2002.

[84] Eugen Tarnow. Properties of the Short Term Memory Structure. *Time*, pages 1–8, 2005.

[85] W. G. Chase and H. A. Simon. Perception in chess. *Cognitive psychology*, 4(1):55–81, 1973.

[86] Mitsuo Kawato. Internal models for motor control and trajectory planning. *Motor systems*, pages 718–727, 1999.

[87] Lisanne Bainbridge. Mathematical equations or processing routines? *Human detection and diagnosis of system failures*, 15:259–286, 1981.

[88] S Suzuki and N Tomomatsu. Skill evaluation based on state-transition model for human adaptive mechatronics (HAM). In *The 30th Annual Conference of the IEEE Industrial Electronics Society*, pages 641–646, 2004.

[89] J. G. March. Exploration and Exploitation in Organizational Learning. *Organization Science*, 2(1):71–87, 1991, z0009.

[90] Robert E. Shaw and Thomas R. Alley. How to draw learning curves: Their use and justification. In *Issues in the ecological study of learning*, pages 275—-304. 1985.

[91] Richard M. Shiffrin and Walter Schneider. Controlled and Automatic Human Information Processing: II. Perceptual Learning, Automatic Attending, and a General Theory. *Psychological Review*, 84(2):127—-190, 1977.

[92] R W Pew. Acquisition of hierarchical control over the temporal organization of a skill. *Journal of Experimental Psychology*, 71(5):764–771, 1966.

[93] David A. Rosenbaum. *Human Motor Control*. Academic press, 2 edition, 2009.

[94] Steven W. Keele. Movement control in skilled motor performance. *Psychological Bulletin*, 70(6):387–403, 1968.

[95] Jack A. Adams. A closed-loop theory of rnotor learning. *Journal of motor behavior*, 3(2):111—-150, 1971.

[96] John R. Anderson. A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, 22(3):261–295, 1983.

[97] KC Anderson and Gaea Leinhardt. Maps as representations: Expert novice comparison of projection understanding. *Cognition and Instruction*, 20(3):283–321, 2002.

[98] Andrea D'Avella, Martin Giese, Yuri P. Ivanenko, Thomas Schack, and Tamar Flash. Editorial: Modularity in motor control: from muscle synergies to cognitive action representation. *Frontiers in Computational Neuroscience*, 9(October):1–6, 2015.

[99] J. E. Laird, P. S. Rosenbloom, and Allen Newell. Towards Chunking as a General Learning Mechanism. *AAAI*, August:188–192, 1984.

[100] Jordi Fonollosa, Emre Neftci, and Mikhail Rabinovich. Learning of Chunking Sequences in Cognition and Behavior. *PLOS Computational Biology*, 11(11):e1004592, 2015.

[101] K. M. Newell and R. E. A. van Emmerik. The acquisition of coordination: Preliminary analysis of learning to write. *Human Movement Science*, 8(1):17–32, feb 1989.

[102] Matthew C. Tresch and Anthony Jarc. The case for and against muscle synergies. *Current Opinion in Neurobiology*, 19(6):601, 2009.

[103] John P. Scholz and Gregor Schöner. The uncontrolled manifold concept: Identifying control variables for a functional task. *Experimental Brain Research*, 126(3):289–306, 1999.

[104] Ml Mark L Latash, Mf Mindy F Levin, John P Scholz, and Gregor Schöner. Motor Control Theories and Their Applications. *Medicina (Kaunas, ...*, 46(6):382–92, 2010.

[105] Francisco J Valero-Cuevas, Madhusudhan Venkadesan, and Emanuel Todorov. Structured variability of muscle activations supports the minimal intervention principle of motor control. *Journal of neurophysiology*, 102:59–68, 2009.

[106] Willam Lowe Bryan and Noble Harter. Studies on the telegraphic language: the acquisition of a hierarchy of habits. *The Psychological Review*, VI(4):345–375, 1899.

[107] Donald G. Mackay and Robert W. Bowman. On producing the meaning in sentences. *The American Journal of Psychology*, pages 23–39, 1969.

[108] Carol E Reiley and Gregory D Hager. Task versus subtask surgical skill evaluation of robotic minimally invasive surgery. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, volume 12, pages 435–42, jan 2009.

[109] Carol E Reiley, Henry C Lin, David D Yuh, and Gregory D Hager. Review of methods for objective surgical skill evaluation. *Surgical endoscopy*, 25(2):356–66, feb 2011.

[110] KW Fischer. A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological review*, 87(6):477–531, 1980.

[111] Jens Rasmussen. Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-13(3):257–266, 1983.

[112] Timothy Kowalewski, Lee White, John Carlis, Erlan Santos, Robert Sweet, Troy Reihsen, and Daniel F Keefe. Exploratory Visualization of Surgical Training Databases for Improving Skill Acquisition. *IEEE Computer Graphics and Applications*, 32(6):71–81, 2012.

[113] A M Derossis, G M Fried, M Abrahamowicz, H H Sigman, J S Barkun, and J L Meakins. Development of a model for training and evaluation of laparoscopic skills. *American journal of surgery*, 175(6):482–487, jun 1998.

[114] Yasushi Yamauchi, Juli Yamashita, Osamu Morikawa, Ryoichi Hashimoto, Masaaki Mochimaru, Yukio Fukui, Hiroshi Uno, and Kazunori Yokoyama. Surgical Skill Evaluation by Force Data for Endoscopic Sinus Surgery Training System. In Takeyoshi Dohi and Ron Kikinis, editors, *Medical Image Computing and Computer-assisted Intervention*, volume 2488 of *Lecture Notes in Computer Science*, pages 44–51. Springer Berlin / Heidelberg, 2002.

[115] Ling Zhang, Beatriz C Abreu, Gary S Seale, Brent Masel, Charles H Christiansen, and Kenneth J Ottenbacher. A virtual reality environment for evaluation of a daily living skill in brain injury rehabilitation: reliability and validity. *Archives of Physical Medicine and Rehabilitation*, 84(8):1118–1124, aug 2003.

[116] Gideon Sroka, Liane S. Feldman, Melina C. Vassiliou, Pepa a. Kaneva, Raad Fayez, and Gerald M. Fried. Fundamentals of Laparoscopic Surgery simulator training to proficiency improves laparoscopic performance in the operating room-a randomized controlled trial. *American Journal of Surgery*, 199(1):115–120, 2010.

[117] Kalevi Tervo, Lauri Palmroth, and Aki Putkonen. A hierarchical fuzzy inference method for skill evaluation of machine operators. In *IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, pages 136–141. Ieee, jul 2009.

[118] J. Durbin. Efficient Estimation of Parameters in Moving-Average Models. *Biometrika*, 46(3/4):306–316, 1959.

[119] Chen, Chennuru, and Zhang. A Language Approach to Modeling Human Behavior. *Citeseer*, (1966):3020–3023, 1998.

[120] W.W. Kong and Surendra Ranganath. Automatic hand trajectory segmentation and phoneme transcription for sign language. In *2008 8th IEEE International Conference on Automatic Face & Gesture Recognition*, pages 1–6. IEEE, sep 2008.

[121] Elizabeth B Torres, Maria Brincker, Robert W Isenhower, Polina Yanovich, Kimberly a Stigler, John I Nurnberger, Dimitris N Metaxas, and Jorge V José. Autism: the micro-movement perspective. *Frontiers in integrative neuroscience*, 7(July):32, jan 2013.

[122] F Meier, E Theodorou, F Stulp, and S Schaal. Movement segmentation using a primitive library. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3407–3412, San Francisco, CA, USA, 2011. IEEE.

[123] Jo a Helmuth, Christoph J Burckhardt, Petros Koumoutsakos, Urs F Greber, and Ivo F Sbalzarini. A novel supervised trajectory segmentation algorithm identifies distinct types of human adenovirus motion in host cells. *Journal of structural biology*, 159(3):347–58, sep 2007.

[124] D R Faria and J Dias. 3D hand trajectory segmentation by curvatures and hand orientation for classification through a probabilistic approach. In *Intelligent Robots and Systems 2009 IROS 2009 IEEERSJ International Conference on*, pages 1284–1289. IEEE, 2009.

[125] R Mann, A D Jepson, and T El-Maraghi. Trajectory segmentation using dynamic programming. *Object recognition supported by user interaction for service robots*, (1):331–334, 2002.

[126] J J Faraway, M P Reed, and J Wang. Modelling three-dimensional trajectories by using Bézier curves with application to hand motion. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 56(5):571–585, 2007.

[127] Giancarlo Ferrari-trecate, Marco Muselli, Diego Liberati, and Manfred Morari. A clustering technique for the identification of piecewise affine systems. *Automatica*, 39:205–217, 2003.

[128] Bin Li and Bérénice Mettler. Application of trajectory segmentation techniques for operator skill evaluation. In *American Control Conference (ACC), 2013*, pages 2680–2686, 2013.

[129] Jaron Blackburn and Eraldo Ribeiro. Human motion recognition using Isomap and dynamic time warping. *Human MotionUnderstanding, Modeling, Capture and Animation*, 4814/2007:285–298, 2007.

[130] J.K. Aggarwal and M.S. Ryoo. Human activity analysis. *ACM Computing Surveys*, 43(3):1–43, apr 2011.

[131] Todd E. Murphy, Chet M. Vignes, David D. Yuh, and Allison M. Okamura. Automatic Motion Recognition and Skill Evaluation for Dynamic Tasks. *Proc. Eurohaptics 2003*, pages 363–373, 2003.

[132] Henry C Lin, Izhak Shafran, David Yuh, and Gregory D Hager. Towards automatic skill evaluation: detection and segmentation of robot-assisted surgical motions. *Computer aided surgery : official journal of the International Society for Computer Aided Surgery*, 11(5):220–30, sep 2006.

[133] Thomas Mackel, Jacob Rosen, and Carla Pugh. Application of hidden markov modeling to objective medical skill evaluation. *Studies in health technology and informatics*, 125:316–8, jan 2007.

[134] Kalevi Tervo, Lauri Palmroth, and Heikki Koivo. Skill evaluation of human operators in partly automated mobile working machines. *IEEE Transactions on Automation Science and Engineering*, 7(1):133–142, 2010.

[135] Bérénice Mettler, Navid Dadkhah, Zhaodan Kong, and Jonathan Andersh. Research Infrastructure for Interactive Human- and Autonomous Guidance. *Journal of Intelligent and Robotic Systems.*, 2012.

[136] Bérénice Mettler, Navid Dadkhah, Zhaodan Kong, and Jonathan Andersh. Infrastructure and Research Framework for Interactive Human and Autonomous Guidance. In *International Conference on Unmanned Aircraft Systems*, 2012.

[137] T.M. Kowalewski. *Real-time Quantitiative Assessment of Surgical Skill*. PhD thesis, May 2012.

[138] Susana Martinez-Conde, Stephen L Macknik, and David H Hubel. The role of fixational eye movements in visual perception. *Nature reviews. Neuroscience*, 5(3):229–240, mar 2004.

[139] DA Robinson. The mechanics of human smooth pursuit eye movement. *The Journal of Physiology*, 180(3):569, 1965.

[140] DA Robinson. The mechanics of human saccadic eye movement. *The Journal of physiology*, 174:245–264, 1964.

[141] Stephen J Huston and Holger G Krapp. Visuomotor transformation in the fly gaze stabilization system. *PLoS biology*, 6(7):1468–1478, jul 2008.

[142] S R Eddy. Hidden Markov models. *Current opinion in structural biology*, 6(3):361–5, jun 1996.

[143] Lydia E. Kavraki. Random Networks In Configuration Space For Fast Path Planning. *IEEE Transactions on Robotics and Automation*, 12(4):566–580, 1994.

[144] LE Kavraki, P Svestka, J.-C. Latombe, and M.H. Overmars. Probabilistic roadmaps for path planning in high-dimensional configuration spaces. *IEEE Transactions on Robotics and Automation*, 12(4):566–580, 1996.

[145] Lydia E. Kavraki, Mihail N. Kolountzakis, and Jean-claude Latombe. Analysis of Probabilistic Roadmaps for Path Planning. *IEEE Transactions on Robotics and Automation*, 14(1):166–171, 1998.

[146] S.S. Ge and Y.J. Cui. New potential functions for mobile robot path planning. *IEEE Transactions on Robotics and Automation*, 16(5):615–620, 2000.

[147] Jerome Barraquand. Numerical potential field techniques for robot path planning. *IEEE Transactions on Systems, Man and Cybernetics*, 22(2):224–241, 1992.

[148] David Hsu, Jean-Claude Latombe, and Rajeev Motwani. Path Planning in Expansive Configuration Spaces. *International Journal of Computational Geometry & Applications*, 9(4 & 5):495–512, aug 1999.

[149] Tomás Lozano-Pérez and Michael A Wesley. An algorithm for planning collision-free paths among polyhedral obstacles. *Communications of the ACM*, 22(10):560–570, 1979.

[150] Franz Aurenhammer. Voronoi diagramsa survey of a fundamental geometric data structure. *ACM Computing Surveys (CSUR)*, 23(3):345–405, 1991.

[151] Shuzhi Sam Ge and Yan Juan Cui. New potential functions for mobile robot path planning. *IEEE Transactions on robotics and automation*, 16(5):615–620, 2000.

[152] Steven M LaValle and James J Kuffner. Randomized kinodynamic planning. *The International Journal of Robotics Research*, 20(5):378–400, 2001.

[153] Kamal Gupta and Angel P Pobil. *Practical motion planning in robotics: Current approaches and future directions*. John Wiley & Sons, Inc., 1998.

[154] Jean-Claude Latombe. *Robot motion planning*, volume 124. Springer Science & Business Media, 2012.

[155] Steven M LaValle. *Planning algorithms*. Cambridge university press, 2006.

[156] Keisuke Sato. Global motion planning using a laplacian potential field. *Journal of the Robotics Society of Japan*, 11(5):702–709, 1993.

[157] Stéphane Ross, Joëlle Pineau, Sébastien Paquet, and Brahim Chaib-draa. Online planning algorithms for POMDPs. *Journal of Artificial Intelligence Research*, 32:663–704, 2008, NIHMS150003.

[158] Bérénice Mettler and Olivier Toupet. Receding Horizon Trajectory Planning with an Environment-Based Cost-to-go Function. *Proceedings of the 44th IEEE Conference on Decision and Control*, pages 4071–4076, 2005.

[159] Bérénice Mettler and Zhaodan Kong. Receding horizon trajectory optimization with a finite-state value function approximation. In *American Control Conference*, pages 3810–3816. Ieee, jun 2008.

[160] James J. Gibson. *The Ecological Approach to Visual Perception.* Houghlon Mifflin, Boston, MA, 1979.

[161] Zhaodan Kong and Bérénice Mettler. Modeling human guidance behavior based on patterns in agent-environment interactions. *IEEE Transaction on Human-Machine Systems*, 43(4):371–384, 2013.

[162] Duane T. McRuer and Henry R. Jex. A Review of Quasi-Linear Pilot Models. *IEEE Transactions on Human Factors in Electronics*, HFE-8(3):231–249, 1967.

[163] François Petitjean, Alain Ketterlin, and Pierre Gançarski. A global averaging method for dynamic time warping, with applications to clustering. *Pattern Recognition*, 44:678–693, 2011.

[164] Taras K Vintsyuk. Speech discrimination by dynamic programming. *Cybernetics and Systems Analysis*, 4(1):52–57, 1968.

[165] J.E. De Camp. Relative distance as a factor in the white rat's selection of a path. *Psychobiology*, 2(3):245–253, 1920.

[166] C. Goerzen, Z. Kong, and Bérénice Mettler. *A Survey of Motion Planning Algorithms from the Perspective of Autonomous UAV Guidance*, volume 57. nov 2009.

[167] Mark De Berg, Marc Van Kreveld, Mark Overmars, and Otfried Cheong Schwarzkopf. Visibility Graphs. In *Computational Geometry*, chapter 15, pages 323–333. 2000.

[168] David Lee. Guiding Movement by Coupling Taus. *Ecological Psychology*, 10(3):221–250, sep 1998.

[169] GD Padfield, DN Lee, and Roy Bradley. How do helicopter pilots know when to stop, turn or pull up?(Developing guidelines for vision aids). *Journal of the American Helicopter Society*, 48(2):108–119, 2003.

[170] Gareth D Padfield, Linghai Lu, and Michael Jump. Tau Guidance in Boundary-Avoidance Tracking - New Perspectives on Pilot-Induced Oscillations. *Journal of Guidance, Control, and Dynamics*, 35(1):80–92, jan 2012.

[171] William Gray. Boundary-Escape Tracking: A New Conception of Hazardous PIO. Technical Report 0704, 2004.

[172] Jonathan Andersh, Bin Li, and Bérénice Mettler. Experimental Investigation of Visuo-Motor Functions in Remote-Control Operation. In *American Control Conference*, 2014.

[173] GD Padfield, Gary Clark, and Armin Taghizad. How Long Do Pilots Look Forward? Prospective Visual Guidance in TerrainHugging Flight. *Journal of the American Helicopter Society*, 52(2):134–145, 2007.

[174] Ian D Loram, Cornelis Van De Kamp, Henrik Gollee, and Peter J Gawthrop. Identification of intermittent control in man and machine. *Journal of the Royal Society Interface*, 9(74):2070–2084, 2012.

[175] Zhaodan Kong and Bérénice Mettler. Foundations of formal language for humans and artificial systems based on intrinsic structure in spatial behavior. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3093–3100. Ieee, sep 2011.

[176] Ryota Terada, Hiroyuki Okuda, and Tatsuya Suzuki. Multi-scale driving behavior modeling using hierarchical PWARX model. In *13th International IEEE Annual Conference on Intelligent Transportation Systems*, pages 1638–1644, 2010.

[177] R Tibshirani, G Walther, and T Hastie. Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B*, 63(2):411–423, 2001.

[178] Benjamin Law, M. Stella Atkins, A. E. Kirkpatrick, Alan J. Lomax, and Christine L. Mackenzie. Eye gaze patterns differentiate novice and experts in a virtual laparoscopic surgery training environment. In *Proceedings of the 2004 symposium on Eye tracking research & applications*, pages 41–48, 2004.

[179] Frouke Hermens, Rhona Flin, and Irfan Ahmed. Eye movements in surgery : A literature review. *Journal of Eye Movement Research*, 6(4):1–11, 2013.

[180] Mary Hayhoe and Dana Ballard. Eye movements in natural behavior. *Trends in cognitive sciences*, 9(4):188–94, apr 2005.

[181] J T Massey, J T Lurito, G Pellizzer, and a P Georgopoulos. Three-dimensional drawings in isometric conditions: relation between geometry and kinematics. *Experimental brain research. Experimentelle Hirnforschung. Experimentation cerebrale*, 88(3):685–690, 1992.