TECHNISCHE UNIVERSITÄT MÜNCHEN
Lehrstuhl für Steuerungs- und Regelungstechnik
Fachgebiet Informationstechnische Regelung

Dynamic Models of Human Perception and Action and Their Application in Telepresence

E. Markus Rank

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs (Dr.-Ing.)

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr.-Ing. habil. Gerhard Rigoll

Prüfer der Dissertation:

1. Univ.-Prof. Dr.-Ing. Sandra Hirche

2. Univ.-Prof. Dr. phil. habil. Hermann Müller
   Ludwig-Maximilians-Universität München

Die Dissertation wurde am 11. Juni 2012 bei der Technischen Universität München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 6. November 2012 angenommen.

Foreword

This thesis summarizes work I have been doing during the last 4 years at the Institute of Automatic Control Engineering (LSR) at Technische Universität München. During this time, I had the pleasure to meet and work with a number of excellent researchers who influenced, inspired, and supported me. I would like to express my gratitude to these people, as without their invaluable help this thesis would not have been possible. First of all I want to thank my advisor Prof. Sandra Hirche for the discussions we had, the discourses about the most promising research directions, and for making me aim at nothing less than the optimum that can be achieved in every aspect of scientific life. In addition, I want to thank Prof. Martin Buss for leading an institute that allows a rich and multidisciplinary scientific exchange between a large group of PhD students, PostDocs, visitors, and guests. Among the guests I enjoyed working with most is Prof. Roberta L. Klatzky (Carnegie Mellon University, Pittsburgh) who helped in the development of the experiments investigating haptic masking discussed in Chapter 5. The collision avoidance algorithm presented in Chapter 6 founds on a method developed by Dr. Matthias Althoff who provided me with MATLAB codes and supported my efforts in this complex field with helpful advices.

The thesis discusses problems arising in the context of telepresence systems which is the main theme of the collaborative research center SFB 453 I was allowed to work in, funded by the Deutsche Forschungsgemeinschaft (DFG). I deeply enjoyed the interdisciplinary work on perceptual phenomena I did together with Prof. Hermann Müller, Dr. Zhuanghua Shi, Heng Zhou, Lihan Chen and other researchers and doctoral students at the General and Experimental Psychology Institute at Ludwig-Maximilians-Universität München. In the follow-up project, the partners I was working with were Dr. Angelika Peer, Dr. Jörg Reisinger, Ken Friedl and Wenliang Zhou – it was a pleasure working with you.

During my PhD times I had the opportunity to visit the labs of Dr. Mandayam Srinivasan (Massachusetts Institute of Technology, Cambridge), and Hong Z. Tan (Purdue University, West Lafayette). This exchange was a valuable step in the development of the main ideas for the thesis as it is now, and I am grateful for the hospitality I was allowed to experience.

I want to thank colleagues and friends which not only helped in scientific questions, but made life at LSR an enjoyable and diverting one. Thomas Schauß, Iason Vittorias, Raphaela Groten, Georg Bätz, Ulrich Unterhinninghofen, Daniel Althoff, Bernhard Weber, Daniela Feth, and my roommates Stefan Klare and Sheraz Khan – thank you for your friendship.

The biggest thanks go to my wife Sophia, my parents and siblings for their loving and enduring support in all concerns.

Munich, December 2012 Markus Rank

to my family...

Abstract

Haptic telepresence systems extend the human workspace to locations that are normally inaccessible, such as under water or in space. For this purpose the operator of a telepresence system physically interacts with a robotic human-system interface capturing his/her actions. Those are transmitted over a communication channel to a remote environment where a second robotic device, the teleoperator, performs the action and collects haptic information. Sent back to the human side, the information is displayed to the human operator by the human-system interface. For ideal telepresence, the human should not be able to tell if he/she interacts directly with a physical environment, or uses a telepresence system. Achieving this ultimate goal requires haptic control technology meeting the human perceptual and motor capabilities exactly enough to keep inevitable imperfectness in sensory feedback imperceptible and natural physical interaction capabilities unaffected. For the development of haptic interaction devices, control and communication algorithms capable of achieving these requirements, the operator's perceptual limits and his/her motor capabilities must be known in the form of quantitative models. Most current perception models are static, neglecting the fact that the characteristics of interaction movement and force can influence the perception limits significantly. In this way, their validity is proven only for the experimental conditions they were determined for. Because one fundamental goal of telepresence is to enable the human operator to freely move and interact, static human models are insufficient to reliably design, parametrize and evaluate haptic telepresence systems in a human-centered way.

In this thesis, dynamic models of human perception and action and their application in telepresence systems are investigated. The prediction capabilities of these novel human models are not limited to static mappings, but can take the physical interaction with the human-system interface explicitly into consideration for the prediction of perceptual limits as differential equations are taken as a modeling form. An innovative mathematical framework allows a systematic development of quantitative dynamic models being suited for a direct application in technical systems. On the basis of this framework, two dynamic perception models are developed. Firstly, the temporal combination of movement and force feedback into a percept of a haptic property, such as inertia or stiffness is investigated in a large number of psychophysical experiments. This process is found to significantly depend on the characteristics of physical interaction. Secondly, haptic masking effects are considered, revealing that perception limits of a haptic property can be influenced by an unrelated haptic stimulus. In both cases, dynamic perception models capture the observed effects with a superior accuracy compared to static mappings. The utilization of dynamic human models in the optimization of a telepresence system's control is demonstrated by means of a novel communication quality control algorithm. An optimal online-regulation scheme of communication time delay is developed, based on a stochastic human behavior model. In extensive user studies, the benefit of this approach for the human operator in terms of improved task performance is demonstrated.

Zusammenfassung

Haptische Telepräsenzsysteme erlauben es einem menschlichen Bediener, Aufgaben in ansonsten unzugänglichen Umgebungen auszuführen, zum Beispiel unter Wasser oder im Weltraum. Im Gegensatz zu einer direkten Interaktion mit der Umgebung operiert der/die Benutzer/in eines Telepräsenzsystems ein robotisches haptisches Eingabegerät, welches seine/ihre Aktionen registriert und über einen Kommunikationskanal in die entfernte Umgebung überträgt. Dort führt ein zweiter robotischer Manipulator die gewünschte Aktion aus und sendet Sensorinformation zurück, welche dem Menschen über die Mensch-Maschine Schnittstelle dargeboten werden. Für ein ideales Telepräsenzerlebnis darf der Mensch nicht mehr unterscheiden können, ob er direkt oder über das Telepräsenzsystem mit der Umgebung verbunden ist. Zum Erreichen dieses Ziels muss das technische System sehr exakt auf die menschlichen Bedürfnisse abgestimmt sein. Einerseits müssen unvermeidbare Störungen in den sensorischen Informationen unterhalb der Wahrnehmungsschwellen bleiben, andererseit muss die natürliche Interaktionsfähigkeit des Operators in vollem Umfang erhalten bleiben. Die meisten heute verfügbaren mathematischen Perzeptionsmodelle, welche die Anforderungen des Menschen an das System quantifizieren können, bestehen lediglich aus statischen Abbildungen von Umgebungsparametern zu einer einfachen Wahrnehmungsgröße. Dadurch bleibt unberücksichtigt, auf welche Weise mit der Umgebung physikalisch interagiert wird, was erwiesenermaßen einen signifikanten Einfluss auf die Wahrnehmung selbst haben kann. Der Geltungsbereich existierender Perzeptionsmodelle ist somit auf spezielle Interaktionsbedingungen limitiert – eine Einschränkung, die für eine zuverlässige und quantitative Berücksichtigung des Menschen in Entwurf, Parametrierung und Evaluation haptischer Telepräsenzsysteme nicht hinnehmbar ist.

Die vorliegende Arbeit befasst sich mit der Beschreibung menschlicher Wahrnehmung und Interaktion durch eine dynamische Modellierungsform, die auf Differentialgleichungen basiert. Dadurch ist die Aussagekraft der neu entwickelten Modelle nicht mehr auf statische Abbildungen limitiert, sondern kann die tatsächliche physikalische Interaktion mit der Umgebung bei der Bestimmung von Wahrnehmungscharakteristika berücksichtigen. Ein innovatives mathematisches Rahmenkonzept wird präsentiert, das die praxisorientierte Entwicklung dynamischer, quantitativer Wahrnehmungs- und Verhaltensmodelle methodisch unterstützt. Auf dieser Basis und untermauert durch zahlreiche psychophysische Untersuchungen am Menschen werden zwei konkrete Wahrnehmungsphänomene untersucht: Zunächst steht die Beschreibung des Kombinationsprozesses von Kraft und Bewegung in das haptische Empfinden eines Umgebungseindrucks von z.B. Trägheit oder Steifigkeit im Vordergrund. Anschließend werden Maskierungseffekte untersucht, die die Verdeckung eines bestimmten haptischen Eindrucks durch einen zweiten Reiz, der in keinem Zusammenhang zum eigentlichen Stimulus steht, modelliert. Beide Phänomene werden durch die neu entwickelten dynamischen Modellen besser und genauer erfasst als durch statische Abbildungen. Die Anwendbarkeit dynamischer Menschmodelle zur Optimierung haptischer Telepräsenzsysteme wird im Zusammenhang mit der Entwicklung eines innovativen Kommunikationskonzepts demonstriert, das die Regulierung der Kanalqualität im laufenden Betrieb erlaubt. Ein dynamisches Optimierungsverfahren, basierend auf einem stochastischen Verhaltensmodell bestimmt die zeitvarianten Kommunikationsparameter, welche es dem/der Bediener/in ermöglichen, eine hohe Leistungsfähigkeit zu erreichen. Dies wird durch ausführliche Nutzerstudien gezeigt.

Contents

Notations

Abbreviations

EEG	electroencephalography
fMRI	functional magnetic resonance imaging
LTIH	linear, time-invariant spatially homogeneous
DoF	degrees of freedom
PD	proportional-derivative
HSI	human-system interface
TO	teleoperator
JND	just noticeable difference
DT	detection threshold
PSE	point of subjective equality
PF	psychometric function

Conventions

Scalars, Vectors, and Matrices
Scalars are denoted by lower case letters in italic type. *Vectors* are denoted by bold lower case letters in italic type, as the vector $\mathbf{x}$ is composed of elements x_i. *Matrices* are denoted by upper case letters in italic type, as the matrix K is composed of elements k_{ij} (i^{th} row, j^{th} column), *frequency transformed variables* are denoted by upper case letters in italic type, followed by a $(j\omega)$.

x or X	scalar
$\mathbf{x}$	vector
X	matrix
X^T	transposed of X
X^{-1}	inverse of X
$f(\cdot)$	scalar function
$\mathbf{f}(\cdot)$	vector function
$\hat{x}$	estimated or predicted value of x
x^*	a specific value of x
$\dot{x}, \ddot{x}, \ldots$	first, second, ... time derivative of x
$X(j\omega)$	frequency transform of $x(t)$

Subscripts and Superscripts

x_0	initial value at time $t = 0$
$x_h(t)$, $\dot{x}_h(t)$, $\ddot{x}_h(t)$	position, velocity, acceleration of interaction point between human and environm
$f_h(t)$	interaction force at interaction point between human and environment
$(\cdot)^a$	associated with the auditive modality
$(\cdot)^v$	associated with the visual modality
$(\cdot)^h$	associated with the haptic modality
$(\cdot)_{env}$	environment-related
$(\cdot)_{phy}$	associated with the physical domain
$(\cdot)_{se}$	associated with the sensory domain
$(\cdot)_{ne}$	associated with the neural domain
$(\cdot)_{con}$	associated with the sensorimotor control domain
$(\cdot)_{perc}$	associated with the perceptual domain
$(\cdot)_{mo}$	associated with the motor domain
$(\cdot)_{und}$	non-delayed
$(\cdot)_{\circlearrowleft}$	closed-loop
$(\cdot)_{sum}$	subordinate
$(\cdot)_{dom}$	dominating
$(\cdot)^{perc}$	perceived value
$(\cdot)_{sim}$	simultaneous
$(\cdot)_{temp}$	temporal
$(\cdot)_{acc}$	accumulated
$(\cdot)_x$	in x-direction
$(\cdot)_y$	in y-direction
$(\cdot)_{perf}$	performance-related
$(\cdot)_{cost}$	cost-related
$(\cdot)^{pos}$	only position components

Symbols

General

$\mathbf{u}(t)$	system input
$\mathbf{y}(t)$	system output
$\mathbf{x}(t)$	system state
t	time
$\Psi(\cdot)$	nonlinear state space equation
$\Phi(\cdot)$	nonlinear output function, sensorimotor observation
A	state matrix
B	input matrix
C	output matrix
D	feedthrough (or feedforward) matrix

s	Laplace operator
$j\omega$	frequency domain operator
$G(s)$	transfer function
$Z(j\omega)$	mechanical impedance
$Y(j\omega)$	mechanical admittance
m	environment inertia
d	environment damping
k	environment stiffness
Ω	set of variables
$E\{\cdot\}$	expected value operator
$R_u(\tau)$	autocovariance function
$R_{yu}(\tau)$	crosscovariance function
J	optimization objective function
T_d	time delay
$\nu(\cdot)$	(joint) sensorimotor transformation

Perceptual and Behavioral Model Foundations

$\gamma(\cdot)$	implicit ordinary differential equation
$\mathcal{L}\{\cdot\}$	Laplace transform
$\zeta(\cdot)$	impedance mapping
K_p	proportional control gain
$u_l(t)$, $u_r(t)$, $v_l(t)$, $v_r(t)$	wave variables
$f_h^d(t)$	desired interaction force
$\dot{x}_t^d(t)$	desired transmitted interaction movement
b	wave impedance
I	stimulus intensity
I_t	target condition stimulus intensity
I_c	control condition stimulus intensity
p_{chance}	chance level
p_{PSE}	probability at PSE
p_{JND}	probability range determining the JND
ΔI_i	step size in step i
δI	initial step size
ΔI	discriminable intensity difference
W	Weber fraction
S	Magnitude of percept
α	exponent in Steven's power law
$G_{enc}(j\omega)$	information encoding transfer function
$G_{acc}(j\omega)$	information accumulation transfer function
$s(t)$	sensory signal
$\mu(t)$	information
$\xi(t)$	noise process
$\bar{\mu}$	mean information content
K	state feedback matrix

Z^N	dataset for identification
$g_0(t)$	impulse response
$v(t)$	noise process
$R_u^n(\tau)$, $R_{yu}^n(\tau)$	autocovariance/crosscovariance matrix, estimated from N samples
$G_{cont}(j\omega)$	sensorimotor control mechanism
$G_{body}(j\omega)$	body dynamics
$G_{sense}(j\omega)$	sensory dynamics

Dynamic Modeling of Sensorimotor Processes

$G_{h,v,a}(j\omega)$	sensory dynamics of haptic, visual, auditive receptors
$\xi(t)$	noise process
$G_m(j\omega)$	muscle dynamics
$Y_{body}(j\omega)$	body admittance
$G_{tissue}(j\omega)$	tissue dynamics
S	sensorimotor domain
θ	feature vector
$\mathbf{u}_{ec}$	efference copy
λ	eigenvalue

Dynamic Combination of Movement and Force

$\delta(\cdot)$	decision criterion
ϵ	decision threshold
$f_{m,res}(t)$	effective muscle force acting on the interacting limb
$G_{inv}(j\omega)$	inverse dynamics of body and environment
m_h	limb inertia
d_h	limb damping
A	movement amplitude
ω	movement frequency
N_{cond}	number of conditions in one experiment

Dynamic Masking in Haptic Discrimination

p_i	physical feature
c_{step}	offset value
t_{resp}	response time

Communication Quality Control

$\mathbf{y}_{perf}(t)$	task performance
$\mathbf{y}_{cost}(t)$	communication cost
bw	bandwidth
θ_T	task parameter vector

θ_C	communication parameter vector
$p_l(t)$	packet loss rate
N_{col}	number of obstacle collisions
t_{com}	completion time
$p_{col}([t, t+T])$	collision probability within a sample time interval
t_p	prediction horizon
$\lambda_{1,2,3}$	weighting factors
c_{max}	maximum cost
$r_{speed}([t, t_T])$	objective for decreasing completion time
$R([0, T])$	reachable set of states in the time interval $[0, T]$
X	discrete state
U^α	input interval
$\Phi^\alpha([0, T])$	state transition matrix, given the input interval U^α
$\Phi([0, T])$	state transition matrix, given all possible inputs
$\mathscr{V}$	volumetric integral operator
$p_i([0, T])$	probability for passing state i
$pos(\cdot)$	operator that extracts position components from a state
N_X	number of discrete states

List of Figures

List of Tables

*"**engineering**, the application of science to the optimum conversion of the resources of nature to the uses of humankind."*

ENCYCLOPÆDIA BRITTANICA [19]

1 Introduction

Engineering, as defined by the renowned Encyclopædia Britannica, is a discipline directed towards a specific goal: All efforts serve the usefulness for humans. In a related way, the sociologist Read Bain in 1937 defined "technology" to include "all tools, machines, utensils, [...] and the skills by which we produce and use them." [2]. Both definitions imply that technology as the product of engineering must suit the human that uses it. In return, the best way to engineer new technology is to include knowledge about the human user into the design of technical systems in order to meet the physical, sensory and cognitive abilities enabling him/her to use the device in the desired manner.

While a human-centered design is essential for all technical systems, telepresence systems take an exceptional position which is discussed in the following.
Telepresence systems enable the human operator to transfer his/her skills to a remote, scaled, or otherwise inaccessible area. This transfer is achieved by operating a robotic human-system interface (HSI) capturing the user actions, transmitting them to a second robotic device, the teleoperator (TO), performing commanded actions and collecting sensory data. These are sent back to the operator and are displayed via the HSI. The design and control of telepresence systems has been subject to optimization towards a significant number of objectives, including the mechanical design [3], transparency [4, 5] and bandwidth [6]. First attempts have been made to include general findings about human perceptual and motor capabilities into account in the design and evaluation of such systems [7, 8]. For a fully user-centered system design, however, accurate knowledge about the human perceptual, physical, and cognitive processes involved in the interaction with the telepresence system, is required.

In general, interactions with our surrounding are characterized by an exchange of information between the perceptual system, the motor system and the environment, as summarized in the well-known action-perception loop [9] (see, e.g., Figure 1.1). While the sensory signal exchange is multimodal in principle, we will restrict our considerations mostly to the haptic modality and add visual information where needed. We do so respecting the fact that most current multimodal telepresence systems provide visual and haptic feedback, where latter is still underdeveloped and subject to current research in many different aspects.

It was recognized a long time ago [9, 10] that the information flow between all components of the action-perception loop is continuous and perception as well as action must process time-varying sensor signals in order to perceive and react. Up to now, little is known about the *dynamics* of these processes, where we refer to the term "dynamics" in the sense of a differential equation. Instead, most effort has been spent on characterizing sensory range and resolution [7, 11–17]. Those measures determine the steady-state of the perception system, and thus neglect the time-varying processes that lead to a specific percept. Most haptic environmental factors, such as stiffness or inertia must be manually interacted with in order to perceive them as their magnitude and properties can be captured only indirectly,

by combining the time-varying movement and force feedback together. We will refer to models that include an excitation with the motor system as *sensorimotor* process models, to emphasize the tight coupling between the perceptual and motor domain. In current, static perception models, the influence of this exploratory movement is largely neglected, implicitly assuming that the environment is sufficiently well excited to reveal all of its properties. When a telepresence system becomes part of the action-perception loop, this fundamental assumption is challenged due to multiple reasons: First, the primary goal during the usage of telepresence systems is the achievement of a specific task in many cases, which requires a certain action pattern, and may not be geared to perceiving the environment properly. Second, the telepresence system may not allow for certain interaction patterns, e.g., due to limitations in mechatronic and control design, or communication effects such as time delay [8, 18]. As a consequence, the validity of static perception models is challenged due to a violation of the fundamental assumption. One solution to explicitly take the interaction movement into account when predicting the human perceptual state is modeling perception with its dynamic properties. A second motivation for the development of dynamic models for sensorimotor processes is the tradeoff between fidelity and performance of a telepresence system on the one side, and the costs for achieving these goals on the other hand which has to be made in every design process for a technical system. Naturally, systems with feedback in higher resolution, more accurate timing, and other high-qualitative features come along with a high price for hardware and associated services, such as a high-performance communication channel. Dimensioning hardware and service quality based on a tradeoff coming from static perceptual and sensorimotor considerations may result in a poor design choice due to the inaccuracy of the model in a real telepresence operation. Dynamic models describe not only the steady-state behavior of perception and motor control, but allow for a time-continuous prediction of perceptual limits, to which the system may adapt online. In this way, algorithms can be developed that provide a good service quality only in situations, where the operator benefits from it and allow a lower quality in the meantime, potentially saving costs.

1.1 Problem Definitions and Challenges

The ultimate goal for the realization of a human-centered design of technical systems is the development of a detailed, quantitative, dynamic model of all processes involved in the sensorimotor loop. Working towards this goal requires solving certain key challenges which arise when multiple research fields come together: Methods from system theory are adopted for describing a biological process which is usually investigated with psychophysical approaches. In return, the application of such models to technical systems is not straightforward as their validity is usually limited to specific operating conditions. Challenges addressed in this thesis are summarized in the following.

Dynamic Modeling of Sensorimotor Processes

The dynamics in mechanical, mechatronic or electronic systems are generally determined by their physical properties. Even if their design is not known a priori, reverse engineering

helps for the development of an initial model which can be refined iteratively later on. For perceptual and sensorimotor processes, such methods fail as the human brain consists of billions to trillions of synapses [20] potentially contributing to the dynamic behavior of the processes under consideration here. Psychological models which are able to address perceptual processes are on the other hand often limited to qualitative statements about the impact of a certain environmental property on the perceptual state, and their applicability in the development of technical systems is limited. In addition, only static relations between the physical world and its perceptual equivalent are usually investigated. The mathematical and methodological foundations for the development of dynamic perceptual and sensorimotor process models applicable to the design of technical systems is missing in current literature.

Parameter Identification of Sensorimotor Models

Many standard methods have been developed for the identification of dynamic technical systems, based on time-series analysis or frequency-domain considerations [21–23]. All these methods require, however, that the input- and output signals can be recorded with a sampling frequency and measuring accuracy sufficient to reveal the underlying dynamic properties. The input signal to the motor system and the output from perception, however, is a perceptual variable which exists only in the brain, thus it is hardly accessible. Even with sophisticated techniques such as electroencephalography (EEG) or functional magnetic resonance imaging (fMRI) allowing to record brain potentials and activations, the meaning of these recordings are hard to interpret till now. Research on perceptual processes thus often must get by with discrete observations of cognitive variables by requesting the human to give responses about the current state at a specific time instance. This strong limitation makes a direct parameter identification of the dynamic sensorimotor system unfeasible but necessitates the development of novel experimental procedures and tools for data analysis. A major problem for these new experimental methods is the relatively low information content of human feedback in a perception task, typically a binary answer in a paired comparison (“difference perceived”/“not perceived”). A careful design of experimental conditions to maximize the information using as less conditions as possible is amongst the most challenging steps in this context.

Application of Dynamic Sensorimotor Models to Telepresence Systems

The declared purpose behind developing and identifying dynamic perceptual and sensorimotor models is to utilize them in the design and evaluation of telepresence systems. Nearly all models of psychological processes are valid only in certain operating conditions. In the development of perceptual and sensorimotor models, a tradeoff between the generality of its prediction capabilities on the one hand and the validity for a specific application on the other must be made. A co-design of application and the dynamic sensorimotor model is thus essential for a successful combination of the two. The problem here lies in a careful review of the essential requirements for an application, and the reduction of the sensorimotor model complexity to a practically feasible degree.

1.2 Main Contributions and Outline of the Thesis

This work provides first, fundamental steps towards the ultimate goal of a dynamic model for the human sensorimotor loop, including human perception. The application of quantitative dynamic models to describe haptic perceptual processes is so far unique. With this contribution, a common language between psychological mechanisms and engineering applications is found which makes an application of such models in technical haptic applications easier. On the one hand, these methods make a significant contribution to the field of engineering, as knowledge about processes within the human operator can be included into the design, optimization, and evaluation of technical systems. On the other hand, a contribution to the psychophysical community is made by introducing novel experimental techniques, models, and experimental design tools from system theory. The focus of this work lies on the haptic modality, as it is the only one that allows physical interaction with the environment, and contains sensory feedback predominantly caused by the reaction to one's own motor behavior. In this way, haptic perception and sensorimotor control are coupled together in the form of a closed control loop, emphasizing the need of a dynamic representation of cognitive and motor processes for an adequate representation of perceptual and task performance. The structure of the thesis can be illustrated around the action-perception loop, as depicted in Figure 1.1.

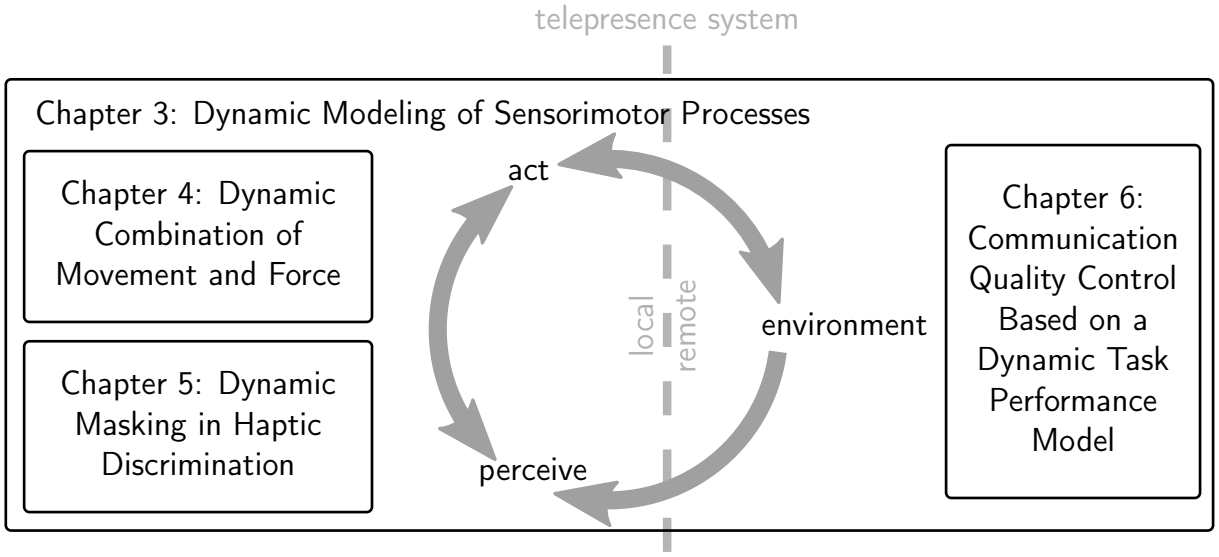


Figure 1.1: The thesis is structured along the action-perception loop.

Dynamic Modeling of Sensorimotor Processes

A successful integration of models describing human sensorimotor and perceptual processes into technical systems requires the establishment of a common modeling language. This motivation in mind, a novel mathematical framework is established in Chapter 3, founding on the notion that all processes within the body are dynamic. This generic tool is introduced as an extension to the currently existing modeling concepts for perception, which statically

relate a physical stimulus to a corresponding percept. The relation between the new concept and the state of the art is discussed, and new application fields utilizing the dynamic formalism are outlined.

Dynamic Combination of Movement and Force

The perception of haptic environmental properties essentially requires the combination of multiple haptic information sources, namely movement and force. Knowledge about the required accuracy with respect to the temporal consistency of these sources is indispensable for a human-centered design of, e.g., communication strategies in telepresence systems. A model for the discrimination of temporal lags in haptic information is developed in Chapter 4, on the basis of the new conceptual framework presented in Chapter 3. A dynamic model realizing a dynamic state observer captures the perception thresholds over several conditions significantly better than static modeling approaches.

Dynamic Masking in Haptic Discrimination

Static perceptual limits for simple environments, consisting of only stiffness, damping, or inertia, are sufficiently known and serve in applications for telepresence systems, e.g., a human-centered transparency analysis [24]. Limits in the perception of more complex environments, consisting of either two or more "components", is though largely unknown. An analysis of the perceptual performance in discriminating damping which is masked with an additional environmental factor – stiffness or inertia – is presented in Chapter 5. Six perceptual models, founded on the mathematical foundation presented in Chapter 3 are proposed to explain the observed discrimination performance. Again, models with an explicitly consideration of a dynamic process are found to resemble human perceptual characteristics better compared to static models. These findings can be utilized for enhancing current approaches for storing and communicating haptic data.

Application of Dynamic Performance Models to Visual-Haptic Telepresence Systems

As an example for the successful integration of dynamic models capturing sensorimotor processes into a technical application, communication quality control as a novel concept in haptic telepresence is considered in Chapter 6. An online-adaptation scheme for the communication channel time delay is proposed on the basis of a dynamic model predicting human task performance in a navigation task, developed by means of the dynamic mathematical framework in Chapter 3. A positive effect of the communication quality control scheme on human task performance can be observed. It should be noted that the online adaptation *requires* a dynamic model formulation, and can not operate reliably on the basis of a static model, emphasizing the need for a dynamic formulation of sensorimotor processes.

2 Perceptual and Behavioral Models with Application in Technical Systems

Summary. *Developing system theoretic models for sensorimotor processes is a highly interdisciplinary endeavor. The fundamental concepts from system theory and psychophysics are introduced here. Aside from basic principles, more specific topics include*

- *quantitative static models of human perception,*
- *technical perception models for dynamic systems,*
- *dynamic models in human motor control.*

The innovative concept of modeling perceptual processes with system theoretic tools is interdisciplinary between system theory and psychophysics. As the audience of this thesis may come from both disciplines, fundamental concepts from each disciplines must be reviewed before topics related to the specific aim of this work can be introduced. A brief introduction to system theoretic modeling forms is presented in Section 2.1. On this basis, haptic environment models are introduced in Section 2.2, along with methods to artificially render a haptic impression using a robotic interface. Haptic telepresence systems as the main application for the methods developed throughout this work are targeted in Section 2.3, where an understanding of the goals and challenges that need to be solved in this context are reviewed and discussed. Existing perception models and experimental techniques for the quantization of these are presented in Section 2.4. While these models are of static nature and methods for the dynamic characterization of perception processes are missing, technical processes can be identified with respect to their dynamics easily. Methods from system theory serving this purpose are given in Section 2.5. Additional inspiration for the development of novel, dynamic models for perception and task performance can be drawn from existing dynamic sensorimotor control models, as reviewed in Section 2.6. The chapter is concluded in Section 2.7 with a discussion of the state-of-the-art, identifying missing pieces that are needed for a truly human-centered design of technical systems, and especially telepresence systems.

2.1 System Theoretic Foundations

In this work, the *dynamic* properties of human perception and sensorimotor control are investigated. Dynamic systems' behavior can require a formulation of the input-output relation in the form of an ordinary differential equation. In this section, three representation forms of

dynamic systems are introduced – the ordinary differential equation, the state space model, and the transfer function.

2.1.1 Ordinary Differential Equation

The fundamental formulation of a dynamic system as we will use it in this thesis is an ordinary differential equation

$$\gamma(\mathbf{u}(t), \dot{\mathbf{u}}(t), \ldots, \mathbf{u}^{(m)}(t), \mathbf{y}(t), \dot{\mathbf{y}}(t), \ldots, \mathbf{y}^{(n)}(t), t) = \mathbf{0} \tag{2.1}$$

where $\mathbf{u}(t) \in U$ is a vector of input values, $\mathbf{y}(t) \in Y$ is a vector of output values. The notation of a superscripted (n) and (m) denotes the nth and mth derivative with respect to time. In contrast to static systems, modeled by

$$\gamma_{stat}(\mathbf{u}(t), \mathbf{y}(t), t) = 0, \tag{2.2}$$

a dynamic systems' output at time t, determined by the solution of (2.1) depends not only on the current input $\mathbf{u}(t)$, but on the history of inputs $\mathbf{u}(\tau)$, $t_0 \leq \tau < t$ and outputs $\mathbf{y}(\tau)$, $t_0 \leq \tau < t$. Applied to modeling the perceptual system, a dynamic perceptual model can be interpreted as one that takes sensory information of the actual time instance and past sensory information into account for determining a perceptual response.

2.1.2 State-Space Formulation

In the case of $m \leq n$, the ordinary differential equation in (2.1) can be transformed into a system of ordinary differential equations of order 1. The resulting system is denoted as state space formulation

$$\mathbf{x}(t) = \psi(\mathbf{x}(t), \mathbf{u}(t), t), \quad \mathbf{x}(0) = \mathbf{x}_0, \tag{2.3}$$
$$\mathbf{y}(t) = \phi(\mathbf{x}(t), \mathbf{u}(t), t), \tag{2.4}$$

where $\mathbf{x}(t)$ is called "state vector". In the case of a linear, time-invariant (LTI) system, (2.3)-(2.4) can be written in matrix form, where

$$\mathbf{x}(t) = A\mathbf{x}(t) + B\mathbf{u}(t), \quad \mathbf{x}(0) = \mathbf{x}_0, \tag{2.5}$$
$$\mathbf{y}(t) = C\mathbf{x}(t) + D\mathbf{u}(t). \tag{2.6}$$

Here, A is the "state matrix", B is the "input matrix", C is the "output matrix", and D is the "feedthrough (or feedforward) matrix".

The state space formulation of dynamic system is appealing from a modeling perspective, as it separates the dynamic part (2.5), containing information about time constants and stability, from the output equation (2.6), (statically) determining the information which is extracted from the state and the input vector.

2.1.3 Transfer Function

In the case of a fully observable and controllable LTI dynamic system with a single input and a single output (the definitions of controllability and observability will be given in Section 2.5.1), all dynamic behavior is captured by its transfer function $G(s)$, where

$$G(s) = \frac{Y(s)}{U(s)} = \frac{\mathscr{L}\{y(t)\}}{\mathscr{L}\{u(t)\}}, \tag{2.7}$$

and $\mathscr{L}\{\cdot\}$ is the Laplace transform. In the case of a stable system, $G(s)$ can be replaced by the Fourier transform $G(j\omega)$, resulting in an equivalence

$$G(j\omega) = G(s). \tag{2.8}$$

The function $G(j\omega)$ is commonly denoted as frequency-domain representation.

2.2 Haptic Environment Models

The haptic modality consists of four sub-modalities: Tactile, kinaesthetic, temperature and pain sensing contributes to the haptic impression of an environment. Besides temperature receptors, humans can make use of position, force, movement, and pressure sensors to perceive the physical world [25]. Tactile sensations include high-frequently changing forces and positions of small amplitude that arise, e.g., in texture perception, as well as shape perception of objects. The kinaesthetic sense is predominantly determining the impression of large-scale movements and forces [26]. Most haptic interfaces, especially those used in current telepresence systems, are not capable of giving thermal and tactile feedback about the environment at the teleoperator's location. Another restriction for the further discussion is to limit our view to single-point contacts between human and the environment. This prevents the precise perception of object shapes that often requires multi-fingered interaction with the environment. As a consequence, the focus of this thesis is on the perception and control of haptic environments that appeal to the kinaesthetic sub-modality.

Definition 2.1 (Dynamic Haptic Environment)**.** A dynamic haptic environment is modeled by a dynamic system defining the relation between interaction forces $\mathbf{f}_{env}(t) \in \mathbb{R}^3$ and a translation movement $\mathbf{x}_h(t) \in \mathbb{R}^3$.

Two formulations of this definition are possible [27]: The first, using mechanical impedances to describe the force feedback resulting from a specific translation in space, which can be written as

$$\mathbf{f}_{env}(t) = \zeta(\mathbf{x}_h(t), \dot{\mathbf{x}}_h(t), \ldots, \mathbf{f}_{env}(t), \dot{\mathbf{f}}_{env}(t), \ldots, t) \tag{2.9}$$

Where $\zeta(\cdot)$ is a mapping function $\mathbb{R}^3 \rightarrow \mathbb{R}^3$. An inverse formulation using an admittance formulation, relating the input forces and torques to a resulting motion, is the second formulation.

We want to discuss the important case of haptic environments, namely linear, time-invariant, spatially homogeneous (LTIH) haptic environments on the example of impedance-type environments. We do so, respecting that in general all haptic environments exhibit nonlinear characteristics, but many practically relevant cases are well-described by linear approximations of them. While equation (2.9) was defined for three-dimensional forces and movement, we will limit our consideration to a one-dimensional translational movement $x_h(t) \in \mathbb{R}$ and a resulting force $f_h(t) \in \mathbb{R}$ in the following for the ease of presentation and notation

$$f_{env} = \zeta(x_h(t), \dot{x}_h(t), \ldots, \dot{f}_{env}, \ldots, t). \tag{2.10}$$

However, the concepts presented in the following are easily extensible to higher dimensions. As a second constraint, only stable environments are taken into consideration as unstable haptic environments do not arise in nature. Although instability can occur in haptic telepresence, as discussed in Section 2.3, haptic interaction with an unstable system is not feasible due to, e.g., safety reasons. Thus, appropriate measures have to be taken to ensure the stability of the haptic environment the human operator is interacting with.

2.2.1 Linear, Time-invariant, Spatially Homogeneous Haptic Environments

LTIH environments are fully described by a linear differential equation with constant parameters.

Remark 2.1. The meaning of the employed term "spatially homogeneous" is that the system behavior does not change with the location it is interacted with, which distinguishes it from the attribute "homogeneous" used for nonlinear systems to state that $f(\alpha x) = \alpha f(x)$.

Due to the linearity of the differential equation, equation (2.10) can be described in the frequency domain, namely

$$F_{env}(j\omega) = Z_{env}(j\omega)X_h(j\omega) \tag{2.11}$$

with $F_{env}(j\omega)$, $X_h(j\omega) \in \mathbb{Z}$ being the Fourier-transform of $f_{env}(t)$, $x_h(t)$, respectively. An important special case of LTIH haptic environments are mass-spring-damper models

$$f_{env}(t) = m\ddot{x}_h(t) + d\dot{x}_h(t) + kx_h(t) \tag{2.12}$$

with m being the inertia, d the damping coefficient and k the spring coefficient. The properties of the system's mechanical impedance

$$Z_{env}(j\omega) = \frac{F_{env}(j\omega)}{\dot{X}_h(j\omega)} = \frac{1}{j\omega}\left((j\omega)^2 m + j\omega d + k\right) \tag{2.13}$$

can be easily divided into a magnitude $|Z(j\omega)|$ and phase component $\angle Z(j\omega)$ by using Euler's formula to

$$Z_{env}(j\omega) = |Z_{env}(j\omega)|\mathrm{e}^{j\angle Z_{env}(j\omega)}. \tag{2.14}$$

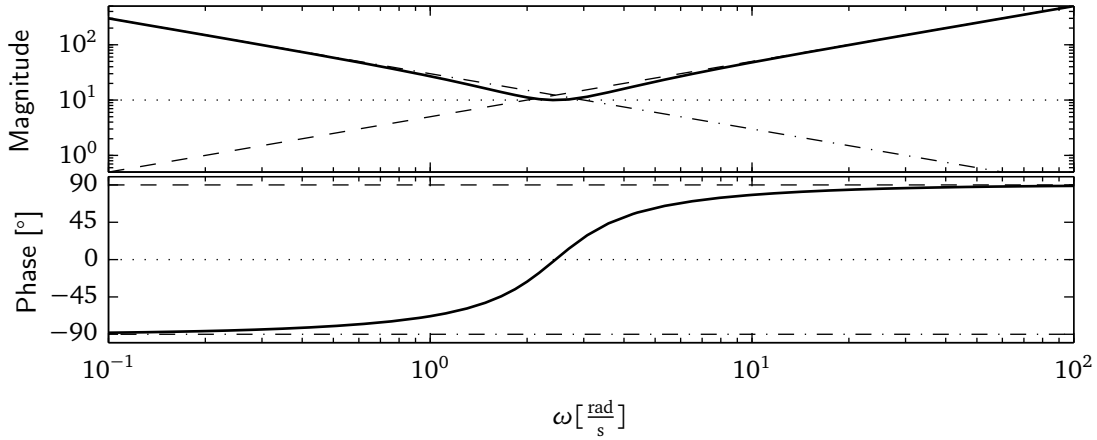


Figure 2.1: An exemplary Bode diagram for a mass-spring-damping system in impedance form. The effect of mass is shown dashed, damping dotted, spring dash-dotted and the overall system in bold solid lines.

While the magnitude states how large the reaction force to the exciting position is, the phase determines the frequency-dependent time delay between input and output. The Bode diagram is a visualization commonly used to visualize amplitude and phase over the logarithmic frequency axis, an example for a mass-spring-damper system is depicted in Figure 2.1. A specific property of LTIH haptic environments is the frequency range of excitation and feedback signal to be in the same order of magnitude, determined by the human force control bandwidth, which is on the order of 10 to 30 Hz [6, 7]. This property is due to the fact that changing environmental features, such as an impact into a stiff wall from free-space is excluded by the limitation to spatially homogeneous environments.

2.2.2 Haptic Rendering

Haptic environments are experiential for humans either by interacting with natural objects, or by emulating an artificial environment using a robotic haptic interface. The methods that are suited for the control of the interface to result in an impression of a haptic environment are called *haptic rendering*. Two control schemes shall be introduced here, serving as haptic rendering algorithms in the experiments presented in this work. One challenge for the rendering with respect to stability and transparency lies in the realistic simulation of stiff contacts, especially during impact situations [28], and sophisticated methods have been developed for this problem. However, the focus of this is on the rendering of continuous and spatially homogeneous environments without considering impacts. All rendering algorithms are discussed for the case of a single motor with one deegree-of-freedom (1DoF). Extensions in more directions in space are possible.

Force Control

A force control scheme capable of displaying an environmental impedance $Z_{env}(j\omega)$ to the human operator is depicted in Figure 2.2. Because of the high noise level that is inherent to

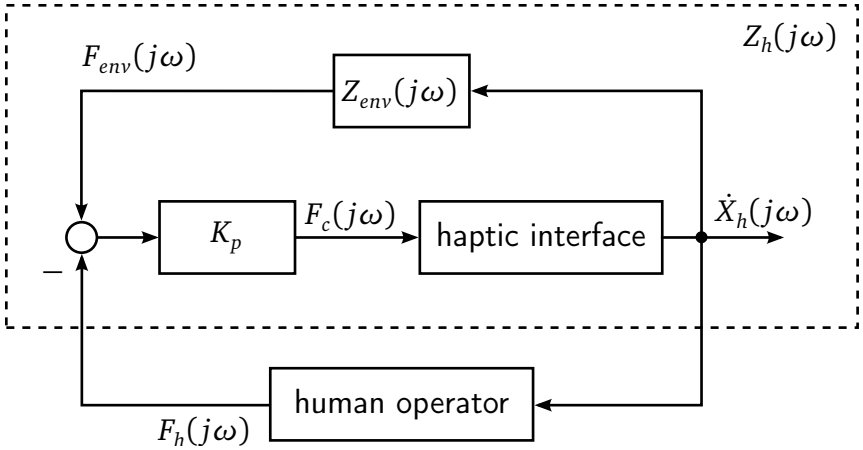


Figure 2.2: Force control can be used to render the desired environment impedance $Z_{env}(j\omega) = \frac{F_{env}(j\omega)}{\dot{X}_h(j\omega)}$. Due to imperfect compensation of the device dynamics, the actually rendered impedance that is displayed to the human operator $Z_h(j\omega)$ can differ.

force sensors, a proportional controller K_p is used to regulate the motor, and no derivative component is considered in the control concept. As a reference force trajectory, the environment force F_{env} is computed, based on the desired environment impedance $Z_{env}(j\omega)$ and the velocity $\dot{X}_h$. Ideally, all dynamical properties of the haptic interface can be compensated by the force controller, such that only the environment impedance is displayed to the human operator, $Z_h(j\omega) = Z_{env}(j\omega)$. In real systems, uncompensated device dynamics, e.g. due to limitations in the controller gain because of existing sensor noise, the impedance $Z_h(j\omega)$ which is displayed to the human may differ from $Z_{env}(j\omega)$. An exemplary analysis of the uncompensated device dynamics is sketched in Appendix A.1.

Position-Based Admittance Control

A position-based admittance control scheme for the rendering of haptic environments is depicted in Figure 2.3. It differs from the force control scheme in several aspects: Firstly, the haptic environment representation to be rendered is given in admittance form $Y_{env}(j\omega)$. Secondly, the low-level controller underlying the position-based admittance control scheme is a position controller instead of the force controller in 2.2.2. Because the device position (or joint angle) is measurable with a high resolution ($\sim (2\pi)/10^5$ rad rotatory, $\sim 10^{-6}$ m translatory) without significant measurement noise, a carefully tuned PD-controller can compensate the device dynamics better compared to the force controller. As a consequence, the rendered environment admittance equals the desired characteristics almost exactly, allowing the stable rendering of a higher stiffness, compared to the force control scheme. However, the formulation of the environment dynamics in admittance form comes along with a disadvantage: Reconsidering a mass-spring-damper system in admittance form transforming an input force into a velocity

$$\ddot{x}_h(t) = \frac{1}{m}(f_h(t) - d\dot{x}_h(t) - kx_h(t)),$$

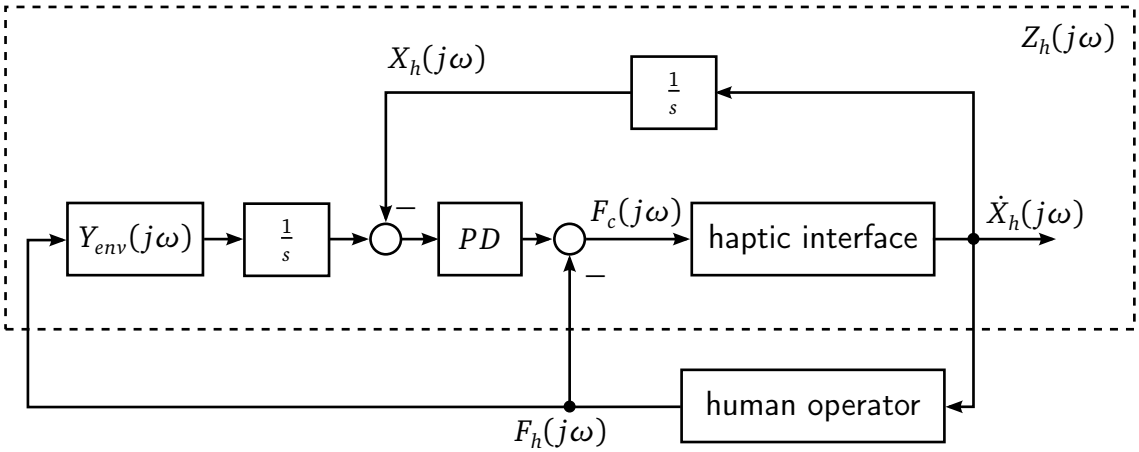


Figure 2.3: The position-based admittance control scheme to render the haptic interface can normally suppress device dynamics better compared to the force control scheme, thus $Z_h(j\omega) \approx Z_{env}(j\omega)$. However, a minimum inertia must be always present.

a minimal inertia must be present at all times, as the acceleration can rise to infinity otherwise.

2.3 Haptic Telepresence Systems

Telepresence systems, also known as teleoperation systems [29–38] in the literature, telerobotic [6,39,40], or telemanipulation systems [41,42] enable a human operator to perform tasks in environments he/she can not access directly with his natural capabilities. While telerobotics and teleoperation are often used synonymously for the principal idea to connect a human to a robotic device which provides arbitrary feedback in the visual, haptic, and/or auditive modality, telemanipulation emphasizes on the haptic aspect of intention transmission and feedback modality. The term *telepresence*, as we will use it throughout this thesis, points out the substantial integration of the robotic system into the human perception/action loop by emphasizing that the operator should feel "present" in the remote environment [43]. The meaning of "presence" shall be discussed briefly besides other topics related to telepresence systems such as the system architecture, control- and operator-related issues.

2.3.1 System Architecture

A prototypical telepresence system is depicted in Figure 2.4. The human operator interacts with a robotic *human-system interface* (HSI) instead of interacting directly with the environment, as the target location of his actions is located out of reach. The user intention, represented exemplary by the hand velocity $\dot{x}_h(t)$ is transmitted over a *communication channel*, which can include some kind of communication network, such as the Internet. The *teleoperator* (TO) receives the transmitted velocity, and drives on the desired interaction trajectory $\dot{x}_t(t)$. The resulting reaction force with the environment $f_{env}(t)$ is measured and transmitted back to the human-system interface where they are displayed to the human operator ($f_h(t)$).

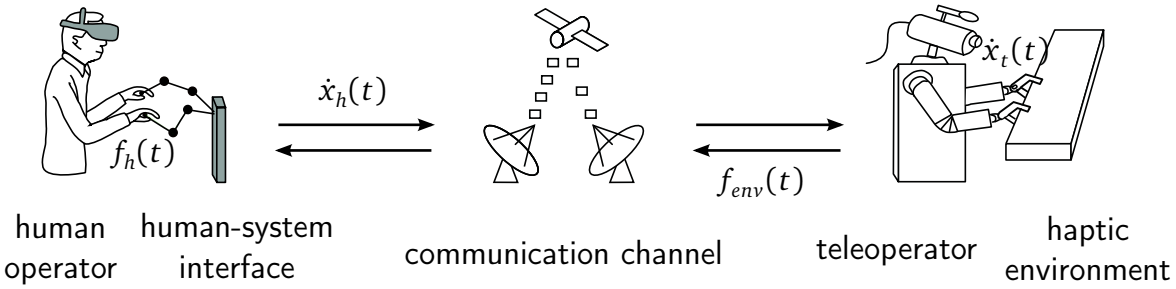


Figure 2.4: A telepresence system with haptic feedback consists of mainly three components: Human-system interface, communication channel, and teleoperator

Remark 2.2. A telepresence system using a two-channel velocity-force control architecture is shown here. There are other approaches using a four-channel architecture, transmitting both velocity and force, or, position and force from HSI to TO and from TO to HSI which are known to have some superior properties. For a review on two- and four-channel control architectures, the reader is referred to the seminal work of Lawrence [5] and a summary of different approaches in [44].

2.3.2 Goals and Challenges

Current telepresence applications include search-and-rescue tasks [45], on-orbit servicing of satellites in space, underwater teleoperation [46] and telesurgery [47](for a survey see [43]). Common to all telepresence applications are three fundamental goals to be achieved – transparency, the feeling of presence, and stability.

Transparency

The original definition of transparency was introduced by Raju in 1989: "An ideal telemanipulator is one that provides complete transparency of the interface. In other words the operator feels as if the task object were being handled directly." [48]. Later, objective measures for transparency have been developed, based on the exact imitation of environment force and teleoperator position at the human-system interface side [49]. A telepresence system is transparent if following conditions hold:

$$\forall t: \; x_h(t) = x_t(t), \qquad f_h(t) = f_{env}(t). \tag{2.15}$$

An alternative transparency measure in the case of a linear, time-invariant environment impedance $Z_{env}(j\omega)$ is the comparison of the latter to the impedance displayed by the HSI to the human operator $Z_h(j\omega)$ [5], where

$$Z_{env}(j\omega) = \frac{F_{env}(j\omega)}{\dot{X}_t(j\omega)}, \qquad Z_h(j\omega) = \frac{F_h(j\omega)}{\dot{X}_h(j\omega)}. \tag{2.16}$$

Ideal transparency is achieved if

$$Z_{env}(j\omega) = Z_h(j\omega). \tag{2.17}$$

Problematic for both transparency definitions is their lacking interpretable meaning with respect to the effect on the human operator. To overcome this problem, transparency measures in the parameter space have been proposed [18, 35]: Based on approximations of the environment impedance as a mass-spring-damper system, mismatches between $Z_{env}(j\omega)$ and $Z_h(j\omega)$ can be expressed as differences in mass, spring constant, and damping coefficient.

It is worth noting that the objective definitions of transparency given in equations (2.15) and (2.17) are somewhat more restrictive than the original definition of Raju [48]. This comes from the fact that humans do not perceive arbitrary small differences between two impedances, forces, or positions. Thus, the operator could have a perfectly transparent feeling of the remote environment even if conditions (2.15) and (2.17) are violated. This mismatch between objective and subjective transparency was accounted for by defining a "perceived transparency" measure [50] that takes perceptual limits into account. However, all perceptual limits taken into consideration for the perceived transparency definition are currently determined *statically*, e.g., without considering the operator's exploration movement or dynamic properties of the perceptual system. Without knowledge about the perceptual dynamics, it is unclear whether the static perception limits are overly conservative, too liberal, or equivalently applicable to dynamic interactions.

Presence

"*Presence* is defined as the subjective experience of being in one place or environment, even when one is physically situated in another" [51]. Following this definition, the human operator must not perceive the telepresence system in any aspect of sensory feedback or limitation in interaction. Presence could thus be looked at as a generalization of transparency to all sensory and motor domains. Two approaches for measuring presence are established in the literature: Subjective measures usually rely on questionnaire data to determine the level of achieved presence, e.g., in [51]. Objective measures focus on a similarity assessment of physiological, psychological, and behavioral responses [52]. The hypothesis there is that the differences between direct environment interaction and telepresent interaction becomes smaller the more the operator feels present. For understanding and enhancing presence in telepresence systems, knowledge about the dynamics of perceptual and behavioral processes is beneficial, as perceptual differences between direct environment interactions and telepresent impressions in tangible interaction situations can be predicted accurately. For a comprehensive summary of approaches, the interested reader is referred to [53].

Stability

By means of the haptic modality, a control loop over the human operator, HSI, communication channel, TO and the environment is closed, exchanging energy in terms of force and movement. Time delay and packet loss, inherent to communication in wide-range teleoperation or and communication channels near their capacity limit, challenge the stability of this control loop. Multiple methods for ensuring stability over communication channels with constant time delay have been developed over the past decades [27, 29–32, 34, 36, 37, 54, 55]. Extensions for time-varying time delay [54] and packet loss [56] have been proposed as

well. The most popular methods for ensuring stability are based on the concept of passivity. A passive system has a negative energy balance, meaning that it produces less energy than it consumes. Taking the energy contained in the system as a Lyapunov function, the passivity criterion ensures that a passive system is stable in the sense of Lyapunov. Furthermore, a connected series of passive systems leads to a globally passive system. In this way, stability can be guaranteed, if it can be ensured that the HSI, communication channel and TO behave passively and the haptic environment and human operator are inherently passive. Widely-used techniques founded on passivity are time-domain passivity [55] and the wave variable transformation [29], which shall be briefly introduced here. The wave variable transformation is defined as

$$\begin{aligned} u_l(t) &= \frac{f_h^d(t) + b\dot{x}_h(t)}{\sqrt{2b}}, \quad u_r(t) = \frac{f_{env}(t) + b\dot{x}_t^d(t)}{\sqrt{2b}} \\ v_l(t) &= \frac{f_h^d(t) - b\dot{x}_h(t)}{\sqrt{2b}}, \quad v_r(t) = \frac{f_{env}(t) - b\dot{x}_t^d(t)}{\sqrt{2b}} \end{aligned} \tag{2.18}$$

where $u_l(t), u_r(t), v_l(t), v_r(t) \in \mathbb{R}$ are denoted wave variables. $f_h(t)$ and $f_{env}(t)$ are the forces displayed to the human operator and the force reflected from the remote environment, respectively. The wave impedance $b > 0$ is a tuning factor and must be chosen in agreement with environment and application. A superscripted d stands for the desired value of a variable that must be enforced by a local controller either on the human-system interface or the telerobot.

In the case that no time delay and packet loss is present in the telepresence system, the communication channel is transparent in the sense of [49], that means $\dot{x}_t^d(t) = \dot{x}_h(t)$ and $f_h^d(t) = f_{env}(t)\,\forall t$. In the case of time delay and/or packet loss, the wave variable transformation on the one hand guarantees safe operation, but results in a change of the haptic environmental properties that are displayed to the human operator, thus a decreased transparency.

2.4 Quantitative Models of Human Perception

The discipline describing the perceptual reaction to a physical stimulus is called *psychophysics* [57]. Within the discipline, two directions can be distinguished: *Inner* psychophysics aims for quantizing the relation between a physiological, neural signal within the body, and the perceptual cause of it. In contrast, *outer* psychophysics, directly relates the physical quantity and the perceptual response. As engineers usually do not have access to biofeedback signals that would allow the application of inner psychophysical models, we will restrict our view to outer psychophysics in the following. Amongst the perceptual phenomena which are of greatest interest to engineers developing technical systems in a human-centered way are those that capture the perceptual thresholds. Two types are to be distinguished: The absolute threshold, capturing the minimal intensity of a stimulus that is required to be perceived at all, and the difference threshold, also called *just noticeable difference* which denotes the minimal amount that two physical quantities have to differ from each other to be distinguishable. Some available measurement techniques and their underlying

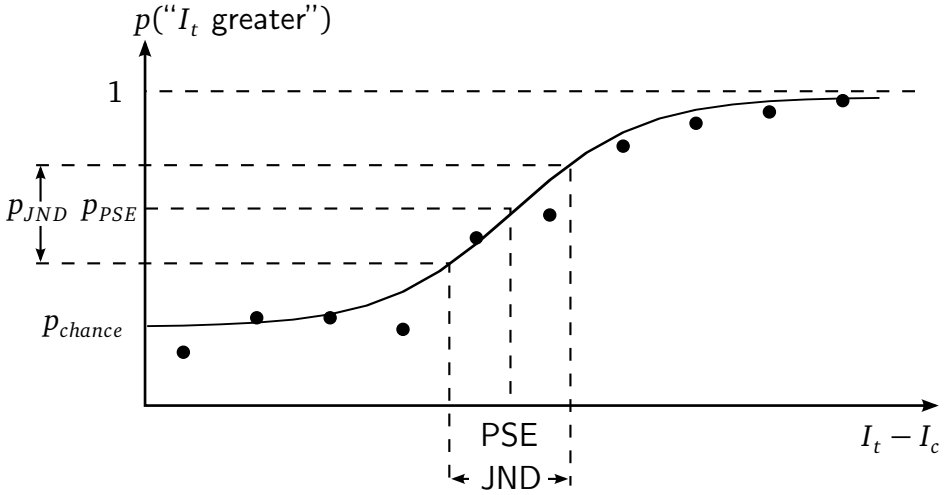


Figure 2.5: The psychometric function relates the difference between a control condition I_c and the target condition I_t to the proportion of "I_t greater" responses. From the psychometric function, the point of subjective equality PSE and the just noticeable difference JND as valuable perceptual performance measures can be deducted.

psychological principles will be outlined before some practically relevant perceptual laws are reviewed which describe the evolving of perceptual thresholds over multiple conditions.

2.4.1 Static Psychophysical Methods

Determining perceptual thresholds requires a toolkit of techniques that allows the extraction of model parameters and threshold values from experimental data. The classical approach for determining these key measures is measuring the psychometric function which captures the statistics of perceptual responses over a range of physical quantities. Two classes of methods that can estimate the absolute and differential thresholds from the psychometric function are discussed – the method of constant stimuli and staircase threshold estimation techniques. We name these methods static as the mapping between the physical quantity and the perceptual response statistics is described by a algebraic function, not considering any dynamics.

The Psychometric Function

The accuracy of human perception is limited, thus physical properties can not be distinguished if they differ too little from each other. In addition, perception is inherently probabilistic, meaning that even if two stimuli $I_t > I_c$ and this difference can normally be perceived correctly, the answer to the question "Is I_t larger than I_c?" is answered incorrectly with "no", e.g., due to a lack of attention. However, with increasing difference, the probability to correctly determine the distinctness p("I_t greater") increases [57].

Definition 2.2 (Psychometric Function [58])**.** The *psychometric function* (PF)

$$P(I) : \Omega_I \rightarrow [0, 1] \tag{2.19}$$

relates the proportion of responses in a sensory task to a physical stimulus value $I \in \Omega_I$. The PF $P(I)$ is centrically symmetric, S-shaped and monotonously increasing.

An exemplary psychometric function is depicted in Figure 2.5. It starts from an offset p_{chance} denoting the chance level and asymptotically approaches 1. The value for p_{chance} is determined by the nature of the experimental task.

Example 2.1. A thought experiment for a perceptual task for the discrimination of tactile stimulus intensity shall be discussed. Consider an experiment where two vibratory stimuli are sequentially applied to the human fingertip. These sequential presentations will be called experimental conditions, one is the control condition, the other is the target condition. Both conditions form one experimental trial. The vibration intensity in the target condition I_t varies from trial to trial, the vibration intensity in the control condition I_c is always at the same level. After each trial, the participant is confronted with a question in which two alternative response possibilities are given, e.g. "Did you experience higher amplitude in the second interval?". In case the target condition with a very low vibration amplitude compared to the control condition, the answer will most likely be "no". The chance level in an experimental task where the target stimuli range from an intensity near zero to a high intensity compared to the control stimulus is $p_{chance} = 0$. In contrast, in the case of a minimum target condition that equals the control condition results in $p_{chance} = 0.5$. The presented experimental procedure introduced here is denoted a 2-interval 2-alternative forced choice (2I2AFC) task.

Two quantities that are particularly important for the design and evaluation of technical systems are the point of subjective simultaneity (PSE) and the just noticeable difference (JND). The PSE is usually defined as the difference between two intensities in the physical quantity that lead to an equal number of correct and incorrect answers, thus $p_{PSE} = 0.5$. The JND denotes human's ability to discriminate between two perceivable stimuli. It is defined as a symmetrical range around the PSE, normally corresponding to the values in intensity differences leading to a probability on the ordinate scale from 0.25 to 0.75.

Remark 2.3. In most cases, the PSE coincides with a difference $I_t - I_c = 0$, however there are exceptions. As an example the point of subjective simultaneity (PSS) that captures the difference in presentation time between two stimuli is nonzero for the case of a visual, and tactile stimulus. Indeed, a probability of 0.5 for responding "simultaneous" is achieved when the visual stimulus is presented approximately 20 ms *earlier* than the tactile [59].

Remark 2.4. A similar procedure as presented here is the determination of the absolute threshold, also denoted as *detection threshold* (DT). In this case, the perceptual question is "Did you perceive a stimulus?" and the psychometric function measures the DT instead of the PSE.

Method of Constant Stimuli

The classical method for measuring the psychometric function which is still used nowadays is the method of constant stimuli [57]. In this experimental procedure, typically five to nine target conditions are selected, distributed usually symmetrically around I_c and covering the whole psychometric function. The range is usually determined using preliminary prestudies, or prior knowledge, e.g., from the literature. A repeated presentation of all conditions over usually 100 or more times allows an estimation of $p(\text{“}I_t \text{ greater”})$ from the fraction

$$p(\text{“}I_t \text{ greater”})\Big|_{I_t^*} \approx \left.\frac{N_{\text{“}I_t \text{ greater”}}}{N_{\Delta I_1}}\right|_{I_t^*}$$

where $N_{\Delta I_1}$ is the number of presentations of the target condition I_t^*. The overall psychometric function is then usually fit towards the experimental data by using a least squares estimation technique. Multiple parametric forms are available as PF candidates. The best fit is usually subject to optimization or known from prior experience.

Staircase Threshold Estimation

Staircase estimation procedures are known as adaptive methods to experimentally estimate perceptual thresholds. Their main advantage over the method of constant stimuli is that thresholds can be estimated using less trials. The idea of staircase procedures is the stepwise approximation of the JND, by adapting the target condition based on the perceptual response of the last presented target stimulus and all prior responses. As an example, the participant in an experiment should judge if an intensity I_t is higher or lower in the target condition compared to a control condition I_c. This judgment is easy if the difference ΔI between stimulus and control condition is large, see Figure 2.6 for an example. Subsequently, if the answer was correct, the difference between stimulus and control condition is reduced. Coming to the discrimination threshold, the participant starts to guess which one is the higher-damped condition. In the case of a wrong answer, the difference between conditions is again increased such that it becomes perceivable again. After a certain number of staircase direction reversals, the procedure terminates and the perception threshold is obtained by taking a weighted average of the last stimulus conditions.

Two variants of staircase procedures are utilized in this thesis – a single staircase procedure to estimate the discrimination thresholds of damping in Section 5.3 [60] and a double-staircase procedure to estimate the time delay detection threshold in Section 4.4.2.

In Section 5.3, the single staircase procedure is combined with an unforced choice paradigm, introducing the possibility to answer “I don’t know” in case no difference between control condition and stimulus condition is perceived. For this procedure, one starts from a surely detectable difference ΔI_1 and the stimulus difference of the following trial $N+1$ is adjusted following the rule

$$\Delta I_{i+1} = \begin{cases} \Delta I_i - \delta I & \text{in the case of a correct answer} \\ \Delta I_i + \delta I & \text{in the case of a “don’t know” answer} \\ \Delta I_i + 3\delta I & \text{in the case of a wrong answer} \end{cases} \tag{2.20}$$

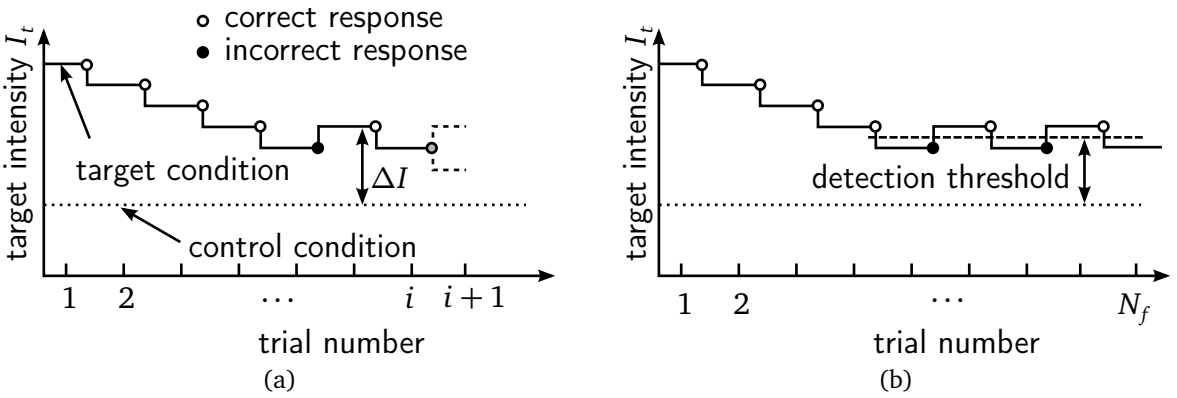


Figure 2.6: (a) The estimation procedure of detection thresholds using a staircase procedure adapts the stimulus condition to be presented in the next trial $i+1$ based on the actual perceptual response (correct/incorrect) and all prior responses. (b) The difference detection threshold is obtained by averaging over the last stimulus conditions.

where δI is an initial step size chosen by the experiment designer. In order to increase the measurement resolution and speeding up the detection threshold estimation procedure, the step size δI can be adapted during the experiment: We halved the step size after the second and forth reversal of the staircase direction, such that the final resolution was four times higher than the initial resolution. The detection threshold is calculated by taking the mean of all δI_i where i takes values between the fourth and eighth reversal of the staircase direction.

The double-staircase procedure utilized in Section 4.4.2 uses two staircases instead of one, where one staircase starts from a surely detectable stimulus difference ΔI_1 and one starts from the control condition, thus $\Delta I_1 = 0$. From which staircase the stimulus condition ΔI_{N+1} is taken is randomly chosen. In contrast to the unforced choice in the single staircase procedure, the participants in this study were required to choose between "difference detected" and "no difference detected". A modified version of the accelerated stochastic approximation method was implemented for each staircase sequence [61]. Following this procedure, the stimulus difference ΔI within the same staircase sequence is calculated according to the adaptation rule

$$\Delta I_{i+1} = \Delta I_i - \frac{\delta I}{1+N_{rev}}(Z_i - 0.5), \tag{2.21}$$

where N_{rev} denotes the number of reversals between the response categories. Z_i encodes the user's response on trial number i that was performed on the respective staircase as $Z_i = 1$ for a "differences detected" judgment and $Z_i = 0$ for a "no difference detected' decision. The procedure stops when (i) both staircases have converged after 5 response reversals and (ii) the mean difference of the last three trials between two sequences is less than δI. Otherwise, the sequence terminates after a total of 40 trials. The detection threshold is estimated by taking the average over the last six trials, 3 trials for each staircase sequence.

2.4.2 Perceptual Laws

Knowing the perceptual limit of one environment property in one specific condition is usually of limited practical relevance. Instead, the change in perception thresholds, e.g., with a shift in the control condition is of greater interest especially for engineers who want to optimize a technical application with respect to human characteristics. Many of these quantitative relations, describing the relation between a specific haptic property and a perceptual response, such as the JND or the intensity of the percept, are available in the literature. In fact, every regression analysis on the results of a psychophysical experiment could be accounted for as a perceptual law. However, two classical laws that were successful in explaining a large number of different psychophysical phenomena should be discussed: The *Weber-Fechner law* and *Stevens' power law*.

Two of the recognized founders of the research field of psychophysics, E.H. Weber and T. Fechner were amongst the first formulating a relation between a physical stimulus and the perceptual response of it - the well-known *Weber-Fechner Law* [62]. It states that the just noticeable difference increases proportional to the magnitude of the stimulus, namely

$$\frac{\Delta I}{I} = W = \text{const.}$$

where ΔI is the JND, I the stimulus intensity and W the so-called "Weber fraction". The Weber-Fechner Law successfully describes a magnitude of JNDs, e.g., mass/inertia [14, 63], damping/viscosity [12, 14], stiffness [11] as well as stimuli in other sensory domains. The law proved its usefulness for technical applications, e.g., in lossy data compression schemes [64] where the amount of haptic data to be transmitted over a telepresence system's communication channel can be significantly reduced without affecting perceived transparency.

Stevens' power law [65] states that the perceived magnitude of a physical stimulus is a power law of the actual magnitude,

$$S = kI^{\alpha},$$

where S is the perceived stimulus magnitude, I the real stimulus intensity, k an arbitrary constant, and α a stimulus-specific exponent. The empirical law was confirmed over a large number of stimuli and domains, e.g., loudness in auditive perception, brightness in visual perception, and muscle force in the haptic domain. A technical application of this law was reported e.g. for the design process of auditive warning signals [66].

2.4.3 Dynamic Diffusion Models

Perceptual laws come short in capturing a) the influence of time-varying sensory signals, e.g., due to different manual exploration, b) accurately predicting the humans' perceptual state over time, e.g., during the operation of a telepresence system. Diffusion models overcome this deficit and have been successfully applied to capture the characteristics of response time and accuracy in modalities other than haptics [67–69]. A diffusion process can be understood as a linear, stochastic dynamical system, as depicted in Figure 2.7. It consists of three parts: The information encoding stage with the dynamical properties of a linear transfer

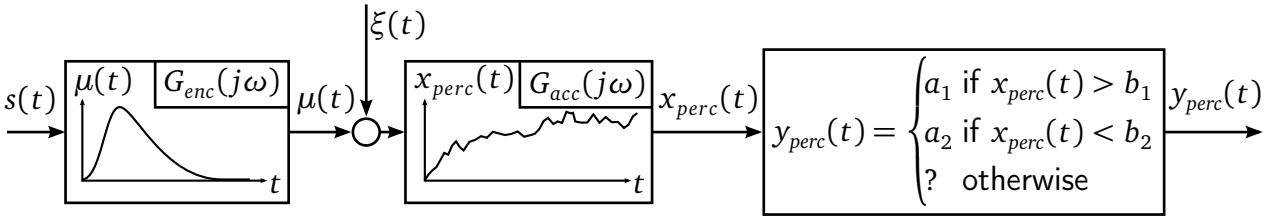


Figure 2.7: A diffusion model for dynamic perceptual processes, consisting of an encoding stage $G_{enc}(j\omega)$, transforming the raw sensory information $s(t)$ into relevant information $\mu(t)$, a noise term $\xi(t)$, an information accumulation process $G_{acc}(j\omega)$ producing the continuous (scalar) perceptual state $x_{perc}(t)$ and the decision stage thereafter, leading to a perceptual response $y_{perc}(t)$.

function $G_{enc}(j\omega)$ extracts relevant information $\mu(t)$ from the sensory signal $s(t)$. The precision of sensory information is limited by a noise process $\xi(t)$ before it is accumulated by an integrator function $G_{acc}(j\omega)$ to the scalar perceptual state variable $x_{perc}(t)$. In the case of a perfect integrator, it can be calculated as

$$x_{perc}(t) = \int_0^t \mu(\tau) + \xi(\tau)\,\mathrm{d}\tau. \tag{2.22}$$

The human detects whether the accumulated information $x_{perc}(t)$ exceeds one of two criteria b_1 and b_2 and chooses between two alternative responses a_1 and a_2. If $x_{perc}(t)$ ends between both boundaries, he/she makes a guess.

Classical psychophysical studies, e.g. the ones concerned with the measurement of JNDs, can be interpreted in the framework of diffusion processes as well. Methods such as the method of constant stimuli, or staircase methods, allow the participant to respond at a fixed time instance t_{resp}. The value for $x_{perc}(t_{resp})$ is thus, taking (2.22) into account,

$$x_{perc}(t_{resp}) = \int_0^{t_{resp}} \mu(\tau)\,\mathrm{d}\tau + \int_0^{t_{resp}} \xi(\tau)\,\mathrm{d}\tau.$$

Repeating the measurement several time minimizes the influence of the second part as

$$E\left\{\int_0^{t_{resp}} \xi(\tau)\,\mathrm{d}\tau\right\} = 0,$$

where $E\{\}$ is the expected value operator. Thus, the classical approach with a fixed response time captures the a measure with is related to the mean information content in the stimulus, $\bar{\mu}$ [69].

2.5 Technical Perception Models

While the dynamic modeling of perceptual and sensorimotor processes is a novel concept, diverse methods and techniques to characterize the dynamics of technical system have been developed in the past. When trying to find an analogy between the human perceptual system on the one, and measurement and estimation techniques for technical systems on the

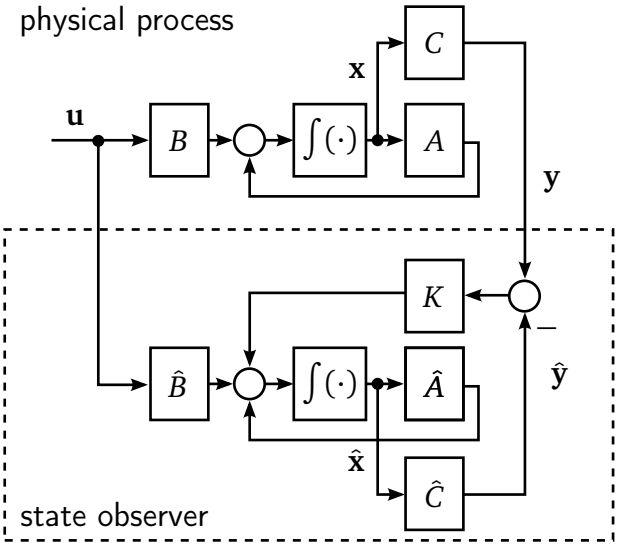


Figure 2.8: A Luenberger observer takes system inputs and outputs into account to compute a state estimate $\hat{\mathbf{x}}(t)$ of the real system state $\mathbf{x}$.

other side, two problem classes can be distinguished: First, the observation of time varying *signals*, which transport energy and are not necessarily directly measurable. Second, the quantification problem of specific *features* within the environment or technical system under investigation, relating two or more signals together. The first class is referred to as state observation problem, the second as system identification problem.

2.5.1 Dynamic State Observation

Definition 2.3 (Dynamic State Observer)**.** A dynamic process or algorithm that allows reconstructing non-measurable states and associate measures of a dynamic system based on a system model is defined as *dynamic state observer*.

A well-known observer structure for linear systems that shall be introduced here is the Luenberger observer [70] which requires measurements of input- and output signals to reconstruct the underlying, non-measurable state vector of the plant system. The observer structure is depicted in Figure 2.8.

The state estimation principle of a Luenberger observer founds on the existence of an internal model of the plant dynamics.

$$\dot{\hat{\mathbf{x}}}(t) = \hat{A}\hat{\mathbf{x}}(t) + \hat{B}\mathbf{u}(t) \tag{2.23}$$

$$\hat{\mathbf{y}}(t) = \hat{C}\hat{\mathbf{x}}(t). \tag{2.24}$$

By feeding back the error between the real, measured system output and the output prediction from the internal model 2.23-2.24, the state estimation asymptotically converges to the true state of the physical system. Given that the system matrices A, B, and C are known with sufficient accuracy, such that

$$\hat{A} \approx A, \quad \hat{B} \approx B, \quad \hat{C} \approx C, \tag{2.25}$$

the state estimation is the solution of the differential equation

$$\dot{\hat{\mathbf{x}}} = A\hat{\mathbf{x}} + B\mathbf{u} + K(\mathbf{y} - C\hat{\mathbf{x}}). \tag{2.26}$$

In technical systems, the feedback matrix K is chosen by the system designer. In this way, it can be ensured that the observer behaves stably, and with a desired behavior in terms of transient behavior.

2.5.2 System Identification

In practical applications, the mathematical description of a physical system which is to be controlled, is often unknown. However, this knowledge is crucial for the design of control algorithms, or for the verification of a design process. In certain cases, the dynamic system model can be developed purely on the basis of physical considerations, such models are called *white-box* models. The class of problem where this kind of modeling technique is suitable is limited in practice, as purely physical models tend to exhibit high complexity. In addition, the system structure must be known, and all parameters either measurable or obtainable from other sources. This is a strong limitation in, e.g., biological systems whose working principles are barely comprehended. Alternatively, system identification techniques allow the development of such mathematical models on the basis of measuring the input and the output to the system under consideration [21]. Two classes of methods can be distinguished: *Black-box* identification techniques do not assume any prior knowledge such as a fundamental physical structure, or initial guesses for parameters to be identified. In contrast, *gray-box* models have some basic knowledge how the system behaves in principle, but no knowledge about the parameters determining the quantitative behavior. We restrict ourselves to the identification of black-box and gray-box models in the following, and some fundamental concepts are reviewed. For a deeper insight into the complex field of system identification, the interested reader is referred to [21–23], and the references therein.

Identification Procedure

For simplicity reasons, all basic concepts will be discussed on the basis of a single-input, single-output, linear, time-invariant, stable dynamic system. Following the notation in [21], three basic entities are required for a successful identification of a given dynamic system:

- A set of data, denoted Z^N,
- a model structure,
- a technique for addressing the model structures with respect to the data.

Due to the fact that system identification is usually accomplished using a computer-aided procedure, all signals contained in the dataset Z^N are available in discrete time steps, thus

$$Z^N = \{u(1), y(1), \ldots, u(N), y(N)\}, \tag{2.27}$$

where $u(i) \in \mathbb{R}$, $i = 1 \dots N$ is a sequence of input values, and $y(i) \in \mathbb{R}$, $i = 1 \dots N$ the corresponding sequence of output values. Input and output sequence are assumed to be stationary random processes. The choice of model structure differentiates black-box models from gray-box models. In the case of a black-box model identification, so-called non-parametric models, e.g., the frequency response of the system, or the impulse response, are considered. In contrast, gray-box models utilize prior knowledge about the process and propose parametric models, e.g., a mass-spring-damper system with unknown mass, spring and damping parameter. The coefficients describing the impulse (or frequency) response, or alternatively the model parameters can be collected in a parameter vector θ which is to be determined by an appropriate identification technique. One exemplary identification method which is practically relevant for this thesis is discussed in the following.

Covariance Method for Black-Box Identification

Considering a dynamic system with impulse response $g_0(t) \in \mathbb{R}$, the response to a discrete-time input signal $u(t) \in \mathbb{R}$ is computed as

$$y(t) = \sum_{k=1}^{\infty} g_0(k)u(t-k) + v(t). \tag{2.28}$$

where $v(t)$ is a noise term accounting for measurement noise [21]. assume the input to be quasi-stationary with

$$E\{u(t)u(t-\tau)\} = R_u(\tau) \tag{2.29}$$

and

$$E\{u(t)v(t-\tau)\} \equiv 0,$$

where $R_u(\tau)$ is called the *autocovariance* function of the input $u(t)$, and τ is denoted as time shift operator. It can be shown [21] that the *crosscovariance* function between input $u(t)$ and output $y(t)$ is computed as

$$E\{u(t)y(t-\tau)\} = R_{yu}(\tau) = \sum_{k=1}^{\infty} g_0(k)R_u(k-\tau). \tag{2.30}$$

In reality, the amount of data samples taken into consideration for the estimation of $R_u(\tau)$ and $R_{yu}(\tau)$ is finite, thus the expected value operator in (2.29) and (2.30) is replaced by the mean, and the sum in (2.28) is finite, namely

$$\hat{R}^N_{yu}(\tau) = \sum_{k=1}^{N} \hat{g}(k)\hat{R}^N_u(k-\tau), \tag{2.31}$$

where $N \in \mathbb{N}$ is the number of samples taken into consideration for the estimation of autocovariance $\hat{R}^N_u(\tau)$ and crosscovariance $\hat{R}^N_{yu}(\tau)$. An optimization technique, e.g., a least squares algorithm, can be used to find the impulse response function $\hat{g}(t)$ from (2.31).

Remark 2.5. The estimation accuracy for $\hat{g}(t)$ depends on the number of samples N taken to estimate $R^N_u(\tau)$ and $R^N_{yu}(\tau)$, respectively. This is evident from the fact that the mean value approximates the expected value only for large N. In addition, N limits the maximum length of the impulse response which can be estimated.

2.6 Quantitative Behavior and Task Performance Models

Everyday-life encounters tell us that humans behavior is not random in many situations, but often tasks are performed stereotypically. This repeatability in motor behavior has inspired researchers to develop behavioral models on the one hand, and models describing the performance in the given task on the other hand.

2.6.1 Empirical Models

One of the first models describing reaching movements towards a specific goal is Fitt's Law [71], describing a tradeoff between the movement speed and the achievable accuracy of the reached destination. Many extensions for the model, e.g. for 2-dimensional [72] and 3-dimensional [73] reaching movements, and for pointing under different feedback conditions such as time-delayed visual feedback [74] have been proposed. A similar tradeoff between the curvature of a trajectory and the angular velocity was found in the 2/3 Power Law [75]. A control theoretic model for the human operator following a trajectory using different input devices was proposed by McRuer and Jex [76] who were seeking for a possibility to describe the human steering behavior of an aircraft. However, the model proved its usefulness in a number of tracking tasks different to steering as well. A model that takes certain physiological properties of the movement apparatus into account is the equilibrium point hypothesis [77, 78], modeling the muscles attached to a moving limb as nonlinear springs. In a resting posture, all springs are in their equilibrium position; movements are caused by voluntarily changing the spring characteristics of an individual "muscle". Although the equilibrium point hypothesis is able to successfully describe certain aspects of human motor control, such as arm stiffness during multi-joint movement [79], one of the main criticism against it is the lack of physiological plausibility for the spring model being representative for a real muscle.

2.6.2 Optimal Control Models

Instead of searching for an empirical law that describes the stereotypical behavior of humans controlling their body and the environment surrounding them, optimal control models try to capture the underlying "biological idea" that leads to a specific task-dependent human behavior. The underlying assumption is that all biological processes converge to a somehow optimal stage, an idea first formulated by Darwin and Spence in 1864 by the concept of the "survival of the fittest" [80]. The optimal control approach is more general than empirical modeling since it is not limited to a specific control law in a certain task under certain circumstances, but can result in different control laws with different parametrization, depending on task, environment, and situation. An optimal control law is though only meaningful with respect to a certain criterion $J(\mathbf{x}_h)$ that is minimized, denoted by

$$\arg\min_{\mathbf{x}_h} J(\mathbf{x}_h). \tag{2.32}$$

A diversity of costs have been proposed, explaining movement stereotypes differently well in different tasks. Among the first criteria introduced as an optimization function for sensorimotor control was energy, proposed by Chow & Jacobsen [81], which is able to explain especially walking patterns well. Another popular optimization functions is the "jerk" as the time-derivative of acceleration, introduced by Hogan [82], or joint torque change [83]. For comprehensive reviews over additional task-specific optimization criteria, and further investigations and generalizations of the presented cost functions, see references [84, 85].

The optimization criteria can be used to explain average observed human behaviors post-hoc on the one hand, and for pre-planning an optimal muscle action pattern in advance on the other hand. These cases are referred to as open-loop optimal control. Such algorithms have the disadvantage that changing task conditions, or unexpected events can not be taken into consideration for the execution of a certain task. Furthermore, the sensorimotor control loop is corrupted by noise on many different levels [86]. Noise and other stochastic uncertainties make open-loop control infeasible but require the feedback of sensory data. Controllers minimizing the expected variance of the hand's final position in a reaching task were proposed to explain human stereotypical behavior and reaction to disturbances [87, 88].

2.6.3 Dynamic Models in Sensorimotor Control

Optimal feedback control algorithms inevitably require dynamic knowledge of the plant to be controlled in order to predict the consequences of a certain control action and choose the appropriate one for optimally achieving a desired goal. The existence of dynamic models of the human body represented internally in the human brain was identified by Ito in 1970 to take place in the cerebellum [89]. The idea that these internal models could be used for motor control was first proposed by Kawato et al. in 1987 [90]. Their motivation of introducing such models in the first place was to find an explanation for highly dynamic human movements facing significant neural and sensory delay. Therefore, they postulated that an inverse model of the human body dynamics, calculating the required motor commands in order to achieve a certain control goal is used, combined with a feedback control law accounting for model uncertainties and changing environment conditions. The resulting structure for the sensorimotor control loop is depicted in Figure 2.9. Such a control concept is indeed able to increase control bandwidth significantly. Although inverse models are difficult to learn from a computational point of view, efficient algorithms utilizing the neural structure of the cerebellum were found that can support their existence [90, 91]. With the help of feedforward and feedback control, a number of findings on motor control [92] and eye movements [93] can be well explained.

Another way to use dynamic models in sensorimotor control is by using forward models, predicting the state of the human body, based on the actual and past motor commands, available in the form of an efference copy. Inspired from optimal control theory with noisy measurements, the Kalman filter theory became popular in recent years to describe the human's ability to estimate his/her internal body state [84, 88, 94–97] and base their control actions upon this estimate. In the case where the human controls a haptic environment or a tool, there are indications that the tool dynamics are also represented in the cerebellum and can be used for motor control as well [98, 99]. A sensorimotor control model based

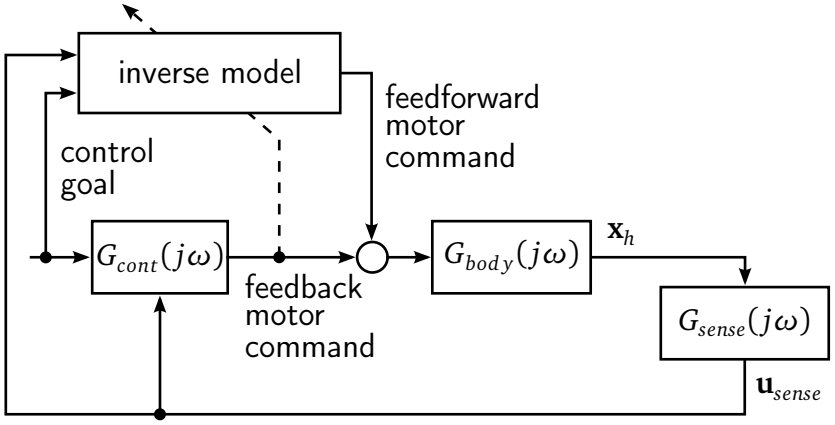


Figure 2.9: A sensorimotor control model including feedforward and feedback control, making use of a forward model of the human body dynamics.

on the Kalman filter approach is depicted in Figure 2.10. Two internal dynamic models are used in the Kalman filter: The first contains a forward model of the human body dynamics to make *state predictions* about the future state of the body $\mathbf{x}_h^*$. Therefore, a neural copy of the actual motor commands, the *efference copy*, is used. Because this state estimate is purely based on a feedforward-estimation it is prone to external disturbances and modeling uncertainties. Thus, a second dynamic *sensor model* is introduced, mimicking the real sensory system $G_{sense}(j\omega)$ and leading to a prediction of the sensory signal $\hat{\mathbf{u}}_{sense}$. By comparing $\hat{\mathbf{u}}_{sense}$ with the real, measured sensory feedback $\mathbf{u}_{sense}$, mismatches between feedforward-prediction and the actual sensor feedback are taken into consideration using the *Kalman gain* to correct the actual state estimate $\hat{\mathbf{x}}_h$ towards the noise-optimal estimate of the actual body state $\mathbf{x}_h$. This state feedback can be used for a control law $G_{cont}(j\omega)$, driving the body dynamics $G_{body}(j\omega)$ to achieve the given *control goal*.

Whether or not inverse dynamic models, the Kalman filter or any kind of model-based control approach is actually encoded in neural substrate and utilized for human sensorimotor control can not be answered with certainty. Proving the secure existence of such dynamic models and their role in motor control requires a much deeper understanding of the organizational processes within the brain, and advanced imaging techniques. However, the fact that the presented models are plausible from an evolutionary point of view and make accurate predictions in multiple experimental and real-life situations justifies further investigations.

2.6.4 Task Performance Models

A task performance model describes the influence of specific environmental properties on the human ability to achieve a specific goal. A comprehensive overview over a number of different performance metrics for task-oriented human-robot interaction is provided in [100]. Although every task aims for a slightly different goal and thus could be evaluated with respect to measures tailored to the specific scenario, some performance measures are used more common than others: The number of failures [101, 102], accuracy [76, 103], and task completion time, also known as time-to-complete [37, 71, 101, 102, 104]. A failure in the

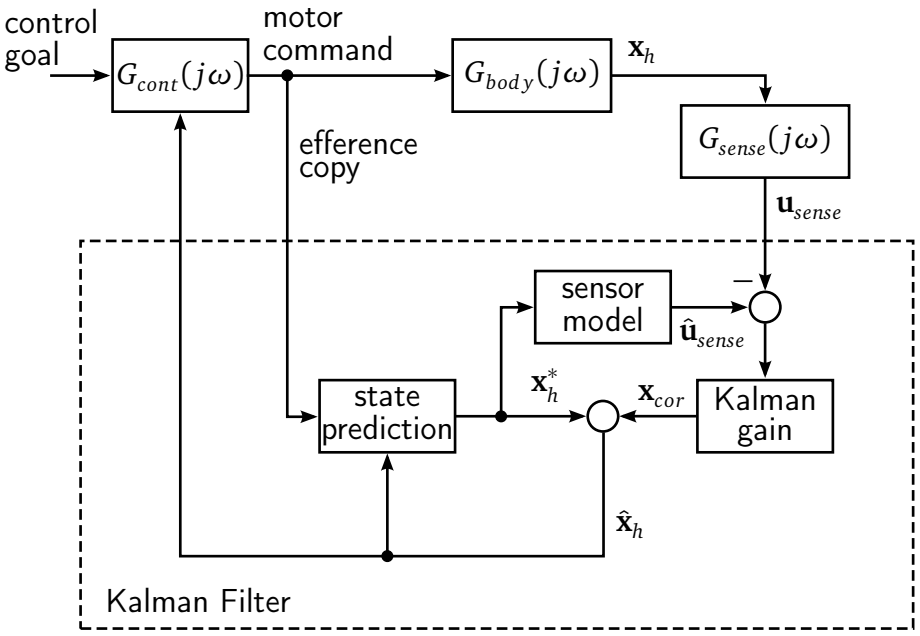


Figure 2.10: A sensorimotor control model based on a Kalman Filter model.

operation of a human-system interface could be characterized, e.g., by a collision with an obstacle, or a damage to a workpiece to be handled. As a failure is characterized by an event, the number of failures is discrete. In contrast, accuracy characterizes a *continuous* measure for the precision that was achieved in the task, e.g., the tracking accuracy in a path following task, measured by the mean squared error. Those performance measures can be used in a variety of telepresence systems and tasks. To just name a few examples, navigation tasks should be performed fast without collisions, pick-and-place tasks are usually required to be accomplished with a high accuracy in little time, and telesurgery needs to be performed precisely.

Task performance is currently primarily described in terms of static relations between an environment property such as the communication time delay [37], or task difficulty [71], and the performance measure. One drawback of this approach is that quantitative conclusions about the performance can be only made after task completion. Static models thus qualify for an application in the optimization of telepresence systems with respect to the average task performance that can be achieved under certain environmental conditions. They can, however, not be used to optimize a telepresence system online, that means, during task execution, to achieve a higher task performance, given the momentary manual interaction and task situation.

2.7 Discussion

Telepresence systems aim for extending the workspace of a human operator to locations that are not accessible to the human by his/her capabilities otherwise, e.g., due to distance, size, or harmful environmental conditions. Because of the strong integration of the technical system into the action-perception loop, the technical system must be designed to suit the

operator's perceptual and motor capabilities. Attempts to include quantitative knowledge about the operator's perceptual and sensorimotor characteristics into a human-centered design of telepresence systems have been made earlier: A perception-optimized design of a data compression algorithm was proposed for communication between human-system interface and teleoperator, transmitting haptic information about movement and force feedback [8, 64, 105, 106]. Not every change in these quantities can be perceived by the human operator. Thus, data packets that contain irrelevant data for the human perceptual system are not transmitted. A selection algorithm based on Weber's law was used to separate relevant from irrelevant data. If all irrelevant packets are discarded, the required amount of data can be reduced significantly without affecting perceptual fidelity. A second example is an algorithm that can decrease task completion time in a telepresent reaching movement with time-delayed visual feedback, based on a minimum-jerk movement model to predict the future movement trajectory [107]. Instead of the actual movement, the prediction is transferred over the communication channel, which virtually decreases the time delay in the case of an accurate trajectory prediction.

These exemplary applications are different in several aspects: First, one application is targeting a perception-optimized and the other a performance-optimized telepresence experience. Furthermore, Weber's Law is a static approximation of the dynamic perceptual apparatus, meaning it can only relate a momentary sensory input to a momentary perceptual response. In contrast, the minimum-jerk model contains the dynamics of the sensorimotor process and results in a prediction of the time series data of the movement trajectory, founded on the momentary and past recordings of the movement. Third, the data compression algorithm continuously adjusts its strategy according to the optimization objective – the perceptual fidelity of the haptic environment. Such continuous adjustments are especially important when a tradeoff has to be made, in this case between the amount of transmitted data and perceptual fidelity. In comparison, the trajectory prediction algorithm does not take the objective of a performance enhancement into direct account for online-adjustments but aims for an increased average task performance. the last difference to be mentioned here lies in the information quantity of each observation: While the perceptual model can predict a discrete, binary decision about the perceptual state (difference detected/not detected), the minimum jerk trajectory is continuous in its values, thus is able to provide richer information, such as position and velocity.

Taking together the advantages of both application examples, an ideal way to design, control, and enhance telepresence systems with respect to the capabilities of the human operator can be outlined: Taking the *dynamic* characteristics of perception and sensorimotor control into consideration to optimize perceptual fidelity and task performance *online*, founded on a *continuous* estimate of the perceptual and performance state. Comparing these requirements with the capabilities of today's perceptual laws and performance measures, summarized in Table 2.1, reveals their shortcomings: Due to the fact that static perception models can not take the influence of movement on the perceptual thresholds into account, their online-prediction capabilities are limited and may be inaccurate. Continuous information about perceptual processes is furthermore limited to static laws such as Steven's Power Law [65], relating stimulus intensity to the equivalent intensity of the percept. This quantity, however, may have only limited practical meaning in the design process of technical systems. On the

	model type		model predictions		information quantity	
	static	dynamic	average	online	discrete	continuous
Ideal perception & performance model		✓		✓		✓
Perception laws	✓		✓	(✓)	✓	(✓)
Task performance measures	✓		✓		✓	✓
Diffusion models		✓	✓	✓	✓	✓
Sensorimotor control models		✓	✓	✓	✓	✓

Tabular 2.1: Overview of the requirements for perceptual and task performance models to be suitable for the optimization of telepresence systems. Current perception laws and task performance measures do not meet the ideal model structure; diffusion processes and sensorimotor control models serve as a starting point for the development of suitable model candidates.

other side, most task performance measures are quantifiable post-hoc, limiting their applicability to an optimization of average task performance. Dynamic task performance models that allow online-predictions of task performance on the basis of current and past inputs and feedback could overcome this shortcoming.

Inspiration for the development of novel dynamic human perception and task performance models can be drawn from diffusion models and sensorimotor control models on the one hand, and technical perception methods on the other: Respecting the three stages in the diffusion model, namely information encoding, accumulation, and decision stage, dynamic models in perception must capture two dynamic aspects as well – the information encoding dynamic and the accumulation dynamics. Models describing sensorimotor control processes make use of internal models, representing the dynamics of body and environment. Such internal models may play a role in perception as well. Task performance results from human behavior, given a certain environment with a specific task setting. The dynamics of the environment, technical system, and the task itself are thus not only to be contained in the sensorimotor control model, but can additionally enhance task performance models, to make performance predictions based on the current and past interactions more accurate.

This thesis aims for starting off a new, dynamic view on haptic perception and task performance, based on the fundamentals presented in this chapter. Two dynamic aspects of haptic perception are developed – the dynamic combination mechanism of movement and force feedback, and masking effects between different environmental properties. For a performance-optimal design of telepresence systems, dynamic performance models are furthermore developed, founded on the environment dynamics and human motor capabilities. These contributions are significant steps towards a fully human-centered design of technical systems that take dynamic perceptual and behavioral properties of the operator into account.

3 Dynamic Modeling of Sensorimotor Processes

Summary. *Describing dynamic properties of haptic perceptual and sensorimotor processes requires a system theoretic view on the sensorimotor loop of information exchange within the human, and with the environment. A mathematical framework on the basis of coupled dynamic systems is proposed as a foundation for the development of such models. This framework allows a*

- *systematic development of dynamic perception and behavioral models,*
- *consideration of existing static perception and action models as a special case of a dynamic sensorimotor process,*
- *direct integration of the developed models into technical applications.*

Considering human knowledge in the design, evaluation and optimization of telepresence systems requires the formulation of quantitative relations between measurable physical quantities and hardly-observable reactions within the human operator. Most currently available perception laws utilized for this purpose (see Section 2.4.2) are limited to a direct, static mapping to describe the perceptual equivalent of a physical property. Perception of haptic environments is though more complex in many cases and such simple models are not suitable for describing human perception sufficiently well: As an example, the just noticeable difference of a mechanical impedance, such as stiffness, is commonly modeled as following Weber's Law, see Section 2.4.2. However, stiffness is not perceived directly since the human lacks receptors for this physical property. Instead, stiffness can only be inferred by exciting the object, e.g., applying a certain force, and observing the deformation. These information are transmitted to the brain where cognitive processes reason about the abstract quantity "stiffness" by combining force and movement together. In this sense, the perceptual equivalent for the physical stiffness estimation does not only depend on the physical stiffness but additionally on the excitation, biological factors such as sensor accuracy, and the performance of the cognitive system to extract the original environmental feature from the available information. Most considerations in this thesis are given to human models describing the direct mapping between haptic stimuli and human perception and behavior. However, different quantitative human models are required in other technical applications: In the development of novel neuroprosthesis, biofeedback is utilized to infer an intended movement which is carried out by the actuated prosthetic limb [108]. For doing so, physiological models describing the mapping between neural information and motor commands must be available. With regard to human perception, inner psychophysical models mapping physical stimuli to neural signals, and the neural response further to a perceptual equiva-

lent [57, 109] are of interest to the haptic community. Physiological knowledge is so taken into consideration for the development of inner psychophysical models.

A physiologically inspired mathematical framework for the systematic development of quantitative human models is proposed in the following. The structure of the framework is flexible enough to capture physiological, perceptual, and behavioral phenomena on different detail levels. While modeling human perception as a set of static mathematical mappings was proposed earlier [110], the fact that human haptic perception is closely linked to the motor system and is indeed a dynamic process [111, 112] was so far neglected. The framework proposed here for a quantitative modeling of processes within the human sensorimotor control loop is able to include dynamics by modeling the closed action-perception loop as a set of coupled differential equations.

This chapter is organized as follows: In Section 3.1, the sensorimotor loop is modeled taking inspiration from the physiological system into consideration, and dynamic behavior is identified on all levels of information transmission and processing required for haptic environment interaction. An abstraction of this physiological view is presented in Section 3.2 where a mathematical framework is introduced in the form of sensorimotor domains, sets, transformations, and observations. The relation between quantitative models based upon the framework introduced here and existing models for perceptual processes and performance models are detailed in Section 3.3 and 3.4. A discussion of the framework's properties with respect to the prediction abilities of the resulting human models and the advantages compared to other approaches concludes the chapter in Section 3.5.

3.1 Dynamic Action-Perception Loop

Modeling processes in human perception and behavior requires a basic understanding of the involved physiological, cognitive, and physical processes. A schematic diagram depicting the exchange of information and energy within the action-perception loop and with the haptic environment is depicted in Figure 3.1. For the benefit of presentation clarity, all dependencies on time and frequency are omitted in the Figure but will be mentioned in the text.

Remark 3.1. All discussions of sensorimotor processes will be discussed in a simplified version using one pair of agonistic/antagonistic muscles acting on a single limb. In reality, the motor apparatus consists of a multitude of muscles acting on different limbs. The discussions in this Section could be extended to the more general case of multiple muscles for one limb by using the forward kinematics of the musculoskeletal apparatus, see e.g. [113].

Six domains and classes of signals and blocks are traditionally distinguished: *Physical* stimuli from the environment are measured by the human senses and encoded into *sensory* signals which are transmitted over *neural* pathways to the brain utilizing them for cognitive actions, namely *perception* and *motor control*. New motor commands are generated, based upon sensory signals and cognitive processes, sent back via *neural* signals to the *motor* system, where they cause a *physical* reaction in the environment. Tangible for the haptic modality, the physical environment is modeled in impedance form, transforming the position of the interacting limb $x_h(t)$ into a reaction force $f_h(t)$. This force is not only recorded by

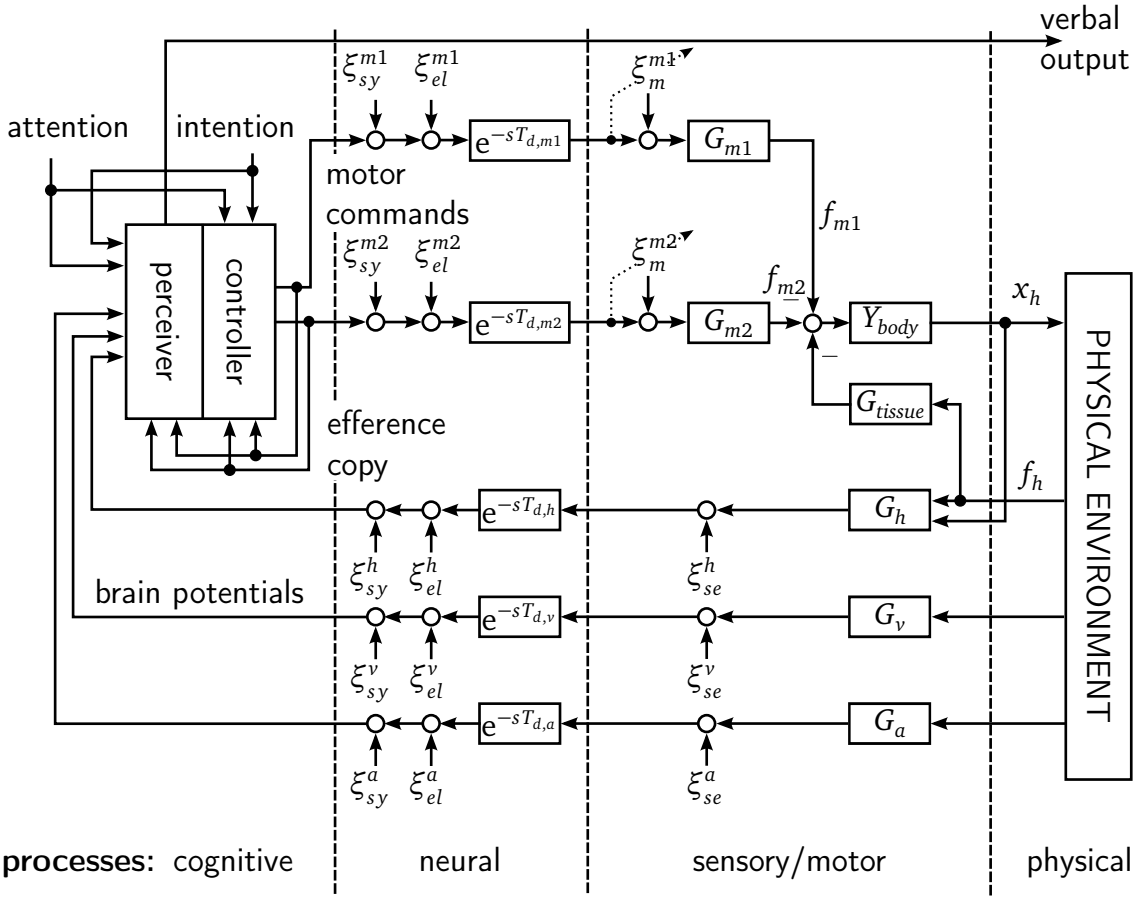


Figure 3.1: A schematic, non-complete information flow diagram of human sensorimotor control processes.

the sensory system but additionally acts physically upon the physiological structure of the body, affecting the effective force that accelerates the body admittance $Y_{body}(j\omega)$. The transfer function modeling tissue dynamics $G_{tissue}(j\omega)$ is part of the body admittance, but plays a particular role in the context of determining the amount of environment force that interferes with the motor actions. In the case of a high interaction force and/or a rigid coupling of environment and body, such as the operation of a kinesthetic haptic interface, the influence of $G_{tissue}(j\omega)$ is negligible, as the tissue is compressed and can be modeled in the compressed state as (quasi-)rigid. In the case of small interaction forces and fast-varying position or force changes, such as in the case of tactile exploration, the tissue dynamics $G_{tissue}(j\omega)$ with its elastic properties reduces the environment force acting back on the bone structure significantly.

Sensors and actors are afflicted with dynamics, represented by transfer functions $G_{m1}(j\omega)$, $G_{m2}(j\omega)$ for the muscle dynamics, $G_h(j\omega)$, $G_v(j\omega)$ and $G_a(j\omega)$ for the dynamics of haptic, visual, and auditive sensing dynamics, respectively. While only three sensor blocks are depicted, the actual number of receptory cells is tremendously higher. Besides these simplifications, the remaining senses of taste and smell are omitted as they play a subordinate role in telepresence. Physiologically inspired models have been developed for describing the dynamics of receptors in the auditory [114], visual [115], and haptic [116]

domain, and muscle activity [117]. Transmission over neural pathways comes with transportation delays $T_{d\{h,v,a\}}$ and $T_{d\{m1,m2\}}$, coming from electrochemical communication. In addition, noise corrupts the signal transfers and processes at various stages [86], accounting for the variability in behavior and perception. We will name the individual noise terms ξ with subscripted noise source and the superscripted sensor modality or involved muscle. We distinguish between sensory noise ξ_{se} which comes from the mechanical (e.g. tactile and audition) or electrical (e.g. optical) transduction and the signal amplification thereafter, cellular and electric noise ξ_{el} due to timing variability of action potentials, and synaptic noise ξ_{sy} that arises at the interface to the brain. For consistency reasons, synaptic and electric noise are subsumed together in the term ξ_{ne} for neural noise. In the motor system, motor noise ξ_m limits the achievable accuracy for an action which is to be performed. In contrast to the other noise sources, this noise intensity is linearly related to the signal level [118], explainable by the number of muscle fibers to be recruited for higher muscle force.

Perception and motor control are closely coupled [111]. There are indications that artifacts such as the attention that is paid to a given task has an influence on both cognitive parts of the sensorimotor loop discussed here [119]. In the field of neurophysiology, the existence of so-called "mirror neurons" is postulated as a common neural substrate for perception and motor action [120]. However, other studies suggest that human behavior is sometimes affected by a distortion in the sensory feedback even if the distortion is below the perception threshold [121]. Due to this inconsistency, the perceptual and motor control system will be treated as individual mechanisms in the following, while we respect explicitly that they interact in many situations. As an additional source of information to the perceptual and motor control system, a neural copy of the commanded motor commands, the so-called *efference copy* is available [84–88, 122, 123]. In contrast to the sensory feedback about the exerted muscle force, the efference copy is nearly undelayed, and assumed to be available without significant noise.

Remark 3.2. All sensory, motor, and body dynamics are depicted and discussed as linear approximations of the nonlinear characteristics of those systems. However, the drawn conclusions and statements made remain valid even if the exact nonlinear behavior is taken into consideration.

3.2 Mathematical Modeling Framework

From a mathematical point of view, the information exchange between the different systems involved in the sensorimotor loop can be expressed as mappings between different sets of parameters and variables, as illustrated in Figure 3.2. Accordingly, the sensorimotor control loop is divided into six sensorimotor domains: The **phy**sical, **se**nsory, **ne**ural, **con**trol, **perc**eptual, and **mo**tor domain[1].

Definition 3.1. A *sensorimotor domain* $a \in \{phy \quad se \quad ne \quad con \quad perc \quad mo\}$ is defined by a triplet (S_a, ψ_a, ϕ_a) where S_a is denoted as *sensorimotor set*, $\psi_a : S_a \to S_a$ contains the domain dynamics, and $\phi_a : S_a \to S_a$ is named *sensorimotor observation*.

[1]The domains are abbreviated *phy*, *se*, *ne*, *con*, *perc*, *mo*, respectively.

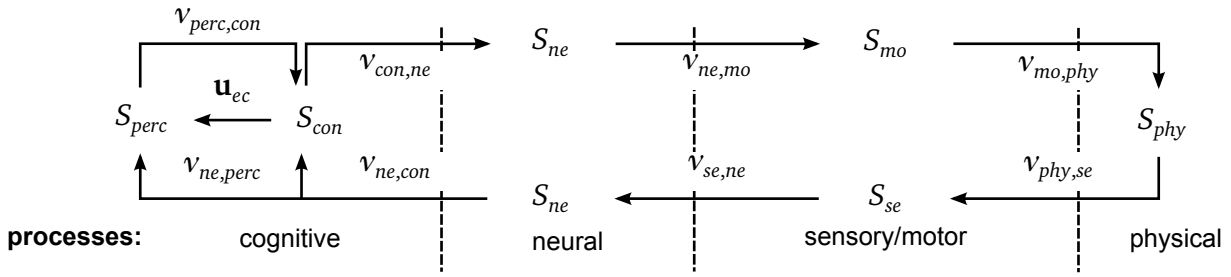


Figure 3.2: Representation of the sensorimotor control loop as a series of mappings between sets of features and states. As a specialty, the perceptual system receives information from the control mechanism about the commanded motor signals – the efference copy $\mathbf{u}_{ec}$.

Definition 3.2. A *sensorimotor set* S_a, $a \in \{phy \quad se \quad ne \quad con \quad perc \quad mo\}$ is defined as

$$S_a = \left\{ \begin{bmatrix} \mathbf{x}_a^T & \theta_a^T & \mathbf{u}_a^T & \mathbf{y}_a^T & \xi_a^T \end{bmatrix}^T \in X_a \times \Theta_a \times U_a \times Y_a \times \chi_a \right\} \tag{3.1}$$

where $\mathbf{x}_a \in X_a$ is a state vector, $\theta_a \in \Theta_a$ is a feature vector, $\mathbf{u}_a \in U_a$ is an vector of input items, $\mathbf{y}_a \in Y_a$ contains quantities that are measurable externally, and $\xi_a \in \chi_a$ contains noise terms and uncertainty in the sensorimotor domain a.

An overview of the different sensorimotor domains can help for an intuitive understanding of their meaning and relevance: The physical world describes the environment surrounding us with its mechanical, optical, acoustic, and other properties. Sensory processes happen within the body, where an accurate modeling is more difficult due to the complexity of the biological system [114–116]. Neural communication is responsible for the signal transmission of sensory signals to the brain. The signal transportation is characterized by spiking electrochemical potentials, varying in intensity and shape. Perceptual and control processes within the human brain are only indirectly measurable by imaging techniques such as EEG and fMRI or observable by verbal responses about the environment perception, which makes it difficult to develop dynamic models of such processes. The motor system contains the actuating muscles as well as the body dynamics. Practical examples for the measures introduced in Definition 3.2 are presented in Table 3.1.

Remark 3.3. Sensorimotor processes are afflicted with time delay. For keeping the notation simple, the time delayed system behavior is not made explicit here, instead we define

$$\mathbf{x}_a(t) = \begin{bmatrix} \mathbf{x}_{a,und}(t)^T & \mathbf{x}_{a,und}(t-T_{d,1})^T & \mathbf{x}_{a,und}(t-T_{d,1})^T & \cdots & \mathbf{x}_{a,und}(t-T_{d,n})^T \end{bmatrix}^T \tag{3.2}$$

$$\theta_a(t) = \begin{bmatrix} \theta_{a,und}(t)^T & \theta_{a,und}(t-T_{d,1})^T & \theta_{a,und}(t-T_{d,1})^T & \cdots & \theta_{a,und}(t-T_{d,n})^T \end{bmatrix}^T \tag{3.3}$$

$$\mathbf{u}_a(t) = \begin{bmatrix} \mathbf{u}_{a,und}(t)^T & \mathbf{u}_{a,und}(t-T_{d,1})^T & \mathbf{u}_{a,und}(t-T_{d,1})^T & \cdots & \mathbf{u}_{a,und}(t-T_{d,n})^T \end{bmatrix}^T \tag{3.4}$$

$$\xi_a(t) = \begin{bmatrix} \xi_{a,und}(t)^T & \xi_{a,und}(t-T_{d,1})^T & \xi_{a,und}(t-T_{d,1})^T & \cdots & \xi_{a,und}(t-T_{d,n})^T \end{bmatrix}^T \tag{3.5}$$

where $\mathbf{x}_{a,und}$, $\theta_{a,und}$, $\mathbf{u}_{a,und}$, $\xi_{a,und}$ is the undelayed state, feature, input, and noise vector, and $T_{d,i}$, $i = 1 \ldots n$ are time delays within the sensorimotor loop.

The processes within each sensorimotor domain are generally dynamic, requiring a description in the form of a stochastic differential equation.

	physical	**se**nsory	**ne**ural	**perc**eptual	**con**trol	**mo**tor
$\mathbf{x}$	position, velocity	lowpass filter	membrane potentials	brain potentials	brain potentials	joint angle, velocity
θ	inertia, damping	amplification	firing threshold	attention	task strategy	muscle fatigue
$\mathbf{u}$	limb position	skin pressure	firing patterns	firing patterns	firing patterns	muscle activation
$\mathbf{y}$	interaction force	EMG signals	EMG signals	verbal response	EEG signals	limb position
ξ	unpredictable behavior	transduction noise	neural noise	neural noise	neural noise	activation noise

Tabular 3.1: Examples for domain inputs, state variables, features, uncertainty sources, and domain outputs in the six sensorimotor domains.

Definition 3.3. *Sensorimotor dynamics* in sensorimotor domain a are captured by a set of nonlinear, first-order stochastic differential equations in the state vector $\mathbf{x}_a$ as

$$\dot{\mathbf{x}}_a(t) = \psi_a\left(\mathbf{x}_a(t), \theta_a(t), \mathbf{u}_a(t), \xi_a(t), t\right), \quad \mathbf{x}_a(0) = \mathbf{x}_{a,0}, \tag{3.6}$$

where $\mathbf{x}_a \in X_a$, $\theta_a \in \Theta_a$, $\mathbf{u}_a \in U_a$, $\xi_a(t) \in \chi_a$.

For sensorimotor control and perception, information must be exchanged between different sensorimotor domains. This communication is abstracted by sensorimotor transformations $\nu_{a,b}$.

Definition 3.4. A *sensorimotor transformation*

$$\nu_{a,b} : S_a \to U_b$$

between the sensorimotor sets S_a and U_b realizing the sensorimotor domains $a, b \in \{phy \quad se \quad ne \quad con \quad perc \quad mo\}$ is defined as a static mapping function $\nu_{a,b}$

$$\mathbf{u}_b(t) = \nu_{a,b}\left(\mathbf{x}_a(t), \theta_a(t), \mathbf{u}_a(t), \xi_a(t), t\right), \tag{3.7}$$

where $\mathbf{u}_b \in U_b$, $\mathbf{x}_a \in X_a$, $\theta_a \in \Theta_a$, $\mathbf{u}_a \in U_a$, $\xi_a(t) \in \chi_a$.

Developing quantitative models crucially requires the measurement of body-internal variables, e.g., the sensorimotor state $\mathbf{x}_a(t)$. This measurement is not always possible by applying an objective, sensor-based method, as technical sensors for specific information may not be available. An example is the perceptual system whose output to the outer world generally is a verbal response, providing only sparse information about the actual process happening in the brain. With respect to this behavior, a sensorimotor observation function is defined that transforms internal variables from the sensorimotor loop into measurable quantities that are available outside of the body.

Definition 3.5. A *sensorimotor observation*

$$\phi_a : S_a \rightarrow Y_a$$

of a sensorimotor domain $a \in \{phy \quad se \quad ne \quad con \quad perc \quad mo\}$ is defined as a differential equation

$$\mathbf{y}_a(t) = \phi_a \left(\mathbf{x}_a(t), \dot{\mathbf{x}}_a(t), \ldots, \theta_a(t), \mathbf{u}_a(t), \dot{\mathbf{u}}_a(t), \ldots, \xi_a(t), t \right), \tag{3.8}$$

where $\mathbf{x}_a \in X_a$, $\theta_a \in \Theta_a$, $\mathbf{u}_a, \in U_a$, $\xi_a \in \chi_a$. The state vector $\mathbf{x}_a(t)$ is influenced by the domain dynamics in (3.6). The output vector $\mathbf{y}_a \in Y_a$ contains all information that can be objectively obtained by measurements from domain a.

Remark 3.4. The standard formulation of state-space models, concentrating all dynamic processes into one differential equation and modeling the output function as a static mapping is limited to modeling causal system behavior. This means that the actual value of any variable depends only on the present and past values of other variables and itself. While this is a valid restriction for closed-loop control processes, captured by sensorimotor transformations, it is too restrictive for the definition of sensorimotor observations, which include the verbal response from the perceptual system. As an example, an anticausal behavior in auditory masking shall be mentioned here: It was found that a sound of low intensity can eventually not be perceived if it is followed by a loud masking stimulus [1, 124]. Thus, the perception of a current stimulus may depend on the future of the sensory input which is an anticausal behavior. Another, more practical reason for choosing a differential equation for $\phi_a(\cdot)$ is a potentially useful non-causal offline-processing for a measured $\mathbf{y}_a(t)$, such as zero-phase filtering.

Remark 3.5. The mathematical transformations ψ_a and ϕ_a, realizing the domain dynamics and sensorimotor observation, were introduced as mappings from S_a into S_a in Definition 3.1. The subsequent Definitions in 3.3 and 3.5 refine the mappings between domain and codomain to $X_a \rightarrow X_a$ for the domain dynamics, and $S_a \rightarrow Y_a$ for the sensorimotor observation, respectively. These refinements are not contradictory to the original statement in Definition 3.1 since $X_a, Y_a \subset S_a$.

Example 3.1. The sensorimotor set, dynamics, observation and transformation functions of the physical domain to the sensory domain shall be discussed for a deterministic haptic mass-spring-damper environment with inertia m, damping d, and stiffness k in admittance form. The state vector $\mathbf{x}_{phy}$ consists of position $x_h(t)$ and velocity $\dot{x}_h(t)$ where we will limit our consideration to a unidirectional movement, such that $\mathbf{x}_{phy}(t) = \begin{bmatrix} x_h(t) & \dot{x}_h(t) \end{bmatrix}^T \in \mathbb{R}^2$. From the definition of an admittance, the input to the environment is the human interaction force $u_{phy}(t) = f_h(t) \in \mathbb{R}$. Assuming a deterministic behavior, no noise and uncertainty terms are required, $\xi_{phy}(t) \in \{\}$. The feature vector θ_{phy} contains all properties that are needed to describe the sensorimotor domain. In this example it contains the admittance parameters $\theta_{phy} = \begin{bmatrix} m & d & k \end{bmatrix}^T \in \mathbb{R}^3$. The sensorimotor observation of the physical model gathers all quantities that are objectively measurable into an observation vector, assuming that we can measure every state and every environment feature $\mathbf{y}_{phy}(t) = \begin{bmatrix} x_h(t) & \dot{x}_h(t) & f_h(t) & m & d & k \end{bmatrix}^T \in \mathbb{R}^6$. Information from the physical domain is gathered by the human senses. The input vector to the sensory domain $\mathbf{u}_{se} \in U_{se}$

must match human capabilities. In our example, position $x_h(t)$, velocity $\dot{x}_h(t)$, and environment force $f_h(t)$ can be acquired, see Section 2.2. The sensorimotor set for the physical domain S_{phy} is thus $S_{phy} = \left\{ \left[\mathbf{x}_{phy}(t)^T \quad \theta_{phy}^T \quad u_{phy}(t) \quad \mathbf{y}_{phy}^T(t) \right]^T \in \mathbb{R}^{12} \right\}$. Based upon these considerations, the functions $\psi_{phy}(\cdot)$, $\nu_{phy,se}(\cdot)$, and $\phi_{phy}(\cdot)$ can be formulated as

$$\dot{\mathbf{x}}_{phy}(t) = \begin{bmatrix} \dot{x}_h(t) \\ \ddot{x}_h(t) \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{d}{m} \end{bmatrix} \mathbf{x}_{phy}(t) + \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} f_h(t)}_{\psi_{phy}(\cdot)}, \qquad \mathbf{x}_{phy}(0) = \begin{bmatrix} x_{h,0} \\ \dot{x}_{h,0} \end{bmatrix},$$

$$\mathbf{u}_{se} = \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \mathbf{x}_{phy}(t) + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} f_h(t)}_{\nu_{phy,se}(\cdot)}, \qquad \mathbf{y}_{phy}(t) = \underbrace{\begin{bmatrix} \mathbf{x}_h(t) \\ f_h(t) \\ \theta_{phy} \end{bmatrix}}_{\phi_{phy}(\cdot)},$$

where $x_{h,0}$ and $\dot{x}_{h,0}$ are constant initial values.

Remark 3.6. The definition of sensorimotor transformations as given in Definition 3.4 restricts the domain dynamics to a state-space formulation of time-continuous differential equations with time-varying parameters. Other modeling forms, such as hybrid models with switching dynamics, are principally valid candidates for sensorimotor transformations as well, but omitted here to reduce the number of model parameters and complexity, thus enhance readability.

3.2.1 Joint Sensorimotor Transformations and Observations

Modeling all components of the sensorimotor loop depicted in Figure 3.1, or the individual transformations and domain dynamics in the dynamic mapping framework depicted in Figure 3.2 is hardly realizable. This is due to the fact that the set of observations is generally only a subset of the full sensorimotor set in domain a: $Y_a \subset S_a$. In the case of an unknown model structure and/or the lack of observability, the full model may not be inferable from the input-output relationship. As a consequence, experimental techniques and methods emerged describing the input-output relation over multiple domains without explicitly considering all intermediate processes. To account for such simplified modeling forms, joint sensorimotor transformations are introduced.

Definition 3.6. A *joint sensorimotor transformation*

$$\nu_{a,b,c} : S_a \rightarrow U_c \tag{3.9}$$

describes the mapping of a sensorimotor set in the sensorimotor domain a into the set of input vectors to the sensorimotor c which is connected to domain a indirectly over a sensorimotor domain b.

It follows that

$$\begin{aligned} \mathbf{u_c} &= \nu_{b,c}\left(\mathbf{x}_b(t), \theta_b(t), \mathbf{u}_b(t), \xi_b(t), t\right) \\ &= \nu_{b,c}\left(\mathbf{x}_b(t), \theta_b(t), \nu_{a,b}(\mathbf{x}_a(t), \theta_a(t), \mathbf{u}_a(t), \xi_a(t), t), \xi_b(t), t\right) \\ &= \nu_{a,b,c}\left(\mathbf{x}_a(t), \mathbf{x}_b(t), \theta_a(t), \theta_b(t), \xi_a(t), \xi_b(t), t\right) \\ &= \nu_{a,b,c}\left(\mathbf{x}_{a,b}(t), \theta_{a,b}(t), \xi_{a,b}(t), t\right) \end{aligned} \tag{3.10}$$

with

$$\mathbf{x}_{a,b}(t) = \begin{bmatrix} \mathbf{x}_a(t)^T & \mathbf{x}_b(t)^T \end{bmatrix}^T \in X_a \times X_b, \qquad \theta_{a,b}(t) = \begin{bmatrix} \theta_a(t)^T & \theta_b(t)^T \end{bmatrix}^T \in \Theta_a \times \Theta_b,$$
$$\xi_{a,b}(t) = \begin{bmatrix} \xi_a(t)^T & \xi_b(t)^T \end{bmatrix}^T \in \chi_a \times \chi_b.$$

Equivalently,

$$\psi_{a,b} : X_a \times X_b \to X_a \times X_b \tag{3.11}$$

can be formulated as

$$\dot{\mathbf{x}}_{a,b}(t) = \psi_{a,b}\left(\mathbf{x}_{a,b}(t), \theta_{a,b}(t), \mathbf{u}_a(t), \xi_{a,b}(t), t\right), \quad \mathbf{x}_{a,b}(0) = \mathbf{x}_{a,b,0} \tag{3.12}$$

with $\psi_{a,b}(\cdot)$ containing dynamic processes which can come from sensorimotor domains a and b.

Outer psychophysical models are an example of founding on joint sensorimotor transformations. The relation between a physical stimulus and the verbal response about a subjective impression of it, e.g., if a force difference ($\Delta f_h \in S_{phy}$) can be perceived ($y_{perc} \in Y_{perc} = \{\text{“yes”} \quad \text{“no”}\}$) is modeled directly, without considering sensory and neural processes [57]. Quantitative relations between environment and a percept can be captured by the transformation $y_{perc} = \phi(\nu_{phy,se,ne,perc}(\cdot), \cdot)$. Another example is the development of task performance models, where joint sensorimotor transformations are involved in modeling closed-loop interactions with a physical environment. These exemplary but practically relevant models are detailed out in the following. Since the focus of this thesis is on the haptic modality, only perceptual and task performance models including haptic feedback are targeted.

Remark 3.7. For simplicity and readability reasons, the notation $\nu_{a,b,c}(\cdot)$ and $\nu_{a,c}(\cdot)$ is used interchangeable.

3.3 Haptic Perception Models

A perception model relates a specific quantity in the physical world, either a physical state variable $x^*_{phy}(t) \in \mathbf{x}_{phy}(t)$, or a specific feature $\theta^*_{phy}(t) \in \theta_{phy}$, to a response about its perceptual equivalent,

$$\begin{aligned} x^*_{phy}(t) \in X_{phy} &\mapsto \mathbf{y}_{perc}(t) \in Y_{perc}, \\ \theta^*_{phy}(t) \in \Theta_{phy} &\mapsto \mathbf{y}_{perc}(t) \in Y_{perc}. \end{aligned} \tag{3.13}$$

Using the formalism of sensorimotor dynamics, transformations and observations, this relation can be expressed as

$$\dot{\mathbf{x}}_{phy,se,ne,perc}(t) = \psi_{phy,se,ne,perc}\left(\mathbf{x}_{phy,se,ne,perc}(t), \theta_{phy,se,ne,perc}(t), \mathbf{u}_{phy}(t), \mathbf{u}_{ec}(t), \xi_{phy,se,ne,perc}(t), t\right), \tag{3.14}$$

$$\begin{aligned} \mathbf{x}_{phy,se,ne,perc}(0) &= \mathbf{x}_{phy,se,ne,perc,0}, \\ \mathbf{y}_{perc}(t) &= \phi_{perc}\left(\dot{\mathbf{y}}_{perc}(t), \ldots, \mathbf{x}_{perc}(t), \ldots, \theta_{perc}(t), \nu_{phy,perc}(\cdot), \mathbf{u}_{ec}(t), \ldots, \xi_{perc}(t), t\right) \end{aligned} \tag{3.15}$$

where

$$\nu_{phy,perc}(\cdot) = \nu_{phy,perc}\left(\mathbf{x}_{phy,se,ne}(t), \theta_{phy,se,ne}(t), \mathbf{u}_{phy}(t), \xi_{phy,se,ne}(t), t\right). \tag{3.16}$$

A specialty of the perceptual system is direct access to the efference copy $\mathbf{u}_{ec}$ from the cognitive motor control process, determining the activation of muscles, as discussed in Sections 3.1 and 3.2. Furthermore, the perceptual system may be aware of the human intention to achieve a certain motor task. Typically, the physical world containing haptic states and features is excited from the motor system. In this way, the input $\mathbf{u}_{phy}(t)$ to the physical sensorimotor domain is determined by the motor system, which itself receives neural signals driven from the cognitive controller mechanism in the brain. Thus, the sensorimotor loop is closed over the haptic environment and described by a closed-loop dynamic process instead,

$$\begin{aligned} \dot{\mathbf{x}}_{\circlearrowleft}(t) &= \psi_{\circlearrowleft}\left(\mathbf{x}_{\circlearrowleft}(t), \theta_{\circlearrowleft}(t), \mathbf{u}_{ne}(t), \xi_{\circlearrowleft}(t), t\right), \\ \mathbf{x}_{\circlearrowleft}(0) &= \mathbf{x}_{\circlearrowleft,0} \end{aligned} \tag{3.17}$$

where

$$\begin{aligned} &\psi_{\circlearrowleft}(\cdot) := \psi_{ne,mo,phy,se,ne,con}(\cdot), \\ &\mathbf{x}_{\circlearrowleft}(t) := \mathbf{x}_{ne,mo,phy,se,ne,con}(t), \quad \theta_{\circlearrowleft}(t) := \theta_{ne,mo,phy,se,ne,con}(t), \quad \xi_{\circlearrowleft}(t) := \xi_{ne,mo,phy,se,ne,con}(t) \end{aligned}$$

and $\mathbf{x}_{\circlearrowleft,0}$ is a constant initial value. Subsequently, equation (3.15) can be rewritten as

$$\mathbf{u}_{perc}(t) = \nu_{\circlearrowleft,perc}(\mathbf{x}_{\circlearrowleft}(t), \theta_{\circlearrowleft}(t), \mathbf{u}_{ne}(t), \xi_{\circlearrowleft}, t), \tag{3.18}$$

$$\dot{\mathbf{x}}_{perc}(t) = \psi_{perc}\left(\mathbf{x}_{perc}(t), \theta_{perc}(t), [\mathbf{u}_{perc}(t)^T,\ \mathbf{u}_{ec}(t)^T]^T, \xi_{perc}(t), t\right), \tag{3.19}$$

$$\mathbf{y}_{perc}(t) = \phi_{perc}\left(\dot{\mathbf{y}}_{perc}(t), \ldots, \mathbf{x}_{perc}(t), \ldots, \theta_{perc}(t), [\mathbf{u}_{perc}(t)^T, \ldots,\ \mathbf{u}_{ec}(t)^T,\ \ldots]^T, \xi_{perc}(t), t\right). \tag{3.20}$$

Remark 3.8. The closed-loop sensorimotor dynamics in equations (3.17) are defined as taking the efferent neural signal $\mathbf{u}_{ne}(t)$ as input to the global closed-loop system. This choice reflects the fact that the independent input to the control loop is actually defined by the cognitive control law, taking the intention for a certain motor action and the sensory feedback from the environment into account.

In the following, our focus will be on the dynamic perception mechanism (3.18)-(3.20). The dynamic perception mechanism (3.18)-(3.20) fulfills two main tasks: The first task is extracting information from the perceptual input signals $\mathbf{u}_{perc}(t)$ and $\mathbf{u}_{ec}(t)$ which is necessary for the percept of the physical world. The second task is formulating a perceptual response $\mathbf{y}_{perc}$. For the extraction of information, two types of problems can be distinguished: Perceiving a physical environment state variable $x^*_{phy}(t) \mapsto \mathbf{y}_{perc}(t)$ requires a reconstruction of $x^*_{phy}(t)$ based on the information available to the perceptual system. Reconsidering from equation (3.20) that $x^*_{phy}(t)$ is encoded into $\mathbf{u}_{ec}(t)$ and $\mathbf{u}_{perc}(t)$, the perception process can be interpreted in a system theoretic sense as a *state estimation problem*. One solution to state estimation problems is a state observer.

In the case of perceiving a physical environment feature $\theta^*_{phy}(t) \mapsto \mathbf{y}_{perc}(t)$, sensory and motor information must be combined together to infer the environmental property. In system theory, this is referred to as a *parameter identification problem*.

3.3.1 Relation to the Diffusion Model

The diffusion model was introduced in Section 2.4.3 as a powerful framework to describe perceptual processes dynamically. Indeed, it can be easily fit into the sensorimotor framework introduced here: The input $s(t)$ to the diffusion model denotes information available to the sensory system. The information being required for the perceptual judgment is extracted by the information encoding stage $G_{enc}(j\omega)$, or using terms from the dynamic perception framework, the perceptual dynamics. The following information accumulation stage $G_{acc}(j\omega)$ is easily included into the perceptual dynamics in the form of an integrator. The decision stage, leading to the perceptual response $\mathbf{y}_{perc}$ is one example of a perceptual observation function $\mathbf{y}_{perc} = \phi_{perc}(\cdot)$.

3.3.2 Relation to Quantitative Perceptual Laws

Static perceptual laws as discussed in Section 2.4.2 directly map features or states in the physical world to their perceptual equivalent. A simplification is made by modeling only the steady state of the perceptual system and neglecting all time-varying influence factors, such as changes in the interaction movement characteristics. In terms of the dynamic sensorimotor framework proposed in this chapter, the steady-state description corresponds to the case $\dot{\mathbf{x}}_{\circlearrowleft}(t) = \psi_{\circlearrowleft}(\cdot) = \mathbf{0}$, such that the mappings

$$x^*_{phy} \in X_{phy} \mapsto \mathbf{y}_{perc} \in Y_{perc},$$
$$\theta^*_{phy} \in \Theta_{phy} \mapsto \mathbf{y}_{perc} \in Y_{perc}.$$

are static.

3.4 Task Performance Models

Task performance describes human's ability to accomplish a certain goal under given circumstances. Classically, it is measured by task performance indexes, such as the task completion time, or the accuracy in achieving a certain goal (see Section 2.6.4). Such static task performance measures are unsuited for an application in algorithms where a prediction about the operator's performance during the execution of a task is required. This could be necessary to proactively assist the operator and increase his/her task performance. An example for an online-adaptation of communication parameters requiring a prediction of the operator's task performance during task execution is discussed in Chapter 6. Predicting the task performance prior to task completion generally requires knowledge about the underlying factors leading to a specific behavior, including

- the given task, represented by a task parameter vector θ_T.
- the dynamic response characteristics of the sensorimotor loop, $\psi_{\circlearrowleft}(\cdot)$, closed over the physical environment.
- a performance state $\mathbf{x}_{perf}$ for capturing cumulative performance measures.

As an example for θ_T, one could think of an obstacle course's length in the context of a navigation task, or the size of the target area in a pick-and-place exercise. Exemplary for the influence of the dynamic response characteristics of $\psi_{\circlearrowleft}(\cdot)$ stands an object's inertia which is to be carried from one place to another and decreases the responsiveness of the sensorimotor loop – a higher inertia requires more time to accelerate and decelerate, thus decreasing performance in the sense of task completion time. An example for the need of a performance state $\mathbf{x}_{perf}$ is the accumulated number of failures in an assembly task that can not be accounted for in a static way.

In the mathematical modeling framework, introduced in Section 3.2, task performance can be considered similarly to a sensorimotor domain, abbreviated *perf*. Equivalently to a dynamic perceptual model (3.19)-(3.20), dynamic task performance can be modeled as

$$\begin{aligned}
\mathbf{u}_{perf}(t) &= \nu_{\circlearrowleft,perf}(\mathbf{x}_{\circlearrowleft}, \theta_{\circlearrowleft}, \mathbf{u}_{ne}, \xi_{\circlearrowleft}, t) \\
\dot{\mathbf{x}}_{perf}(t) &= \psi_{perf}\left(\mathbf{x}_{perf}(t), \theta_{perf}(t), \mathbf{u}_{perf}(t), t\right), \qquad \mathbf{x}_{perc}(0) = \mathbf{x}_{perc,0}, \\
\mathbf{y}_{perf}(t) &= \phi_{perf}\left(\dot{\mathbf{y}}_{perf}(t), \ldots, \mathbf{x}_{perf}(t), \ldots, \theta_{perf}(t), \mathbf{u}_{perf}(t), \ldots, t\right).
\end{aligned}$$

In a similar manner as for other sensorimotor domains, $\psi_{perf}(\cdot)$ captures the performance dynamics, while $\phi_{perf}(\cdot)$ is the *performance observation*, resulting in a continuous task performance observation $\mathbf{y}_{perf}(t)$. Because the task performance model is defined by the system designer, noise terms are not considered, as those constitute no additional benefit.

3.4.1 Relation to Static Performance Measures

Most classical performance measures capture the task performance after the completion of a task. Thus, these measures correspond to the value $\mathbf{y}_{perf}(t_{com})$ where t_{com} is the task completion time.

3.5 Discussion

The biggest progress beyond the state of the art in the mathematical framework for sensorimotor processes presented in this chapter is the consideration of dynamic processes in the description of sensorimotor processes. As a consequence, haptic perception and task performance can be modeled with a greater reliability and with a higher accuracy in a larger variety of experimental conditions. The framework is founded on a set of coupled differential equations, modeling behavior in the physical, sensory, neural, cognitive, and motor sensorimotor domain. The measurable quantities of each sensorimotor domain are contained in sensorimotor observations, formulated as filter functions with respect to all available information in this domain.

The fundamental advantages of this dynamic framework compared to static perceptual laws and performance measures are:

1. Physiological factors, e.g., the availability of sensors for capturing physical properties, can be taken into account for explaining perceptual characteristics and task performance.

2. Models on different abstraction levels can be developed using the same formalism, e.g., physiological muscle models, inner psychophysics models, and input-output models for outer psychophysics.

3. Not only perceptual responses and task performance indexes can be modeled, but also hidden measures which are not measurable otherwise, such as the perceptual state $\mathbf{x}_{perc}(t)$ and the task performance state $\mathbf{x}_{perf}(t)$. Based on these measures, the perception of a specific scene and the performance in a specific task can be potentially predicted even before completion.

These properties are useful when systematically developing quantitative perceptual and task performance models which are applicable not only in a static, human-centered evaluation and optimization of technical systems, but even in algorithms capable of online-optimizing these goals.

3.6 Conclusions and Open Problems

We presented a mathematical framework for the dynamic description of sensorimotor processes, specifically perception and task performance models. As the fundamental structure of the framework is based on stochastic differential equations in the form of state space models, the models developed on this foundation can be easily applied to a model-based experimental design and a utilization in human-centered control algorithms, interface design, and evaluation of telepresence systems.

An open issue is a simple and intuitive representation of such models in a form that can be accepted not only by engineers familiar with differential equations but also for psychologists, experts in quantifying human perceptual phenomena.

4 Dynamic Combination of Movement and Force

Summary. *Perception of a haptic feature such as stiffness, damping, or inertia requires the combination of movement and force information. Temporal limits of this combination process are investigated, and modeled quantitatively in the form of a dynamic process. Comparing the predictions of multiple model candidates with respect to their ability to capture the perception thresholds of time delay between movement and force, we find that*

- *movement characteristics significantly influence the perception threshold,*
- *dynamic models capture the perception threshold better than static models,*
- *differences between two environments may be discriminated by a comparison of feedback force or limb movement rather than the actual environment feature.*

Haptic telepresence systems are designed to convey a specific environmental impression to the human operating the system. Usually, this impression is not directly measurable by mechanoreceptors, e.g., limb movement and force feedback, but is encoded into an environmental property such as a mechanical impedance (see Sec. 2.2). To perceive such a feature, information from the interaction movement must be related to a specific force feedback. In telepresence systems, haptic information from the remote side are often affected by the technical system, e.g., due to communication unreliabilities such as time delay in the communication channel of a telepresence system, where the temporal relation between movement and force is distorted. For the design of high-fidelity telepresence systems that aim for perfect haptic transparency, it is crucial to know the precision which is required for the presentation of interaction movement and force feedback in order to not affect haptic perception.

In this chapter, three different mechanisms are introduced that are suited for combining movement and force into a unified percept. These hypothetical mechanisms are both motivated from a system theoretic point of view, and from findings in the sensorimotor control literature and perceptual models from other domains than haptics. Conceptually new compared to existing models in the literature for the perception of haptic environments, we include dynamic models of body and environment to describe the combination process of limb movement and force. A series of three experiments measuring the detection thresholds of time delay in the haptic feedback is presented, where the participants' sinusoidal interaction movement, the magnitude of force feedback, and the type of haptic environment is varied systematically. A parameter comparison mechanism, a matched filter model, and a state observer model of body and environment are evaluated with respect to their capabil-

ity to reflect the observed perceptual behavior. From these candidates, the observer model shows superior prediction properties in comparison to its alternatives.

Theoretical models for the combination of interaction movement and feedback force are discussed in Section 4.1. Dynamic model candidates predicting perception thresholds for temporal misalignment between limb movement and force feedback are introduced in Section 4.2. The model-guided experimental design of three psychophysical studies for haptic time delay perception, explicitly considering the different model predictions in the choice of experimental conditions, is described in Section 4.3. Experimental methods and results are presented in Section 4.4. A general discussion of the summarized results from all experiments with respect to their applicability to haptic telepresence systems in Section 4.5 finishes the chapter.

4.1 Background

Investigations on perceptual limits of haptic environments, determining the relation between an interaction movement and the force feedback, have a long tradition in psychophysical research. Indeed, many perceptual laws discussed in Section 2.4.2 describe humans' abilities to detect variations in haptic environment features, such as stiffness, damping, or inertia. As an example, Weber's Law originally describes the just noticeable difference of weight and its dependency on the weight magnitude [62]. The quantization of perceptual thresholds in terms of discriminable environmental features is appealing from an engineering point of view due to its practicability. As an example from a telepresence application, the difference between the haptic environment impedance $Z_{env}(j\omega)$ on the remote side and the impedance $Z_h(j\omega)$ which is actually displayed to the human operator can be determined analytically [18, 35]. Using perceptual thresholds of these specific environment features, it can be assessed if the difference between the two stimuli can be perceived, thus if the telepresence system is judged as transparent or not. A fundamental problem of perception models for environment impedances is that the human does not possess receptors for impedances. Instead, haptic properties such as stiffness, inertia, or damping are inferred indirectly by extracting temporal and magnitude information from the sensed interaction movement and force feedback. Most current perception laws are static with respect to the exploration actions. It is though known, that limb movement can significantly influence the perception of a haptic environment, e.g., the perception of hardness. Instead of an object's stiffness, a movement-dependent measure, the "rate-hardness" determines how hard an object is felt [125]. Another example for the dependencies between movement and haptic environment perception is estimation of a moved mass that was found to depend on factors such as movement amplitude and frequency [126].

One solution to overcome this modeling inaccuracy is to take the time series data of movement and force directly into account to explain the perceptual performance of environment features. For spring environments, such models have been proposed earlier: Stiffness difference perception can be modeled by a simple comparison mechanism of the maximum forces that are generated in an instructed interaction movement up to a specified position, exploring a virtual spring [16, 127]. A direct comparison of time-series sensory signals, sensed in two successive comparison trials is though problematic. Slight changes in the interaction

movement, result in a in a difference of the force signal, which can lead to potentially wrong conclusions. Instead, an internal representation of the haptic characteristics could be used to predict the force of the memorized environment on-line and, perception of differences can be detected based on the simulated time series.

Reconsider from Section 3.3 that the problem of perceiving an environment feature is a parameter identification problem from a system theoretic point of view. Against this background, a decision whether two environments are perceived as the being the same or different can be made in at least two ways: Either, two independent parameter identification processes can be performed, and the results are compared to each other. Perceptual limits can so be explained with uncertainties in the identification result, e.g., due to a limited exposure time, and internal noise processes. Alternatively, a model verification procedure can be used to judge whether two environments are the same or different. This approach founds on the assessment of how good new sensory data is explained with an internal, previously identified model. The model can be learned, e.g., by previous encounters with a haptic characteristic of a certain type, or by directly comparing two environments successively.

In the following, we want to introduce three alternative hypothetical mechanisms, principally capable of capturing the limits in the combination process of movement and force into a percept of an environmental feature. In contrast to existing perception models, dynamic models are utilized to capture perceptual phenomena, allowing the explicitly consideration of varying movement conditions. As a result, extensions to the classical, static view on perception thresholds to dynamic threshold models are made. The models found on the sensorimotor framework introduced in Chapter 3 and are used for a model-based experimental design of a psychophysical study.

4.2 Perception Model Representations

Models for the combination process of movement and force into an estimate of an environment feature can be embedded into the mathematical formalism of perception processes proposed in Chapter 3 as

$$\mathbf{y}_{perc}(t) = \phi_{perc}\left(\mathbf{x}_{perc}(t), \ldots, \theta_{perc}(t), \mathbf{u}_{perc}(t), \ldots, \mathbf{u}_{ec}(t), \ldots, \xi_{perc}(t), t\right), \tag{4.1}$$

see equation (3.15). Here, $\mathbf{y}_{perc}(t) \in Y_{perc}$ is the perceptual response, taking the perceptual state variable $\mathbf{x}_{perc}(t) \in X_{perc}$, the input information available from the sensory feedback (and transmitted over neural pathways) $\mathbf{u}_{perc}(t) \in U_{perc}$, and the efference copy $\mathbf{u}_{ec}(t) \in U_{ec}$ into account. The parameter vector $\theta_{perc}(t) \in \Theta_{perc}$ influences the characteristics of the (generally time-varying) perceptual model. The nature of the perceptual response $\mathbf{y}_{perc}(t)$ is diverse and depends on the specific perceptual task: In the case of a free response about the perception of a specific environment, y_{perc} could be "perceiving feature x with intensity y". Technical applications though often require only human's difference discrimination abilities. These abilities could be captured by a perception model that models the response if two stimuli can be discriminated or not as

$$y_{perc}(t_{resp}) = \begin{cases} \text{“different”} & \text{if } \exists 0 \leq t \leq t_{resp} : |\delta(\cdot)| > \epsilon, \\ \text{“same”} & \text{if } \forall 0 \leq t \leq t_{resp} : |\delta(\cdot)| < \epsilon, \end{cases} \tag{4.2}$$

where

$$\delta(\cdot) = \delta\left(\mathbf{x}_{perc}(t), \ldots, \theta_{perc}(t), \mathbf{u}_{perc}(t), \ldots, \mathbf{u}_{ec}(t), \ldots, \xi_{perc}(t), t\right)$$

and t_{resp} is the time instance where the response is given. The function $\delta(\cdot)$ is denoted as *decision criterion*, and ϵ as *decision threshold*, determining the maximum allowable value for $\delta(\cdot)$ to be not perceivable.

Remark 4.1. This model assumes the existence of a specific, constant threshold value ϵ which determines the perceptual response based on noisy perceptual input information. While deterministic threshold models are inappropriate for the modeling of perceptual phenomena, the existence of a decision threshold in the context of a stochastic decision variable $\delta(\cdot)$ is common in psychophysical modeling, e.g., in signal detection theory [128], or diffusion models [68, 69].

In the following, three candidates for the decision criterion $\delta(\cdot)$ in (4.2) are proposed and embedded into a reduced and simplified version of the sensorimotor loop sketched in Figure 4.1a. Body admittance $Y_{body}(j\omega)$ and environment impedance $Z_{env}(j\omega)$ are linearly approximated and only a single limb is moved, actuated by a single force $f_{m,res}(t) \in \mathbb{R}$ from the muscles acting on that limb in one direction, see also Chapter 3. The tissue dynamics, filtering the environment force acting on the moved limb are neglected, assuming that the tissue is in a compressed state for the time of interaction, thus approximately rigid. In the following, we will discuss perception models which take a single human limb's state

$$\mathbf{x}_h(t) = \begin{bmatrix} x_h(t) & \dot{x}_h(t) \end{bmatrix}^T \in \mathbb{R}^2, \tag{4.3}$$

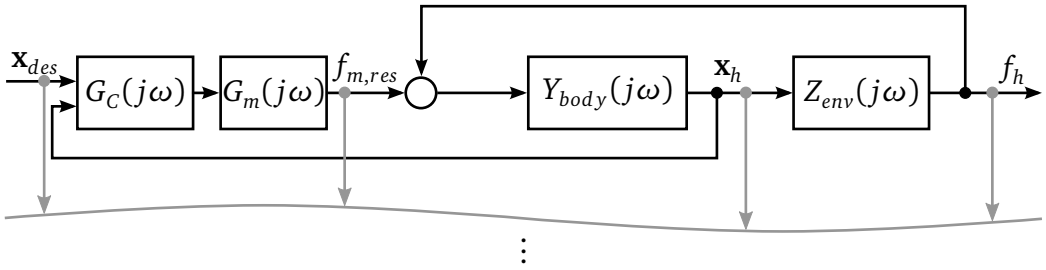
moving in a haptic environment as an input, provoking a environmental force $f_h(t) \in \mathbb{R}$ reacting back on the motor force $f_{m,res}(t)$. While the interpretation of force and state in terms of input and output variables is interchangeable, we will limit our considerations to the aforementioned case. We do not make assumptions about the structure of the control mechanism responsible for the specific movement but assume that desired trajectories can be executed with sufficient accuracy. The noise terms, transmission time delay, and sensor transfer functions mentioned in Figure 3.1 are not depicted and assumed to be compensated for in the control algorithm, but their influence will be discussed in the subsequent sections.

Different hypothetical perception processes are presented in the following. Evidence from psychophysical and sensorimotor literature for the validity of the model candidates are provided.

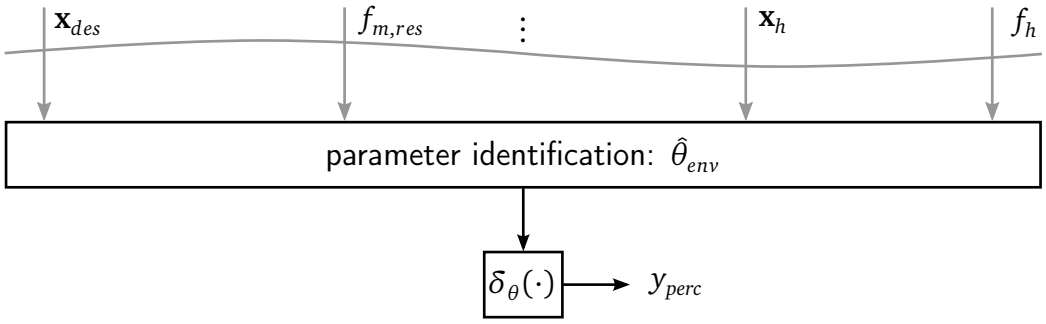
4.2.1 Feature Comparison

Current perception models capturing the discrimination abilities of a human perceiving a haptic environment assume the existence of a threshold in terms of environmental properties [11, 12, 62, 63]. A decision criterion based on differences in the environment features is proposed as a representative of this classical view. For comparing two haptic environments with physical features θ_1 and θ_2, respectively, perceptual estimates $\hat{\theta}_1$, $\hat{\theta}_2$ must be obtained by haptic exploration. Based upon, a perception mechanism following equation (4.2) can be formulated as

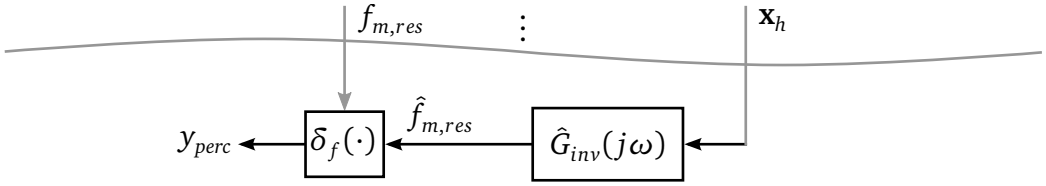
$$\delta(\cdot) = \delta_\theta\left(\hat{\theta}_1, \hat{\theta}_2\right), \quad \epsilon = \epsilon_\theta. \tag{4.4}$$



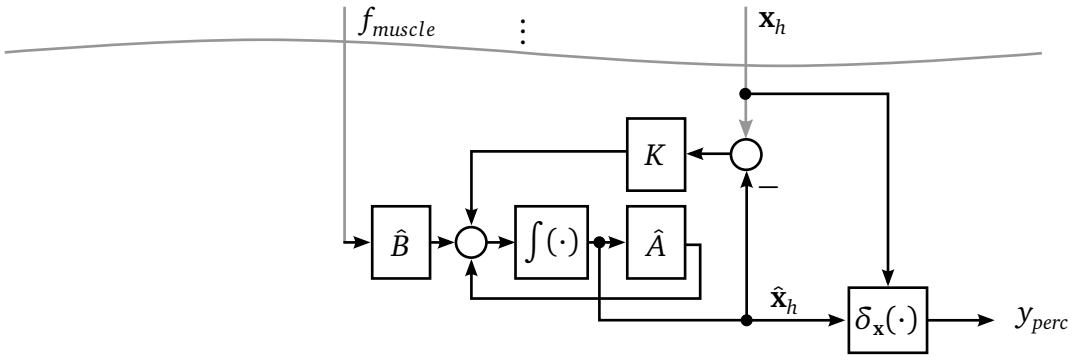
(a) A subset of the sensorimotor control loop, abstracting a position-controlled limb.



(b) Difference perception model by comparing identification results.



(c) Difference detection using model validation, based on a matched filter.



(d) Difference detection using model validation, based on a state observer.

Figure 4.1: Three alternative discrimination mechanisms to detect differences between two haptic environment features.

From a control theoretic point of view, the perceptual process of determining the perceptual estimates $\hat{\theta}_{phy,1,2}(t)$ requires a parameter identification method which is to be encoded into the perceptual apparatus.

One specific black-box system identification technique, based on the estimation of the input-output covariance matrix and the autocovariance matrix of the input has been discussed in Section 2.5.2. In studies on monkeys' brains, correlation methods as a normalized form of covariance methods have been found to explain brain activity in selected regions well if the animal attends to a certain visual stimulus [129]. This could be taken as evidence for the existence of a neural substrate for performing correlations efficiently in certain perceptual brain regions. Correlation mechanisms can furthermore explain humans' performance in detecting temporal differences in audio-visual signals [130]. A perceptual model on the basis of a parameter identification technique is depicted in Figure 4.1b.

4.2.2 Model Verification Using Matched Filtering

An alternative approach to judge whether two haptic environments have the same or different properties is the usage of a model verification technique. In practical system identification procedures, verification methods are a standard procedure to check whether an identified has good generalization capabilities. A model of the first haptic environment is identified on the informational basis of the first exploration. During the exploration of the comparison environment, sensory information is gathered and compared to a prediction of the sensory output, using an internal model of the first environment's property. If prediction and sensory evidence matches, the environments are the same. If there is a mismatch between prediction and feedback, the two environments are classified as different. Diverse verification methods are utilized in various technical applications, differing in the criterion which is taken into consideration for classification [131].

Here, a verification problem utilizing a comparison of the input information is proposed, namely

$$\delta(\cdot) = \delta_f \left(f_{m,res}(t), \hat{f}_{m,res}(t) \right), \quad \epsilon = \epsilon_f, \tag{4.5}$$

where $\hat{f}_{m,res}(t)$ is a prediction of the force to move in the previously identified haptic environment, and $f_{m,res}(t)$ is the measured force. The optimal solution for the reconstruction with respect to the signal-to-noise ration is the usage of a *matched filter* model. The optimal filter that reconstructs the motor action from measuring the state $\mathbf{x}_h(t)$ is a dynamic model containing the inverse body dynamics and the environment impedance, denoted by its transfer function $G_{inv}(j\omega)$.

There is experimental evidence for the usage of inverse dynamic models in sensorimotor control by predicting the motor actions from the sensed state of the body, see Section 2.6.3. Similarly, the model $\hat{G}_{inv}(j\omega)$ capturing body dynamics, sensor- and environmental dynamics potentially plays a role in perception as well, see Figure 4.1c. The aforementioned stiffness estimation model on the basis of comparing the maximum forces can be seen as a representative of a perception model using inverse dynamics [16, 127].

4.2.3 Model Verification Using State Observers

Instead of the exerted force, the body state $\mathbf{x}_h(t)$ can be taken into consideration for model veification, equivalent to a perceptual model candidate

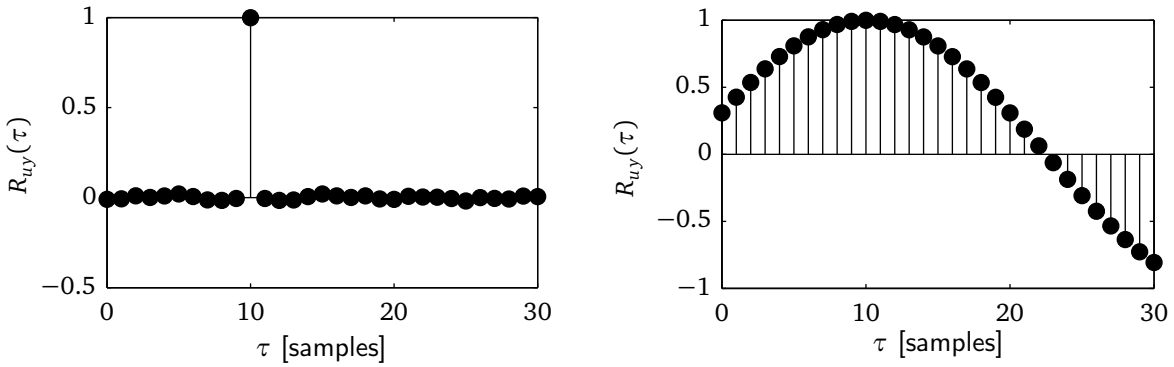
$$\delta(\cdot) = \delta_{\mathbf{x}}\left(\hat{\mathbf{x}}_h(t), \mathbf{x}_h(t)\right), \quad \epsilon = \epsilon_{\mathbf{x}}. \tag{4.6}$$

In contrast to matched filters, predicting the motor action from a sensory estimate of the body state, state-space observers use a forward model of body and environment dynamics to predict the body state based on the current motor action. The Luenberger observer, a representative of linear state observers, is depicted in Figure 4.1d. The state matrix $\hat{A} \in \mathbb{R}^{2\times 2}$ contains all dynamic processes of body and environment, the input matrix $\hat{B} \in \mathbb{R}^2$ defines the impact of the input to each state. The error between the feedforward prediction of limb state $\hat{\mathbf{x}}_h(t)$ and the measured limb state $\mathbf{x}_h(t)$ is taken into consideration to correct the state estimate by multiplication with the gain matrix $K \in \mathbb{R}^{2\times 2}$. The influence of the gain matrix K should be briefly discussed in the following: First, the case of $K = 0$ is considered: If $\hat{A}$ and $\hat{B}$ capture the body and environment behavior exactly and the initial state estimate $\hat{\mathbf{x}}_h(0)$ is correct, the state estimate over time $\hat{\mathbf{x}}_h(t)$ equals the real state $\mathbf{x}_h(t)$. If the model is imperfect, the estimated state deviates from the real state. This error is attenuated by choosing an appropriate $K \neq 0$, however, it does not vanish. In technical applications, K is used to assign arbitrary dynamics to the observer's behavior, depending on the application's demands, see Section 2.5.1. In the case of white noise affecting the output measurement and states, the noise-optimal choice for K is the Kalman Gain. This choice turns the Luenberger observer into a stationary Kalman filter. Kalman filters were indeed found to describe sensorimotor control processes well, see Section 2.6.3. This is a motivation to consider such a structure as a candidate for perceptual processes as well.

4.3 Model-Guided Experimental Design

Two haptic environments which are discriminable from each other can differ in various ways: On the one hand, differences in the magnitude of the force feedback, given a specific exploration movement can be noticed. Such differences can be due to a variation in the spring constant, damping, or inertia. On the other hand, temporal distortions such as time delay between movement and force feedback is capable of changing the impression of the environment. Although time delay in haptic feedback is not a natural phenomenon in everyday-life haptic interactions, it is a problem in the operation of telepresence systems over large distances [42, 132], space [37], or under water [46]. We will focus on the investigation of distortions in the haptic combination process due to temporal faults for two reasons: First, perception of time delay in haptic interaction with an environment is poorly researched, but such knowledge is helpful to provide guidelines and specifications of haptic telepresence systems. Second, temporal distortions are better suited to dissociate between the three model candidates proposed in Sections 4.2.1 - 4.2.3, as will be detailed out later.

Three hypothetical perception mechanisms that could lead to detecting time delay in the haptic feedback have been proposed by the author in earlier work [133, 134]: An *internal clock* and a perceptual mechanism sensitive to the *force discrepancy* between the real sensory



(a) Crosscovariance with white noise signals.

(b) Crosscovariance with sinusoidal signals.

Figure 4.2: The crosscovariance function R_{uy} of two discrete-time signals with $y(k) = u(k - 10)$, depicted for the cases of white noise (a) and sinusoidal signals (b). Both have a peak at the lag corresponding to the time delay.

force feedback and a predicted, non-delayed force reference have been considered. Furthermore, a perception mechanism of the *own exploration* action has been proposed. These model candidates for time delay detection are related to the hypothesized combination models for movement and force, derived in Section 4.2 and their predictions are compared in the following.

Remark 4.2. For simplicity reasons, and as commonly done in the literature [49, 135, 136], the different approaches are discussed using a reduced dynamic model of the human motor apparatus, with a single joint and one muscle pair attached, approximated by a mass-damper system with human-like parameters ($m_h = 2$kg, $d_h = 2$Ns/m), taken from [49].

4.3.1 Feature Comparison and the Internal Clock

A discrimination mechanism for time delay can be based on an identification procedure. We will limit our considerations to methods founding on the covariance between input and output data. The crosscovariance function between a white noise input signal $u(t)$ and an output signal $y(t)$ which is the time-delayed version of $u(t)$ has a peak at the location of the time delay, see Figure 4.2. In the case of autocorrelated input signals, or non-equally-shaped input and output signals, e.g. the input and output of a dynamic system, the crosscovariance function is non-zero for lags others than the time delay as well, the example of sinusoidal signals is illustrated in Figure 4.2b. Crosscovariance estimates for the time delay $\hat{T}_{d,\{1,2\}} \in \hat{\theta}_{phy,\{1,2\}}$ are a standard tool for identifying the amount of time-delay in system theory [137].

For the perception of haptic time delay, it was hypothesized in [134] that the time between a limb movement and the resulting time-delayed environment force may be perceived using an internal clock mechanism. The origins of this hypothesis lie in temporal perception theory [138], according to which time delay, as a special type of duration, is estimated from an internal clock. Assuming that a simple manual movement does not influence the internal clock, the detection threshold $T^{\theta}_{d,thresh}$ of time delay in haptic feedback is independent of the

operator's manual movement and force feedback, that is, it is constant.

$$T_{d,thresh} = \text{const.}$$

In this way, the predictions from the clock mechanism hypothesized in [134] coincident with a perceptual threshold model founding on identified estimates of the time delay $\hat{T}_{d,\{1,2\}} \in \hat{\theta}_{phy,\{1,2\}}$

$$y_{perc}(t_{resp}) = \begin{cases} \text{“different”} & \text{if } |\hat{T}_{d,1} - \hat{T}_{d,2}| > T^{\theta}_{d,thresh} \\ \text{“same”} & \text{if } |\hat{T}_{d,1} - \hat{T}_{d,2}| < T^{\theta}_{d,thresh} \end{cases}. \tag{4.7}$$

4.3.2 Matched Filter Model and Force Discrepancy

Instead of a direct comparison between identified estimates, differences in the environmental time delay can be perceived through prediction errors, e.g., by using a matched filter. Matched filter predictions found on a sensory measurement of the current body and environment state $\mathbf{x}_h(t)$ and estimate the motor command that is needed to move in a haptic environment. In the frequency domain, this mechanism is given by

$$\hat{F}_{m,res}(j\omega) = \hat{G}_{inv}\mathbf{X}_h(j\omega),$$

where $\hat{F}_{m,res}(j\omega)$ and $\mathbf{X}_h(j\omega)$ are the Fourier-transforms of $\hat{f}_{m,res}(t)$ and $\mathbf{x}_h(t)$. Considering the sensorimotor system abstracted in Figure 4.1a, $\hat{G}_{inv}(j\omega)$ can be easily calculated to

$$\hat{G}_{inv}(j\omega) = \hat{Z}_{env}(j\omega) + \frac{1}{Y_{body}(j\omega)}, \tag{4.8}$$

where $\hat{Z}_{env}(j\omega)$ is a previously identified environment impedance which is to be verified with the actual sensory and motor data. We assume the body admittance $Y_{body}(t)$ to be sufficiently well-known from everyday-experience, thus no estimation of this dynamic system is needed in every sensory experience of the environment. If the filter model is not exact, e.g., due to the fact that $Z_{env}(j\omega)$ is afflicted with time delay but the internal model representation $\hat{Z}_{env}(j\omega)$ is not, the actual resulting muscle force $f_{m,res}(t)$ that is required to move the limb, differs from the prediction. The required force $f_{m,res}(t)$ to move in a haptic environment with time delay, the filter prediction $\hat{f}_{m,res}(t)$ without considering the time delay, and the error between both, is exemplary depicted in Figure 4.3.

Perceiving the time delay T_d in haptic feedback over the difference in environmental force feedback and predicted, non-delayed force feedback is also the second hypothesis proposed in [134], named as “force discrepancy model”. It is inspired from experiments reporting that stiff environments are perceived softer in the case of time-delayed force feedback [127, 139, 140]. The mechanism proposed in [127] capturing the observed stiffness estimation behavior best, relies on the force at the maximum position, which is smaller with increasing time delay. The model in the form of equation (4.2) can be formulated as

$$y_{perc}(t_{resp}) = \begin{cases} \text{“different”} & \text{if } \exists 0 \le t \le t_{resp} : \ |\hat{f}_{m,res}(t) - f_{m,res}(t)| > \Delta f_{thresh} \\ \text{“same”} & \text{if } \forall 0 \le t \le t_{resp} : \ |\hat{f}_{m,res}(t) - f_{m,res}(t)| < \Delta f_{thresh} \end{cases}. \tag{4.9}$$

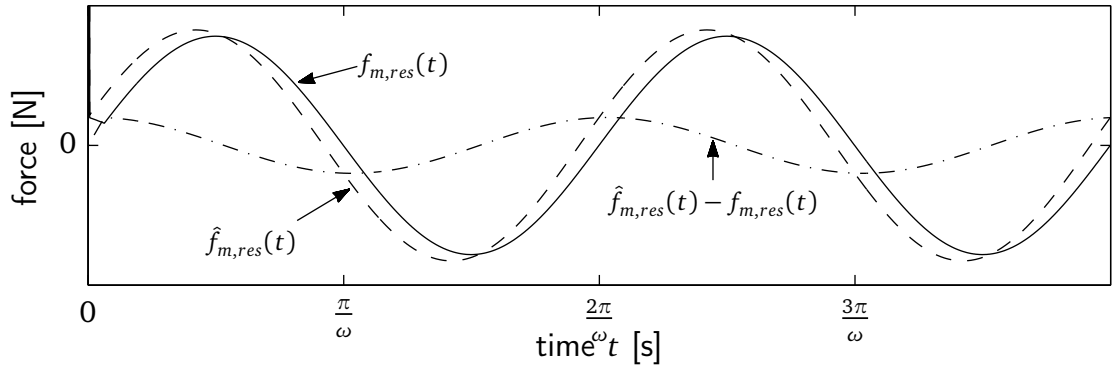


Figure 4.3: The force which is required to move in a haptic environment with time delay $f_{m,res}(t)$ (solid) differs from the filter prediction $\hat{f}_{m,res}(t)$ (dashed) that does not consider time delays in the filter dynamics. The resulting force error (dash-dotted) is the decision criterion in the second perception model.

In addition, we propose a prediction model based on Weber's Law, as force difference detection was found to follow these characteristics [7]. In this case, time delay can be perceived if the fraction of force error and force magnitude exceeds the Weber fraction W:

$$y_{perc}(t_{resp}) = \begin{cases} \text{“different”} & \text{if } \exists 0 \leq t \leq t_{resp} : \frac{|\hat{f}_{m,res}(t) - f_{m,res}(t)|}{f_{m,res}(t)} > W \\ \text{“same”} & \text{if } \forall 0 \leq t \leq t_{resp} : \frac{|\hat{f}_{m,res}(t) - f_{m,res}(t)|}{f_{m,res}(t)} < W \end{cases}. \tag{4.10}$$

Remark 4.3. The perceptual model candidates formulated in equations 4.9 and 4.10 do not make the time delay detection threshold explicit. The dependency is instead encoded into the difference in force feedback: While the actual force comes from the exploration of a time-delayed environment, the “reference force” $\hat{f}_{m,res}(t)$ contains the non-delayed environment dynamics. In the following, the amount of time delay resulting in a difference in force is referred to as $T^{f}_{d,thresh}$. The equivalent for the Weber fraction is denoted as $T^{W}_{d,thresh}$.

4.3.3 State Observer and Active Exploration Model

Instead of differences in the exerted force, the body's state could affect the discrimination of time delay as well, resulting in a perception model candidate

$$y_{perc}(t_{resp}) = \begin{cases} \text{“different”} & \text{if } \exists 0 \leq t \leq t_{resp} : |\hat{\mathbf{x}}_h(t) - \mathbf{x}_h(t)| > \Delta\mathbf{x}_{thresh} \\ \text{“same”} & \text{if } \forall 0 \leq t \leq t_{resp} : |\hat{\mathbf{x}}_h(t) - \mathbf{x}_h(t)| < \Delta\mathbf{x}_{thresh} \end{cases}. \tag{4.11}$$

In this model, the state estimate $\hat{\mathbf{x}}_h(t)$ come from a dynamic state observer. Compared to the inverse model simulation, the state observer has one more degree of freedom – the feedback matrix K. In our case, where the state vector consisting of the limb position and velocity is completely measurable since the human has receptors for those quantities, $K \in \mathbb{R}^{2\times 2}$ has four entries k_{ij}, with $i, j = 1, 2$. The observer behavior is described by

$$\dot{\hat{\mathbf{x}}}_h(t) = \hat{A}\hat{\mathbf{x}}_h(t) + \hat{B}f_{m,res} + K(\mathbf{x}_h(t) - \hat{\mathbf{x}}_h(t)), \tag{4.12}$$

thus the magnitude of the observation error depends on the feedback matrix K and the exploration pattern $\mathbf{x}_h(t)$.

The application of forward models was implicitly considered in [134] by the hypothetical time delay perception mechanism in the form of an "active exploration model". It states that the detection of time delay in a haptic system depends on the characteristics of the exploration movement $x_h(t)$ as one component of the state vector $\mathbf{x}_h(t)$. The motivation of this candidate is drawn from results reported for experiments in visual virtual environments using a head-mounted display. It has been demonstrated that the detection of asynchrony between head motion and visual feedback is dependent on the turning speed of the head [141]. Higher speeds result in lower detection thresholds for visual delays. The user of such systems may infer latency in the head tracking system based on the spatial discrepancy between the delayed and non-delayed object positions [142] that is, the displacement of an object from its originally expected position introduced by the latency between head motion and visual feedback. In impedance-type haptic environments, a similar mechanism based on deviations of the limb and environment state $\mathbf{x}_h(t)$ from the predicted state that is obtained with non-delayed environment dynamics $\hat{\mathbf{x}}_h(t)$ may support the detection of delayed feedback.

Remark 4.4. Equivalent to the force-dependent time delay discrimination thresholds discussed in Remark 4.3, the state-dependent time delay discrimination threshold is denoted as $T^{\mathbf{x}}_{d,thresh}$.

4.3.4 Model-Guided Stimulus Selection

Testing the capabilities of the model structures described in Sections 4.3.1-4.3.3 to predict the perceptual limits of the process combining movement and force together requires psychophysical experiments measuring these limits under different conditions. To be able to favor one hypothetical perception mechanism over another, the differences between the predicted perception limits should be maximized by an optimal selection of experimental conditions. Many haptic environments are based on the simple haptic "building block" stiffness, damping and inertia with their specific parameters k, d, and m. The individual components are equivalently described by their mechanical impedance and by a differential equation model,

$$Z_{env}(j\omega) = \frac{F_h(j\omega)}{\dot{X}_h(j\omega)} = \left\{ \frac{k}{j\omega},\ d,\ mj\omega \right\} \quad \text{and} \quad f_h(t) \ = \left\{ kx_h(t),\ d\dot{x}_h(t),\ m\ddot{x}_h(t) \right\}. \tag{4.13}$$

In the following, we will discuss predictions for the time delay detection detection threshold in force feedback from the environment on the example of a spring environment with spring constant k. Without loss of generality, the equilibrium point of the spring is set to the position $x_h = 0$.

In contrast to the perception model candidate founding on parameter identification, the predicted perception limits of time delay on the basis of the matched filter model and the state observer model depend on the interaction with the haptic environment. To compare latter hypothetical models and determine the one with superior accuracy, the influence of movement characteristics on time delay detection shall be investigated explicitly here. The

force feedback from a spring afflicted with time delay T_d, is calculated by

$$f_h(t) = kx_h(t - T_d). \tag{4.14}$$

Respecting tissue dynamics which can be approximated by a mass-spring-damper model, the overall motor action that is required to move the limb in contact with the environment is

$$f_{m,res}(t) = m_h\ddot{x}_h(t) + d_h\dot{x}_h(t) + kx_h(t - T_d). \tag{4.15}$$

Without loss of generality, we consider the case that a non-delayed environment model is learned and utilized in the matched filter and state observation models, and the delayed feedback is perceived. Consequently, the predicted force feedback $\hat{f}_{m,res}(t)$, utilizing a model of body and environment *without* time delay is

$$\hat{f}_{m,res}(t) = m_h\ddot{x}_h(t) + d_h\dot{x}_h(t) + \hat{k}x_h(t). \tag{4.16}$$

Assuming further that the stiffness k is sufficiently accurately represented in the internal model, thus $\hat{k} \approx k$, the prediction error between model simulation and sensory feedback is calculated to

$$\Delta f_{m,res}(t) = \hat{f}_{m,res}(t) - f_{m,res}(t) = k(x(t) - x(t - T_d)). \tag{4.17}$$

With the same assumptions made for modeling time delay detection thresholds following the state observer theory, the decision criterion is based on the state observation

$$\hat{\mathbf{x}}_h(t) = \begin{bmatrix} \hat{x}_h(t) & \dot{\hat{x}}_h(t) \end{bmatrix}^T, \tag{4.18}$$

where a forward model of body and non-delayed environment dynamics is utilized. In the case of a time-delayed spring force feedback, the state predictions are determined by the system of differential equations

$$\begin{bmatrix} \dot{\hat{x}}_h(t) \\ \ddot{\hat{x}}_h(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{d_h}{m_h} & -\frac{k}{m_h} \end{bmatrix} \begin{bmatrix} \hat{x}_h(t) \\ \dot{\hat{x}}_h(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{m_h} \end{bmatrix} f_{m,res}(t) + \begin{bmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{bmatrix} \left(\begin{bmatrix} x_h(t) \\ \dot{x}_h(t) \end{bmatrix} - \begin{bmatrix} \hat{x}_h(t) \\ \dot{\hat{x}}_h(t) \end{bmatrix} \right). \tag{4.19}$$

A sinusoidal movement

$$x(t) = A\sin(\omega t) \tag{4.20}$$

with amplitude A and frequency ω is chosen as interaction pattern. This choice has the following advantages: First, a sinusoidal movement it is easy to understand and perform for participants in a psychophysical experiment. Second, by variations in the amplitude and frequency, the predicted time delay detection thresholds vary with model hypotheses.

For the matched filter model, the error in the motor command for the case of a compliant environment can be computed to

$$\Delta f_{m,res}(t) = kA(\sin(\omega t) - \sin(\omega(t - T_d))). \tag{4.21}$$

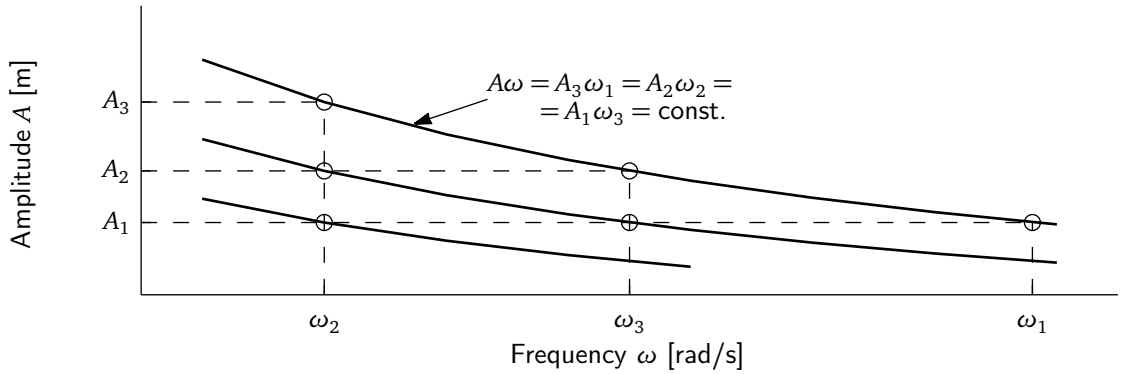


Figure 4.4: Six pairs of movement amplitudes and frequencies were chosen in such a way that ω, A and $A\omega$ have three different levels respectively.

Figure 4.3 depicts the time series of the filter result with non-delayed environment dynamics, the actual motor command required to move in the desired way, given the delayed sensory feedback, and the deviation between them. Analog to an "image slip" mechanism which was found to model time delay in visual perception well [142], the maximum force difference between expected non-delayed system behavior and sensory feedback could be a key factor in the detection of the time delay. It can be easily derived that the force difference reaches its maximum at time $\frac{1}{2}T_d$ after the zero-crossings of the predicted (non-delayed) force reference, which is expressed by

$$\Delta f_{m,res,max} = \Delta f_{m,res}(t)|_{t=\frac{1}{2}T_d} = kA2\sin(\frac{1}{2}\omega T_d) \approx kA\omega T_d. \tag{4.22}$$

The last step in the calculation holds for small values of ωT_d, which is a valid assumption for the practically relevant range of time delays in telepresence applications and the movement frequencies considered in the experiments.

Similarly, the state observation error can be computed by solving equation (4.19) for the specific interaction movement from equation (4.20) and the motor action from (4.15). In contrast to the solution for the maximum force error in equation (4.22), the maximum state error depends on the entries of the feedback matrix K. These values are not known a priori. Thus, the experimental conditions are optimized for the matched filter model, and the prediction capabilities of the state observer model are tested post-hoc with an appropriate feedback matrix K that is identified based on experimental data.

Keeping time delay T_d at a constant level, the maximum force error as the prediction criterion for time delay detection is higher with a greater amplitude A, and/or higher movement frequency ω. This means in return, that time delay needed to exceed a hypothesized perception threshold on force error is smaller with larger A and/or higher ω. Notably, the maximum force error as introduced in equation (4.22) depends on the product of A and ω, predicting that choosing values of A and ω such that their product is constant ($A\omega = \text{const.}$) results in the same detection threshold. For testing the influence of movement amplitude, frequency and their product, a systematic experimental design with three levels for A, three levels for ω and three levels of $A\omega$ as depicted in Figure 4.4 is chosen.

Another factor in the computation of the maximum force error according to equation (4.22) is the stiffness coefficient k. The perception model predicts a lower time delay detection threshold in the case where stiffness is higher.

In order to test the prediction model for its generality with respect to other environments than stiffness, the prediction capabilities in damping and inertia environments are tested as well. The magnitude d_1 of damping d, and the magnitude m_1 of inertia m are chosen to fulfill

$$\left.\frac{\Delta f_{m,res,max}}{f_{m,res}(t)|_{\Delta f_{m,res}(t)=\Delta f_{m,res,max}}}\right|_{d=d_1} = \left.\frac{\Delta f_{m,res,max}}{f_{m,res}(t)|_{\Delta f_{m,res}(t)=\Delta f_{m,res,max}}}\right|_{m=m_1},$$

such that the Weber fraction is equal in both conditions, resulting in a constant time delay detection threshold in the case of a perception criterion based on Weber's Law.

4.4 Experimental Investigations

Due to the large number of influence factors that help to determine a valid perception model for the combination of movement and force into a percept of a time-delayed haptic feedback, the measurement of detection levels of time delay in haptic feedback is divided into three experiments. In the first experiment, a simple stiffness environment with constant spring coefficient is used as haptic environment and the movement pattern is varied systematically, using the conditions depicted in Figure 4.4. Second, the stiffness coefficient is varied while the experimental idea of keeping the product of A and ω constant is kept. Third, the time delay detection thresholds obtained for stiffness are compared to those in damping and inertia environments, keeping the interaction movement constant.

4.4.1 Experimental Methods

A linear haptic interface with one degree of freedom was used in the experiments and described in Appendix A.1. Visual information was displayed on a 22" LCD Monitor with a refresh rate of 60 Hz. White noise was replayed as background sound during the experiment via KOSS QZ99 Headphones to mask any auditive cues from the haptic device. Participants were required to sit in an upright position centered towards the equilibrium point of the virtual spring, and the forcefield was rendered in the participant's transverse plane within a comfortable manual reachable range. Participants' perceptual responses were collected using a joystick. The force control scheme described in Section 2.2.2 was applied to render all stimuli. Applying this algorithm, the environment characteristics are displayed imperfectly, affected by the actuator dynamics. The deviation from the commanded and impedance is discussed in Appendix A.1. All drawn conclusions in the following sections are based on the actually rendered environment impedance, instead of the commanded impedances.

4.4.2 Variation of Movement

In order to examine the influence of the movement frequency ω, amplitude A, and force discrepancy which is directly dependent on $A\omega$, six experimental conditions were selected

as illustrated in Figure 4.4. The three movement amplitudes $A_{1,2,3}$ were set to 11.25 cm, 15 cm, and 20 cm, the frequencies $\omega_{1,2,3}$ to 0.75 Hz, 1.0 Hz, and 1.33 Hz, respectively. The spring stiffness k was kept constant at 65 $\frac{\mathrm{N}}{\mathrm{m}}$.

Participants

Fifteen university students (3 male; age range 21 - 37 years) participated in the experiment. All were right-handed and had normal or corrected-to-normal vision; none of them reported any history of somato-sensory disorders. Informed consent was obtained from all participants prior to the experiment.

Procedure

A 1-interval 2-alternative forced choice synchrony/asynchrony judgment was selected as the experimental task. We applied an adaptive double staircase method to keep testing brief and so avoid fatigue effects, as described in Section 2.4.1. One staircase started from an initial time delay $T_d = 100$ms, the second from $T_d = 0$ms. The initial step size for time delay differences between two conditions within one staircase was set to $\delta T_d = 20$ms. The response alternatives were “delayed haptic environment” and “non-delayed haptic environment”.

To train the participants’ memory of the non-delayed spring target environment, a pre-experimental practice block was inserted, where the subjects could explore a spring environment as long as they wanted. In addition, three different haptic latencies (0, 70, and 100 ms) were presented randomly, along with the value of the latency on the screen. In order to practice the experimental task, a dot on the screen moving in the desired sinusoidal way with a given frequency and amplitude indicated the required movement trajectory. The participant was asked to move the haptic device so as to follow the moving visual dot. Two vertical bars marked the movement boundaries and rhythmic click sounds indicated the reversal time of the dot. In the formal experiment, there was no indication of the delay level and no visual guidance of the movement on the screen. The guidance cue was removed since a pilot study had shown that tracking the visual movement was a rather attention-demanding task, potentially interfering with the required perceptual judgment. However, the click sounds were preserved to help users move in the right rhythm. The experiment was divided into six blocks, each block contained one of the six experimental conditions shown in Figure 4.4. At the beginning of each block, 3 practice trials (without visual guidance) with random latencies were presented before starting a double staircase procedure. The participants explored the system for 10 s and then indicated whether or not there was a force delay in the system by pushing the joystick to either the left (“delayed system”) or the right (“undelayed system”). The whole experiment took approximately one hour to complete.

Results

In total, 12 valid data sets were further analyzed; three participants’ data had to be excluded due to failure of convergence of the double staircase sequences in two or more blocks. In 90% of all cases, the staircase procedure converged within 30 trials. An analysis of

$\frac{1}{2\pi} \cdot \omega\,[\frac{1}{s}]/A\,[\text{cm}]$	$\frac{1}{2\pi} \cdot \hat{\omega}/\hat{A}$	$T_d \pm$ SE [ms]
0.75/11.25	0.71/8.89	46 ± 4.5
0.75/15	0.70/10.45	47 ± 7.3
0.75/20	0.68/13.19	37 ± 6.3
1/11.25	0.93/8.62	41 ± 5.0
1/15	0.92/10.74	37 ± 4.2
1.33/11.25	1.24/8.65	36 ± 5.8

Tabular 4.1: Key measures of the actual movement and detection thresholds with the corresponding standard errors observed in Experiment 1. In column 2, the mean of the actually observed movement amplitude $\hat{A}$ and frequency $\hat{\omega}$ are summarized; the last column presents mean thresholds with standard errors.

the movement trajectories revealed that all participants in the experiment made reasonably sinusoidal-shaped movements, although the amplitudes and frequencies deviated slightly from the desired motion trajectory (see Figure 4.5a). Mean values of real amplitudes and frequencies are summarized along with mean threshold values and standard errors in Table 4.1.

The estimated mean thresholds for two factors are shown in Figure 4.5b. All datasets (i.e., threshold estimates) are examined by a univariate analysis of variance (ANOVA) with amplitude A, frequency ω as fixed factors, and subjects as random factor. The analysis reveals the main effects of amplitude A and frequency ω to be significant, $F(2,22) = 4.38, p < 0.05$ and $F(2,22) = 4.79, p < 0.05$, respectively. A further contrast test for the factor of amplitude shows that the threshold is only significantly lower for $A = 20$ cm compared to the other two amplitude levels, $F(2,56) = 3.83, p < 0.05$. Another contrast test for the factor frequency reveals that the thresholds decrease significantly when the movement frequency increases, $F(2,56) = 5.87, p < 0.01$. A separate univariate ANOVA is conducted for the factor of $A\omega$, which reveals the main effect to be significant, $F(2,22) = 6.42, p < 0.01$. A contrast test shows that the largest value of $A\omega$ results in the lowest threshold.

To further examine the relationship between detection thresholds and the actual movements, we conduct a linear regression for the mean thresholds with $\hat{A}$ and $\hat{\omega}$,

$$T_d = 84.5 - 2.32\hat{A} - 23.67\hat{\omega}, r^2 = 0.833, \tag{4.23}$$

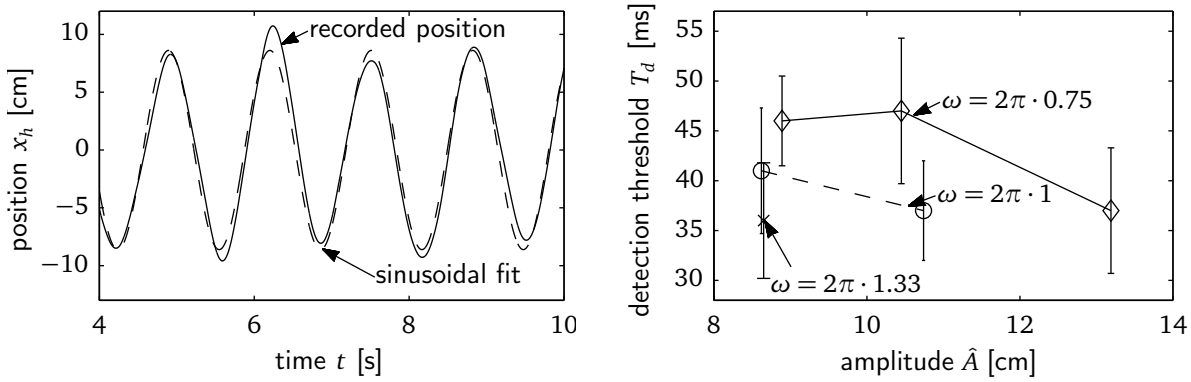
and a linear regression for the mean thresholds with $\hat{A}\hat{\omega}$,

$$T_d = 63.75 - 2.69\hat{A}\hat{\omega}, r^2 = 0.815. \tag{4.24}$$

Both linear regressions suggest that the independent movement parameters $\hat{A}$ and $\hat{\omega}$ and their product $\hat{A}\hat{\omega}$ are potentially influential factors for haptic feedback delay.

4.4.3 Variation of Environmental Stiffness

Participants of this experiment were required to move in a sinusoidal way similar to the previous experimental task, described in Section 4.4.2 Two pairs of A and ω values were



(a) (b)

Figure 4.5: (a) An example of a manual movement with a given frequency and amplitude. The dashed curve denotes the best sinusoidal fit to the actual movement which is plotted solid. (b) Mean detection thresholds as a function of actual movement amplitude $\hat{A}$ and instructed frequency ω. Solid stars correspond to frequency $\omega = 0.75$ Hz, dashed circles to $\omega = 1$ Hz, and the cross to $\omega = 1.33$ Hz. Error bars indicate 1 standard error.

selected: $[A_1 = 20$ cm, $\omega_1 = 0.75$ Hz$]$ and $[A_2 = 15$ cm, $\omega_2 = 1$ Hz$]$, thus the product $A_i\omega_i$, $i \in \{1, 2\}$ was constant. In addition, three discriminable levels of spring stiffness were examined: $k_1 = 40\frac{\text{N}}{\text{m}}$, $k_2 = 50\frac{\text{N}}{\text{m}}$, and $k_3 = 65\frac{\text{N}}{\text{m}}$. k_1 and k_2 were selected to be lower than k_3 (used in Experiment 1) in order to avoid fatigue during the Experiment.

Participants

Ten participants took part in this experiment (two of them had taken part in the previous experiment; 4 male; age range 21 - 29 years). All had normal or corrected-to-normal vision and were right-handed; none of them reported any history of somato-sensory disorders. Informed consent was obtained from all participants prior to the experiment.

Procedure

A 2-interval 2-alternative forced-choice (2I2AFC) paradigm was used here. We chose this method because several participants in the previous experiment found it reportedly hard to remember the baseline ("non-delayed") condition, making them adopt a conservative response strategy during the experiment. On each trial, two intervals were presented, one standard interval with non-delayed force feedback and one comparison interval with delayed force feedback. By providing a standard stimulus on each trial, the 2I2AFC procedure helps to reduce response bias and variability among subjects. The order of the standard and comparison (target) stimuli was randomized across trials. Each stimulus was presented for 7 seconds, with a transition phase between them. In the transition phase, the system latency was linearly transferred from one state to another to avoid cues arising from abrupt changes of system latency. After participants explored the two stimuli, they were prompted to respond to the question "In which interval of the trial did you experience a delayed

$\frac{1}{2\pi} \cdot \omega\,[\frac{1}{s}]/A\,[\mathrm{cm}]/k\,[\frac{\mathrm{N}}{\mathrm{m}}]$	$\frac{1}{2\pi} \cdot \hat{\omega}/\hat{A}$	$T_d \pm$ SE [ms]
1/15/40	1.06/14.8	24 ± 4.5
1/15/50	1.08/14.8	25 ± 6.5
1/15/65	1.12/14.4	28 ± 9.3
0.75/20/40	0.84/18.5	34 ± 9.3
0.75/20/50	0.85/18.4	31 ± 4.5
0.75/20/65	0.84/18.7	37 ± 9.5

Tabular 4.2: Key measures of the actual movement and detection thresholds with the corresponding standard errors observed in Experiment 2. In column 2, the actually observed movement amplitude $\hat{A}$ and frequency $\hat{\omega}$ are summarized; the last column presents mean thresholds with standard errors.

force feedback?", to which they made a 2I2AFC decision. The adaptive double-staircase method described in Section 2.4.1 was modified according to the changed experimental paradigm. While the detection threshold has been 50% in Experiment 1, it is raised to 75% in the 2I2AFC paradigm of Experiment 2. Therefore, the adaption rule for the step size was changed to

$$T_{d,i+1} = T_{d,i} - 2\frac{20\text{ ms}}{1+N_{rev}}(Z_i - 0.75), \tag{4.25}$$

with the same terms as in equation (2.21). After every 3 experimental blocks, there was a break permitting participants to take a rest. The remaining procedure, including the familiarization with time delay in the haptic feedback, was the same as in Experiment 1, and the whole experiment took about 2 hours to complete.

Results

One data set was excluded from the further analysis due to the participant having failed to follow the required movement frequency. The mean discrimination thresholds, along with their standard errors are summarized in Table 4.2.

Estimated from the actual movements, the mean values of the product factor, $\hat{A}\hat{\omega}$, were 15.9 and 15.6 for two movement pairs [$A_1 = 20$ cm, $\omega_1 = 0.75$ Hz] and [$A_2 = 15$ cm, $\omega_2 = 1$ Hz], respectively - that is, not significantly different from each other (paired t-test: $p > 0.1$). The individual threshold estimates were further examined by a univariate ANOVA with fixed factors ω (same as for A) and stiffness k and subjects as random factor. This analysis revealed the factor ω (i.e., pair of $A\omega$) to be significant, $F(1,8) = 9.46, p < 0.05$, while the factor k (stiffness) failed to reach significance, $F(2,16) = 0.798, p > 0.1$. There was no significant interaction, $F(2,16) = 0.245, p > 0.1$.

4.4.4 Variation of Environment Type

Participants of this experiment were required to make sinusoidal arm movements of fixed amplitude $A = 15$ cm and frequency $\omega = 2\pi$ rad/s. Three haptic environments were used as

experimental stimuli: Besides the stiffness condition used in Sections 4.4.2-4.4.3, perceptual limits for time delay discrimination were measured in a damping and inertia environment additionally. The environment-specific constants were chosen to result in constant time delay detection thresholds in case of the Weber-based prediction model. For comparison reasons, the stiffness value is set equal to the previous conditions, $k = 65\frac{\mathrm{N}}{\mathrm{m}}$. The damping condition is determined to $d_1 = 65/(2\pi)\frac{\mathrm{Ns}}{\mathrm{m}}$. As a pure inertia can not be rendered in a stable way if time delay is added, the inertia $m = 22/(2\pi)^2\frac{\mathrm{Ns}^2}{\mathrm{m}} \approx 560$ g is combined with a damping $d_2 = 43/(2\pi)\frac{\mathrm{Ns}}{\mathrm{m}}$.

Participants

Ten university students (7 male; age range 20-27 years; 9 right-handed) participated in the experiment; they were paid at a rate of 10 Euros per hour. None of them reported any history of somatosensory disorders. Informed consent was obtained from all participants prior to the experiment.

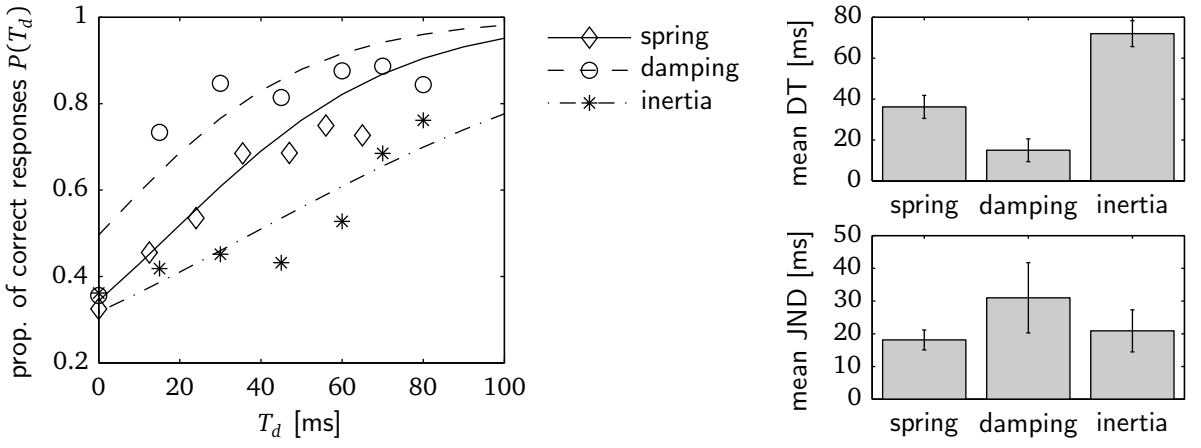
Procedure

An odd-one-out 3-alternative forced-choice (3AFC) was used in this experiment. Each trial was subdivided into three intervals of 3 seconds; one randomly chosen interval contained the stimulus condition with a specific level of time delay introduced into the force feedback. In the remaining two intervals, the system responded with a non-delayed force feedback. In order to remove any abrupt-onset cues due to switching delay conditions, a transition phase of 1.5 seconds was added between delayed and non-delayed force feedback intervals. No feedback about the correctness of the answer was provided. Using the method of constant stimuli, described in Section 2.4.1, seven levels of time delay T_d, segmented equally between 0 ms and 50 ms, were tested for the spring environment, and seven delays between 0 ms and 80 ms for the mass and damping environments. These levels were determined in a pilot study and respected stability conditions for the experimental setup. Each delay level was tested 21 times, yielding a total of 441 trials. The experiment was conducted over three sessions, where each session contained 7 repetitions of all conditions in a random order.

Results

Three participants failed to follow the movement instructions accurately enough, so their data had to be excluded from further analysis. With the remaining data sets, an adjusted logistic function [58]

$$P(T_d) = \gamma + (1-\gamma)\cdot\frac{1}{1+e^{\frac{\alpha-T_d}{\beta}}} \tag{4.26}$$



(a) (b)

Figure 4.6: (a) Estimated psychometric curves for the three environments for one typical participant. (b) Mean thresholds (± standard error, $n = 7$) for the different environments.

is used for estimating the psychometric function, where the guessing rate γ is set to 1/3 according to the 3AFC paradigm. Figure 4.6a shows typical correct responses, produced by one participant, for the three different environments. Using equation (4.26), the discrimination threshold DT and the just noticeable difference JND at response level 67% can be easily obtained as DT= α and JND= $\beta \log 3$. In Figure 4.6b, the mean DTs and JNDs for spring, damper and inertia are presented. The results indicate that the discrimination threshold is the largest for the inertia, with a mean of 72 ms. The threshold associated with the damper is the lowest, with a mean of 15 ms. In the spring condition, time delay can be discriminated from a non-delayed spring starting at a threshold of 36 ms. A repeated-measures ANOVA reveals the discrimination thresholds to be sensitive to the different environments, $F(2,12) = 14.17, p < 0.01$. Follow-on comparison tests show the discrimination thresholds to differ reliably from each other ($p < 0.05$). This indicates that the different environments do indeed influence the subjective judgment of haptic feedback time delay. A further repeated-measures ANOVA for the JNDs of discrimination fails to reveal a significant effect of the environment, $F(2,12) = 1.03, p > 0.1$.

4.5 Discussion

Four substantial findings can be summarized from the experiments described in Sections 4.4.2-4.4.4:

1. The detection thresholds for time delay in the force feedback are negatively correlated with movement frequency and movement amplitude.

2. Movement amplitude and frequency influence the delay detection separately.

3. Within a comfortable force range, scaling of stiffness does not affect discrimination of time delay in force feedback.

4. Time delay in damping force feedback can be detected easiest, followed by time delay in spring force feedback and inertial force.

Recollecting the model predictions for potential influence factors on the detection limits for time delay in haptic feedback from Sections 4.3.1-4.3.3, none of the hypothesized perception mechanism can account for all phenomena observed in experimental data. For technical applications such as telepresence systems, however, a quantitative model for a perceptual phenomenon which does not account for all observed phenomena is still valuable as it allows an educated design choice for systems and control algorithms. As a consequence, the model candidate explaining the experimental findings best, though not perfectly, is identified, before consequences from the findings for the design and control of telepresence systems are discussed.

4.5.1 Model Predictions

To compare the prediction capabilities of each perception model candidate, specific parameterizations have to be found by fitting the predictions of each model to experimental data. The chosen parameter identification procedure which is required for this purpose founds on the minimization of the squared prediction error. The model fits of models from Sections 4.2.1-4.2.3 are summarized in Figure 4.7. As we used a different experimental method in each study, predicted detection thresholds are calculated for each experiment individually. Furthermore, as not all participants performed all three experiments, only conclusions about the mean detection thresholds can be drawn. The mean squared error (MSE) between model predictions and experimental data is 127.34 ms^2 for the feature comparison model, 96.0 ms^2 for the matched filter model, 127.7 ms^2 for the matched filter model with considering Weber's Law, 98.3 ms^2 for the state observer using the position error as decision variable, and 85.7 ms^2 for the state observer model with the limb velocity as decision variable, respectively. In the following, the individual identification procedures and the models' prediction characteristics are discussed more detailed.

Feature Comparison Model

The perceptual identification process for haptic environment features which is discussed here is based on an estimation of the covariance matrix. For this representation, an accurate estimate of the temporal relation between a dynamic system's input and output signal is achieved. Time delays in the environment are easily detectable by finding the peak in the input-output cross-correlation, see Section 2.5.2. A perceptual threshold for time delay detection could thus be explained by an uncertainty in identifying the corresponding lag operator τ, or the peak detection. As there is no apparent reason why uncertainty in τ should change with input amplitude, frequency, magnitude, or the type of environment, and the peak in the cross-covariance of time-shifted sinusoids is symmetric (see Figure 4.2), the predicted time delay detection threshold based a correlation model is constant. Finding the

k	d	m	$\hat{A}$ [cm]	$\frac{\hat{\omega}}{2\pi}$
65	0	0	8.89	0.71
65	0	0	10.45	0.70
65	0	0	13.19	0.68
65	0	0	8.62	0.93
65	0	0	10.74	0.92
65	0	0	8.65	1.24
40	0	0	14.80	1.06
50	0	0	14.80	1.08
65	0	0	14.40	1.12
40	0	0	18.50	0.84
50	0	0	18.40	0.85
65	0	0	18.70	0.84
65	0	0	11.10	1.03
0	$\frac{65}{2\pi}$	0	11.30	1.05
0	$\frac{43}{2\pi}$	$\frac{22}{(2\pi)^2}$	11.20	1.08

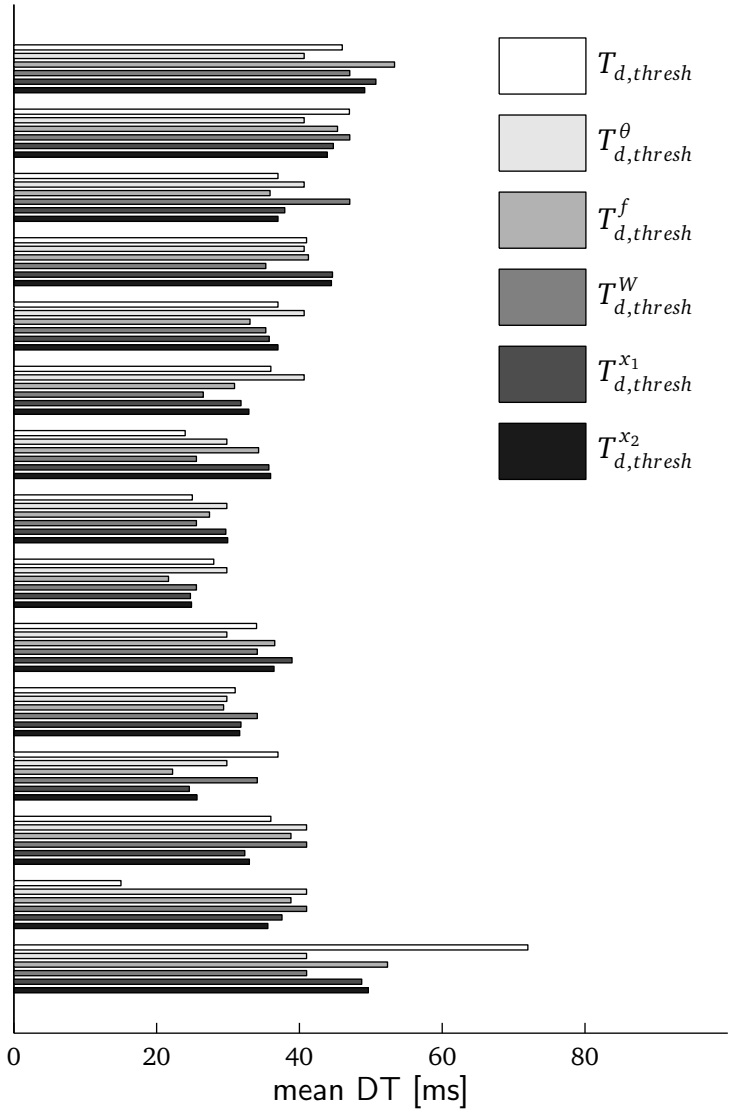


Figure 4.7: The three model candidates for the perceptual process combining motion and force feedback together predict different mean detection thresholds for time delay in the experimental conditions investigated in Sections 4.4.2-4.4.4. The measured detection thresholds are depicted as the top bar. Below bars illustrate (from top to bottom) the feature comparison model, matched filter model, matched filter model with Weber's law, and the state observer model using errors in position and velocity as threshold variable.

covariance model that leads to the best predictions with respect to the mean squared error corresponds to solving an optimization problem

$$\arg\min_{T^{\theta}_{d,thresh}} \frac{1}{N_{cond}} \sum_{i=1}^{N_{cond}} (T_{d,thresh,i} - T^{\theta}_{d,thresh})^2 \tag{4.27}$$

where N_{cond} is the number of conditions in the respective experiment, and $T^{\theta}_{d,thresh}$ is the (constant) time delay detection threshold. The solution to this optimization problem is the mean time delay over all conditions within one experiment. The individual detection thresholds for time delay vary significantly within subjects, depending on the movement and the environment, and the mean squared deviation of model predictions from the mean detection thresholds is highest (127.34 ms^2) in the assumption of a constant detection limit. In this sense, the average time delay detection threshold is a baseline for evaluating the other model candidates.

Matched Filter Model

The matched filter prediction gives an estimation of the motor signal which is required to move the limb in a desired way, founding on the basis of an internal representation of a haptic environment. A detection mechanism for time delay that is proportional to the error between actual muscle force and the prediction is based on inverse models for human body and environment dynamics. Indeed, the results of the movement variation experiment in Section 4.4.2 support a force-based detection threshold, as the product between movement amplitude and frequency, proportional to the maximum force error, is found to be a significant influence factor. The detection thresholds for the conditions with equal maximum force does not differ significantly.

The parameterization of a matched filter model on the basis of the experimental results requires solving a nonlinear constrained optimization problem

$$\arg\min_{T_{d,thresh},\Delta f_{thresh}} \frac{1}{N_{cond}} \sum_{i=1}^{N_{cond}} (T_{d,thresh,i} - T^{f}_{d,thresh,i})^2 \tag{4.28}$$
$$s.t. \max \Delta f_{m,res,i}(t) = \max |f_{m,res,i}(t) - \hat{f}_{m,res,i}(t)| = \Delta f_{thresh} \; \forall i \in (1, \ldots, N_{cond})$$

where Δf_{thresh} is the (constant) detection threshold for the difference between the delayed and non-delayed exerted force. The predicted motor action on the basis of the measured state $\mathbf{x}_h(t)$ is computed for each individual experimental condition, indexed by i, and denoted $\hat{f}_{m,res,i}(t)$. A numeric optimization algorithm based on the interior-point method was used to find the optimal parameterization fitting all experimental conditions [143]. Using the dynamic inverse model to explain average detection thresholds for time delay perception results in lower prediction errors (96.7 ms^2) compared to the feature comparison model prediction. The mean force difference thresholds for the experiments are 1.4 N for the first, 1.2 N for the second, and 1.7 N for the third experiment. The differences between these thresholds may be explained by the fact that different experimental paradigms were used: While the first experiment relies on the discrimination of time delayed haptic feedback without giving the reference of an non-delayed environment, the other two provide one or two phases in each trial where the control condition is presented.

Force difference perception for experiments with slowly-changing forces is known to follow Weber's Law, introduced in Section 2.4.2 [7]. The Weber fraction of $\Delta f_h(t)$ could be an explaning model for the perception limits with changing environmental stiffness, as discussed in Section 4.4.3. The optimization problem to be solved is similar to equation (4.28)

$$\begin{aligned} &\underset{T^W_{d,thresh,i},W}{\arg\min} \sum_{i=1}^{N_{cond}} (T_{d,thresh,i} - T^W_{d,thresh,i})^2 \\ &s.t. \max \frac{\Delta f_{m,res,i}(t)}{f_{m,res,i}(t)} = W \; \forall i \in (1,\ldots,N_{cond}) \end{aligned} \tag{4.29}$$

with W the Weber fraction. Indeed, the model fit for the experiment with different stiffness levels is admittedly good, with a MSE of only 4.5 ms^2, but the model performs poor in all other conditions, yielding to a total MSE of 127.7 ms^2. This makes the prediction model based on a matched filter model and a threshold based on Weber's Law about equally successful as the thresholds predicted by a covariance model.

State Observer Model

In contrast to the matched filter perception model, the state observer model utilizes an estimation of the body state for the decision about the environment time delay. The difference between the observed state and actual state heavily depends on the choice of the feedback matrix K, as discussed in Section 4.2.3. The model predicts perception limits based on a threshold in the state estimation error. Because the state $\mathbf{x}_h(t)$ consists of two components, namely the limb position $x_h(t)$ and velocity $\dot{x}_h(t)$, two versions of the optimization problem of finding the minimum mean squared error between time delay detection threshold and the model predictions are discussed:

$$\begin{aligned} &\underset{T^{x_1}_{d,thresh,i},\Delta\mathbf{x}_h,K}{\arg\min} \frac{1}{N_{cond}} \sum_{i=1}^{N_{cond}} (T_{d,thresh,i} - T^{x_1}_{d,thresh,i})^2 \\ &s.t. \max \Delta x_h(t) = \max |x_h(t) - \hat{x}_h(t)| = \Delta x_{h,thresh} \; \forall i \in (1,\ldots,N_{cond}) \end{aligned} \tag{4.30}$$

and

$$\begin{aligned} &\underset{T^{x_2}_{d,thresh,i},\Delta\mathbf{x}_h,K}{\arg\min} \frac{1}{N_{cond}} \sum_{i=1}^{N_{cond}} (T_{d,i} - T^{x_2}_{d,thresh,i})^2 \\ &s.t. \max \Delta \dot{x}_h(t) = \max |\dot{x}_h(t) - \hat{\dot{x}}_h(t)| = \Delta \dot{x}_{h,thresh} \; \forall i \in (1,\ldots,N_{cond}). \end{aligned} \tag{4.31}$$

The problems formulated in (4.30) and (4.31) have five free parameters which must be optimized. Due to the comparably low number of experimental conditions which are available for model fitting, the solution can depend on the chosen initial values. Suitable values are found from a simulation of the state space observer for a number of different feedback matrices K and the computation of the state observation errors $\Delta\mathbf{x}_h(t)$ that result from the experimentally obtained time delay detection thresholds. The matrix K that results in the lowest variance for $\Delta\mathbf{x}_h(t)$ between all conditions of each experiment is taken as initial values for the optimization problems stated in equations (4.30) and (4.31).

As a result, the state observers with feedback matrix

$$K_1 = \begin{bmatrix} 11.8 & 36.3 \\ 33.3 & 31.1 \end{bmatrix}, \text{ and } K_2 = \begin{bmatrix} 0 & 9.8 \\ 9.4 & 11.4 \end{bmatrix} \tag{4.32}$$

for predictions based on x_h and $\dot{x}_h$, respectively, make predictions with the lowest mean squared error.

Synopsis

Comparing the predictions from all models introduced in Sections 4.2.1-4.2.3 leads to the conclusion that the state observer model with a detection mechanism on the observation error in limb velocity is most successful in capturing the observed perceptual behavior. It is noteworthy that a state observer has the greatest number of degrees of freedom available to fit the experimental data. A criticism frequently raised when searching for computational models resembling a perceptual, or cognitive process is that "with 8 [...] parameters you can fit any pattern of results" [68]. This argument is admittedly true and without further experimental investigations, the capabilities of all models proposed here to predict perception thresholds for time delay in the combination process of movement and force feedback are limited to the small, but practically relevant set of movement stimuli and haptic environments presented here.

4.5.2 Implications for Telepresence Systems

Time delay is a critical issue for haptic telepresence systems operating over long distances. Challenges to be dealt with include technical issues such as system instability and, on the side of the human operator, impaired perception of the environment's haptic properties.

The experiments presented in this chapter point towards important factors that must be taken into consideration when designing a haptic telepresence system, or developing haptic communication protocols: First of all, the operator's movement must be taken into consideration to evaluate whether a time delay in the communication channel is perceived or not. A haptic task which requires only slow movements can tolerate longer delays in the feedback than a highly dynamic task requiring fast movements. Not only the task can limit the velocity and allowable movement frequency, but also the haptic interface. A smaller workspace on the one hand, and high friction or uncompensated inertia on the other hand can influence the detection thresholds. The workspace dimensions of the local haptic interface determine the maximum movement amplitude, and detection thresholds increase. With larger inertia and damping of the local haptic interface, the achievable human movement frequency decreases – resulting in a higher detection threshold for time delay.

The finding that a scaling of the stiffness coefficient within the investigated range does not influence the sensitivity of temporal perception is interesting for the application in a specific telepresent application, namely micromanipulation. In this area, small forces arising in a micro-scale environment must be augmented for the user to provide a perceptible haptic impression [144]. For the case of delayed haptic feedback, our finding suggests that the scaling factor can be chosen irrespective of haptic latency. Note, however, that we only

validated this hypothesis for a limited range of stiffnesses. In extreme scenarios, such as stiff contact with a rigid object, an infinitesimally small time delay may result in an unstable system, which completely changes the characteristics of the system. The human operator may then be able to infer the time delay from increasing oscillations in the force feedback.

The identification of dynamic models capturing the detection thresholds for time delay in haptic feedback has direct application for the design of communication algorithms, or haptic rendering systems as well: The greatest benefit of these models lies in the possibility to consider the influence of interaction movements on the perceptual threshold explicitly. In this way, more accurate predictions whether time delay in haptic feedback is perceived or not can be utilized during the execution of a task, and appropriate measures can be taken. As an example, communication quality control algorithms that provide the possibility to adjust the transmission quality of a communication channel, e.g., the time delay, can be optimized based on quantitative online-predictions from a perception model. For a more detailed view on communication quality control, we refer to Chapter 6 where such techniques are utilized to attenuate a drop in human task performance in a time-delayed simulated telepresence task.

4.6 Conclusions and Open Problems

Humans do not possess senses for haptic environment properties such as stiffness, damping, or inertia. Instead, temporal and magnitude information from movement and force feedback must be combined together to infer such measures. A dynamic model for the perceptual process of combining movement and force feedback into a unified percept of a haptic environment's feature has been developed. In specific, the so far unknown temporal limits of this process have been investigated. Taking together the results of six psychophysical experiments on time delay perception thresholds, a dynamic model in the form of a state observer is identified as the model capturing human discrimination performance best, when movement and force feedback are temporally misaligned.

Although all model candidates have been tested for a number of different movements, their shape was so far restricted to sinusoids of different amplitudes and frequencies. For a more general applicability to haptic telepresence systems, other movements must be considered as well. Ultimately, perceptual responses for time-delayed feedback from arbitrary voluntary explorations shall be predictable. Furthermore, the modeling performance in the third experiment, considering time delay perception levels in stiff, damped and inertial environments have not been captured well by either model proposed so far. Alternative models with other decision criteria could further improve the prediction performance. Together with a dynamic perception model for the influence of magnitude information on the combination of movement and force, conclusions about perception mechanisms for abstract environmental features such as stiffness, damping and inertia could be drawn.

5 Dynamic Masking in Haptic Discrimination

Summary. *Many practically relevant haptic environments can be approximated by combining inertia, damping, and stiffness. It is largely unknown how differences in the environmental properties are perceived by the operator of a telepresence system, specifically how the perception of deviations in one property is influenced by the overall environment composition. Such knowledge could though significantly enhance evaluation techniques for haptic transparency. Identifying a dynamic perception model for discriminating damping in a combination with inertia or stiffness, we find that*

- *the just noticeable difference of damping increases linearly with the magnitude of the masking property,*
- *differences in damping may be judged by a change in force feedback magnitude,*
- *dynamic perception models with an accumulating behavior achieve a higher modeling precision.*

A large class of haptic environments can be modeled as a combination of inertia, damping, and stiffness components. While the just noticeable difference (JND) for each component is known [7, 11, 12, 14], the perception of environments with more than one integrant has received little attention. Specifically, it is unknown how well differences in one component are perceivable in the presence of other, potentially distracting parts. Knowledge of such effects is though crucial for the design of perception-optimized haptic rendering algorithms, mechanisms, or telepresence systems: In cases where haptic characteristics can not be rendered exactly, due to, e.g., design or energy constraints, tolerances for the required accuracy of the environment characteristics containing inertia, damping, and stiffness is required. In the tactile and auditive modality, non-relevant signals are known to deteriorate the perception of a stimulus signal, which is known as *masking effect*. Two types of masking effects are known in these domains: Simultaneous masking, meaning that a loud tone or vibration enlarges the perceptual thresholds in sound and tactile perception, such that soft tones or low vibrations are not detectable, but are perceivable if presented alone. In contrast, temporal masking describes the phenomenon that stimuli can be perceived worse if an intense vibration or sound is presented temporally close to it.

In this Chapter, masking effects in the perception of haptic environments are investigated. The just noticeable difference of damping is measured under different masking conditions, meaning, additional inertia or stiffness. It is found that the JND for damping is significantly larger when a masking stimulus is present, compared to the discrimination of pure damping. Inspired from findings in auditive and tactile perception, different perception models are tested for their ability to capture the discrimination thresholds under different masking

conditions. A simultaneous masking model for the haptic domain based on Weber's Law, and a temporal masking model on the basis of a nonlinear filter algorithm temporally extending the masking stimulus are developed. In addition, and inspired from the structure of a diffusion model which is able to capture perceptual phenomena in multiple domains others than haptics, the inclusion of an information accumulation stage into the perception model is investigated. Overall, six quantitative dynamic model candidates for the discrimination of damping differences are evaluated and compared with respect to their internal and external validity. While no temporal masking effect can be found in the experimental data, the addition of an information accumulation stage into the perception model improves the prediction capabilities significantly. The best model capturing the discrimination performance of damping masked by stiffness or inertia consists of a simultaneous masking model and an accumulating component.

The remainder of the Chapter is organized as follows: After reviewing fundamentals of masking effects in other domains than haptics in Section 5.1, two experiments are presented. In Section 5.2, a pilot study for selecting stimuli for the main experiment is performed. The results from the main experiment, measuring damping JNDs in the presence of masking stimuli in Section 5.3 are given in Section 5.4. Different dynamic haptic masking models are motivated and developed in Section 5.5, and simulation results presented in Section 5.6. A discussion of all results and application examples of the novel findings conclude the chapter in Section 5.7.

5.1 Background

The detection or discrimination of a physical stimulus can be influenced by the context it is embedded into. So-called masking effects have been observed in tactile perception, as well as the auditory domain: Superimposing a vibration signal to a target stimulus in the tactile domain makes a discrimination of the target stimulus significantly harder [145, 146]. In auditory research, two kinds of masking effects are known – simultaneous masking and temporal masking [1, 124]. Simultaneous masking describes the changes in human ability to discriminate a sound with a given intensity in the simultaneous presence or absence of a masking sound. As an illustrative practical example, the sound of the doorbell is not heard when the vacuum cleaner is turned on, even though the doorbell's ring could be easily registered without additional noise. Due to the temporal masking phenomenon, sounds with a low intensity can not be perceived even in the case of a masking stimulus with high intensity being only in its temporal vicinity, although it is not present at the same time instance. Interestingly, this masking effect is observable for a time interval before and after the masking stimulus. In the case of a masking stimulus preceding the target stimulus, the effect is called *forward masking*, in the case of the masking stimulus following the target, the effect is referred to as *backward* masking. Both, simultaneous and temporal masking effects are widely used to reduce the amount of data to be stored in psychoacoustic lossy data storage formats, e.g., MP3 [147]. To the best knowledge of the author, there are no findings on masking phenomena when perceiving a haptic impedance such as combinations of inertia, damping, and stiffness. However, force and torque were found to interact in perceptual dis-

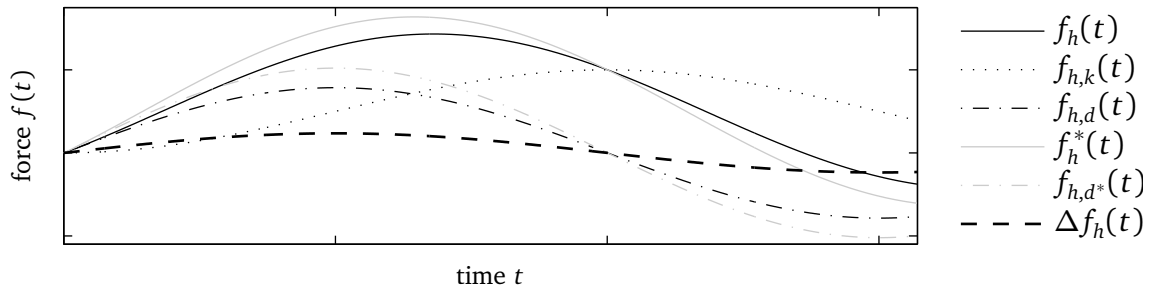


Figure 5.1: An interaction movement with a haptic environment consisting of a damping d and stiffness k component results in an interaction force $f_h(t)$ which is composed of a damping force $f_{h,d}(t)$ and stiffness component $f_{h,k}(t)$. A different damping coefficient d^* results in a change in force feedback.

crimination [148], indicating that unrelated stimuli can also affect discrimination abilities in kinesthetic stimuli.

Knowing the perceptual limits of haptic environment properties can help to decide the best alternative among different designs that introduce deviations from an ideal behavior. In the field of telepresence systems, haptic perception limits have been taken into account to analyze *perceived transparency* [24]. Hereby, the error between displayed and remote environment is evaluated in a human-oriented way. In this context, interactions between different environment parameters can be used to define perceptual limits more exactly. So far, only pure inertia, damping, or stiffness have been considered. An extension to more general telepresence architectures and environment impedances is the evaluation of telepresence systems based on transparency errors of environment parameters [35]. Additional application examples are found in the area of haptic rendering algorithms.

Humans do not possess a dedicated sense for environment parameters. Instead, inertia, damping, and stiffness must be inferred indirectly by perceiving movement and force feedback. For an intuitive understanding of the discrimination of a specific haptic environment feature in the presence of one or multiple other features, an interaction with a linear, time-invariant homogeneous haptic environment shall be discussed. In this example, the human interacts with the environment using a specific trajectory.

Example 5.1. We consider a linear, time-invariant spatially homogeneous haptic environment which is modeled by a stiffness k combined with a damping d. The environment is explored by following a specific trajectory $x_h(t)$, resulting in an interaction force $f_h(t)$ which can be decomposed into force components from the stiffness $f_{h,k}(t)$ and the damping $f_{h,d}(t)$, such that $f_h(t) = f_{h,k}(t) + f_{h,d}(t)$. Exemplary forces from a sinusoidal interaction movement are depicted in Figure 5.1. Keeping the interaction movement but changing one environmental parameter, e.g., the damping d to another value d^* results in a change in force feedback due to damping, from $f_{h,d}(t)$ to $f_{h,d^*}(t)$, resulting in a force difference $\Delta f_h(t) = f_{h,d^*}(t) - f_{h,d}(t)$. Perceiving this difference is crucial for the detection of differences in the damping, as the interaction trajectory is held constant, thus does not contain relevant feedback information. While the difference $\Delta f_h(t)$ only depends on the difference in damping, the role of the force component related to stiffness $f_{h,k}(t)$ for the perception of $\Delta f_h(t)$ remains unclear. It seems reasonable that a high force which is unrelated to the actual perception task can influence

the discrimination capabilities of a human interacting with a haptic environment. In the following, we will refer to this potential effect as *haptic masking*.

Remark 5.1. While position-controlled interactions have been discussed here, a similar argumentation for interactions with a constant force profile can be found.

We present two experiments to study masking effects in haptic environments [149]. As a model for this environment, a mass-spring-damper system is considered. Measuring the JND for each environment parameter masked by the other two environment parameters would result in a huge number of conditions, thus we limit our view to the JND for damping. Damping is chosen for two reasons: On the one hand, it can be assumed that the largest masking effect between two stimuli is achieved if they are as similar as possible. The frequency responses of stiffness and inertia differ from that of damping by a phase shift of only $\pm 90°$, whereas the phase shift between stiffness and inertia is $180°$. On the other hand, damping is of significant importance for controller design, as higher damping can enlarge stability margins, thereby guaranteeing safe operation of a system.

5.2 Pilot Study for Stimulus Selection

Two different damping levels $d_1 = 10$ Ns/m and $d_2 = 20$ Ns/m which are above the absolute detection threshold and small enough to prevent fatigue were chosen. These stimuli are easily discriminable, as JNDs for viscosity are reported between 13.6% [14] and 34% [12]. In this pilot study, we separately determine parameter values for stiffness and inertia that will be used in the main experiment to mask damping. For each parameter, we further distinguish between two cases: In the first case, humans should perceive the overall environment as a combination of damping and stiffness or damping and inertia, where neither dominates the other. For this case, the values for stiffness and inertia are denoted as k_1 and m_1. In the second case, humans should perceive an environment consisting mainly of stiffness or inertia. Damping should still be perceivable but should be subordinate. In this case, stiffness and inertia are denoted as k_2 and m_2.

Participants

The study was performed by 10 subjects including the author. One of the subjects was female and all were PhD students. Their mean age was 27.0 years. All gave their informed consent to participate in the study. Prior to the experiment, all participants were familiarized with the setup and the conditions by exploring environments consisting of pure inertia, pure damping, and pure stiffness. All participants completed the experiment in less than 20 minutes.

Procedure

To find the values for $k_{\{1,2\}}$ and $m_{\{1,2\}}$, participants were asked to modify an environment consisting of damping and a minimal inertia $m_0 = 0.5$ kg, which was necessary to ensure system stability. The participants could add stiffness or inertia by using a turning knob. In the conditions where stiffness had to be added, one full turn of the knob corresponded

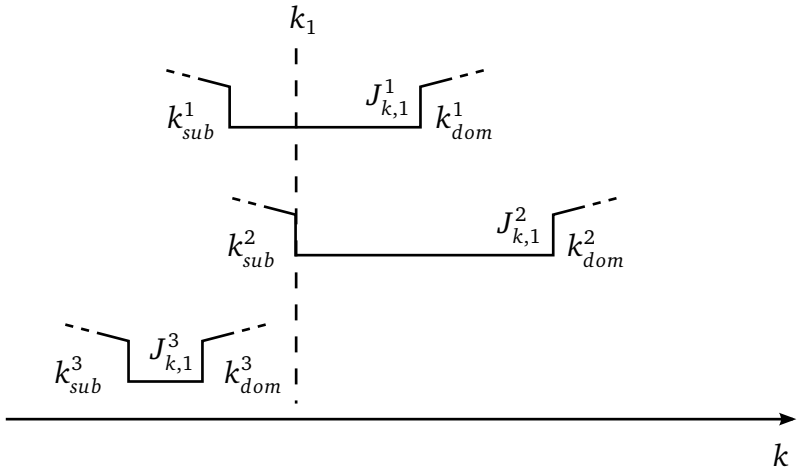


Figure 5.2: The magnitude of the masking stimulus conditions is identified using optimization techniques, illustrated on the example of the environment stiffness k_1 that feel neither subordinate nor dominating over the damping impression. Different cost functions $J^{1...3}$ are associated with individual participants.

to 10 N/m; for inertia, one turn was equivalent to a change of 1 kg. Subjects were asked to produce two different conditions: One, where the impression of stiffness or inertia was slightly subordinate to damping (k_{sub}, m_{sub}), and a second, where stiffness or inertia was the slightly more dominating stimulus (k_{dom}, m_{dom}).

Experimental Setup

The haptic environments were rendered on the 1 DoF linear haptic interface described in Appendix A.1, using a position-based admittance control scheme (see Section 2.2.2) with a sample rate of 1 kHz. Participants' ears were covered by EX-29 headphones playing pink noise to cancel out the sound from the haptic device, and their sight was blocked by eye-masks to eliminate visual cues.

5.2.1 Results and Stimulus Selection

The experimental conditions for the main experiment were determined from the recorded dataset as follows: In principle, every value of stiffness in the interval $\left[k_{sub}, k_{dom}\right]$ is perceived as approximately equally dominating by the individual subject. Generally, it is not possible to find one value k_1 that is within $\left[k_{sub}, k_{dom}\right]$ for all subjects due to the large between-subject variance. Therefore, the number of participants that felt k_1 as equally dominating as damping is to be maximized. In addition, as this solution may not be unique, we minimize the mean distance between k_1 and the individual intervals $\left[k_{sub}, k_{dom}\right]$. The corresponding optimization problem can be written as

$$\arg\min_{k_1} \sum_{i=1}^{n} \sum_{j=1}^{2} J^i_{k,j}$$

with

$$J^i_{k,j} = \begin{cases} 0 & \text{if } k_1 \in \left[k^i_{sub,j},\ \ k^i_{dom,j}\right] \\ c_{step} + k_1 - k^i_{dom,j} & \text{if } k_1 > k^i_{dom,j} \\ c_{step} + k^i_{sub,j} - k_1 & \text{if } k_1 < k^i_{sub,j} \end{cases}$$

is a cost function for determining the stiffness value k_1 for participant i, given the damping level d_j and n denotes the number of participants. Subsequently, $k^i_{sub,j}$ and $k^i_{dom,j}$ are the subordinate and dominant stiffness value that is chosen by the participant, and c_{step} is a constant of large value to penalize solutions where $k_1 < k^i_{sub,j}$ or $k_1 > k^i_{dom,j}$. By setting c_{step} appropriately, it can be guaranteed that the optimization goal of choosing a value for k_1 which is between k_{sub} and k_{dom} is prioritized over minimizing the distance of the "outliers" in all cases. For our purpose, we identify $c_{step} = 1000$ as appropriate value. An analogous problem can be formulated to determine the inertia m_1. Examples of perceptual intervals and the corresponding cost functions are depicted in Fig 5.2. Using this procedure, a stiffness $k_1 = 19.0$ N/m and an inertia $m_1 = 2.8$ kg were determined.

For k_2 and m_2 the maximum value for k_{dom} and m_{dom} that were reported by the participants was taken. One dataset was excluded for k_2 as an outlier because it was outside a band of two standard deviations around the mean. Following this procedure, $k_2 = 42.4$ N/m and $m_2 = 5.3$ kg were determined.

5.3 Damping Discrimination Experiment

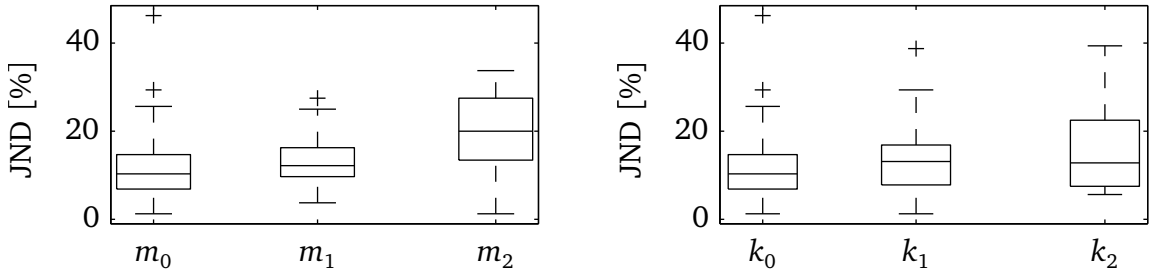
In the second experiment, discriminable differences of damping in haptic environments that consist of damping along with inertia or stiffness is measured. The aim of this design is to identify potential masking properties of those distractive stimuli. In total, 10 conditions are tested: JNDs for the damping parameters $d_{1,2}$ alone and with masking stimuli of either $m_{1,2}$ or $k_{1,2}$. To assure stability of the low-level position-based admittance controller, a minimum inertia m_0 of 0.5kg is always present.

Participants

The experiment was performed by 8 paid subjects from different disciplines with a mean age of 28.5 years. All of them gave their informed consent before participation. Two of them had experience with haptic devices, three were female and all were right-handed. Before starting the main experiment, all participants were familiarized with the stimuli and procedure: pure inertia, damping, stiffness, and combinations of damping together with inertia and stiffness were presented, each followed by information about the specific environment.

Procedure

In each experimental trial, the control condition and a stimulus condition were presented for 4 seconds each. The order of the two conditions was randomized. The two conditions were separated by a one-second break during which the device moved to the initial position. The



(a) (b)

Figure 5.3: Damping JNDs differ depending on (a) inertia and (b) stiffness.

participant was notified of the break by an audibly beep replayed over the headphones. The subjects had to decide which condition was more damped. In addition, they were allowed to respond that they did not know. Based on the participants' answer, the stimulus was modified after each trial using an adaptive-staircase threshold estimation procedure [60], discussed in Section 2.4.1. The adaptive estimation of the damping JND for one control condition took 20-30 trials. All 10 control conditions were presented twice, once starting the staircase from below and once from above the control condition. The order of control conditions was fully intermixed. Five staircase procedures were combined into one experimental session which was completed in about 30 minutes. After three JND estimation procedures, a break of five minutes was inserted. In total, four experimental sessions were performed, separated by a break of at least 30 minutes to avoid fatigue. The experimental setup was the same as described in Section 5.2.

After successful completion of all experimental conditions, participants were required to answer five questions:

1. How did you judge about the environments? By differences in movement or differences in force?

2. Was there a sensory feature on which you based your judgment in the conditions with inertia?

3. Was there a sensory feature on which you based your judgment in the conditions with stiffness?

4. Please describe how damping felt to you.

5. Was the judgment easier in the stiffness conditions or the inertia conditions?

5.4 Results

Percentual damping JNDs (relative to the control conditions) for different masking stimuli are depicted in Fig. 5.3. In order to determine effects of both types of masking stimuli, two 3-factor, repeated-measures ANOVAs for $r \times d \times m$ and $r \times d \times k$ were performed, where r is the

repetition (1,2). The main effect of inertia m shows a significant influence ($F(2,14) = 5.37$, $p < .05$, $\eta_p = .43$) as does stiffness k ($F(2,14) = 6.26$, $p < .05$, $\eta_p = .47$). The main effect of d, Greenhouse-Geisser corrected for sphericity, is not significant in either ANOVA ($F(1.00,7.00) = 1.79$, $p = .22$ and $F(1.00,7.00) = 1.44$, $p = .27$). The effect of repetitions, also Greenhouse-Geisser corrected, is insignificant as well ($F(1.00,7.00) = 0.66$, $p = .45$ and $F(1.00,7.00) = 1.74$, $p = .23$). No interactions were found in either ANOVA.

Polynomial trend analyses showed significant linear trends relating the damping JND to both inertia and stiffness while quadratic trends were not significant. For inertia, 85% of the main effect was accounted for by the linear trend ($SS_m = 1017.9$, $SS_{m,\text{linear}} = 862.9$, $p < .05$)[1]. For stiffness, 99% of the main effect was accounted for by linearity ($SS_k = 239.5$, $SS_{k,\text{linear}} = 236.9.9$, $p < .05$).

Remark 5.2. As the different levels of inertia and stiffness are only approximately equally spaced, the linear trend is approximate. The significant linear trend indicates at least a monotonic relationship between damping JND and value of m/k. As linear regressions fit to the exact values of m and k only minimally improved the fits, the noise introduced by unequal spacing appears to be minimal.

The results indicate that damping JNDs depend on the overall composition of the environment. In the cases where inertia or stiffness dominate the perception of the environment, the discrimination of damping apparently becomes harder.

5.4.1 Analysis of Questionnaire Data

Only two participants reported that they were perceiving differences in movement, all others felt as if they attended to differences in the force feedback. In conditions where inertia was added as masking stimulus, all of the participants reported that they paid attention to the higher force (or resistance) they needed to move the higher-damped environment. In addition, three participants concretized that they only attended to the force when they felt that the velocity was constant. For the stiffness conditions, five subjects reported similar properties as for the inertia condition. Two participants responded that they were paying more attention to the movement where the stiffness is pushing into the moving direction, one was paying most attention to the onset force when he/she started to move. The feeling of damping was specified as “force-demanding”, “breaking” and a “resistance” or “viscous fluid”. All participants used terms expressing that damping was hindering them of performing their interaction unopposed. Only two participants felt it easier to discriminate damping that was masked by stiffness rather than inertia, although the JND for d masked by m_2 was the largest of all conditions.

5.5 Towards a Model for Dynamic Haptic Masking

Quantitative knowledge about the performance of humans perceiving damping differences during exploration of a haptic environment containing damping and stiffness, or damping and inertia is of great practical relevance. Thus, a quantitative model of the masking effects

[1] SS stands for *Type III Sum of Squares*.

observed in this experiment is formulated, founded on the dynamic perception and task performance framework introduced in Chapter 3. The perception of mechanical damping d is understood as a mapping

$$d \mapsto \mathbf{y}_{perc} \tag{5.1}$$

where the damping $d \in \Theta_{phy}$ is from the set of physical features and $\mathbf{y}_{perc} \in$ {"difference perceived" "no difference perceived" "I don't know"}. With respect to the dynamic perception framework (3.18)-(3.20), this mapping depends on the closed-loop dynamics

$$\dot{\mathbf{x}}_{\circlearrowleft}(t) = \psi_{\circlearrowleft}(\mathbf{x}_{\circlearrowleft}(t), \theta_{\circlearrowleft}(t), \mathbf{u}_{ne}(t), \xi_{\circlearrowleft}(t), t), \quad \mathbf{x}_{\circlearrowleft}(0) = \mathbf{x}_{\circlearrowleft,0} \tag{5.2}$$

where $\mathbf{x}_{\circlearrowleft}(t) \in X_{\circlearrowleft}$ contains all states of the closed sensorimotor loop, including the physical states of the environment $\mathbf{x}_{phy}(t) \in X_{phy} \subset X_{\circlearrowleft}$. Similarly, the physical features, including the environment parameters inertia, damping, and stiffness are contained in the environment feature vector $\theta_{phy}(t) \in \Theta_{phy} \subset \Theta_{\circlearrowleft}$. Noise terms from, e.g., sensory, neural, cognitive, and motor processes, are summarized in the noise vector $\xi_{\circlearrowleft} \in \chi_{\circlearrowleft}$.

For the development of a concrete perceptual model, we discuss some assumptions that help in reducing the model complexity.

Assumption 5.1. The cognitive mechanisms controlling the motor actions and the perception of environment movement and feedback force have accurate knowledge about the own body configuration and dynamics.

By assuming accurate body knowledge, the control loop between the human operator and the environment can be modeled in the coordinate system of the physical environment, neglecting nonlinear transformations from each involved muscle to the interaction point and back. On the control side, an inverse kinematic transformation is needed to calculate the joint angles and muscle forces that allow reaching a certain location and exerting a specific force. It is reasonable to assume that such a transformation is available to the human, as body control to achieve a specific task in the world coordinate system is achieved without a significant conscious effort. On the perceptual side, a forward kinematic transformation is necessary to determine the position and force at the interaction point, based on the position (muscle length) and force of the motor system. In return, the inverse kinematics are needed when interactive forces, introduced by physical interaction with an environment shall be determined based on the muscle tension of a group of muscles attached to the interacting limb. The existence of internal representations of body dynamics in sensorimotor control is widely accepted [84, 96, 98, 111, 122, 123].

Assumption 5.2. The time constants in the sensory and neural system are small in comparison to the mechanical time constants of body and environment.

We show the validity of this assumption by reconsidering that the motor system is able to operate in a range of about 10 to 30 Hz [6, 7], but the somatosensory system can recognize vibrations and force changes up to a frequency of 1000 Hz [7], making perception much faster in terms of response characteristics. The neural pathways simplify to a noisy transmission channel in the case of neglecting their dynamics. As an additional simplification,

we neglect the influence of attention and all other information sources apart from haptics, and furthermore assume all dynamic processes to be time-invariant. These assumptions are warrantable by our careful experimental design.

Considering assumptions 5.1 and 5.2, the state vector in the closed-loop dynamics (5.2), normally considering states in the physical, sensory, neural, control, and motor sensorimotor domain (see Section 3.1) simplifies to

$$\mathbf{x}_{\circlearrowleft}(t) = \begin{bmatrix} \mathbf{x}_{phy}(t) & \mathbf{x}_{mo}(t) & \mathbf{x}_{con}(t) \end{bmatrix}^T \tag{5.3}$$

with the physical $\mathbf{x}_{phy}(t)$, state motor state $\mathbf{x}_{mo}(t)$, and controller state $\mathbf{x}_{con}(t)$. The state variable describing the physical world is of specific importance for the development of the dynamic masking model and consists of the limb position and velocity, $\mathbf{x}_{phy}(t) = \begin{bmatrix} x_h(t) \, \dot{x}_h(t) \end{bmatrix}^T$. The information that is available to the perceptual system $\mathbf{u}_{perc}(t)$ is based on the quantities the sensory system is capable to capture. Within the sensorimotor framework, we propose

$$\begin{aligned} \mathbf{u}_{perc}(t) &= \nu_{\circlearrowleft,perc}(\mathbf{x}_{\circlearrowleft}(t), \theta_{\circlearrowleft}(t), \mathbf{u}_{ne}(t), \xi_{\circlearrowleft}(t)) \\ &= \begin{bmatrix} x_h^{perc}(t) & \dot{x}_h^{perc}(t) & f_h^{perc}(t) \end{bmatrix}^T \in \mathbb{R}^3. \end{aligned} \tag{5.4}$$

The perceptual input quantities $x_h^{perc}(t)$, $\dot{x}_h^{perc}(t)$ and $f_h^{perc}(t)$ are denoted as perceptual estimates of the (unidirectional) limb position $x_h(t)$ and velocity $\dot{x}_h(t)$, and the interaction force $f_h(t)$, respectively. Due to sensory, neural, and motor noise, $\mathbf{u}_{perc}(t)$ is afflicted with uncertainty. Apart from that, we assume

$$E\{\mathbf{u}_{perc}(t)\} \sim \begin{bmatrix} x_h(t) & \dot{x}_h(t) & f_h(t) \end{bmatrix}^T, \tag{5.5}$$

where $E\{\cdot\}$ is the expected value operator.

Remark 5.3. We do not explicitly consider the time delays in the neural transmission of information here, due to the fact that they are hardly measurable and can thus not be taken into consideration for the development of quantitative perceptual models. Instead, it is assumed that the human has sufficient experience with the control of his/her body and the sensory observation of it that he/she can compensate internal delay processes and ensure temporal consistency.

5.5.1 Model Candidates

The analysis of the questionnaire data in Section 5.4.1 shows that the majority of participants base their perceptual feedback on force instead of movement. Therefore, models of damping discrimination based on interaction force as decision variable are considered here. An additional reason for this choice is that the matched filter model introduced in Chapter 4 explains time delay discrimination performance with differences in force feedback reasonably well while requiring less tuning parameters than, e.g., the state observer model.

In order to detect differences in the physical damping of the environment based on force feedback, the perceptual system must have access to an estimate of some kind of force reference to which the actual sensory signal can be related to. This reference can be produced using an internal representation of body and environment dynamics.

Assumption 5.3. The human haptic perceptual system can store dynamic descriptions of his/her body and the environment he/she interacts with and use them for the generation of reference signals.

The justification for this assumption is given in Section 4.2. With the procedure used in the experiments for damping discrimination in mind where two stimuli s_1 and s_2 are sequentially explored and compared, a perception model based on differences in force feedback can be formulated as a sensorimotor observation

$$\mathbf{y}_{perc}(t) = \phi_{perc}\left(\dot{\mathbf{y}}_{perc}(t), \ldots, \mathbf{x}_{perc}(t), \dot{\mathbf{x}}_{perc}(t), \ldots, \theta_{perc}(t), \mathbf{u}_{perc}(t), \dot{\mathbf{u}}_{perc}(t), \ldots, \xi_{perc}(t), t\right).$$

Without loss of generality, we assume from now on that s_1 starts at time $t = 0$ and contains the stimulus with smaller damping. The decision which stimulus contains the larger damping is made at a time instance t_{resp} during the exploration of s_2. A perceptual model for damping discrimination is proposed as

$$y_{perc}(t_{resp}) = \begin{cases} \text{“}s_1 \text{ greater”} & \text{if } \exists 0 \le t \le t_{resp} : \delta(f_h^{perc}(t), \hat{f}_h(t), \theta_{perc}, t_{resp}) < \epsilon_1, \\ \text{“don't know”} & \text{if } \forall 0 \le t \le t_{resp} : \epsilon_1 \le \delta(f_h^{perc}(t), \hat{f}_h(t), \theta_{perc}, t_{resp}) \le \epsilon_2, \\ \text{“}s_2 \text{ greater”} & \text{if } \exists 0 \le t \le t_{resp} : \delta(f_h^{perc}(t), \hat{f}_h(t), \theta_{perc}, t_{resp}) > \epsilon_2, \end{cases} \tag{5.6}$$

where $\delta(\cdot) : \mathbb{R} \to \mathbb{R}$, $\epsilon_1, \epsilon_2 \in \mathbb{R}$, $\epsilon_1 < \epsilon_2$ and

$$\hat{f}_h(t) = \hat{m}\ddot{x}_h^{perc}(t) + \hat{d}\dot{x}_h^{perc}(t) + \hat{k}x_h^{perc}(t), \tag{5.7}$$

is an estimate of the environment force that could be perceived if the stimulus condition s_1 with the perceived environment inertia $\hat{m}$, damping $\hat{d}$, and stiffness $\hat{k}$ was presented instead of s_2. The function $\delta(f_h^{perc}(t), \hat{f}_h(t), \theta_{perc}, t_{resp})$ extracts a criterion from the perceptual input $f_h^{perc}(t)$ and $\hat{f}_h(t)$ with the help of perception parameters θ_{perc}. Specific to verbal output, the value of y_{perc} is only available at the response time t_{resp}. The parameters ϵ_1 and ϵ_2 are thresholds that cause the decision for one of three response alternatives. Without loss of generality, we assume that the thresholds are symmetric around 0, thus $\epsilon_2 = -\epsilon_1 = \epsilon$.

Remark 5.4. Sensory feedback from the receptors within the muscles presumably contribute significantly to the perception of environment force. The alternative sensory feedback f_m which is the force equivalent of all contributing muscles that lead to a movement in the environment could instead be taken into consideration to formulate a force estimate with the environment from stimulus s_1 similar to (5.7) as

$$\hat{f}_m(t) = (m_{body} + \hat{m})\ddot{x}_h^{perc}(t) + (d_{body} + \hat{d})\dot{x}_h^{perc}(t) + (k_{body} + \hat{k})x_h^{perc}(t). \tag{5.8}$$

The body is approximated as a mass-spring-damper system with parameters m_{body}, d_{body}, and k_{body}, as commonly done in the literature [49, 135, 136]. The tissue dynamics are neglected, assuming a stiff coupling between body and the haptic device (e.g., by a tight grip). Based on Assumption 5.1 that humans are aware of their body dynamics by sufficient experience in everyday-life situations, the additional terms in (5.8) are supposed to be internally compensated, thus (5.7) and (5.8) can be considered to make identical predictions.

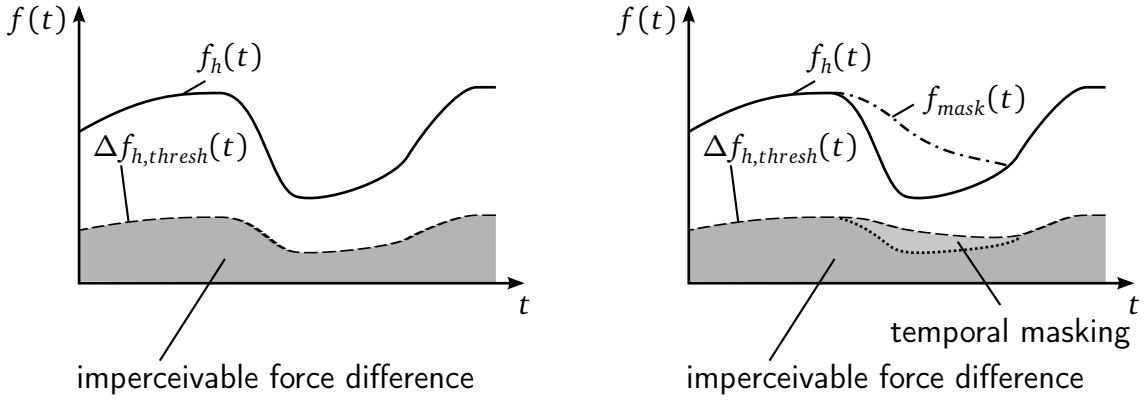


Figure 5.4: Masking effects based on (a) simultaneous masking, and (b) a combination of simultaneous and temporal masking.

Force Threshold Model

The simplest mechanism for the decision criterion $\delta(\cdot)$ in the perception model (5.6) is based on the deviation between the sensory feedback force $f_h^{perc}(t)$ and the internal estimate $\hat{f}_h(t)$, thus

$$\delta_{thresh}(f_h^{perc}(t), \hat{f}_h(t)) = \Delta f_h(t) = f_h^{perc}(t) - \hat{f}_h(t) \tag{5.9}$$

and

$$\epsilon_{thresh,1} = -\Delta f_{h,thresh} = \text{ const.}, \quad \epsilon_{thresh,2} = \Delta f_{h,thresh} = \text{ const.}. \tag{5.10}$$

This model does not explicitly consider a masking mechanism that is due to additional environment components, such as stiffness, or inertia. However, masking effects could be implicit due to, e.g., a change in the movement $x_h(t)$, $\dot{x}_h(t)$ which itself determines the environment force feedback $f_h(t)$, and the estimate thereof, $\hat{f}_h(t)$.

Alternatives to this model candidate are motivated by known masking effects in the auditory domain and the model structure of a diffusion process.

Simultaneous and Temporal Masking

A masking phenomenon similar to the simultaneous masking effect in auditory research described in Section 5.1 for force discrimination is Weber's law. It states that the just noticeable force difference depends on the magnitude of the environmental force feedback. With respect to the time-varying nature of the force profile and considering the fact that it can cross zero, a correction for Weber's law at low intensities is added, by increasing the denominator by a quantity f_{sat} [57]

$$\Delta f_{h,thresh}(t) = W_f(|f_h^{perc}(t)| + f_{sat}) \tag{5.11}$$

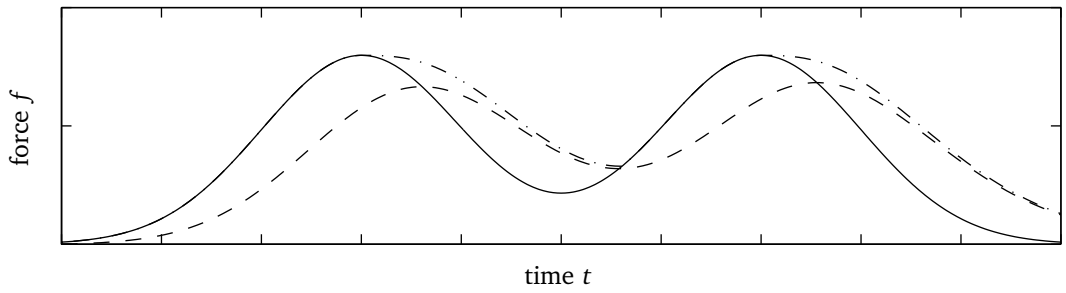


Figure 5.5: For a model of simultaneous and temporal masking, a nonlinear filter algorithm is applied to the masking stimulus f_h (solid), extending f_h in the temporal domain after a peak occurs (dash-dotted). In the case of a lowpass filtering procedure as in [1], the simultaneous masking effect is decreased in an undesirable way (dashed).

with $W_f = \text{const.}$ the Weber coefficient for force. In this way, the masking stimulus is the force magnitude, and the target stimulus is the force difference. An illustration of the simultaneous masking effect based on Weber's law is depicted in Figure 5.4a. An integration of (5.11) into the perception model (5.6) as a second candidate can be formulated as

$$\delta_{sim}(f_h^{perc}(t), \hat{f}_h(t), \theta_{perc}) = \frac{f_h^{perc}(t) - \hat{f}_h(t)}{|f_h^{perc}(t)| + f_{sat}}, \tag{5.12}$$

where $f_{sat} \in \theta_{perc}$, and

$$\epsilon_{sim,1} = -W_f = \text{const.}, \quad \epsilon_{sim,2} = W_f = \text{const.}.$$

Temporal masking phenomena have not been investigated for the haptic modality so far. In principle, the masking stimulus is stretched in time such that the simultaneous masking effect for the perception of a difference in force feedback is extended to times before, and after the time instance of a high masking stimulus intensity. We propose a temporal masking model on the basis of a first-order lowpass filter, related to a model that can capture auditory forward masking reasonably well [1]. For the moment, only forward masking will be considered. The lowpass filter causes the magnitude of a high masking stimulus to decay slower, which is a desired behavior when modeling temporal masking. On the other hand, it counteracts the effect of simultaneous masking as the masking stimulus' amplitude is decreased, as depicted in Figure 5.5. Thus, a nonlinear filter algorithm for forward masking is suggested as

$$f_{mask}(t) = |f_h^{perc}(t)| - \min(0, T_{mask}\dot{f}_{mask}(t)), \tag{5.13}$$

realizing a temporal masking behavior without affecting the simultaneous masking effect of Weber's law. The effect of (5.13) on $f_h(t)$ is depicted in Figure 5.5. With $f_{mask}(t)$, the perceptual threshold for force differences in equation (5.11) can be further extended to

$$\delta_{temp}(f_h^{perc}(t), \hat{f}_h(t)) = \frac{f_h^{perc}(t) - \hat{f}_h(t)}{f_{mask}(T_{mask}, t) + f_{sat}(t)}, \tag{5.14}$$

where $f_{sat}, T_{mask} \in \theta_{perc}$, and

$$\epsilon_{temp,1} = -W_f = \text{ const.}, \quad \epsilon_{temp,2} = W_f = \text{ const.},$$

and serve as a third model candidate to capture human's damping discrimination performance. The combination of temporal masking and simultaneous masking is illustrated in Figure 5.4b.

Accumulation Model Structure

Perceptual models in the form of diffusion processes successfully describe humans' performance in detection and discrimination tasks in other modalities than the haptic one [67–69], as discussed in Section 2.4.3. The structure of these models is divided into three stages: an information encoding function, followed by an accumulation stage and the decision process, as depicted in Figure 2.7. The perception criteria in (5.9), (5.11) and (5.14) inhibit an information extraction function $\delta(f_h^{perc}(t), \hat{f}_h(t), \theta_{perc})$ and a decision process as formulated in in (5.6) but lack the information accumulation stage. Three additional candidates for $\delta(\cdot)$ in (5.6) are proposed, based upon (5.9), (5.11), and (5.14) but completed with an accumulation stage as

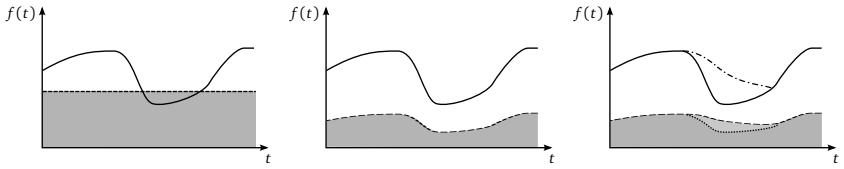
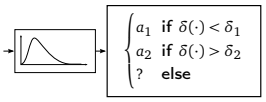
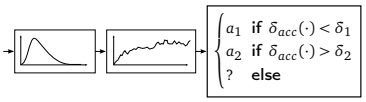
$$\delta_{acc,thresh}(f_h^{perc}(t), \hat{f}_h(t), \theta_{perc}, t_{resp}) = \int_0^{t_{resp}} \text{sign}\left(\dot{x}_h^{perc}(t)\right)\left(f_h^{perc}(t) - \hat{f}_h(t)\right) \text{dt},$$
$$\epsilon_{acc,thresh,1} = -\Delta f_{acc,thresh} = \text{ const.}, \quad \epsilon_{acc,thresh,2} = \Delta f_{acc,thresh} = \text{ const.}, \tag{5.15}$$

$$\delta_{acc,sim}(f_h^{perc}(t), \hat{f}_h(t), \theta_{perc}, t_{resp}) = \int_0^{t_{resp}} \frac{\text{sign}\left(\dot{x}_h^{perc}(t)\right)\left(f_h^{perc}(t) - \hat{f}_h(t)\right)}{|f_h^{perc}(t)| + f_{sat}(t)} \text{dt},$$
$$\epsilon_{acc,sim,1} = -W_{acc,f} = \text{ const.}, \quad \epsilon_{acc,sim,2} = W_{acc,f} = \text{ const.}, \tag{5.16}$$

$$\delta_{acc,temp}(f_h^{perc}(t), \hat{f}_h(t), \theta_{perc}, t_{resp}) = \int_0^{t_{resp}} \frac{\text{sign}\left(\dot{x}_h^{perc}(t)\right)\left(f_h^{perc}(t) - \hat{f}_h(t)\right)}{f_{mask}(t) + f_{sat}(t)} \text{dt},$$
$$\epsilon_{acc,temp,1} = -W_{acc,f} = \text{ const.}, \quad \epsilon_{acc,temp,2} = W_{acc,f} = \text{ const..} \tag{5.17}$$

All six model candidates for human damping discrimination performance are summarized in Table 5.1.

Remark 5.5. The reason for considering the term $\text{sign}\left(\dot{x}_h^{perc}(t)\right)$ in equations (5.15)-(5.17) is the following: Consider the case that the sensory feedback $f_h^{perc}(t)$ comes from a haptic environment with larger damping than the previously explored and now simulated force reference $\hat{f}_h(t)$, $d > \hat{d}$. In this case, the difference $f_h^{perc}(t) - \hat{f}_h(t) = (d - \hat{d})\dot{x}_h^{perc}(t)$ has the sign of $\dot{x}_h^{perc}(t)$. Movements with positive and negative velocities thus cancel out when taking the integral, in the extreme case it is zero in case the movement starts and stops at the same location. A multiplication with $\text{sign}\left(\dot{x}_h^{perc}(t)\right)$ results in a positive values for the integrand in (5.15)-(5.17), regardless of the movement direction. Similarly, in the case $d < \hat{d}$, the integrand is all negative.

			
	$\delta_{thresh}(\cdot)$	$\delta_{sim}(\cdot)$	$\delta_{temp}(\cdot)$
	$\delta_{acc,thresh}(\cdot)$	$\delta_{acc,sim}(\cdot)$	$\delta_{acc,temp}(\cdot)$

Tabular 5.1: Overview table of model candidates for human damping discrimination.

5.6 Model Predictions

With the introduction of dynamic perception model candidates on the basis of the dynamic modeling framework introduced in Chapter 3, the problem of determining a damping threshold is transformed into a threshold estimation on the basis of values for $\delta(\cdot)$, as summarized in Table 5.1. Because $\delta(\cdot)$ depends not only on d, but on the perceived force feedback, and the interaction movement which changes from trial to trial, the transformation does not preserve the shape of a staircase, as can be seen from Figure 5.6. Thus, the threshold value can not be estimated solely from the last four trials of the staircase which is the standard procedure for this experimental method, as discussed in Section 2.4.1, but must be determined with respect to the complete dataset. As a consequence, the optimal threshold value which divides the perceptual responses into "s_1/s_2 greater" and "don't know" must be found using an optimization procedure. In addition to those variables, the model parameters f_{sat} and T_{mask} are determined using a model identification procedure.

5.6.1 Parameter Identification

A perfect perception model predicts the actually given perceptual response whether s_1 or s_2 contains more damping, given the respective damping parameters, and the interaction movement $x_h(t)$, $\dot{x}_h(t)$ during the exploration of s_2 where the judgment is made. Finding the optimal parametrization yielding to the best prediction performance for the perception models summarized in Table 5.1 thus corresponds to minimizing the classification error rate

$$\underset{\theta_{perc},\epsilon_1,\epsilon_2}{\arg\min}\ r_{mis}, \tag{5.18}$$

where

$$r_{mis} = \frac{N\left(y_{perc}(t_{resp}) = \text{“}s_1 \text{ greater”} \mid \delta(\cdot) > \epsilon_1\right) + N\left(y_{perc}(t_{resp}) = \text{“}s_2 \text{ greater”} \mid \delta(\cdot) < \epsilon_2\right)}{N\left(y_{perc}(t_{resp}) = \text{“}s_1 \text{ greater”}\ \right) + N\left(y_{perc}(t_{resp}) = \text{“}s_2 \text{ greater”}\ \right)}. \tag{5.19}$$

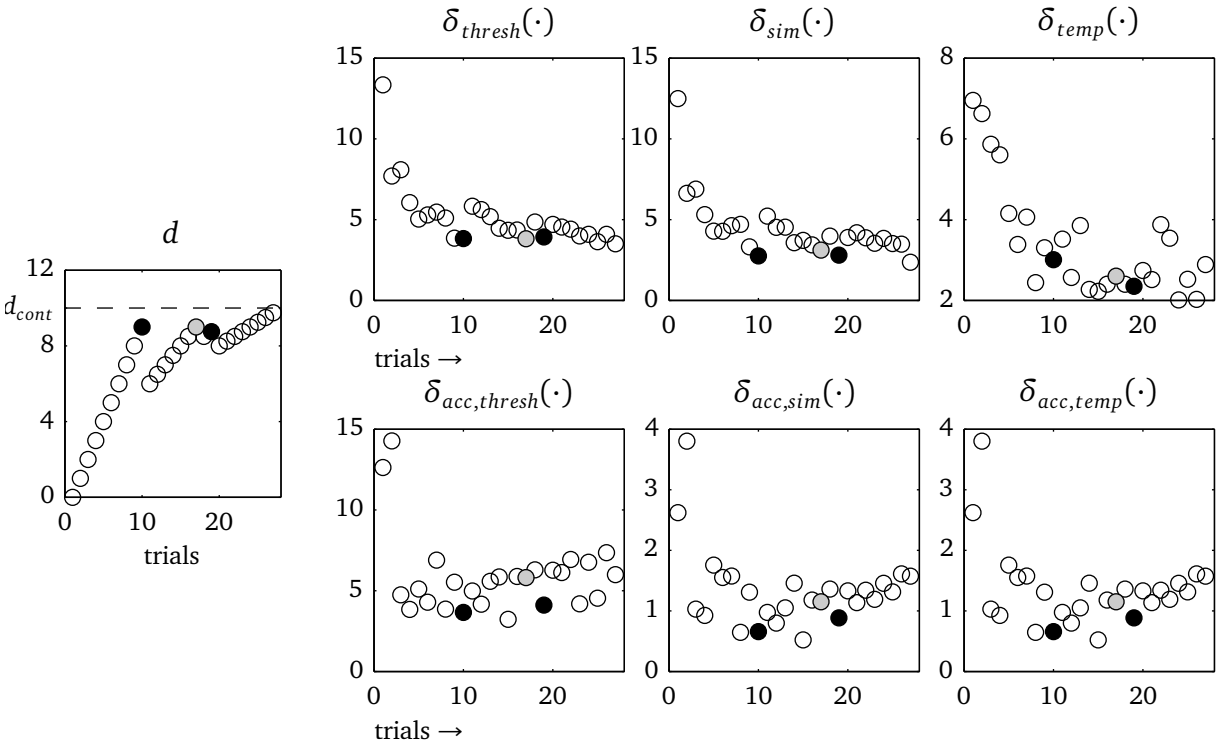


Figure 5.6: The estimation of a perceptual threshold for damping utilizing a staircase estimation procedure in (a) is transformed into a threshold estimation problem on the basis of different judgment criteria summarized in (b). Correct answers are marked as empty circles, "don't know" answers gray, wrong answers black.

Here, $N\left(y_{perc}(t_{resp}) = \text{“}s_1 \text{ greater”} \,|\delta(\cdot) > \epsilon_1\right) \in \mathbb{N}^0$ denotes the number of misclassified trials where the actual response is "s_1 contained more damping" given the decision criterion $\delta(\cdot)$ is above the threshold, thus predicting a "don't know" or "s_2 contained more damping" answer. The other terms in numerator and denominator are defined accordingly. Directly taking (5.19) as cost function in the minimization is infeasible for most optimization algorithms as it can only attain discrete values and thus, no gradient can be determined. Instead, an optimization criterion is based on the squared error from misclassified trials ($\delta(\cdot) > \epsilon_1$ if the answer was "s_1 greater" or $\delta(\cdot) < \epsilon_2$ if the answer was "s_2 greater"). The parameters for the perception model candidates in Table 5.1 are thus calculated by solving

$$\underset{\theta_{perc},\epsilon_1,\epsilon_2}{\arg\min} \sum_{i\in\Omega_{N_1}} \left(\delta^i(\cdot) - \epsilon_1\right)^2 + \sum_{i\in\Omega_{N_2}} \left(\delta^i(\cdot) - \epsilon_2\right)^2 \tag{5.20}$$

where

$$\Omega_{N_1} = \left\{y_{perc}(t_{resp}) = \text{“}s_1 \text{ greater”} \,|\delta(\cdot) > \epsilon_1\right\}$$
$$\Omega_{N_2} = \left\{y_{perc}(t_{resp}) = \text{“}s_2 \text{ greater”} \,|\delta(\cdot) < \epsilon_2\right\}.$$

The notation $\delta^i(\cdot)$ and $y^i_{perc}(t_{resp})$ denotes the values of $\delta(\cdot)$ and $y_{perc}(t_{resp})$ in trial i, respectively. Although (5.18) and (5.20) are not equivalent, the continuous squared distance inhibits similar behavior: Both optimization functions have a lower bound at 0, and their value increases with the number of misclassified trials. However, the distance metric in (5.20) is sensitive to large misclassification errors which can lead to an estimate of ϵ resulting in a higher misclassification rate than the optimum with respect to (5.19). Due to the computational advantages of (5.20), the identification of model parameters θ_{perc} and ϵ are taken as an approximation for the globally optimal solution of problem (5.18).

Leave-one-out Crossvalidation

If all data were taken into consideration to estimate θ_{perc} and ϵ, it would be unclear how well the identified model can generalize to new trials, not contained in the dataset. This problem is well-known in machine learning techniques as overfitting, and cross-validation techniques have been developed to investigate the generalization capabilities of a learned model [150]. The available dataset is divided into training data and validation data. While training data is used to find the optimal values for the model parameters and threshold, the validation data is taken to assess the value of the optimization criterion, in our case the squared distance of misclassified samples, calculated by the objective function in (5.20). In the case of a high squared error using the validation data with the model parametrization identified with the training data, the generalization ability of the model is poor. If the squared error with training and validation data are approximately equal, the model can generalize well. In our case, the dataset is limited to relatively few trials within one condition, thus we apply a leave-one-out crossvalidation technique here. With this method, the validation data consists of only one sample, and the model is trained with all remaining trials. This procedure is repeated with every sample being the validation data once. The resulting value for the objective function (5.20) is the mean over all validation rounds.

5.6.2 Results

Similar to the experimental evaluation in Section 5.3, individual thresholds for every condition are identified, applying the leave-one-out crossvalidation technique to the optimization problem described in (5.20). The model parameters f_{sat} and T_{mask} are though considered to be inherent to the participant, thus only one value per subject is optimized. The quality of the model fit is evaluated with respect to two metrics: The mean misclassification rate within the experimental conditions is a measure for the *internal validity* of the model and is depicted in Figure 5.7. Two 2-way ANOVAs found no significant effect of the decision criterion $\delta_i(\cdot)$ where $i \in \{$*thresh; acc,thresh; sim; acc,sim; temp; acc,temp*$\}$ ($F(5,467) = 1.16$, $p > .05$ and $F(1,471) = 3,84$, $p > 0.05$, respectively). It is interesting to mention that the parameter identification for the temporal masking model with accumulation stage results in an optimal estimate of $T_{mask} = 0$ for all participants, and otherwise in the same threshold values and f_{sat} as for the simultaneous masking model with accumulation stage. The temporal masking model with $T_{mask} = 0$ thus simplifies to the simultaneous masking model. All model parameters and threshold estimates are summarized in Table B.1 in the Appendix.

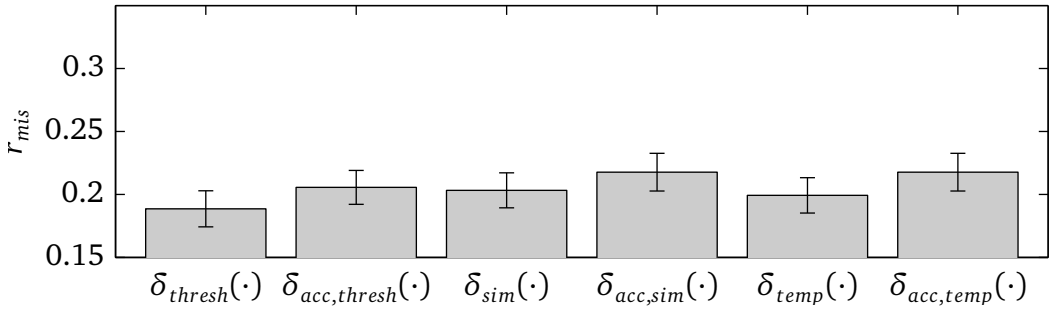


Figure 5.7: The mean misclassification rate within each experimental condition.

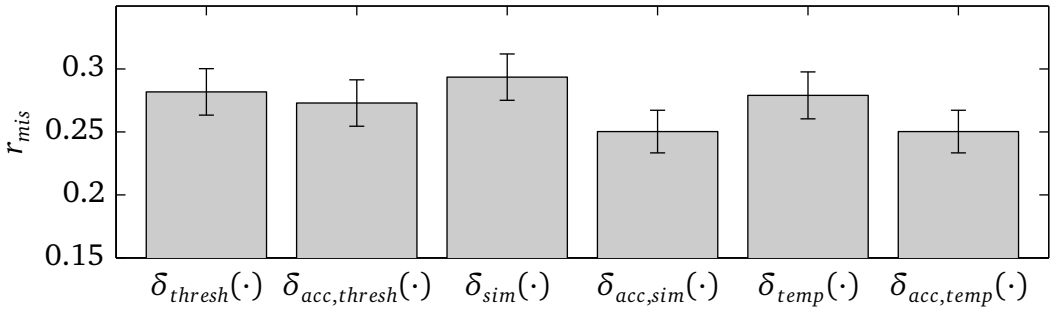


Figure 5.8: The mean misclassification rate between the experimental conditions is significantly lower in the model candidates including an accumulation stage.

Thresholds models in the form of (5.6) which can be used in a human-centered design, optimization, and evaluation of technical systems must not only explain single conditions with individual threshold estimates, but must generalize over multiple conditions, such as multiple masking stimuli. Thus, we test the different model candidates for their *external validity*, capturing the ability to predict the trial data of other conditions, given the threshold of one condition. Similar to the internal validity estimation we apply a crossvalidation technique, where the set of all trials is divided into training set and validation set. Taking the training data of one experimental condition results in one threshold criterion estimate ϵ as discussed above. The classification error with all remaining trials from different conditions is a measure for the external validity. Crossvalidation is done for all threshold estimates, and the mean of all misclassification rates is reported in Figure 5.8. It can be concluded that the error rate between conditions is significantly higher than within one condition (paired t-test, $p \approx 0$). The classification error rate is found to be significantly higher with damping masked by k_1 and m_1 as in those conditions where k_2 and m_2 are used as masking stimuli (paired t-test, $p < .01$). Furthermore, perception models including an accumulation stage result in a significantly lower classification error rate compared to those without information accumulation (paired t-test, $p < 0.05$). Overall, the simultaneous masking model with information accumulation stage leads to the lowest classification error rate between conditions.

5.7 Discussion

Masking effects exist not only in the tactile and auditory modality but also in the perception of haptic environments. An addition of inertia or stiffness to a damping environment results in a significant increase of damping JND. This finding further supports the conclusion of Chapter 4 that static mappings from environmental parameters to the perceptual response are insufficient for characterizing human perception of a physical feature. Furthermore, the results indicate that perceptual thresholds for individual environment components are of limited expressiveness when exploring a physical environment other than pure stiffness, damping, or inertia.

Among the most important results of the experiment presented here is the improved prediction accuracy in the cases where an accumulation stage is included in the model structure of computational damping discrimination models. The superior prediction quality can be seen as an indicator of the applicability of diffusion models to capture haptic phenomena as well.

The experimental conditions for all trials in the experiment for damping discrimination were determined using a staircase method. In this way, the number of correctly recognized conditions and false classifications is unequal, likewise is the number of conditions above the detection threshold and below. For increasing validity of identification results, a symmetric distribution of stimulus conditions is desirable. The impact of biased distributions in experimental conditions on the obtained results must be investigated in future research.

It was concluded from questionnaire data that participants control their limb movement and attend to the resulting interaction force for making a perceptual judgment. Alternative control strategies could also result in a perception mechanism which include the interaction force in their decision criterion. The influence of a prior motor plan [84], based on the dynamics of the internal environment model, and other hypothetical mechanisms such as internal forward models in sensorimotor control during the exploration [88] are yet to be examined.

5.7.1 Implications for Telepresence Systems

The current findings have major implications for a human-centered design, optimization, and evaluation of teleoperation systems and other technical systems. In [24] the concept of *perceived transparency* is introduced which extends the classical transparency evaluation in a human-oriented way by including perceptual limits. Whether or not communication time delay and control parameter changes in a two-channel teleoperation system are perceived by the human user has been predicted over their effects for the operator, exploring an isolated environment inertia, damping, and stiffness. Without considering masking effects in haptic environment perception, such predictions about the perceptual state are overly conservative if they are generalized to combinations of the basic building blocks. In a related way, perceptual models considering masking effects can also extend a new method for the analysis of four-channel teleoperation systems [35], which approximates the transparency error using a mass-spring-damper model.

The design of mechanisms with multiple predefined haptic properties and haptic rendering algorithms are other applications to which our results can be applied. When developing

a mechanism with large inertia or stiffness, a design-related deviation in damping from the ideal value may be tolerable since it is not perceivable. Similarly, haptic rendering algorithms cannot always reflect all environmental properties in sufficient accuracy. For stiff environments, e.g., a certain amount of damping is necessary to assure stability. Therefore, considering the obtained results, damping could be adjusted in a way that ensures stability while not being perceivable to the operator.

The quantitative dynamic perceptual models developed here could be easily integrated into an online-optimization of diverse haptic applications, such as model-mediated haptic telepresence [151, 152]. In this example, a simple model of the remote environment is identified on the teleoperator side, and model parameters instead of force and movement are transmitted over the communication channel. On the human-system interface side, a local model is parameterized with the identified parameters and rendered locally. In this way, instabilities due to time delay could be overcome, and in the case of a perfectly known environment, perfect transparency could be achieved. Changes in the environment must be transmitted to update the local haptic model. The novel dynamic model developed for environmental damping can help to determine which updates must be sent to the human-system interface, and which ones can be left out, as they are not perceivable by the human operator.

More generally, the finding that the perceptual process behind discriminating damping is better approximated by including an information accumulation stage has a large potential for enhancing a similar application: A perception-based haptic data reduction algorithm can drastically reduce the amount of data to be sent over the communication channel in a telepresence system. The algorithm discussed in [64, 105, 106] decides whether or not a packet containing haptic data should be transmitted, based on whether or not the difference to the previously transmitted packet is perceivable or not. In the past, this decision has been based on the Weber Law, applied to each sample. With the finding that accumulative processes reflect force-based difference perception better than instantaneous thresholds, this approach may be overly conservative. A further reduction of the transmitted data can be expected if short violations of the Weber threshold are tolerated, until the accumulated decision variable reaches the perception threshold. For a successful application in such advanced applications, further investigations into a potential accumulating behavior in other haptic perception phenomena are strongly encouraged.

5.8 Conclusions and Open Problems

In this chapter, masking effects influencing the discrimination ability of haptic environmental features which are embedded into an environmental context with unrelated and potentially disturbing features have been investigated. It could be shown that the JND of damping increases linearly with the magnitude of additional inertia and stiffness “masking” the damping impression. Modeling this effect by means of a dynamic perception model reveals that a model structure considering an accumulation of information performs significantly better than a structure without such behavior.

So far, only the JND for damping has been investigated. It is interesting to see whether the perceptual bounds for inertia and stiffness exhibit similar behavior or not. Furthermore,

a fully crossed experimental design with masking stimuli containing a inertia *and* stiffness component are an interesting open problem. The application of the experimental findings to a real transparency evaluation is left for future research.

6 Communication Quality Control Based on a Dynamic Task Performance Model

Summary. *Achieving a high task performance is among the main objectives in the design and control of telepresence systems. With current static task performance measures, an optimization of the system with respect to task performance can only aim for an average enhancement. On the methodological foundation of describing sensorimotor processes dynamically, this Chapter demonstrates*

- *the development of a dynamic task performance model based on stochastic reachable sets,*
- *its optimal consideration in a novel communication quality control algorithm,*
- *a significantly improved task performance attributed to the new technique.*

Telepresence systems are designed to serve a specific purpose, e.g., telesurgery systems for accomplishing a medical procedure [47] or telepresence systems in space for maintenance tasks [153]. Besides taking objectives such as achieving a high degree of presence or transparency into account for the design of such systems, one key target of telepresence systems is often to obtain a high task performance. Techniques for achieving this goal are until now limited to a static optimization of design and control, allowing to achieve an improved performance *on average*.

In this chapter, a novel dynamic task performance model for an abstracted visual-haptic telepresence task is developed, founded on stochastic reachable sets. On this basis a communication quality control algorithm is developed, regulating the transmission time delay between the human-system interface and a virtual teleoperator *online*. In this way the task performance can not only be optimized on average, but on a trial-by-trial basis, respecting the operator's system interactions. This is, to the knowledge of the author, the first time that such a technique has been applied to telepresence systems. Two evaluation studies considering a navigation task through a course of obstacles show a significantly improved task performance compared to the case without communication quality control, but at the same time reveal limitations associated with the choice of performance criterion.

Existing static task performance measures for visual-haptic teleoperation and their application in optimizing visual-haptic telepresence systems are discussed in Section 6.1. Methods for communication quality control and the development of a dynamic task performance model on the basis of the methodological framework in Chapter 3 are in the focus of Section 6.2. The development of an exemplary, task-specific performance model for a navigation task with time-varying haptic and visual feedback properties is described in Section 6.3. The

evaluation of the optimal quality control algorithm in two application examples with respect to performance enhancement and cost savings are discussed in Sections 6.5 and 6.6. Methods utilized for the modeling of task performance based on the theory of stochastic reachable sets are furthermore outlined in Appendix C.

6.1 Background

A human-centered design of technical system must take the perceptual capabilities into consideration, as discussed in Chapters 4 and 5, but must not forget that apparatus such as telepresence systems are usually designed to help humans in performing certain tasks. Task performance considerations and usability issues have been taken into consideration for the design of mechanisms, especially haptic human-system interfaces earlier [7, 13, 154, 155], and task performance is often used as an evaluation criterion of devices and algorithms [121, 156–158]. All of these optimization techniques influence either the design of the system *before* task execution or serve as a benchmark criterion *after* task completion. There are less approaches taking task performance into consideration for a system optimization *during* the actual execution of a specific task, e.g., by adapting control algorithms to the specific task situation. An online-adaptation has significant benefits compared to an a priori optimization: While optimizing parameters or the design of a system before task execution can only lead to an optimized average task performance over multiple tasks and operators, online-optimization can take the current task situation into account and adapt the system in an optimal way during every task execution.

An important factor affecting the task performance of a human operating a telepresence system is the quality of feedback information. Communication between human-system interface and teleoperator is of particular importance in this context. Real communication channels involving wireless transmission channels in air, space and underwater suffer from time delay and data loss [159]. The operator notices time delay by visually perceiving the feedback from his/her performed movements later. In the haptic modality, stabilizing techniques display the mechanical environment impedance and haptic events which occur distorted, e.g., during impacts [8, 30, 49]. Stabilization is nevertheless inevitable since a global control loop with the remote environment is closed via the haptic modality, where time delay decreases the stability margins of the overall telepresence system, and packet loss must be handled appropriately to avoid instability. Multiple studies report a consistently lower task performance in many different conditions with visual [37, 160, 161] and haptic [37, 104] time delay than without, even though different experimental setups, performance measures, and performed tasks were investigated. Lost packets are as well capable of affecting the operator of a telepresence system: In the visual modality, rather smooth movements can become unsmooth and jumpy. For haptics, the corruption of haptic feedback depends on the environment impedance, packet rate, and loss burst length [8]. The perception of visual-haptic stimuli under the influence of dropped packets [162] has been studied earlier. However, there are only very few studies investigating the impact of packet loss on task performance, not coming to a consistent conclusion. While there is a noticeable effect on movement time in certain cases where packets are lost due to time-varying

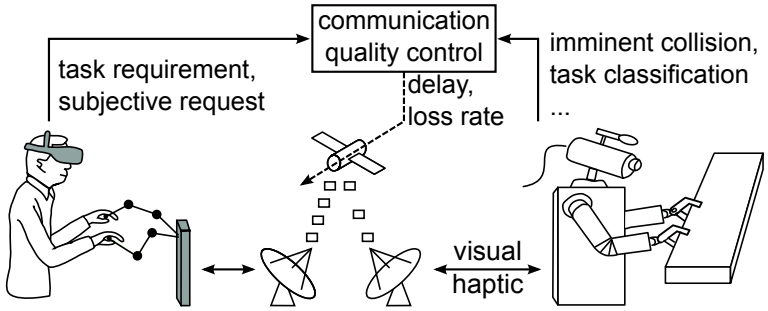


Figure 6.1: Communication quality control in telepresence can consider various quality requests from the operator and teleoperator.

delay [163], no significant influence of lost information on task performance is observable in other cases [156, 164].

Low task performance caused by communication artifacts can be addressed in several ways. Assistive functions are developed to improve task performance, however, a priori knowledge about the task is required [103, 165]. In general, this is in conflict with one of the fundamental goals of telepresence of setting the human operator into full control as his expert, presumably a priori unknown knowledge is needed to perform the specific task. Other approaches utilize knowledge about the specific time delay of the communication channel, amongst them are predictive displays [41] and the predictive wave variable transformation [33]. These predictions though require strong assumptions about the remote environment in the form of computational models, imposing significant challenges in the case of a dynamically changing and heterogeneous scenery.

In the present approach, communication quality control as a novel, conceptually different approach to networked telepresence systems is introduced, see Figure 6.1 for an illustration. Recent developments in communication protocols, e.g., IPv6 together with specialized hardware provide the ability to regulate the quality of the communication channel on-line in terms of transmission time delay and packet loss probability [166]. In telepresence systems, quality requirements for visual and haptic data transmission can come from the operator's subjective solicitation or objective imminent threats such as a potential obstacle collision.

6.2 Methods

The advent of versatile applications requiring network resources with different needs of network quality e.g. in terms of transmission latency, jitter, and packet loss rate, led to the development of communication quality control techniques, also known as quality-of-service (QoS) control. These techniques allow an adjustment of network parameters defining the transmission quality. The easiest communication control mechanism is an agreement with a network service provider offering a guaranteed quality-of-service [167] over a time period. Alternatively, specific protocols and hardware e.g. in wireless communication have been developed, allowing a quality control on shorter time scales, in the extreme case on a packet-to-packet base [168]. This technique has the advantage that only the actually needed resources have to be provided for the main purpose, giving the opportunity for additional

services to use the same communication channel when the communication demand is low. In the following, we will restrict our considerations to the latter case of a packet-by-packet QoS control scheme.

An adjustment of communication parameters during the actual task execution process can be understood as an *adaptive controller*, which was defined by Åström to be "a controller with adjustable parameters and a mechanism for adjusting the parameters." [169]. The development of an adaptive control scheme can be divided into four steps:

- Determine the desired system behavior.
- Find a suitable control law with adjustable parameters.
- Develop a mechanism for adjusting the parameters.
- Implement the control law.

In the next Sections, these four steps will be discussed for communication quality control during the usage of a telepresence system.

6.2.1 Optimal Performance/Cost Tradeoff

Enhancing human task performance in achieving a specific goal using a telepresence system is amongst the ultimate goals targeted in this chapter. In most practical cases, however, task performance can only be enhanced on charge of another, concurring goal which is keeping the costs for such systems low. As an example, the usage of a dedicated communication channel with low latency and data loss rate which allows the human operator to achieve high task performance, but is costly compared to other, shared communication solutions (e.g., the internet). The "desired system behavior" for telepresence systems helping humans to achieve a high task performance while keeping costs to be spent to a minimum can be formulated as a multi-objective optimization problem

$$\max \left[\mathbf{y}_{perf}(t)^T \; -\mathbf{y}_{cost}(t)^T \right]^T, \; 0 \leq t \leq t_{com}, \tag{6.1}$$

where $\mathbf{y}_{perf}(t)$ is a specific task performance criterion evaluated at time t, and $\mathbf{y}_{cost}(t)$ is the corresponding cost which has to be spent.

6.2.2 Suitable Adaptive Control Law

The second step towards the development of a communication quality control scheme enhancing human task performance is to find a "suitable control law with adjustable parameters" realizing this goal. In contrast to most technical control problems, the control mechanism determining the interaction of the human operator with the telepresence system and the remote environment can not be designed, but is a biological process within the human brain, see Section 3.1. The cognitive control process is modeled to take the operator's intention, sensory feedback and other factors into account to determine the neural signal $\mathbf{u}_{ne}(t)$ leading to the desired motor action. The dynamic system which is under control is characterized by the state-space equation

$$\dot{\mathbf{x}}_{\circlearrowleft}(t) = \psi_{\circlearrowleft}\left(\mathbf{x}_{\circlearrowleft}(t), \theta_{\circlearrowleft}(t), \mathbf{u}_{ne}(t), \xi_{\circlearrowleft}(t), t\right), \tag{6.2}$$

where the state variable $\mathbf{x}_{\circlearrowleft}(t)$ contains all states in the closed sensorimotor loop, describing physical, sensory, neural, cognitive, and motor processes. In the here considered case that a telepresence system is part of the sensorimotor loop, dynamic processes in human-system interface, communication, and remote environment are also contained. The parameter vector $\theta_{\circlearrowleft}(t)$ and the noise process $\xi_{\circlearrowleft}(t)$ are defined accordingly.

Since there is no direct way of tuning the control parameters in the cognitive controller itself, alternative ways of optimizing human task performance must be considered. It is clear from (6.2), that the communication quality $\theta_C(t) \in \theta_{\circlearrowleft}(t)$ has an influence on the closed-loop dynamics. Instead of tuning the cognitive controller directly, the environment which is to be moved is under control of the system designer. Considering the task is represented by a set of task parameters $\theta_T(t) \in \Theta_T \subset \Theta_{\circlearrowleft}$, we define communication-independent parameters $\theta_T^i(t) \subset \Theta_T$, such as the distance from start to goal in a pick-and-place task, and communication-dependent parameters $\theta_T^d(\theta_C(t), t) \subset \Theta_T$, e.g., a change in the impedance due to changes in time delay [18].

6.2.3 Parameter Adjustment Mechanism

Achieving the control goal formulated in (6.1) requires quantitative knowledge about the behavior of $\mathbf{y}_{perf}(t)$ and $\mathbf{y}_{cost}(t)$. Taking the closed-loop dynamics and the control behavior discussed above into consideration, $\mathbf{y}_{perf}(t)$ is the output of a task performance model, following the argumentation in Section 3.4

$$\mathbf{y}_{perf}(t) = \phi_{perf}\left(\mathbf{x}_{perf}(t), \ldots, \theta_{perf}(t), \mathbf{u}_{perf}(t), \ldots, t\right), \tag{6.3}$$

where

$$\mathbf{u}_{perf}(t) = \nu_{\circlearrowleft,perf}(\mathbf{x}_{\circlearrowleft}, \theta_{\circlearrowleft}, \mathbf{u}_{ne}, \xi_{\circlearrowleft}, t),$$
$$\dot{\mathbf{x}}_{perf}(t) = \psi_{perf}\left(\mathbf{x}_{perf}(t), \theta_{perf}(t), \mathbf{u}_{perf}(t), t\right).$$

The consideration of a performance state vector $\mathbf{x}_{perf}(t)$ is required for modeling cumulative performance measures, among others. It is important to acknowledge that the relevant parameters which can be adjusted by the communication quality control algorithm must be contained in the performance model as well, thus $\theta_C(t) \in \theta_{perf}$. Similar to the description of task performance, a cost model can be developed as

$$\mathbf{y}_{cost}(t) = \phi_{perf}\left(\mathbf{x}_{cost}(t), \ldots, \theta_{cost}(t), \mathbf{u}_{cost}(t), \ldots, t\right). \tag{6.4}$$

Finding the optimal communication quality parameter vector $\theta_C(t)$ in the sense of equation (6.1) is formulated as a multi-objective, dynamic optimization problem,

$$\underset{\theta_C(t)\in\Theta_C}{\arg\max} \left[\mathbf{y}_{perf}(t)^T \; -\mathbf{y}_{cost}(t)^T\right]^T, \; 0 \le t \le t_{com}. \tag{6.5}$$

The multi-objective nature of the optimization problem can be addressed using a variety of approaches [170]. It can be distinguished between solutions resulting in a set of Pareto-optimal values for $\theta_C(t)$ where one suitable solution is selected and applied to regulate the communication quality. Alternatively, only a single compromise solution between the concurring terms can be computed. For an overview of tangible solvers for both classes of solutions, the interested reader is referred to [170].

The globally optimal, time-varying communication parameters $\theta_C^*(t)$ can be determined by using a dynamic programming algorithm and value iteration. The prerequisites for applying these methods are admittedly strong: All models, including the environment, task, and the human behavior must be known exactly and a priori. These premises are often not satisfiable in a telepresence scenario because the environment as well as the human operator inhibit unknown, and often stochastic behavior. Model predictive control (MPC) is a promising alternative to dynamic programming [171], although the solution may not result in the globally optimal solution. MPC approaches repeatedly solve the dynamic optimization process in every time step but consider only a limited prediction horizon. The advantages of using a model predictive approach are diverse: The environment as well as the future behavior of system and human must only be known for a limited time horizon, making it more robust to modeling uncertainties. Furthermore, MPC can handle higher-order system dynamics while dynamic programming suffers from the "curse of dimensionality", thus is suitable only for low-dimensional systems. For an overview of different MPC algorithms, the interested reader is referred to [171].

6.2.4 Implementation

A large class of algorithms solving (6.5) are operating in discrete time. Without an explicit proof, we will seamlessly switch between continuous-time models and discrete-time models and assume their predictions and outputs to be equivalent. The prerequisite of this equivalence is a sufficiently high sampling rate, approximating the continuous time t with discrete time steps $t, t+1, \ldots$. As implementations are application-specific in all cases, an illustrative example for an adaptive control scheme in telepresence systems, where task performance is optimized shall be given here.

Example 6.1. A telepresent pick-and-place task with visual feedback and an adjustable communication bandwidth $bw(t) \in \mathbb{R}^+$, directly affecting the resolution of the transmitted video stream is considered. The task performance $y_{perf}(t) \in \mathbb{R}^+$ in this task is the achieved positioning accuracy during the placement procedure when using the telepresence system. As a scalar communication cost criterion $y_{cost}(t) \in \mathbb{R}^+$, we want to consider the transmission cost over one sample time of the dynamic task performance model. Assume that the following is known about communication system, task, and human operator either from experimentation or from prior knowledge:

- The cost $y_{cost}(t)$ for reserving communication bandwidth $bw(t)$ for one time unit is $y_{cost} = \gamma\, bw(t)$ where $\gamma \in \mathbb{R}^+$ is the (constant) parameter determining the network's cost policy.
- The operator task performance $y_{perf}(t)$ scales linearly with communication bandwidth.
- In the placing phase where it is important to achieve high task performance, the operator's absolute velocity in one direction $|\dot{x}_h(t)| \in \mathbb{R}$ is low compared to his maximum speed $\dot{x}_{h,max}$.
- Due to a large inertia of the object to be placed and a limited actuator force, the velocity stays approximately constant within a prediction horizon of 100 time steps.

Based on the experimental observations and considerations about the network quality, exemplary performance and cost models can be derived with the bandwidth as the only adjustable control parameter $\theta_C(t) = bw(t)$ where $bw(t) \in BW = [bw_{min}, bw_{max}]$.

$$\begin{aligned} y_{perf}(t) &= \phi_{perf}(bw(t), \dot{x}_h(t)) = (\dot{x}_{h,max} - |\dot{x}_h(t)|)\, bw(t), \\ \dot{x}_h(t+1) &= \dot{x}_h(t). \\ y_{\text{cost}}(t) &= \phi_{\text{cost}}(bw(t)) = \gamma\, bw(t). \end{aligned}$$

Resolving the multicriteria optimization problem by a weighted sum with weighting factors $\lambda_1, \lambda_2 > 0$, an exemplary MPC problem with a prediction horizon of 100 samples can be formulated as

$$\underset{bw(t)\in BW}{\arg\max} \sum_{j=0}^{100} \lambda_1 \phi_{perf}(bw(t+j), \dot{x}_h(t+j)) - \sum_{j=0}^{100} \lambda_2 \phi_{\text{cost}}(bw(t+j)). \tag{6.6}$$

Because of the system dynamics

$$\dot{x}_h(t+1) = \dot{x}_h(t+2) = \ldots = \dot{x}_h(t+100) = \dot{x}_h(t)$$

the solution to the MPC problem posed in (6.6) is the simple control law

$$bw(t) = \begin{cases} bw_{max} & \text{if } (\lambda_1(\dot{x}_{h,max} - |\dot{x}_h(t)|) - \lambda_2\gamma) > 0 \\ bw_{min} & \text{otherwise} \end{cases}$$

The choice of the performance model is the most challenging part in the development of an optimal communication quality control scheme as it implicitly assumes the operator to behave only based on velocity and video quality, and ignoring all other features. Because of this uncertainty about the operator, any controller developed based on the model should be iteratively evaluated and refined in subsequent experiments.

Over the next sections, we want to introduce a communication quality control algorithm aiming at maximizing task performance in a simulated telepresence task over real communication channels, while keeping the cost to be spent for communication as low as possible. The communication quality parameters under investigation are time delay and packet loss, as these factors play an important role in the context of long-distance and wireless telepresence applications. The approach of model development, algorithm design and implementation is maintained.

6.3 Dynamic Task Performance Model for Collision Prediction

Because task performance is a quantity which is only meaningful in the context of a specific assignment, a particular application scenario is introduced here. A navigation task through a course of obstacles is considered, where visual and haptic feedback is provided. For the ease of experimentation, evaluation and comparison of the results, the telepresence system

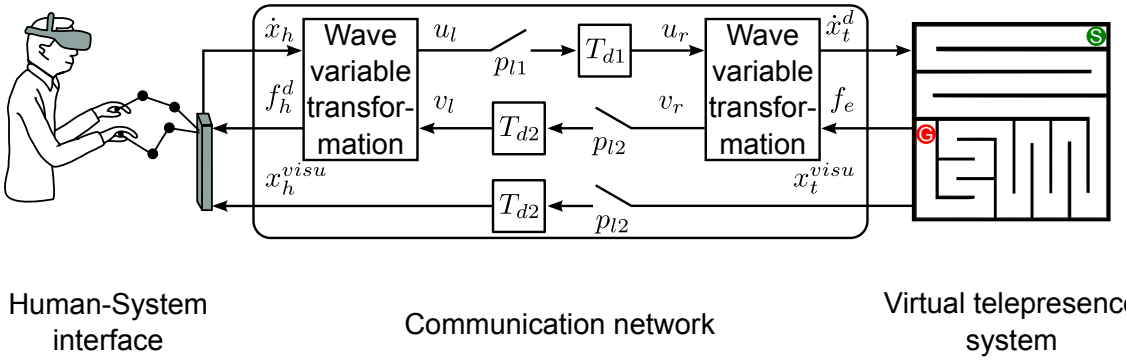


Figure 6.2: The experimental setup, consisting of a 2DoF haptic interface with monitor for visual feedback, and a virtual environment with emulated network characteristics, wave variable transformation, and a virtual labyrinth. The experimental task is going from start S to goal G without colliding with the walls while being as fast as possible.

is abstracted by a virtual environment and the teleoperator replaced by a mass-less avatar, represented by a circle in the visual feedback. Furthermore, the operator's allowable motions are constrained to a plane by using a haptic interface with two degrees of freedom. The obstacle course is represented by a labyrinth with an unique path similar to previous studies [101,102]. The virtual environment and the experimental setup is depicted in Figure 6.2. Time delay $T_d(t)$ and packet loss are recognized as challenging problems for telepresence systems operating over long distances [8, 30, 34, 36–38, 54, 55, 133]. As a consequence, we examine a communication quality control algorithm capable of varying the communication time delay $T_d(t)$ and the packet loss rate $p_l(t)$ in the communication quality parameter vector $\theta_C(t) = [T_d(t)\ p_l(t)]^T$. To guarantee stability also in the case of contact with an obstacle, the haptic signal exchange between human-system interface and virtual environment is mediated by the wave variable transformation introduced in Section 2.3.2, ensuring stability for arbitrary large time delays. Natural performance metrics $y_{perf}(t)$ for this task are [101]

1. the number of obstacle collisions N_{col} at time t,
2. task completion time t_{com}.

In the following, we will focus on the reduction of obstacle collisions first and sketch an extension for task completion time subsequently. Our professed goal is a task performance model which can be used in a MPC algorithm for communication quality control. Thus, the components $\phi_{perf}(\cdot)$ and $\psi_{perf}(\cdot)$ and their dependency on communication-sensitive task parameters are derived explicitly to describe the collision probability as a prediction model for obstacle collisions, calculated from stochastic reachable sets. To reduce the model complexity, $\phi_{perf}(\cdot)$ and $\phi_{\text{cost}}(\cdot)$ are considered to exhibit no time-varying parameters.

6.3.1 Dynamic Performance Model

The minimization of obstacle collisions requires a model that is able to predict $N_{col}(t)$ over a time horizon in the future. A particular problem in this context is the quantization

of $N_{col}(t) \in \mathbb{N}$: Due to the fact that the prediction horizon in MPC algorithms is normally reasonably short, the predicted collision count will be either 0 or 1 in most cases. This low "resolution" of the objective function may lead to a control algorithm, similar to the bang-bang control scheme discussed in Section 6.2.4. Instead of $N_{col}(t)$, the probability $p_{col}(t)$ capturing the momentary probability of a collision will be considered instead. This measure can return continuous values, where a collision probability $p_{col}(t) = 1$ is equivalent to an inevitable obstacle collision.

From intuition, multiple factors influence the collision probability at a given time: In situations where the operator maintains a small distance to an obstacle, or approaches it with a high velocity increases the likelihood that a collision will occur. Furthermore, the system dynamics influence the collision probability. As an example, if the telerobot's avatar moves rapidly towards an obstacle, it is easy to decelerate it and avoid a collision if its inertia is low. In the case of a high inertia, the operator has to put a huge amount of effort into deceleration, making it more unlikely that a collision can be avoided. Relatedly, the operator's physical abilities, e.g., the maximum force he/she is able to exert, affects the collision probability. In addition to these factors in the haptic modality, time delay in the visual feedback is known to influence human behavior as well [37, 160, 161]. Due to the high complexity of the sensorimotor system coupled with the visual-haptic telepresence system, we propose a drastically reduced sensorimotor model for the computation of collision probability here: Utilizing nothing more than the maximum and minimum bound of the operator's force excitation capabilities which we assume to be independent of the communication parameters, a purely mechanical model with bounded inputs is developed. The mechanics that the human interacts with is modeled by its impedance $Z_h(j\omega)$ and includes the human-system interface and the transmission of the remote environment. As an imperfect transmission of haptic data affects haptic transparency, thus the system dynamics, an analysis of task parameters affected by communication is performed first, before the dynamic model part $\psi_{perf}(\cdot)$ and the performance output $y_{perf}(t) = \phi_{perf}(\cdot)$ are discussed.

Remark 6.1. Without considering any model of the human interaction *strategy*, but only his/her physical *abilities*, the model is presumably conservative with respect to collision probability. This is due to the fact that actions not helping for task completion (e.g., accelerating in a wrong direction, towards an obstacle), are considered being equally likely as goal-directed actions. On the other side, the fact that task knowledge is unnecessary for the prediction of task performance is a strength of this approach: Telepresence systems often operate in highly unstructured environments without a predefined goal, at least none that is easily quantifiable in an algorithm.

6.3.2 Communication-Sensitive Task Parameters $\theta_T^d(\theta_C(t))$

The task of navigating through a course of obstacles is represented by a task-associated feature vector $\theta_T(t)$, consisting of communication-independent factors $\theta_T^i(t)$ and features depending on communication $\theta_T^d(t)$. Independent of time delay and packet loss are, e.g., the path length and obstacle position. Haptic transparency in contrast is affected by imperfect data transmission, and is analyzed here. We will discuss transparency in the case of a time-delayed communication channel and the wave-variable transformation, introduced in

Section 2.3.2 for guaranteeing stable closed-loop control, following an approach described in [8].

Considering the definition of wave variables in equation (2.18) and the important special cases of constant time delay, the impedance felt by the human operator is approximately

$$Z_h(j\omega) \approx Z_h^{\text{app}}(j\omega) = b\frac{2Z_{env}(j\omega) + bT_d j\omega}{2b + T_d Z_{env}(j\omega)j\omega}. \tag{6.7}$$

Note that this approximation holds only for frequencies $\omega < \frac{1}{3T_d}$. In the considered case, we will assume for now and show later experimentally that this inequality holds and the approximation is valid. As our task consists of free-space motion in a virtual environment without a teleoperator having own dynamics, the environmental impedance is $Z_{env}(j\omega) = 0$. Since an impact into the walls is considered as a failure in task execution and only happening occasionally, it will not be analyzed any further. For free-space movement, equation (6.7) simplifies to an inertia proportional to time delay,

$$m_{T_d} = \frac{bT_d}{2}. \tag{6.8}$$

While there exists an analytic solution for the displayed impedance in the case of constant time delay, the effect of packet loss is non-deterministic as packet dropouts generally are of stochastic nature. Furthermore, their influence on system transparency strongly depends on the reconstruction strategy for data loss. Replacing a lost package containing wave variable information with zero values is able to preserve passivity, thus stability, but results in significant position drift and decreases transparency drastically [8]. Opposing, a hold-last-sample (HLS) technique repeating the last transmitted value in the case of missing information leads to better transparency but provokes active, thus potentially unstable behavior. We consider an energy supervising algorithm [8] here, ensuring passivity while preserving transparency as much as possible.

6.3.3 State Dynamics $\dot{\mathbf{x}}_{perf}(t) = \psi_{perf}(\cdot)$

The first step in the development of a dynamic model for task performance is the definition of the state vector $\mathbf{x}_{perf}(t)$. Recalling that we base our considerations purely on the mechanical properties of the human-system interface with 2 degrees of freedom and the transmitted environment, the states

$$\mathbf{x}_{perf}(t) = \begin{bmatrix} x_{h,x}(t) & \dot{x}_{h,y}(t) & x_{h,y}(t) & \dot{x}_{h,y}(t) \end{bmatrix}^T \tag{6.9}$$

suffice for describing the dynamics of the system. Here, $x_{h,x}(t)$ and $x_{h,y}(t)$ denote the positions in the coordinate system of the haptic interface, the velocities $\dot{x}_{h,y}(t)$ and $\dot{x}_{h,y}(t)$ are defined accordingly. The task of navigating through a course of obstacles considered here can be abstracted as a free-space motion as long as no obstacle collision occurs. Due to the necessity of knowing the collision probability *before* a collision occurs to have the ability of regulating the communication parameters and prevent an impact, the performance model must be most accurate in the free-space movement. Since the operator is in contact with a real robotic device, the ideal case where no haptic feedback, thus $Z_h(j\omega) = 0$ can often

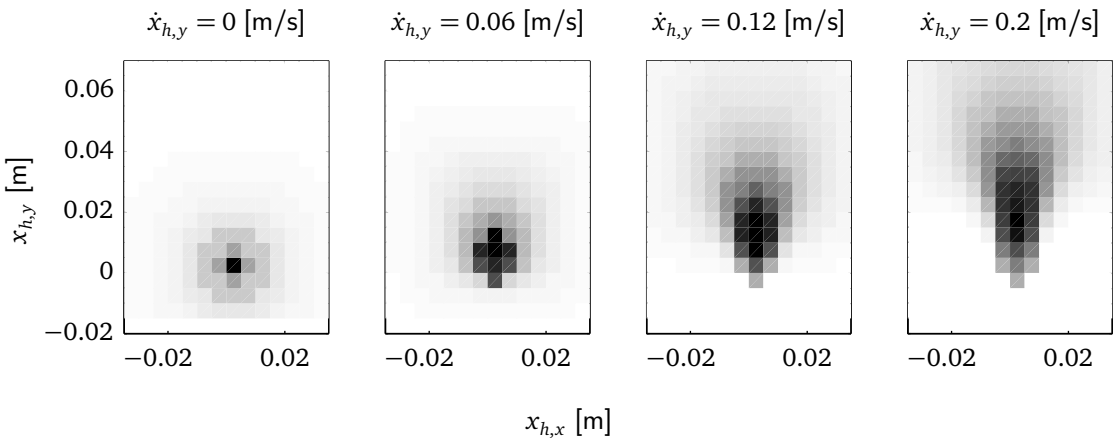


Figure 6.3: The probabilities $p_i([0, 500\text{ms}])$ of being in a specific discrete state X_i are projected to the x/y-plane. Four initial states X_j with different initial velocities $\dot{x}_{h,y}(0)$ are depicted. Darker states are more likely to be reached.

not be achieved. Instead, uncompensated dynamics introduce a small remaining impedance which is often modeled as a mass-damper model [8]. With the state variable given in (6.9), the linear model determining the performance dynamics is

$$\dot{\mathbf{x}}_{perf}(t) = A\mathbf{x}_{perf}(t) + B\mathbf{u}_{perf}(t) \tag{6.10}$$

with

$$\mathbf{u}_{perf}(t) = \begin{bmatrix} f_{h,x}(t) \\ f_{h,y}(t) \end{bmatrix}, \quad A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -\frac{d}{m} & 0 & 0 \\ 0 & 0 & 0 & -\frac{d}{m} \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{6.11}$$

The input vector $\mathbf{u}_{perf}(t)$ contains the operator's force $f_h(t)$ in the two dimensions x and y. The overall inertia m is determined on the one hand by the uncompensated dynamics from the human-system interface m_{HSI}, on the other hand by the communication-dependent inertia m_{T_d} as introduced in (6.8), $m = m_{HSI} + m_{T_d}$. The damping $d = d_{HSI}$ is only determined by the interface dynamics, as the environment has no additional damping component.

6.3.4 Performance Output Function $y_{perf} = \phi_{perf}(\cdot)$

The performance output function $\phi_{perf}(\cdot)$ relates the state and relevant communication and task parameters to the performance measure $y_{perf}(t)$ which is the probability to collide with an obstacle within the next time step $p_{col}([t, t+T])$. We consider the probability over the time interval $[t, t+T]$ where t is the current time index and T is the sample time, instead of a single time instance, such that collisions that occur between t and $t+T$ are not missed. For the calculation $p_{col}([t,\ t+T])$, an approach based on probabilistic reachable sets [172,173] is used. A reachable set denotes the set of states that can be reached from the given initial state at time t within a time interval $[t,\ t+T]$, considering certain bounds on

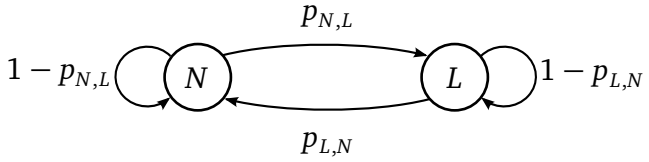


Figure 6.4: The Gilbert-Elliot model is a 2-state Markov process and approximates the network characteristics in packet-based data transmission networks. 'N' and 'L' denote the states 'no packet loss' and 'packet loss' state, respectively.

the input. To calculate the probability of a collision with an obstacle, the probability of being at a location within the reachable set must be furthermore known. This can be achieved by a quantization of the state space, and the introduction of a Markov chain with the state transition matrix $\Phi([t,\, t+T])$, containing the probabilities of reaching specific (discrete) states. Together with the known or measured positions of obstacles, the collision probability can be calculated. Projections to the $x-y$ plane of the probabilities to be in different states within a time interval $[0,\, 500\text{ms}]$ are depicted in Figure 6.3 for illustrative purposes. To the knowledge of the author, a similar method has never been applied to human motion prediction in telepresence systems before. To enhance readability, all side issues related to the computation of reachable sets, collision probability, and the implementation issues thereof are discussed in Appendix C. For a deeper methodological discussion of the principles applied, the interested reader is referred to [172, 173] and the references therein.

6.4 Experimental Evaluation

An experiment to evaluate the dynamic performance model developed in Section 6.3.1 is performed. To limit the number of degrees-of-freedom in the experimental design, the communication quality parameters are kept constant over time, thus θ_C is time-invariant. The impact of different communication quality parameters levels $\theta_C = [T_d \;\; p_l]^T$ are related to the number of collisions as the original performance measure and the predictions from the collision probability model.

6.4.1 Network Model

The network used in the experiment is emulated instead of using a real, time-delayed, and lossy communication channel to ensure a reliable presentation of experimental conditions. For simulating packet loss characteristics, either a Bernoulli process or the Gilbert-Elliot model [174, 175] are popular choices where latter was chosen for the current experiment. The Gilbert-Elliot model is a two-state Markov process and is depicted in Figure 6.4. In contrast to a Bernoulli process where each sample is independent of all other samples, this model can resemble specific characteristics of packet loss in packet-based networks such as the Internet [176] because the probability of a packet being lost depends on the transmission state of the previous packet. In this way, bursts of lost packets can be generated which are indeed part of the observed real network characteristic. Lost packets must be reconstructed on the receiving side of the telepresence system: Visual feedback does not affect system

stability, therefor we choose a hold-last-sample reconstruction for the visual datastream. This makes feedback freeze during loss bursts. In the haptic modality we take the energy-supervised hold-last-sample algorithm as discussed in Section 6.3.2. The effects of loss bursts on haptic feedback are hard to capture in general, but will be analyzed in the specific context of the task considered here.

6.4.2 Experimental Design

The experimental setup is described in Appendix A.2. For the control of the haptic interface, a position-based admittance controller was utilized, as described in Section 2.2.2. The inertia m_{HSI} was set to 9 kg, and a damping $d_{HSI} = 8\,\frac{\text{Ns}}{\text{m}}$ was present to increase the stability margins.

Three levels of round-trip time delay $T_d = \{0, 0.04, 0.1\}$s and packet loss rates $p_l = \{0, 0.1, 0.2\}$ with a fixed mean burst length of 60 ms at a packet rate of 1000 packets/sec for the haptic modality and 60 frames/sec for the visual modality were tested in an orthogonal experiment design. For simplicity reasons, we set time delays and loss rates in the send- and receive channel to $T_{d1} = T_{d2} = T_d/2$ and $p_{l1} = p_{l2} = p_l/2$, respectively. The wave impedance parameter b was set to $b = 100$ as a tradeoff between transparency in free-space and oscillations in impact situations. These parameter choices result in an inertia felt by the human operator of $m = 9$kg for $T_d = 0$s, $m = 11$kg for $T_d = 0.04$s, and $m = 14$kg for $T_d = 0.1$s, based on the inertia from the admittance controller and the time delay-dependent inertia m_{T_d} calculated from equation (6.8). All conditions were randomly presented to a human operator with 20 repetitions, leading to a total of 180 trials per subject where one trial was completed in approximately 45 seconds. The instruction for all participants was to go through the labyrinth without touching any wall while being as fast as possible. The instruction of not touching a wall was prioritized verbally over the speed of the movement.

Six university students from different disciplines participated in the experiment. All were right-handed and had normal or corrected-to-normal vision. Because of the large number of conditions, the experiment was split into two sessions with 10 repetitions of each condition. One session was completed in approximately one hour.

6.4.3 Results and Discussion

Influence of Communication Quality on $\theta_T^d(\theta_C)$

In order to draw valid conclusions about the observed results, we validate the assumptions made about the effect of time delay and packet loss on task parameters $\theta_T(\theta_C)$, in our case the time delay-dependent inertia m_{T_d}. A mass-damper admittance model with estimated inertia $\hat{m}$ and estimated damping $\hat{d}$

$$\hat{\dot{x}}_{h,x}(t) = \frac{1}{\hat{d}}\left(f_{h,x} - \hat{m}\ddot{x}_{h,x}\right)$$

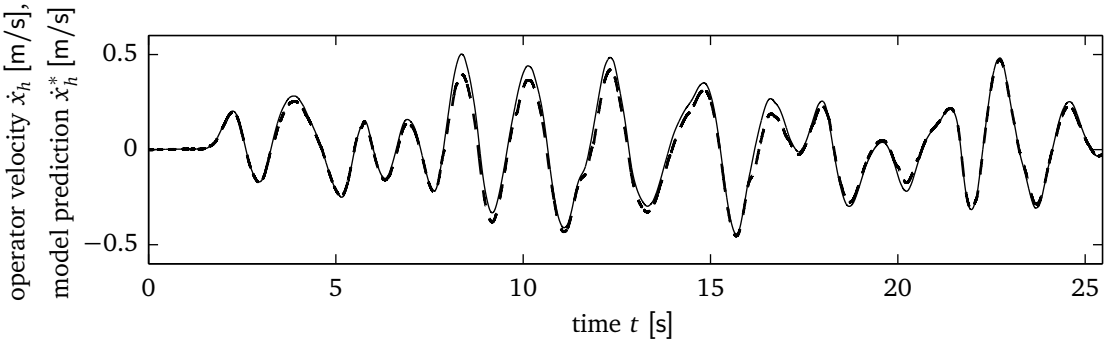


Figure 6.5: Predictions of human velocity $\hat{\dot{x}}_h$ from real force data and an identified mass-damper model (solid), compared to the actual velocity $\dot{x}_h$ (dashed) which was obtained with a packet loss rate of 20%.

is identified using a system identification algorithm, minimizing the squared error between the recorded velocity and model simulations,

$$\arg\min_{\hat{m},\hat{d}} (\dot{x}_{h,x}(t) - \hat{\dot{x}}_{h,x}(t))^2.$$

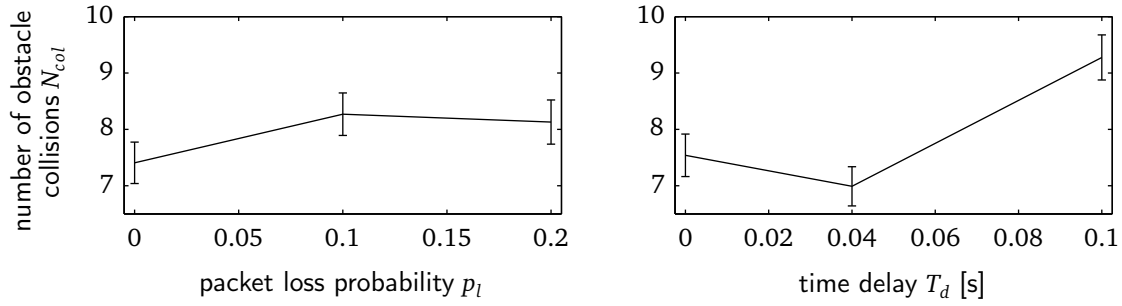
A measure of fit for the identified model is calculated by

$$\xi = 1 - \sqrt{\frac{\int_0^{t_{com}} \dot{x}_h(t) - \hat{\dot{x}}_h(t) \mathrm{dt}}{\int_0^{t_{com}} \dot{x}_h(t) - \bar{\dot{x}}_h(t) \mathrm{dt}}}$$

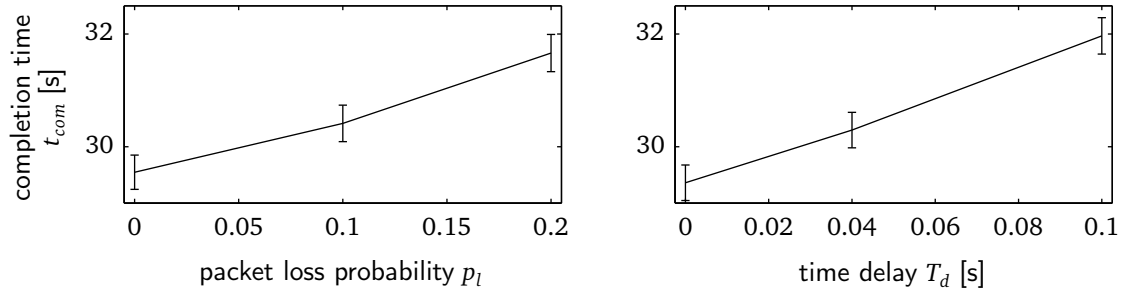
where $\bar{\dot{x}}_h$ is the mean handle velocity, calculated over data taken from a whole trial of the experimental procedure, respectively. A value of 1 for the fit measure ξ stands for a perfect match between model predictions and actual data while 0 equals a simulation behavior that is no better than simulating the mean value of all observations over a whole trial, regardless of the input. Predictions from the model and real data with no time delay and a packet loss rate of 20% are depicted in Figure 6.5.

In the case of pure time delay, the model fit with measured velocity and force data from one experimental trial is $\xi \geq 0.98$ over all trials and time delay conditions. The magnitude of identified damping and mass for the different time delay conditions is consistent with the values of the admittance controller and the additional inertia due to time delay and the wave variable transformation as computed from equation (6.8). The identified models for the time delay conditions are also capable of fitting data from trials with 20% of packet loss with a goodness-of-fit measure of $\xi \geq 0.80$. Deviations between measured and simulated handle velocity in the time instances of packet loss bursts result from a change in the mechanical impedance during lost packets which are responsible for the lower consilience of model and data. The magnitude of this impedance change depends on the burst length, and the force and velocity profiles during loss bursts.

Taking together the high goodness-of-fit for a mass-damper model in the case of time delay and packet losses, we can confirm the model for time-delayed free-space motion telepresence in equation (6.8). Furthermore, these results qualify the mass-damper model in equation (6.11) as suitable to compute the collision probability.



(a) The number of collisions with an obstacle N_{col} is significantly increased with time delay.



(b) The experimentally observed completion time t_{com} depends nearly linearly on time delay T_d and packet loss rate p_l.

Figure 6.6: Time delay has significant influence on the number of collisions N_{col} and completion time t_{com}, whereas packet loss shows only significant impact on completion time.

Influence of Communication Quality on $N_{col}(t_{com})$ and t_{com}

The influence of time delay and packet loss rate on the total number of obstacle collisions during task execution $N_{col}(t_{com})$ is depicted in Figure 6.6. Findings on the task completion time t_{com} are reported for informative reasons as well. Two 2-way repeated-measure ANOVAs were performed to reveal significant effects of packet loss rate and communication time delay on $N_{col}(t_{com})$ and t_{com}, respectively. Packet loss rate shows no significant influence on the number of obstacle collisions ($F(2,8) = 2.8, p = 0.12$) while time delay is of significant influence ($F(2,8) = 17.03, p < 0.01$). Contrast tests indicate that only the largest time delay level $T_d = 0.1$s results in significantly more collisions while the collision count in the conditions $T_d = 0$s and $T_d = 0.04$s are not significantly different ($p = 0.13$). Packet loss rate as well as time delay showed to influence completion time significantly ($F(2,8) = 18.51$, $p < 0.001$ and $F(2,8) = 48.25$, $p < 0.001$, respectively). The relation of completion time and the communication effects T_d and p_l can be described by a linear regression within the investigated range of parameters. The norm of residuals to the regression model is 0.18, indicating a good fit.

These results confirm previous findings that time delay has negative influence on task accuracy and completion time [37, 160, 161]. The negative influence of packet loss on task completion time though contradicts the previous finding that video game players are unaffected by information dropouts [177]. This can be explained by the differences in loss burst length which was presumably longer in our experiment and the consequences of losses

on the sensory feedback – while it is unknown how losses affect the game experience, in this experiment they cause the visual feedback to freeze and change the mechanical impedance in the instance of time delay.

Generally, humans are capable of adapting their control strategy to a variety of task situations [76, 111, 178]. Loss of task performance, however, indicates imperfect compensation for the change in environment dynamics as accuracy (more obstacle collisions) degrades as well as speed decreases (longer time for task completion). This can be partially explained by the changed environment impedance discussed in Section 6.4.3: The time delay-dependent inertia that must be moved by the human operator additionally to the inertia and damping from the position-based admittance controlled haptic interface requires more work to be spent for acceleration. As humans' natural behavior is known to exhibit energy-saving properties [84], the task completion time t_{com} achievable with a control strategy using similar energy levels with low inertia is lower than in the case of a high inertia. An equivalent argument can be found for the number of collisions as the amount of energy to be spent for deceleration in order to prevent a collision is also higher with a larger inertia.

The effect of the packet loss rate p_l on task performance in terms of the number of collisions and the task completion time can not be explained by changes in the impedance as p_l does not influence the mechanical properties, see Section 6.4.3. Recalling that corruption of sensory feedback due to packet dropouts is manifested in a stagnation of visually perceived motion, followed by a position jump and an instantaneous varying haptic impedance during loss bursts, the system behavior may appear less predictable for the operator, thus the risk to collide with an obstacle increases. Such unpredictable system behavior was found to affect operator's control strategy [179] in two possible ways: Operators using a risk-averse control strategy are found to lower their control gain while risk-seeking participants adjust their gain to rise with the amount of uncertainty in the sensory feedback. A risk-averse strategy with respect to the number of collisions was induced in this experiment by the explicit instruction to focus more on the avoidance of collisions than on execution speed. A lower control gain lets the human operator react with smaller control inputs, e.g., smaller forces. This fact can lead to a lower speed during navigation and thus to longer completion time. This could explain why the task completion time is significantly affected by the packet loss rate but not the number of collisions.

Influence of Communication Quality on $p_{col}([t,\, t+T])$

To serve as a dynamic task performance model, the relation between the collision probability $p_{col}([t,\, t+T])$ and the number of collisions $N_{col}(t)$ as the original performance measure must be known. The sampling time T used for computing the collision probability must be chosen in agreement with the dynamics of the input to the dynamic task performance model, in our case the exerting force $f_h(t)$ in x and y-direction. This is due to the fact that within one sampling interval $[0,\, T]$ the input to the dynamic system is assumed to be constant, see Appendix C.2. Human force exerting capabilities are known to be limited to a bandwidth of about 10Hz [6], thus T is set to 100ms. For each time delay level, an individual state transition matrix $\Phi([0,\, 100\text{ms}])|_{T_d}$ must be computed due to the change in system dynamics. Intuitively, a higher time delay results in a higher inertia m_{T_d} to be moved by the operator. From a resting state, the reachable area around the initial position is thus

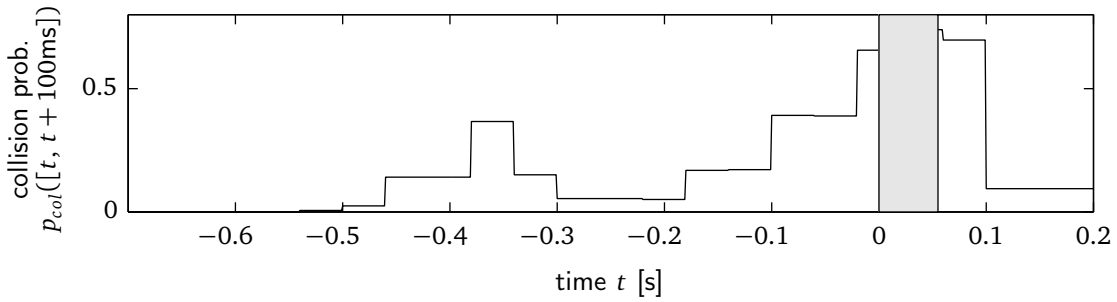


Figure 6.7: An example for the raise in collision probability $p_{col}([t,\, t+100\text{ms}])$ before a collision occurs, indicated in gray.

smaller if only a limited force is available, compared to the case where the effective inertia is small. The time-course of the collision probability $p_{col}([t,\, t+100\text{ms}])$, computed off-line for a trajectory recorded in the experiment, is exemplarily depicted in Figure 6.7.

A raise in collision probability can be noticed before the actual impact at collision time $t = 0$. The fact that $p_{col}([t,\, t+100\text{ ms}])$ does not increase to 1 is due to the fact that the dynamic performance model solely relies on the system dynamics and the operator's physical abilities. This makes it possible, though unlikely, that the collision could be avoided. Nevertheless, the highly elevated value of the collision probability up to ≈ 0.7 in comparison to situations where no collision occurs indicates that $p_{col}([t,\, t+100\text{ ms}])$ is capable of predicting a collision within the next time step when it actually occurs.

In addition to the capability of predicting collisions in general, time delay and packet loss rate must influence the collision probability $p_{col}([t,\, t+100\text{ ms}])$ similarly to the number of collisions $N_{col}(t)$. We test this model property by computing the collision probability for different levels of time delay and packet loss, right before a collision actually occurs. The difference in collision probability, calculated over all occurred collisions in the dataset for multiple levels of time delay is depicted in Figure 6.8. It can be noted that with an increasing time delay level, the rise in collision probability is highly significant ($F(6, 49987) = 765.83$, $p < 0.001$). On the other hand, packet loss rate shows no effect on the collision probability. This was expected as we did not include any other dynamics than the mechanical impedance and assumed the force range to be unaffected by changes in communication quality. Recalling from Section 6.4.3 that the influence of packet losses on the impedance is negligible, the collision probability $p_{col}([t,\, t+100\text{ ms}])$ does not depend on p_l.

6.5 Communication Quality Control for Improved Collision Avoidance Performance

The main goal of the communication quality control scheme targeted here is the minimization of collisions and network costs by an adjustment of network quality. Since the effect of packet loss on the number of collisions is insignificant, only time delay will be taken into consideration when developing a communication quality MPC algorithm.

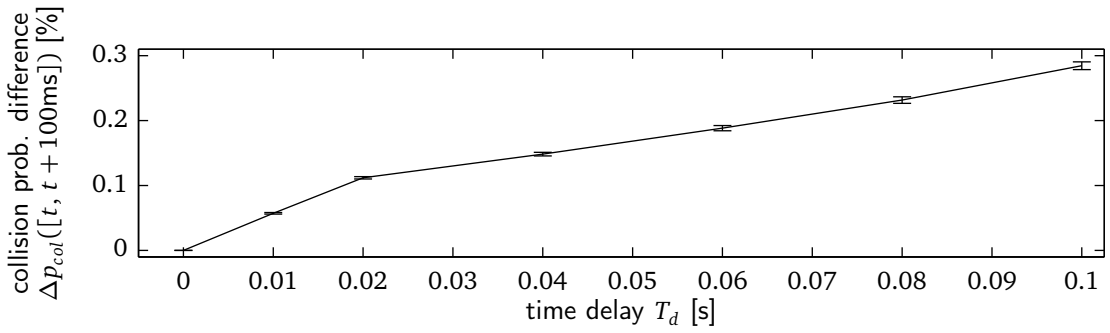


Figure 6.8: The difference in collision probability within the next sample $\Delta p_{col}([t,\, t+100\text{ms}]) = p_{col}([t,\, t+100\text{ms}]) - p_{col}([t,\, t+100\text{ms}])|_{T_d=0}$ is higher with larger time delay.

6.5.1 Model Predictive Control Algorithm

The original problem formulation for increasing task performance while keeping costs low which was posed in equation 6.5 is a multi-objective dynamic optimization problem

$$\underset{\theta_C(t)\in\Theta_C}{\arg\max} \left[\mathbf{y}_{perf}(t)^T \; -\mathbf{y}_{cost}(t)^T\right]^T,\; 0 \le t \le t_{com}.$$

In the concrete example of aiming for a lower collision probability while maintaining a low communication cost, we resolve the multi-objective nature of this problem by a weighted sum, using weighting factors λ_1 and λ_2. A model predictive controller for solving this optimization problem can be formulated as

$$\underset{T_d([t,\,t+t_p])\in\Gamma_d}{\arg\min} \underbrace{\lambda_1 p_{col}([t,\, t+t_p]) + \lambda_2 \phi_{\text{cost}}([t,\, t+t_p])}_{J_1} \tag{6.12}$$

where t_p is the prediction horizon, and Γ_d is a set of possible time delay values the control algorithm can choose from.

Remark 6.2. The prediction horizon of the MPC algorithm proposed in equation (6.12) is longer than the sample time of $T = 100$ ms in general. The collision probability prediction can be extended over the whole horizon, by using an iterative procedure, as described in Appendix C.2.2. In addition to this extension, modifications are necessary to make the collision probability computation fast enough for a real-time application like communication quality control. These modifications are summarized in Appendix C.2.3.

6.5.2 Communication Cost Model $y_{cost}(t) = \phi_{\text{cost}}(\cdot)$

There are several possibilities defining network costs, amongst them the cost per packet and the overall cost at task completion. For teleoperation, the task and thus its duration is usually unknown, making a consideration of the overall cost unsuitable, thus we will use the cost per packet. The relation between network cost and time delay is usually defined to be monotonically decreasing with increasing time delay [180, 181], such as exponential or

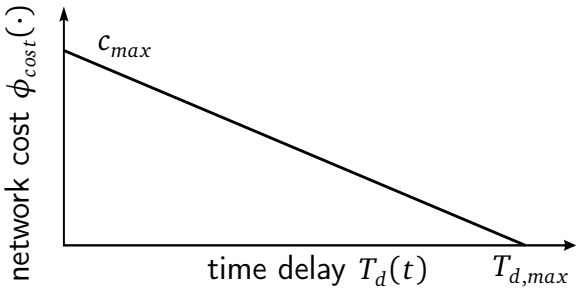


Figure 6.9: The cost for transmitting a packet is monotonically decreasing with communication time delay. In the current experiment, the relation is modeled to be linear.

rational functions. We set the cost function to a first-order rational function in agreement with current literature [181],

$$y_{cost}(t) = \phi_{\text{cost}}(T_d(t)) = c_{max} - \frac{c_{max}}{T_{d,max}} T_d(t) \tag{6.13}$$

such that a time delay $T_d = 0$ corresponds to a maximum cost c_{max} and the highest possible time delay $T_{d,max}$ comes free-of-charge as illustrated in Figure 6.9. Without loss of generality, we set the maximum cost $c_{max} = 1$. For the experiment, we will assume that enough network resources are available such that every quality request can be fulfilled and is handled immediately.

6.5.3 Stability considering Time-varying Time Delay

The online-control of time delay poses problems for guaranteeing stability. The wave variable transformation is known to provoke active, thus potentially unstable behavior of the communication channel if time delay is time-varying. As a solution to this problem a time-varying scaling factor

$$k(t) = \sqrt{1 - \dot{T}_d(t)/2}$$

is inserted into the communication channel [54] as shown in Figure 6.10. This factor depends on the temporal derivative of time delay and can fully prohibit communication in the case of a rapidly rising value of $T_d(t)$. Due to the large impact of this phenomenon on the displayed impedance, the effect is attenuated by constraining the rate of the variable time delay. This is achieved by pre-processing the time delay request using a first order lowpass filter with cutoff frequency 5Hz.

6.5.4 Experimental Design

The network emulator described in Section 6.4.1 is complemented with the ability to regulate the time delay online, furthermore a virtual environment with different obstacle positions as depicted in Figure 6.10 is considered. The number of different time delay levels the communication quality controller can choose from is set to $\Gamma_d = \{0 \quad 0.05 \quad 0.1 \quad 0.15 \quad 0.2\}$s. Multiple paths are freely decidable for the human

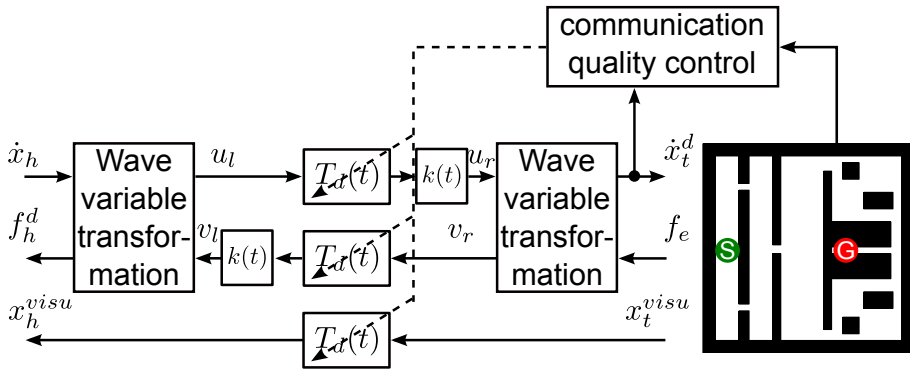


Figure 6.10: The experimental setup used in Section 6.4 is complemented with a communication quality control mechanism and a time-varying gain to ensure stability.

operator, while all paths have the same length and structure, as they are symmetric. A sampling rate of 25Hz on the same experimental PC that executed the haptic rendering can be achieved. The weighting factors λ_1 and λ_2 in the MPC algorithm posed in equation (6.12) is tuned to weight collision probability higher than network costs. This means that maximum latency is applied in non-critical situations but the time delay is lowered in the case of an imminent collision where lower latency produces lower collision probability $p_{col}([0, t_p])$. The actual values for λ_1 and λ_2 have been found in piloting experiments.

After familiarization with the system and training of at least 5 trials with quality-controlled time delay and 5 constant-delay trials, each participant had to perform 40 trials with communication quality control using the cost function from equation (6.12) and 40 trials using a cost function without considering the collision probability, resulting in a constant time delay of $T_d = 0.2$.

5 engineering students participated in this experiment, none of them participated in the Experiment described in Section 6.4. All were right-handed. All experimental conditions, quality-controlled and constant time delay conditions, were fully intermixed to avoid adaptation effects. The duration of one trial was less than 30 seconds, thus all 80 trials could be finished within one hour.

The experiment was performed using the same hardware described in Appendix A.2.

6.5.5 Results and Discussion

Paired t-tests reveal that communication quality control has a significant positive effect on the task performance in terms of the number of obstacle collisions ($p < 0.05$), as depicted in Figure 6.11. On the other side, completion time is not significantly changed by active control of time delay in the communication channel ($p = 0.96$). In Figure 6.12, the online-control of communication quality using the developed model predictive controller is depicted exemplary. In the case of no imminent collision, the time delay is on its maximum level. In cases where a better cost/task performance tradeoff can be achieved with a lower time delay level, the channel quality is switched accordingly. The success of the communication quality control scheme can be explained by considering that in the case of an imminent collision, the operator usually decelerates movement to not collide with the obstacle. Lowering the

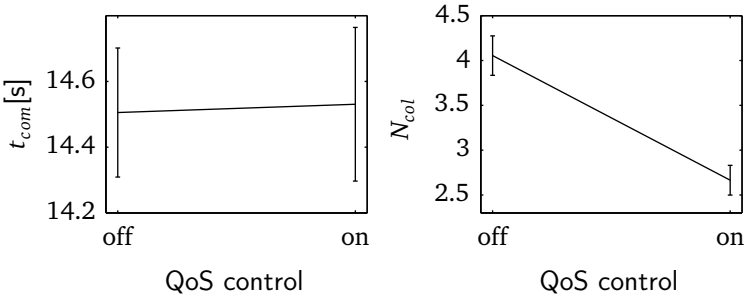


Figure 6.11: While completion time is not significantly affected by communication quality control, the number of collisions with the walls are lower when time delay is controlled according to the collision probability.

inertia during this phase leads to a shorter breaking distance, helping the operator to avoid an impact.

It can be additionally noted that completion time is not significantly influenced by the proposed quality adjustment algorithm. This is despite the fact that time-varying time delay also causes packets to be lost [8] and data loss was found to increase task completion time, see Section 6.4.3. The difference between the lost packets in the quality-controlled condition and simulated packet dropout investigated earlier lies in the fact that time delay is lowered not randomly, but only when the operator approaches an obstacle. Reconsidering that these phases are generally dominated by breaking actions, the effect of the lost information on completion time may be negligible.

6.6 Towards Communication Quality Control for Task Completion Time Improvement

While the previous experiment showed improved task performance in the sense of avoiding collisions, we aim to influence completion time as well as accuracy in the next step. For this purpose, a prediction mechanism allowing to consider completion time in $y_{perf}(t)$ is needed. Based on the reachability computation used to predict collisions, we propose a prediction term $r_{speed}([t,\, t+T])$ similar to the collision probability $p_{col}([t,\, t+T])$. The term

$$r_{speed}([t,\, t+t_p]) = \sum_{i=1\ldots N} p_i^{pos}([t,\, t+t_p])\sqrt{x_{h,x,i}^2 + x_{h,y,i}^2}.$$

where p_i^{pos} is the probability to be at a specific position $x_{h,x,i}$, $x_{h,y,i}$ and N is the number of discrete state variables taken into consideration for the computation of the collision probability. The motivation for this definition of $r_{speed}([0, t_p])$ is the following: Large values of r_{speed} are achieved when the probability to reach far distances from the starting point is high. This, in turn, can lead to a larger likelihood for making faster movements, capable of leading to a smaller completion time. Extending the MPC problem in (6.12) to

$$\arg\min_{T_d \in \Gamma_d} J_2 = \arg\min_{T_d \in \Gamma_d} \lambda_1 p_{col}([t,\, t+t_p]) + \lambda_2 \phi_{\text{cost}}([t,\, t+t_p]) + \lambda_3 r_{speed}([t,\, t+t_p]) \quad (6.14)$$

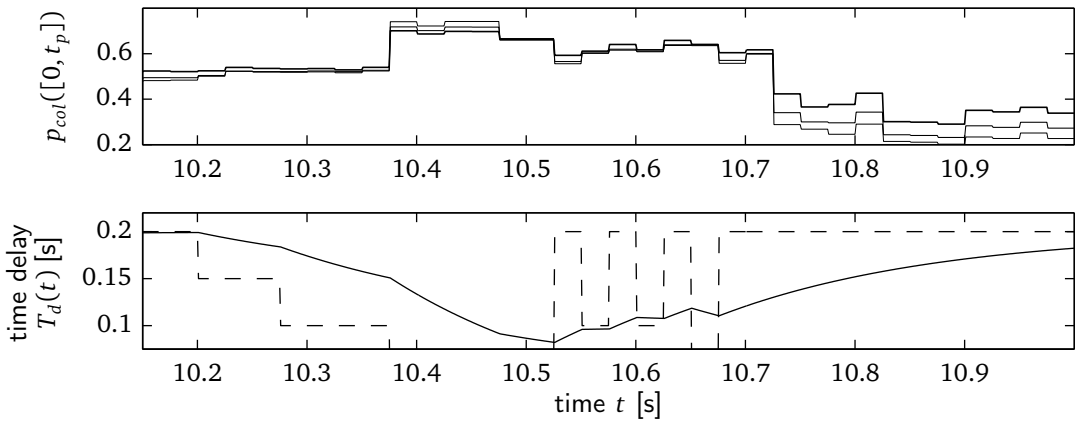


Figure 6.12: The value of the cost function $J_1(t)$ is displayed exemplarily for three time delay levels ($T_{d,max}$ dotted, $T_{d,min}$ solid, $T_{d,med}$ dash-dotted). The optimal time delay $T_{d,opt}$ with minimum value of $J_1(t)$ (dashed) is lowpass-filtered (solid) and applied to the communication channel.

respects both collisions and completion time. The practical impact of quality control using this specific cost function on the haptic properties of the system is ambiguous. Considering the operator's maximum force capabilities together with the environment dynamics changing with time delay can help for an intuitive understanding of the algorithm's properties: From a resting position, the reachable area around the initial point is symmetric in every direction. Its width is reciprocal to the communication channel's time delay as it affects the inertia to be accelerated. With the same force exertion capabilities, a smaller inertia can be moved further in the same time. The value for the objective function r_{speed}, can be maximized in this case with choosing small time delay. In contrast, when the initial velocity is high, the inertial mass helps maintaining a high velocity which is reflected in high probability values for locations in the movement direction, see Figure 6.3. In this case, r_{speed} is large with large inertia, equivalent to large time delay.

The value of the weights $\lambda_{1,2,3}$ have to be chosen carefully in order to let the collision avoidance term overrule r_{speed} in the case of an imminent collision as both terms are working in opposite directions. λ_2 can be used to adjust the average time delay level as it makes it more or less expensive to use the communication channel with small latency.

6.6.1 Apparatus and Procedure

The experiment for testing the effectiveness of the proposed communication quality control algorithm was performed using the same setup and virtual environment as in Section 6.5. All participants underwent the same training procedure with at least 5 quality-controlled trials and 5 trials with constant time delay, followed by the main experiment consisting of 40 trials with varying time delay and 40 trials with constant time delay. In contrast to the previous experiment, conditions including communication quality control were paired with constant time delay conditions that result in equal network costs. Corresponding to the linear network cost model in (6.13), equal costs are produced by trials with equal mean

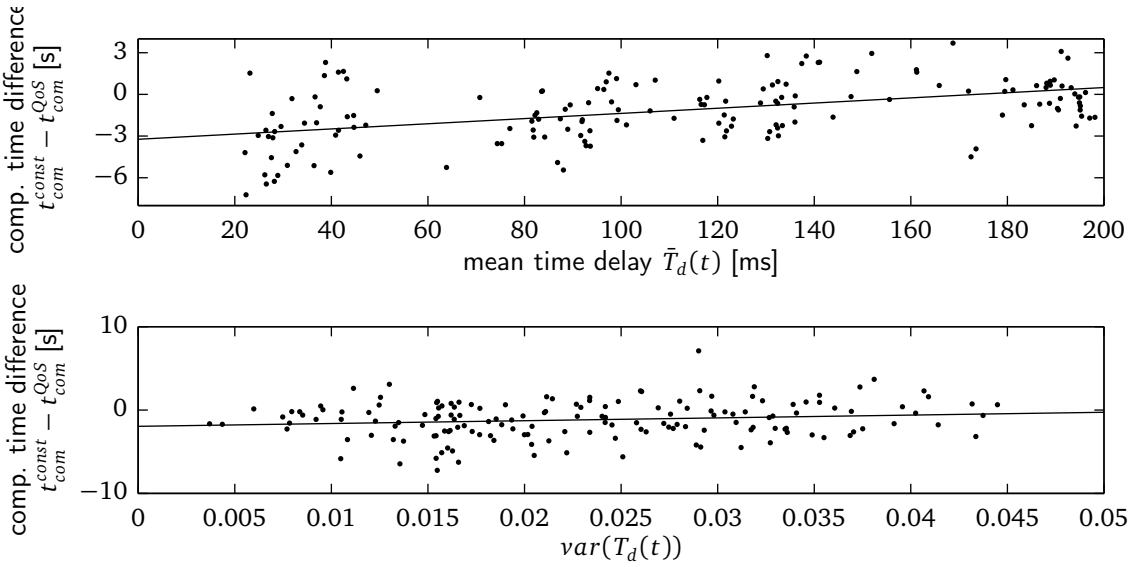


Figure 6.13: Correlations of the change in completion time between trials with constant velocity t_{com}^{const} and with active communication quality control t_{com}^{QoS}, over mean time delay $\bar{T}_d$ and time delay variance $var(T_d)$.

time delay, given that completion time is the same. Though completion time is not under control of the experimenter, it stays within small bounds, making it reasonable to apply the mean time delay in the constant time delay condition.

Each participant was tested with a different weighting factor of λ_3 in cost function (6.14) – λ_2 was fixed to assure that collision avoidance is able to overrule agility in terms of an imminent collision. This experimental design resulted in different means and variances of $T_d(t)$.

6.6.2 Results and Discussion

Performance Evaluation

There was no common trend between participants and conditions: Paired t-tests revealed that two subjects showed significantly longer completion time ($p < 0.05$) compared to the equal-cost condition with fixed time delay. One subject had less wall collisions with quality control than without ($p < 0.05$). None of the participants showed a significant decrease in completion time. To quantify the influence of mean time delay, a regression analysis was performed using all trials from all subjects. The variance in time delay as a measure for the number and magnitude of communication quality adjustments was analyzed similarly. The change in completion time between constant time delay condition and the quality-controlled channel is significantly lower with increasing mean time delay ($F = 43.0, p < 0.001$), and higher with increasing variance in time delay ($F = 8.00, p < 0.01$) as depicted in Figure 6.13.

Discussion

While the communication quality control algorithm aiming for lowering the number of obstacle collisions demonstrated a general benefit of controlling time delay on the basis of predictions from a human task performance model, considering an additional term for improving execution speed can neither consistently improve task completion time nor collision avoidance. This applies for the case that the performance achieved with communication quality control is compared to a equal-cost policy with constant time delay. An explanation for the fact that both task performance indicators can not be improved by communication quality control can be searched in the difference between the optimization objective of solely focusing on collision avoidance and the current objective, especially regarding the movement phases when time delay is lowered and increased. Time delay is normally high during regular, non-critical movements when using an algorithm with cost function J_1 from equation (6.12) and adjustments are made before imminent collisions only. Minimizing J_2 from equation (6.14) instead results in a low time delay (inertia) during slow motions and high delay (inertia) with high velocity. Thus, the change in time delay is noticeable during the acceleration phase of the movement which has potential to disturb the operator. It is furthermore known that humans are able to adjust their arm impedance to the actual load condition in the sense that arm stiffness increases when the load on the hand is higher [182]. Together with the previous consideration it can be concluded that the arm stiffness is low when a small inertia is to be accelerated from a resting state. When switching to a high inertia during acceleration, the arm impedance is no longer appropriate, potentially impeding the desired effect to speed up task completion.

Influence of the Communication Cost Policy

Performance evaluation heavily depends on the network cost policy. In our case where cost are assumed to scale linearly with requested channel time delay, the effort of communication quality control could not even pay off in terms of completion time benefits. Given that network quality control produces less cost than average time delay and considering the influence of time delay on completion time found in Section 6.4, a task performance benefit of controlling the quality online compared to the equal-cost policy is expected.

6.7 Conclusions and Open Problems

A dynamic task performance model has been developed, founded on the dynamic sensorimotor framework introduced in Chapter 3. The human motor capabilities have been taken into consideration for a calculation of the probability to collide with an obstacle, based on stochastic reachable sets. An innovative human-centered application of communication quality control to haptic telepresence demonstrates a potential for improving task performance in this way.

The fact that it was not possible to demonstrate a cost-effective adaptation of communication quality based on the current task performance model indicates room for further improvements: First, a task performance model capturing more details of human behavior, such as the control strategy, could be used to make predictions about future movements,

thus collisions and speed metrics more accurate. In addition, the influence of visual feedback and its impact on the human operator was neglected so far. By including, e.g., perceptual knowledge about the discrimination of time delay into the parameter adaptation algorithm, further improvements could be made.

7 Conclusions and Future Directions

The focus of this thesis is on the development of dynamic models of human perception and action and their application in telepresence. This interdisciplinary endeavor requires the seamless integration of system theoretic modeling considerations into psychophysical experimental techniques for the determination of perception properties on the one hand, and the development of engineering methods that are optimized with respect to human sensorimotor capabilities on the other hand. The question of haptic perception's dynamic behavior, which is addressed in this thesis for the first time, is a significant contribution to both disciplines involved: Current haptic psychophysical models capture only the steady-state of perception, loosing information about the transient reaction leading to it and are inaccurate as they do not take time-varying influence factors such as the interaction movement into account. On the other hand, today's integration of quantitative human requirements into the design of technical systems is still in its infancy.

A starting point towards a dynamic formulation of the complete sensorimotor loop is a system theoretic analysis of the information exchange between the environment, the sensory, neural, cognitive, and motor system which are involved in the haptic interaction with the physical world. A mathematical framework, capable of capturing this information exchange is developed in Chapter 3 on the basis of coupled dynamic systems in state-space form. With this system theoretic view on the sensorimotor system, the development of quantitative perception models for the dedicated purpose of being integrated into the design, control, and evaluation of a technical system is possible in a systematic way. Besides a discussion of phenomena which can be modeled using this framework, the relation to existing perceptual, behavioral, and task performance models is discussed.

Due to the fact that little is known about the dynamic processes determining haptic perception, the first steps in this new field must aim for a basic understanding of fundamental mechanisms in haptic perception: One of these foundations is the combination of movement and force feedback to a unified percept of environmental features, such as the feel of an object's inertia, or stiffness. In Chapter 4, experimental work on the temporal perception limits in this combination process is presented. A model-based experimental design considering three particular dynamic model candidates in the choice of conditions is the first innovation in this context and demonstrates the usefulness of the dynamic modeling framework developed in Chapter 3. The resulting model, capturing the perception thresholds for temporal inconsistencies between movement and force best is a state observer model, based on an internal representation of body and environment dynamics. Remarkably, all models considering including a dynamic process result in a lower prediction error compared to a static perception model for time delay. The obtained results are of significant importance for a human-centered transparency evaluation, and the design of novel communication concepts.

A second example for the need of a dynamic formulation for human perceptual processes is the modeling of haptic masking effects. In an experimental study on the perception

of damping in Chapter 5, difference thresholds for the damping coefficient are measured. These are found to be higher when additional features, such as inertia or stiffness, are combined with the damping. Explaining such masking effects requires a consideration of the movement's and force feedback time-series data. A comparison between predictions from different model candidates reveals that perception mechanisms exhibiting accumulating behavior resemble human perception better compared to models without an integrator. This finding adds evidence to the need of a dynamic formulation in order to capture perceptual effects accurately, and furthermore has an impact on the design, and evaluation of haptic mechanisms and algorithms: As a certain amount of information must come together before it can be perceived, algorithms which aim to keep distortions below the detection threshold could allow short violations of "static" thresholds which avoids an overly conservative design. In addition, the finding of high imperceptible damping differences in environments containing additional stiffness or inertia can help in the development of stable haptic rendering algorithms, and aid the design of haptic mechanism.

Besides perception, the ability to perform tasks with a high performance is an objective to be taken into consideration in the design of telepresence systems. With static definitions of task performance such as the collision count in a navigation task, systems can be optimized to allow the operator achieve a higher task performance on average. The measures taken to achieve this enhancement may though be overly expensive if their positive effect on task performance is limited to only specific situations, e.g., before an imminent obstacle collision, and unnecessary in the rest of the time. The usage of dynamic task performance models on the basis of the mathematical sensorimotor framework presented in Chapter 3 allows a more accurate utilization of resources, by predicting the task performance online during task execution. In Chapter 6, an parameter adaptation scheme is investigated on the example of communication quality control. By lowering the time delay of the communication channel in an abstracted telepresence task on the basis of predictions of the collision probability from a dynamic task performance model, a significant improvement of task performance could be observed. At the same time, the usage of costly, high-quality communication has been minimized.

In summary, the introduction of dynamic models for the quantification of haptic perceptual and behavioral phenomena, allows for

- predictions of higher accuracy for perceptual limits compared to static perception laws,
- a simple possibility for integrating human knowledge into a technical system,
- dynamic optimization of haptic telepresence with respect to task performance.

7.1 Outlook

Dynamic modeling of haptic perception is still in its infancy. However, based on the first promising results presented in this thesis, research directed towards the development and identification of the system dynamics in human perception and sensorimotor control is expected to fall on fertile ground in the psychophysical *and* engineering domain. Some specific

challenges and interesting research questions, directly deducted from the considerations presented in this thesis are discussed briefly in the following.

Haptic Diffusion Models

The dynamic properties of perceptual processes in other sensory domains than haptics have been found to be captured reasonably well by a diffusion process. In the current experiments, an increase in prediction accuracy can be observed when an accumulating stage is included into the model structure for the discrimination of damping, see Section 5.6. This structural similarity to the diffusion model motivates further investigation into a more generic modeling of haptic perception on the basis of the three-stage diffusion process, capturing information extraction, accumulation, and decision.

Dynamic Just Noticeable Differences

The JND of environment features as perceptual performance indicator has been successful not least because technical applications can be optimized with respect to these properties, without explicitly considering the human operator. The current model formulation of dynamic perceptual processes can be used to extend the classical JND to dynamic JND models. A formal definition and intuitive forms of expression could make the new notion of dynamic haptic perception accessible to a wider audience of researchers from the psychophysical and system theoretic field.

Integration of Dynamic Perception Models into Telepresence Systems

While the success of dynamic models enhancing task performance in a telepresence task has been demonstrated, an integration of dynamic haptic perception models into real applications have been only sketched so far. One innovative direction is the perceptual optimization of time-delayed telepresence systems by means of a communication quality control algorithms, as proposed in Section 6.5. In a similar way as the collision probability, a model of perceptual fidelity with respect to time delay can be considered in the cost function taken into consideration for determining the optimal time delay level.

Extension to Other Technical Systems

It has been mentioned on several occasions in the thesis that the developed models and obtained results may be useful not only in the context of telepresence, but expendable to other technical systems as well. Potential application fields include the stable and perception-optimized rendering of virtual haptic environments, and the design of haptic mechanisms, e.g., for the optimization of the perceived quality of automotive control elements such as push buttons.

A Apparatus

A.1 1 DoF Linear Actuator

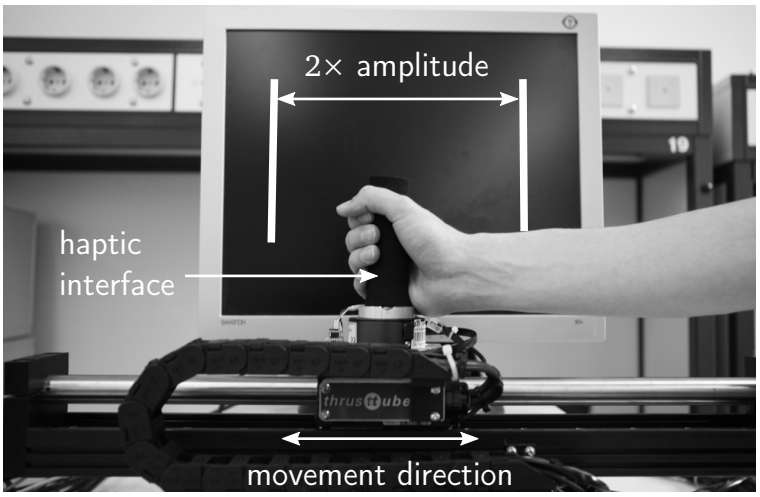


Figure A.1: A 1DoF linear actuator served as haptic device for experiments performed in Chapters 4 and 5.

The experimental setups utilized in the experiments in Chapters 4 and 5 are principally equivalent and base on a ServoTube linear motor module (Copley Controls Corp.). The setup in Chapter 4, depicted in Figure A.1 has a rubber-coated handle for user interaction mounted on top of a a 6 deegree-of-freedom (DoF) force-torque sensor (JR3, Inc.). The other setup used in Chapter 5 has a wooden endeffector as interaction handle, and a linear 1DoF force sensor (Burster). The devices are controlled by a PC, equipped with a Sensoray 626 DAQ Card running Gentoo Linux in the case of all experiments in Chapter 4 and Ubuntu Linux in the experiment presented in Chapter 5. The haptic environments are rendered in real-time using the RealTime Application Interface (RTAI) [183] for Gentoo Linux and the CONFIG_PREEMPT_RT kernel patch [184] for Ubuntu Linux.

For an evaluation of the actually presented impedances $Z_h(j\omega) = \frac{F_h(j\omega)}{\dot{X}(j\omega)}$ in the case of the force control scheme utilized in Chapter 4, the haptic interface dynamics was identified in the frequency domain [3]. The uncontrolled linear actuator dynamics were found to be sufficiently well captured by a linear second-order system with the commanded force $f_c(t)$ as input and the endeffector position $x_h(t)$ as output. To determine the specific parameters, a standard least-squares system identification procedure was applied. By employing the implemented controller K_p used in the experimental procedure and the environmental dynamic equations in (4.13) as they were used in the actual experimental procedure, the respective frequency responses for $Z_h(j\omega)$ were calculated – see Fig. A.2.

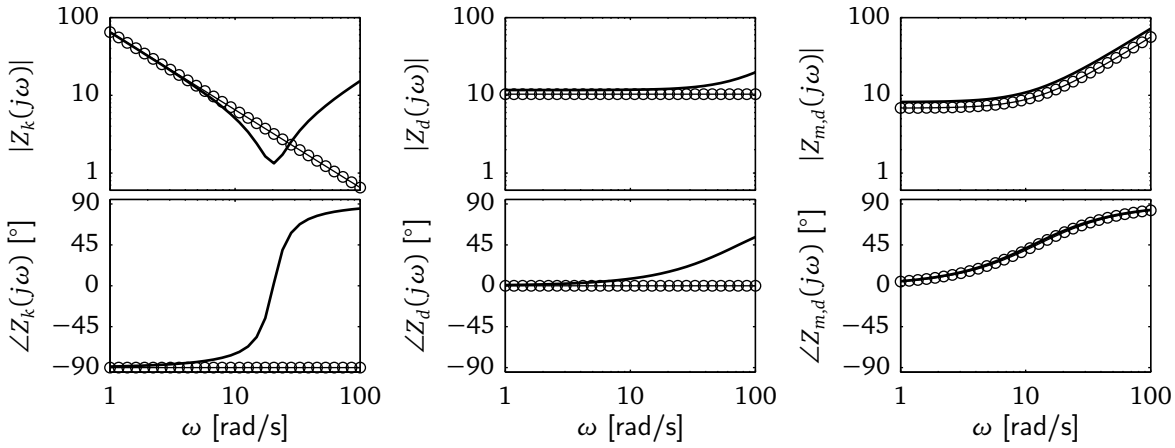


Figure A.2: Amplitude (upper) and phase characteristics (lower) of the ideal environment impedances for stiffness $Z_k(j\omega)$, damping $Z_d(j\omega)$, and the mass-damper $Z_{m,d}(j\omega)$ (from left to right, circle markers) in the experiments in Chapter 4 and the actually rendered environmental characteristics (bold, without marker).

From the Bode diagrams, it can be deducted that inertia and damper were rendered quite accurately around the movement frequencies that are in the range of the experiments described in Chapter 4, both in terms of amplitude and phase characteristic. The influence of the actuator's inertia and friction, however, change the phase characteristic of the spring to approximately -45° at the movement frequency of 2π rad/s, instead of an ideal phase of -90°.

The stimuli used in Chapter 5 were rendered using a position-based admittance control scheme. As this rendering method is known to be very reliable and accurate, the impedances displayed to the human operator are almost exactly the desired characteristics, $Z_h(j\omega) \approx Z_{env}(j\omega)$.

A.2 2 DoF Linear Actuator

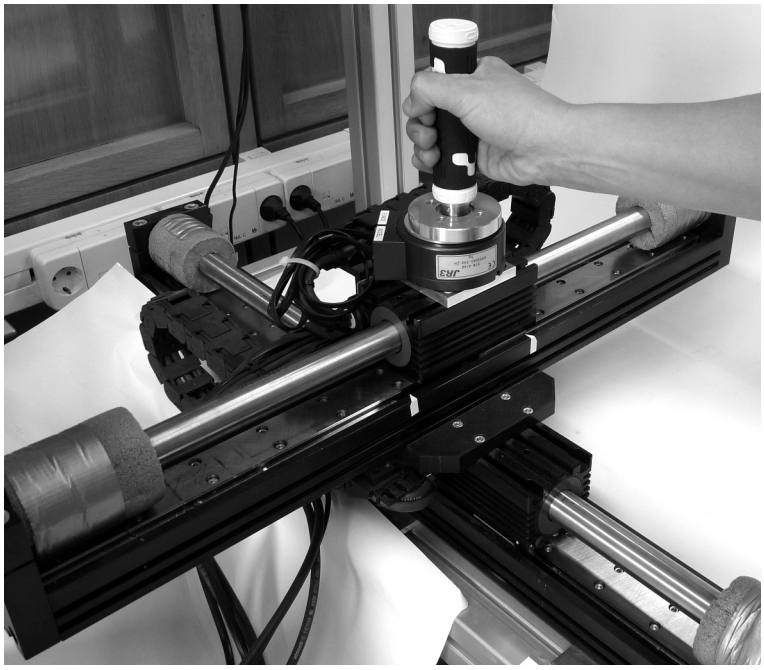


Figure A.3: The haptic interface with two actuated degrees of freedom was used in the experiments in Chapter 6.

The custom-made 2DoF experimental setup used in the experiments performed in Section 6 consists of a Thrusttube module 2504 (Copley Controls Corp.), mounted on top of a Thrusttube module 2510 in a right angle. Each actuator is equipped with an optical position encoder with resolution 1 μm. A 6DoF JR3 force-torque sensor together with a handle provides force feedback. The haptic interface is controlled in real-time at a sampling rate of 1kHz using a Quad-Core AMD Phenom desktop PC equipped with a Sensoray 626 DAQ card running Gentoo Linux with the RTAI kernel patch for maximum timing reliability. Visual stimuli are displayed on a 42inch flat screen TV at a refresh rate of 60Hz. The inherent time delay between visual and haptic stimuli was measured using a luminance sensor to be within one refresh cycle, thus varying between 0 and 16ms. A picture of the setup is provided in Figure A.3

B Tables

B.1 Masking Thresholds and Model Parameters

model	threshold values $\epsilon = -\epsilon_1 = \epsilon_2$										f_{sat}	T_{mask}
	d_1 m_0 0	d_2 m_0 0	d_1 m_0 k_1	d_1 m_0 k_2	d_1 m_1 0	d_1 m_2 0	d_2 m_0 0	d_2 m_0 k_1	d_2 m_1 k_2	d_2 m_2		
$\delta_{thresh}(\cdot)$	3.5	6.3	2.9	7.2	4.1	4.3	2.0	6.0	6.6	6.7	n.a.	n.a.
$\delta_{acc,thresh}(\cdot)$	5.7	10.6	4.1	10.6	6.8	6.0	3.0	9.0	10.5	11.3	n.a.	n.a.
$\delta_{sim}(\cdot)$	2.9	5.2	1.0	1.8	3.2	3.3	1.1	1.9	5.4	5.5	1.00	n.a.
$\delta_{acc,sim}(\cdot)$	1.9	2.6	0.9	1.9	1.7	1.2	0.7	1.7	2.3	2.3	1.00	n.a.
$\delta_{temp}(\cdot)$	2.7	4.4	1.0	1.8	3.0	2.9	1.1	1.9	4.7	4.9	0.99	0.01
$\delta_{acc,temp}(\cdot)$	1.9	2.6	0.9	1.9	1.7	1.2	0.7	1.7	2.3	2.3	1.00	0.00

Tabular B.1: Summary of indentified model parameters and threshold values of the dynamic masking model developed in Section 5.5.

C Collision Probability Computation Based on Reachable Sets

C.1 Probabilistic Reachable Sets

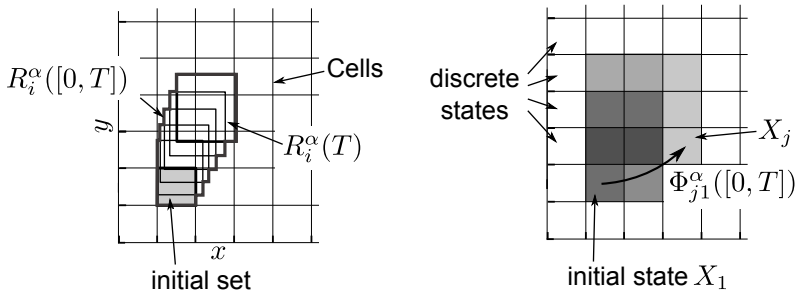
Without loss of generality, we will constraint our considerations to the time interval $[0, T]$. We denote $R([0, T]) \in \mathbb{R}^4$ the reachable set of states from an initial set of states X within the time interval $[0, T]$ given that the system input is bounded between $[u_1, u_2]$. An illustration of the reachable set is depicted in Figure C.1a. The exact reachable set can be calculated analytically for this problem in principle [185], but only for a limited class of input signals that are to be known in advance. This assumption is violated as our input signal is determined by the (unknown) behavior of the human operator. Instead, an overapproximation of the reachable set is calculated using zonotopes [172]. While information about the reachability of a certain position at a certain velocity is binary ("yes"/"no"), we are more interested in the probability of the avatar actually being in that specific state. By abstracting the continuous dynamics with a discrete time Markov chain, these probability distribution can be calculated. Discrete time Markov chains require, besides time discretization, state space and input discretization in order to calculate the entries $\Phi^{\alpha}_{ji}([0, T])$ of the transition matrix $\Phi^{\alpha}([0, T])$. The state space is discretized into N_X equally-sized discrete states X_i, $i = 1 \ldots N_X$, $X_i \in \mathbb{R}^4$ and the input space into N_u equally-sized inputs U_α, $\alpha = 1 \ldots N_u$, $U_\alpha \in \mathbb{R}^2$. We denote $\Phi^{\alpha}_{ji}([0, T])$ as the transition probability of a discrete state X_j to a discrete state X_i within the time interval $[0, T]$, given the discrete input U_α =const.. It is calculated as the volumetric fraction of the reachable set $R^{\alpha}_{i}([0, T])$ intersecting with the cell X_j as $\mathcal{V}(R^{\alpha}_{i}([0, T]) \cap X_j)/\mathcal{V}(X_j)$, where $\mathcal{V}$ is the volumetric integral operator, $R^{\alpha}_{i}([0, T])$ denotes the reachable set, starting from $\mathbf{x}_{perf}(0) \in X_i$ under the effect of $u \in U_\alpha$.

An illustration of the reachable set and the corresponding Markov chain is shown in Figure C.1.

C.2 Collision Probability Calculation

The probability for passing a state X_i in the time interval $[0,\ T]$ when starting from an initial state X_0 can be computed as

$$p_i([0,\ T]) = \sum_{j=1}^{N_X} \Phi^{\alpha}_{ji}([0,\ T])\, p_j(0)\, q_i, \tag{C.1}$$



(a) 2D-projection of the reachable set.

(b) Abstraction by a Markov chain.

Figure C.1: (a) A 2-dimensional projection of the 4D reachable set $R_i^\alpha([0,T])$ for the dynamical system in (6.10) under the influence of an input $u \in U_\alpha$. (b) Abstraction by a Markov chain with transition probabilities $\Phi_{ji}^\alpha([0,T])$. Darker states are more probable to be reached from initial state X_1.

where

$$p_j(0) = \begin{cases} 0 & \text{if } X_j \neq X_0 \\ 1 & \text{if } X_j = X_0 \end{cases} \text{ and } q_i = \begin{cases} 0 & \text{if the obstacle intersects with } pos(X_i), \\ 1 & \text{otherwise.} \end{cases} \tag{C.2}$$

The $pos(\cdot)$ operator extracts the position components out of a discrete state. As the probability to be in a state that intersects with an obstacle is 0 due to the definition of q_i, the actual collision probability within the time interval can be computed as

$$p_{col}([0,T]) = 1 - \sum_{i=1}^{N_X} p_i([0,T]). \tag{C.3}$$

To ensure that positions on the opposite of an obstacle are still recognized as resulting in a collision if the predicted path goes through it, the size of the obstacles is virtually enlarged in direction of movement using a raytracing algorithm. Note that the definition of q_i results in an overapproximation of the collision probability $p_{col}([0,T])$ as it considers the state X_i to result in a collision if the wall position intersects with the discrete state regardless of the intersection area.

C.2.1 Parametrization

For the calculation of $p_{col}([0,T])$ in the experiments described in Sections 6.4, 6.5 and 6.6, a concrete parametrization of the dynamic performance model is required: The position space is discretized between the maximum and minimum reachable distance in x and y direction into 41 intervals, respectively. The number of velocity states is set to 21 in each direction, leading to a total number of $41 \times 41 \times 21 \times 21 = 741321$ discrete states X. For the input vector, consisting of the forces exerted by the operator in both directions, one interval $[-f_{h,max}, f_{h,max}]$ where $f_{h,max} = 8\text{N}$ is chosen based on an analysis of the maximum force in the recorded datasets of all participants.

C.2.2 Collision Probability Computation over Larger Time Intervals

Within a prediction time interval $[0,\ T]$ the force exerted by the human operator is assumed to be constant. To allow for changing forces in the case of predictions beyond T, it is necessary to make predictions over more than one sample time. Again, without loss of generality, we restrict our considerations to the case $t = 0$. The prediction horizon t_p is chosen to be a multiple of T as $t_p = N_p T$. The probability that the human operator collides with an obstacle within the prediction horizon $[0, N_p T]$ is computed over probabilities of being in a state X_i within a number of future time intervals $[kT,\ (k+1)T]$ where $k = 0 \dots N_p - 1$ as

$$p_i([kT,\ (k+1)T])|_{T_d([0,kT])} = \sum_{j=1}^{N_X} \Phi^{\alpha}_{ji}([0,\ T])|_{T_d(kT)}\, p_j(kT)|_{T_d([0,(k-1)T])}\, q_i, \tag{C.4}$$

$$p_i((k+1)T)|_{T_d([0,kT])} = \sum_{j=1}^{N_X} \Phi^{\alpha}_{ji}(T)|_{T_d(kT)}\, p_j(kT)|_{T_d([0,(k-1)T])}\, q_i.$$

The probability q_i is defined by equation (C.2). It can be noted that $p_i([kT,\ (k+1)T])$ does not only depend on the initial state but also on the time delay in all prior time intervals $T_d([0,\ kT])$. This is due to the fact that time delay changes the effective inertia and with it the state transition matrix $\Phi^{\alpha}([0,\ T])$. The collision probability for future time intervals is then

$$p_{col}([kT,(k+1)T])|_{T_d([0,kT])} = 1 - \sum_i p_i([kT,(k+1)T])|_{T_d([0,kT])}. \tag{C.5}$$

C.2.3 Online-Computation

The state space is quantized as described in Section 6.4.3 and the set of possible time delay levels is defined as $\Gamma_d = \{0 \quad 0.05 \quad 0.1 \quad 0.15 \quad 0.2\}$s. The resulting size of the transition matrices $\Phi^{\alpha}([0,\ T])|_{T_d \in \Gamma_d}$ and $\Phi^{\alpha}(T)|_{T_d \in \Gamma_d}$, and the number of probability distributions $p([kT,\ (k+1)T])$ required to compute all collision probabilities $p_{col}([kT,(k+1)T])|_{T_d([0,kT])}$ within the prediction horizon $[0,\ t_p]$ is too large to be feasible for online-computation of the collision probability. Thus, a number of modifications are made to the algorithm to provide $p_{col}([0,\ t+t_p])$ fast enough for an online control of network quality.

Constant Time Delay $T_d([0,\ t_p]) = T_d(0)$

The time delay is constrained to be constant over the whole prediction horizon: $T_d([0,\ t_p]) = T_d(0)$.

This modification reduces the number of probabilities $p_i([kT,\ (k+1)T])|_{T_d(0)}$ by a factor $(||\Gamma_d|| - 1)^{N_p}$ where $||\Gamma_d|| = 5$ denotes the number of allowed time delay levels. This simplification may though result in a suboptimal solution to the optimization problem posed in equation (6.12). However, an offline-simulation of the experimental data recorded in the experiment described in Section 6.4 with the original method from equations (C.4)-(C.5) revealed that the optimal network quality control policy before an actual collision occurs is

applying the minimum time delay over all prediction steps. As this marginal case is included into the reduced set of permissible delay conditions, suboptimal solutions may be primarily computed in the case of no imminent collision, resulting in eventually higher network costs.

Precalculation of $p_i([kT,\,(k+1)T])|_{T_d(0)}$

The probabilities $p_i([kT,\,(k+1)T])|_{T_d(0)}$ in equation (C.4) are precalculated for all possible time delay levels $T_d(0) \in \Gamma_d$, for all prediction intervals $k = 0 \ldots N_p - 1$, and for all possible initial states leading to different distributions $p(0)$. It is furthermore set that no obstacle is within the reachable area $q_i = 1 \forall i = 1 \ldots N_X$.

Precalculating the probability for being in a specific state while neglecting the environment and potential collisions brings the advantage that the large state transition matrices $\Phi([0,\,T])|_{T_d \in \Gamma_d}$ and $\Phi(T)|_{T_d \in \Gamma_d}$ must not be known at runtime. Furthermore, the number of initial probability distributions $p(0)$ can be significantly reduced by considering only states containing the origin of the $x-y$-plane, and those that have only velocity components in positive y-direction. By choosing a coordinate system that is fixed to the inertial frame of the telerobot moving in positive y-direction, all other initial conditions can be ignored. The disadvantage of this procedure is that the real probability distribution considering obstacle collisions can not be calculated exactly from the non-colliding probabilities. This is due to the fact that q_i influences the value for $p_i((k+1)T)$ which determines all subsequent values of $p_i([kT,\,(k+1)T])|_{T_d(0)}$.

Discarding Velocity Components

Velocities are not needed to detect a collision as we consider the collision speed as irrelevant. The probability to reach a specific position interval in a future time interval $p_i^{pos}([kT,\,(k+1)T])|_{T_d(0)}$ is calculated, removing velocity components from the states as they are extraneous for a collision. The probability to pass a specific position interval within the time frame $[kT,\,(k+1)T]$ is calculated as

$$p_i^{pos}([kT,\,(k+1)T])|_{T_d(0)} = \sum_{j \in \Omega_i^{pos}} p_j([kT,\,(k+1)T])|_{T_d(0)} \tag{C.6}$$

where Ω_i^{pos} contains all states X_j that have the same range of position states x and y.

Approximation of $p_{col}([kT,\,(k+1)T])$

The collision probability is estimated based on the resulting distribution $p^{pos}([kT,\,(k+1)T])|_{T_d(0)}$, taking account of inaccuracies introduced by the modifications above. The collision probability in a time interval is approximated by

$$p_{col}([kT,\,(k+1)T])|_{T_d(0) \in \Gamma_d} \approx 1 - \sum_i p_i^{pos}([kT,\,(k+1)T])|_{T_d(0) \in \Gamma_d} \hat{q}_i^{pos} \tag{C.7}$$

where $\hat{q}_i^{pos}$ contains the position information of the obstacles within the reachable range, similar to the definition in equation (C.2). The main difference to q_i is, despite that only position information is contained, that all obstacles are enlarged to compensate for the approximation inaccuracy introduced by the precalculation of $p_i^{pos}([kT, (k+1)T])$. The magnitude of enlargement is estimated from a worst-case consideration, respecting the system dynamics and the operator's maximum force.

Estimation of $p_{col}([0, N_p T])$

As we are only interested in the collision probability over the whole prediction horizon, we can reduce the information to be taken into consideration for the minimization problem in equation (6.12) by selecting the maximum value of $p_{col}([kT, (k+1)T])|_{T_d(0)}$, $k = 0 \dots N_p - 1$. This reduces the optimization problem to choosing the minimum of $||\Gamma_d|| = 5$ values.

Taken all modifications together, the optimal communication quality parameter $\theta_C(t)$ for each time instance t is calculated following these steps:

- Compute $\hat{q}^{pos}$ by measurement of the obstacle location in the $x-y$-plane, discretization and enlargement of the obstacles.
- Compute $p_{col}([kT, (k+1)T])$ for all values of $T_d \in \Gamma_d$ and all $k = 0 \dots N_p - 1$, using precalculated probability distributions $p^{pos}([kT, (k+1)T])$.
- Compute $p_{col}([0, t_p]) = \max_{k=0 \dots N_p - 1} p_{col}([kT, (k+1)T])$ for all values of $T_d \in \Gamma_d$.
- Solve the minimization problem in equation (6.12).

Bibliography

[1] K. Park and S. Lee, "An Engineering Model of the Masking for the Noise-Robust Speech Recognition," *Neurocomputing*, vol. 52, pp. 615–620, 2003.

[2] R. Bain, "Technology and State Government," *American Sociological Review*, vol. 2, no. 6, pp. 860–874, 1937.

[3] A. Peer, *Design and Control of Admittance-Type Telemanipulation Systems*. Dissertation, Technische Universität München, München, 2008.

[4] J. Colgate and J. Brown, "Factors Affecting the z-Width of a Haptic Display," in *Proceedings of the 1994 IEEE International Conference on Robotics and Automation*, pp. 3205–3210, 1994.

[5] D. Lawrence, "Stability and Transparency in Bilateral Teleoperation," *IEEE Transactions on Robotics and Automation*, vol. 9, no. 5, pp. 624–637, 1993.

[6] T. Brooks, "Telerobotic Response Requirements," in *Proceedings of the 1990 IEEE International Conference on Systems, Man and Cybernetics*, pp. 113–120, 1990.

[7] H. Tan, M. Srinivasan, B. Eberman, and B. Cheng, "Human Factors for the Design of Force-Reflecting Haptic Interfaces," *Dynamic Systems and Control*, vol. 55, no. 1, pp. 353–359, 1994.

[8] S. Hirche, *Haptic Telepresence in Packet Switched Communication Networks*. Dissertation, Technische Universität München, München, 2005.

[9] W. Warren, "Action Modes and Laws of Control for the Visual Guidance Of Action," *Advances in Psychology*, vol. 50, pp. 339–379, 1988.

[10] P. Kugler and M. Turvey, *Information, Natural Law, and the Self-Assembly of Rhythmic Movement*. Lawrence Erlbaum Associates, Inc, 1987.

[11] L. Jones and I. Hunter, "A Perceptual Analysis of Stiffness," *Experimental Brain Research*, vol. 79, no. 1, pp. 150–156, 1990.

[12] L. Jones and I. Hunter, "A Perceptual Analysis of Viscosity," *Experimental Brain Research*, vol. 94, no. 2, pp. 343–351, 1993.

[13] B. Deml, "Human Factors Issues on the Design of Telepresence Systems," *Presence: Teleoperators and Virtual Environments*, vol. 16, no. 5, pp. 471–487, 2007.

[14] G. Beauregard, M. Srinivasan, and N. Durlach, "The Manual Resolution of Viscosity and Mass," in *Proceedings of the ASME Dynamic Systems and Control Division*, vol. 1, pp. 657–662, 1995.

[15] X.-D. Yang, W. Bischof, and P. Boulanger, "Perception of Haptic Force Magnitude During Hand Movements," in *Proceedings of the 2008 IEEE International Conference on Robotics and Automation*, pp. 2061–2066, 2008.

[16] H. Tan, N. Durlach, G. Beauregard, and M. Srinivasan, "Manual Discrimination of Compliance Using Active Pinch Grasp: The roles of Force and Work Cues," *Attention, Perception, & Psychophysics*, vol. 57, no. 4, pp. 495–510, 1995.

[17] A. Kaas, M. Stoeckel, and R. Goebel, *Human Haptic Perception: Basics and Applications*, ch. The neural bases of haptic working memory, pp. 113–129. Springer, 2008.

[18] S. Hirche, A. Bauer, and M. Buss, "Transparency of Haptic Telepresence Systems With Constant Time Delay," in *Proceedings of the 2005 IEEE Conference on Control Applications*, pp. 328–333, 2005.

[19] Encyclopædia Britannica, ""engineering"." http://www.britannica.com/EBchecked/topic/187549/engineering, 2012.

[20] R. Williams and K. Herrup, "The Control of Neuron Number," *Annual Review of Neuroscience*, vol. 11, no. 1, pp. 423–453, 1988.

[21] L. Ljung, *System Identification*. PTR Prentice Hall, Upper Saddle River, N.J., 2nd ed., 1999.

[22] L. Ljung, *The Control Handbook*, ch. 58 - System Identification, pp. 1033–1054. CRC, 1996.

[23] J.-N. Juang, *Applied System Identification*. Prentice HallEnglewood Cliffs, New Jersey, 1994.

[24] S. Hirche and M. Buss, *Advances in Telerobotics*, ch. 12, Human Perceived Transparency with Time Delay, pp. 191–209. Springer, 2007.

[25] Z. Halata and K. Baumann, *Human Haptic Perception: Basics and Applications*, ch. Anatomy of receptors, pp. 85–92. Springer, 2008.

[26] S. Gandevia, D. McCloskey, and D. Burke, "Kinaesthetic Signals and Muscle Contraction," *Trends in neurosciences*, vol. 15, no. 2, pp. 62–65, 1992.

[27] R. Adams and B. Hannaford, "Stable Haptic Interaction With Virtual Environments," *IEEE Transactions on Robotics and Automation*, vol. 15, no. 3, pp. 465–474, 1999.

[28] M. Lin and M. Otaduy, *Haptic Rendering: Foundations, Algorithms and Applications*. AK Peters, Ltd., 2008.

[29] R. J. Anderson and M. W. Spong, "Bilateral Control of Teleoperators With Time Delay," *IEEE Transactions on Automatic Control*, vol. AC-34, no. 5, pp. 494–501, 1989.

[30] P. Arcara and C. Melchiorri, "Control Schemes for Teleoperation With Time Delay: A Comparative study," *Robotics and Autonomous Systems*, vol. 38, no. 1, pp. 49–64, 2002.

[31] K. Hashtrudi-Zaad and S. E. Salcudean, "Analysis of Control Architectures for Teleoperation Systems With Impedance/Admittance Master and Slave Manipulators," *The International Journal of Robotics Research*, vol. 20, no. 6, pp. 419–445, 2001.

[32] P. F. Hokayem and M. W. Spong, "Bilateral Teleoperation: An Historical Survey," *Automatica*, vol. 42, no. 12, pp. 2035–2057, 2006.

[33] S. Munir and W. Book, "Internet-Based Teleoperation Using Wave Variables With Prediction," *IEEE/ASME Transactions on Mechatronics*, vol. 7, no. 2, pp. 124–133, 2002.

[34] G. Niemeyer and J. J. E. Slotline, "Stable Adaptive Teleoperation," *IEEE Journal of Oceanic Engineering*, vol. 16, no. 1, pp. 152–162, 1991.

[35] T. Schauß and A. Peer, "Parameter-Space Transparency Analysis of Teleoperation Systems," in *Proceedings of the 2012 IEEE Haptics Symposium*, pp. 111–116, 2012.

[36] A. Shahdi and S. Sirouspour, "Adaptive/Robust Control for Time-Delay Teleoperation," *IEEE Transactions on Robotics*, vol. 25, no. 1, pp. 196–205, 2009.

[37] T. Sheridan, "Space Teleoperation Through Time Delay: Review and Prognosis," *IEEE Transactions on Robotics and Automation*, vol. 9, no. 5, pp. 592 –606, 1993.

[38] I. Vittorias and S. Hirche, "Stable Teleoperation With Communication Unreliabilities and Approximate human/environment Dynamics Knowledge," in *Proceedings of the 2010 American Control Conference*, pp. 2791–2796, 2010.

[39] G. Hirzinger, B. Brunner, J. Dietrich, and J. Heindl, "Sensor-Based Space Robotics-ROTEX and its Telerobotic Features," *IEEE Transactions on Robotics and Automation*, vol. 9, no. 5, pp. 649–663, 1993.

[40] C. J. Lane, C. R. Carignan, B. R. Sullivan, D. L. Akin, T. Hunt, and R. Cohen, "Effects of Time Delay on Telerobotic Control of Neutral Buoyancy vehicles," in *Proceedings of the 2002 IEEE International Conference on Robotics and Automation*, vol. 3, pp. 2874–2879, 2002.

[41] A. Bejczy, W. Kim, and S. Venema, "The Phantom Robot: Predictive Displays for Teleoperation With Time delay," in *Proceedings of the 1990 IEEE International Conference on Robotics and Automation*, vol. 1, pp. 546–551, 1990.

[42] A. Peer, S. Hirche, C. Weber, I. Krause, M. Buss, S. Miossec, P. Evrard, O. Stasse, E. S. Neo, A. Kheddar, and K. Yokoi, "Intercontinental Cooperative Telemanipulation Between Germany and Japan," in *Proceedings of the 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2715–2716, 2008.

[43] G. Niemeyer, C. Preusche, and G. Hirzinger, *Handbook of Robotics*, ch. 31 Telerobotics, pp. 741–757. Springer, 2008.

[44] S. Hirche, M. Ferre, J. Barrio, C. Melchiorri, and M. Buss, *Advances in Telerobotics*, ch. Bilateral Control Architectures for Telerobotics, pp. 163–176. Springer, 2007.

[45] A. Jacoff, E. Messina, B. Weiss, S. Tadokoro, and Y. Nakagawa, "Test Arenas and Performance Metrics for Urban Search and Rescue Robots," in *Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems. IROS 2003*, vol. 4, pp. 3396–3403, 2003.

[46] L. Whitcomb, "Underwater Robotics: Out of the Research Laboratory and Into the Field," in *Proceedings of the 2000 IEEE International Conference on Robotics and Automation*, vol. 1, pp. 709–716, 2000.

[47] G. Guthart and J. Salisbury Jr, "The Intuitive™Telesurgery System: Overview and Application," in *Proceedings of the 2000 IEEE International on Robotics and Automation. ICRA'00*, vol. 1, pp. 618–621, 2000.

[48] G. Raju, G. Verghese, and T. Sheridan, "Design Issues in 2-Port Network Models of Bilateral Remote Manipulation," in *Proceedings of the 1989 IEEE International Conference on Robotics and Automation*, pp. 1316–1321, 1989.

[49] Y. Yokokohji and T. Yoshikawa, "Bilateral Control of Master-Slave Manipulators for Ideal Kinesthetic coupling-Formulation and Experiment," *IEEE Transactions on Robotics and Automation*, vol. 10, no. 5, pp. 605–620, 1994.

[50] S. Hirche and M. Buss, "Human-Oriented Control for Haptic Teleoperation," *Proceedings of the IEEE*, vol. 100, no. 3, pp. 623–647, 2012.

[51] B. Witmer and M. Singer, "Measuring Presence in Virtual Environments: A Presence Questionnaire," *Presence*, vol. 7, no. 3, pp. 225–240, 1998.

[52] J. Freeman, S. Avons, R. Meddis, D. Pearson, and W. IJsselsteijn, "Using Behavioral Realism to Estimate Presence: A Study of the Utility of Postural Responses to Motion Stimuli," *Presence: Teleoperators & Virtual Environments*, vol. 9, no. 2, pp. 149–164, 2000.

[53] J. Van Baren and W. IJsselsteijn, "Measuring Presence: A Guide to Current Measurement Approaches," tech. rep., OmniPres project IST-2001-39237, 2004.

[54] P. Berestesky, N. Chopra, and M. Spong, "Discrete Time Passivity in Bilateral Teleoperation Over the Internet," in *Proceedings of the 2004 IEEE International Conference on Robotics and Automation*, vol. 5, pp. 4557–4564, 2004.

[55] B. Hannaford and J. H. Ryu, "Time-Domain Passivity Control of Haptic Interfaces," *IEEE Transactions on Robotics and Automation*, vol. 18, no. 1, pp. 1–10, 2002.

[56] S. Hirche and M. Buss, "Packet Loss Effects in Passive Telepresence Systems," in *Proceedings of the 43rd IEEE Conference on Decision and Control, 2004. CDC.*, vol. 4, pp. 4010–4015, 2004.

[57] G. A. Gescheider, *Psychophysics: Method, Theory, and Application*. Lawrence Erlbaum, 1985.

[58] B. Treutwein and H. Strasburger, "Fitting the Psychometric Function.," *Attention, Perception, & Psychophysics*, vol. 61, no. 1, pp. 87–106, 1999.

[59] Z. Shi, S. Hirche, W. X. Schneider, and H. Muller, "Influence of Visuomotor Action on Visual-Haptic Simultaneous Perception: A Psychophysical Study," in *Proceedings of the 2008 Symposium on Haptic interfaces for Virtual Environment and Teleoperator Systems*, pp. 65–70, 2008.

[60] C. Kaernbach, "Adaptive Threshold Estimation With Unforced-Choice Tasks," *Perception & Psychophysics*, vol. 63, no. 8, pp. 1377–1388, 2001.

[61] H. Kesten, "Accelerated Stochastic Approximation," *Annals of Mathematical Statistics*, vol. 29, pp. 41–59, 1958.

[62] E. Weber, *Die Lehre vom Tastsinne und Gemeingefühle auf Versuche Gegründet*. Friedrich Vieweg und Sohn, 1851.

[63] J. Kreifeldt and M. Chuang, "Moment of Inertia: Psychophysical Study of an Overlooked Sensation," *Science*, vol. 206, no. 4418, pp. 588–590, 1979.

[64] P. Hinterseer, S. Hirche, S. Chaudhuri, E. Steinbach, and M. Buss, "Perception-Based Data Reduction and Transmission of Haptic Data in Telepresence and Teleaction Systems," *IEEE Transactions on Signal Processing*, vol. 56, no. 2, pp. 588–597, 2008.

[65] S. Stevens, "On the Psychophysical Law.," *Psychological Review*, vol. 64, no. 3, pp. 153–181, 1957.

[66] E. Hellier and J. Edworthy, "On Using Psychophysical Techniques to Achieve Urgency Mapping in Auditory Warnings," *Applied Ergonomics*, vol. 30, no. 2, pp. 167–171, 1999.

[67] P. Smith, "Stochastic Dynamic Models of Response Time and Accuracy: A Foundational Primer," *Journal of Mathematical Psychology*, vol. 44, no. 3, pp. 408–463, 2000.

[68] R. Ratcliff, "A Diffusion Model Account of Response Time and Accuracy in a Brightness discrimination Task: Fitting Real Data and Failing to fit Fake but plausible Data," *Psychonomic Bulletin & Review*, vol. 9, pp. 278–291, 2002.

[69] T. Pleskac and J. Busemeyer, "Two-Stage Dynamic Signal Detection: A Theory of Choice, Decision time, and Confidence.," *Psychological Review*, vol. 117, no. 3, pp. 864–901, 2010.

[70] D. Luenberger, "Observing the State of a Linear System," *IEEE Transactions on Military Electronics*, vol. 8, no. 2, pp. 74–80, 1964.

[71] P. Fitts, "The Information Capacity of the Human Motor System in Controlling the Amplitude of Movement.," *Journal of Experimental Psychology*, vol. 47, no. 6, p. 381, 1954.

[72] I. MacKenzie and W. Buxton, "Extending Fitts' law to two-Dimensional Tasks," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 219–226, 1992.

[73] A. Murata and H. Iwase, "Extending Fitts' law to a Three-Dimensional Pointing Task," *Human movement science*, vol. 20, no. 6, pp. 791–805, 2001.

[74] E. Hoffmann, "Fitts' Law With Transmission Delay.," *Ergonomics*, vol. Jan, pp. 37–48, 1992.

[75] F. Lacquaniti, C. Terzuolo, and P. Viviani, "The law Relating the Kinematic and Figural Aspects of Drawing Movements," *Acta Psychologica*, vol. 54, no. 1-3, pp. 115–130, 1983.

[76] D. McRuer and H. Jex, "A Review of Quasi-Linear Pilot Models," *IEEE Transactions on Human Factors in Electronics*, vol. HFE-8, no. 3, pp. 231–249, 1967.

[77] E. Bizzi, N. Hogan, F. Mussa-Ivaldi, and S. Giszter, "Does the Nervous System use Equilibrium-Point Control to Guide Single and Multiple Joint Movements?," in *Proceedings of the Controversies in Neuroscience Conference: I. Movement control (1990, Portland, Oregon).*, pp. 603–613, 1992.

[78] A. Feldman and M. Levin, "The Origin and use of Positional Frames of Reference in Motor Control," *Behavioral and Brain Sciences*, vol. 18, no. 4, pp. 723–744, 1995.

[79] H. Gomi and M. Kawato, "Equilibrium-Point Control Hypothesis Examined by Measured arm Stiffness during Multijoint Movement," *Science*, vol. 272, no. 5258, pp. 117–120, 1996.

[80] H. Spence, *The Principles of Biology*. D. Appleton and company, 1864.

[81] C. Chow and D. Jacobson, "Studies of Human Locomotion via Optimal Programming," *Mathematical Biosciences*, vol. 10, no. 3-4, pp. 239–306, 1971.

[82] N. Hogan, "An Organizing Principle for a Class of Voluntary Movements," *The Journal of Neuroscience*, vol. 11, no. 11, pp. 2745–2754, 1984.

[83] Y. Uno, M. Kawato, and R. Suzuki, "Formation and Control of Optimal Trajectory in Human Multijoint arm Movement," *Biological Cybernetics*, vol. 61, no. 2, pp. 89–101, 1989.

[84] E. Todorov, "Optimality Principles in Sensorimotor Control," *Nature Neuroscience*, vol. 7, no. 9, pp. 907–915, 2004.

[85] P. Bays and D. Wolpert, "Computational Principles of Sensorimotor Control That Minimize Uncertainty and Variability," *The Journal of Physiology*, vol. 578, no. 2, pp. 387–396, 2007.

[86] A. Faisal, L. Selen, and D. Wolpert, "Noise in the Nervous System," *Nature Reviews Neuroscience*, vol. 9, no. 4, pp. 292–303, 2008.

[87] C. Harris and D. Wolpert, "Signal-Dependent Noise Determines Motor Planning," *Nature*, vol. 394, no. 6695, pp. 780–784, 1998.

[88] D. M. Wolpert and Z. Ghahramani, "Computational Principles of Movement in Neuroscience," *Nature Neuroscience Supplement*, vol. 3, pp. 1212–1217, 2000.

[89] M. Ito, "Neurophysiological Aspects of the Cerebellar Motor Control System.," *International Journal of Neurology*, vol. 7, no. 2, pp. 162–176, 1970.

[90] M. Kawato, K. Furukawa, and R. Suzuki, "A Hierarchical Neural-Network Model for Control and Learning of voluntary Movement," *Biological Cybernetics*, vol. 57, no. 3, pp. 169–185, 1987.

[91] M. Kawato and H. Gomi, "The Cerebellum and VOR/OKR Learning Models," *Trends in Neurosciences*, vol. 15, no. 11, pp. 445–453, 1992.

[92] S. Kitazawa, T. Kimura, and P. Yin, "Cerebellar Complex Spikes Encode Both Destinations and Errors in arm Movements," *Nature*, vol. 392, no. 6675, pp. 494–497, 1998.

[93] M. Shidara, K. Kawano, H. Gomi, and M. Kawato, "Inverse-Dynamics Model eye Movement Control by Purkinje Cells in the Cerebellum," *Nature*, vol. 365, no. 2, pp. 50–52, 1993.

[94] E. Todorov and M. Jordan, "Optimal Feedback Control as a Theory of Motor Coordination," *Nature neuroscience*, vol. 5, no. 11, pp. 1226–1235, 2002.

[95] J. Saunders and D. Knill, "Visual Feedback Control of Hand Movements," *The Journal of Neuroscience*, vol. 24, no. 13, pp. 3223–3234, 2004.

[96] K. Körding and D. Wolpert, "Bayesian Decision Theory in Sensorimotor Control," *Trends in Cognitive Sciences*, vol. 10, no. 7, pp. 319–326, 2006.

[97] M. Sherback and R. D'Andrea, "An Optimal Control Approach to Sensorimotor Parametrization," in *Proceedings of the 2007 American Control Conference*, pp. 2036–2041, 2007.

[98] H. Imamizu, S. Miyauchi, T. Tamada, Y. Sasaki, R. Takino, B. Putz, T. Yoshioka, and M. Kawato, "Human Cerebellar Activity Reflecting an Acquired Internal Model of a new Tool," *Nature*, vol. 403, no. 6766, pp. 192–195, 2000.

[99] R. Shadmehr and F. Mussa-Ivaldi, "Adaptive Representation of Dynamics During Learning of a Motor Task," *The Journal of Neuroscience*, vol. 14, no. 5, pp. 3208–3224, 1994.

[100] A. Steinfeld, T. Fong, D. Kaber, M. Lewis, J. Scholtz, A. Schultz, and M. Goodrich, "Common Metrics for Human-Robot Interaction," in *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-robot Interaction*, pp. 33–40, 2006.

[101] M. F. Zäh and S. M. Clarke, "Flexible and Configurable Haptic Environments Through Virtual Graphical Force Sensors," in *Proceedings of the 2004 EuroHaptics Conference*, pp. 486–489, 2004.

[102] A. Lecuyer, J.-M. Burkhardt, J. Le Biller, and M. Congedo, ""A4": a Technique to Improve Perception of Contacts With Under-Actuated haptic Devices in Virtual Reality," in *Proceedings of the 2005 First Joint Eurohaptics Conference, and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics*, pp. 316–322, 2005.

[103] C. Passenberg, A. Peer, and M. Buss, "A Survey of Environment-, Operator-, and Task-Adapted Controllers for Teleoperation Systems," *Mechatronics*, vol. 20, no. 7, pp. 787–801, 2010. Special Issue on Design and Control Methodologies in Telerobotics.

[104] C. Lawn and B. Hannaford, "Performance Testing of Passive Communication and Control in Teleoperation with Time Delay," in *Proceedings of the 1993 IEEE International Conference on Robotics and Automation*, vol. 3, pp. 776–783, 1993.

[105] E. Steinbach, S. Hirche, J. Kammerl, I. Vittorias, and R. Chaudhari, "Haptic Data Compression and Communication," *Signal Processing Magazine, IEEE*, vol. 28, no. 1, pp. 87–96, 2011.

[106] I. Vittorias, H. Ben Rached, and S. Hirche, "Haptic Data Reduction in Multi-DoF Teleoperation Systems," in *Proceedings of the 2010 IEEE International Symposium on Haptic Audio-Visual Environments and Games (HAVE)*, pp. 1–6, 2010.

[107] C. Smith and P. Jensfelt, "A Predictor for Operator Input for Time-Delayed Teleoperation," *Mechatronics*, vol. 20, no. 7, pp. 778–786, 2010.

[108] M. B. I. Raez, M. S. Hussain, and F. Mohd-Yasin, "Techniques of EMG Signal Analysis: Detection, Processing, Classification and Applications.," *Biological Procedures Online*, vol. 8, no. 1, pp. 11–35, 2006.

[109] M. C. Stüttgen, C. Schwarz, and F. Jäkel, "Mapping Spikes to Sensations," *Frontiers in Neuroscience*, vol. 5, pp. 125/1–17, 2011.

[110] P. Kammermeier, M. Buss, and G. Schmidt, "A Systems Theoretical Model for Human Perception in Muti-Modal Presence Systems," *IEEE/ASME Transactions on Mechatronics*, vol. 6, no. 3, pp. 1–11, 2001.

[111] W. Warren, "The Dynamics of Perception and Action," *Psychological Review*, vol. 113, no. 2, pp. 358–389, 2006.

[112] J. A. S. Kelso, *Dynamic Patterns: The Self-Organization of Brain and Behavior*. MIT Press, 1995.

[113] O. Khatib, E. Demircan, V. De Sapio, L. Sentis, T. Besier, and S. Delp, "Robotics-Based Synthesis of Human Motion," *Journal of Physiology*, vol. 103, no. 3, pp. 211–219, 2009.

[114] B. Delgutte, *Springer Handbook of Auditory Research*, vol. 6, ch. Physiological models for basic auditory percepts, pp. 157–220. Springer-Verlag, 1996.

[115] E. Simoncelli and D. Heeger, "A Model of Neuronal Responses in Visual Area MT," *Vision Research*, vol. 38, no. 5, pp. 743–761, 1998.

[116] W. Loewenstein and R. Skalak, "Mechanical Transmission in a Pacinian Corpuscle. An Analysis and a Theory," *The Journal of Physiology*, vol. 182, no. 2, pp. 346–378, 1966.

[117] A. Hof and J. Van Den Berg, "EMG to Force Processing I: an Electrical Analogue of the Hill Muscle model," *Journal of Biomechanics*, vol. 14, no. 11, pp. 747–753, 1981.

[118] K. Jones, A. Hamilton, and D. Wolpert, "Sources of Signal-Dependent Noise During Isometric Force Production," *Journal of Neurophysiology*, vol. 88, no. 3, pp. 1533–1544, 2002.

[119] W. Prinz and B. Hommel, *Common Mechanisms in Perception and Action: Attention and Performance XIX*, vol. 67. Oxford University Press, 2001.

[120] G. Rizzolatti and L. Craighero, "The mirror-neuron system," *Annu. Rev. Neurosci.*, vol. 27, pp. 169–192, 2004.

[121] C. Jay, M. Glencross, and R. Hubbold, "Modeling the Effects of Delayed Haptic and Visual Feedback in a Collaborative Virtual Environment," *ACM Transactions on Computer-Human Interaction*, vol. 14, no. 2, pp. 8/1–31, 2007.

[122] S. J. Blakemore, S. J. Goodbody, and D. M. Wolpert, "Predicting the Consequences of our own Actions: the Role of Sensorimotor context Estimation," *Journal of Neuroscience*, vol. 18, no. 18, pp. 7511–7518, 1998.

[123] R. C. Miall and D. M. Wolpert, "Forward Models for Physiological Motor Control," *Neural Networks*, vol. 9, pp. 1265–1279, 1996.

[124] T. Dau, D. Püschel, and A. Kohlrausch, "A Quantitative Model of the 'effective' Signal Processing in the Auditory System. I. Model Structure," *The Journal of the Acoustical Society of America*, vol. 99, p. 3615, 1996.

[125] D. Lawrence, L. Pao, A. Dougherty, M. Salada, and Y. Pavlou, "Rate-Hardness: A new Performance Metric for Haptic Interfaces," *IEEE Transactions on Robotics and Automation*, vol. 16, no. 4, pp. 357–371, 2000.

[126] B.-B. Gabriel and L. Scocchia, "Is Mass Invariant? Effects of Movement Amplitude and Duration," in *Proceedings of the 25th Meeting of the International Society for Psychophysics, Fechner Day 2009*, 2009.

[127] A. Pressman, L. Welty, A. Karniel, and F. Mussa-Ivaldi, "Perception of Delayed Stiffness," *The International Journal of Robotics Research*, vol. 26, no. 11-12, pp. 1191–1203, 2007.

[128] N. Macmillan and C. Creelman, *Detection Theory: A User's Guide*. Lawrence Erlbaum, 2005.

[129] E. Niebur and C. Koch, "A Model for the Neuronal Implementation of Selective Visual Attention based on Temporal Correlation Among Neurons," *Journal of Computational Neuroscience*, vol. 1, no. 1, pp. 141–158, 1994.

[130] W. Fujisaki and S. Nishida, "Temporal Frequency Characteristics of Synchrony–Asynchrony Discrimination of Audio-Visual Signals," *Experimental Brain Research*, vol. 166, pp. 455–464, 2005.

[131] Åström, K.J., and Eykhoff,P., "System Identification A Survey," *Automatica*, vol. 7, pp. 123–162, 1971.

[132] A. Peer, S. Hirche, C. Weber, I. Krause, M. Buss, S. Miossec, P. Evrard, O. Stasse, E. S. Neo, A. Kheddar, and K. Yokoi, "Intercontinental Multimodal Tele-Cooperation Using a Humanoid Robot," in *Proceedings of the 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 405–411, 2008.

[133] M. Rank, Z. Shi, H. Müller, and S. Hirche, "Delay Discrimination in Continuous Haptic Environments," in *Proceedings of the EuroHaptics Conference 2010, Lecture Notes in Computer Science*, vol. 6191, pp. 205–212, 2010.

[134] M. Rank, Z. Shi, H. Müller, and S. Hirche, "Perception of Delay in Haptic Telepresence Systems," *Presence: Teleoperators & Virtual Environments*, vol. 19, no. 5, pp. 389–399, 2010.

[135] J. Gil, A. Avello, A. Rubio, and J. Florez, "Stability Analysis of a 1 dof Haptic Interface Using the Routh-Hurwitz criterion," *IEEE Transactions on Control Systems Technology*, vol. 12, no. 4, pp. 583–588, 2004.

[136] N. Hogan, "Controlling Impedance at the Man/machine Interface," in *Proceedings of the 1989 IEEE International Conference on Robotics and Automation*, pp. 1626–1631, 1989.

[137] K. Biswas and G. Singh, "Identification of Stochastic Time-Delay Systems," *IEEE Transactions on Automatic Control*, vol. 23, no. 3, pp. 504–505, 1978.

[138] P. Fraisse, "Perception and Estimation of Time," *Annual Review of Psychology*, vol. 35, pp. 1–36, 1984.

[139] H. Ohnishi and K. Mochizuki, "Effect of Delay of Feedback Force on Perception of Elastic Force: a Psychophysical Approach," *IEICE Transactions on Communications*, vol. 90, no. 1, pp. 12–20, 2007.

[140] M. Di Luca, B. Knörlein, M. Ernst, and M. Harders, "Effects of Visual-Haptic Asynchronies and Loading-Unloading Movements on Compliance Perception," *Brain Research Bulletin*, vol. 85, no. 5, pp. 245–259, 2011.

[141] R. S. Allison, L. R. Harris, M. Jenkin, U. Jasiobedzka, and J. E. Zacher, "Tolerance of Temporal Delay in Virtual Environments," in *Proceedings of the Virtual Reality 2001 Conference (VR'01)*, 2001.

[142] B. D. Adelstein, T. G. Lee, and S. R. Ellis, "Head Tracking Latency in Virtual Environments: Psychophysics and a Model," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, pp. 2083–2087, 2003.

[143] R. Byrd, M. Hribar, and J. Nocedal, "An Interior Point Algorithm for Large-Scale Nonlinear Programming," *SIAM Journal on Optimization*, vol. 9, no. 4, pp. 877–900, 1999.

[144] N. Ando, P. Korondi, and H. Hashimoto, "Development of Micromanipulator and Haptic Interface for Networked micromanipulation," *IEEE/ASME Transactions on Mechatronics*, vol. 6, no. 4, pp. 417–427, 2001.

[145] G. Gescheider, S. Bolanowski, and R. Verrillo, "Vibrotactile Masking: Effects of Stimulus Onset Asynchrony and Stimulus frequency.," *Journal of the Acoustical Society of America*, vol. 85, no. 5, pp. 2059–2064, 1989.

[146] J. Craig, "Vibrotactile Difference Thresholds for Intensity and the Effect of a Masking Stimulus," *Attention, Perception, & Psychophysics*, vol. 15, no. 1, pp. 123–127, 1974.

[147] K. Brandenburg, "MP3 and AAC Explained," in *AES 17th International Conference on High-Quality Audio Coding*, 1999.

[148] B. Wu, R. L. Klatzky, and R. L. Hollis, "Force, Torque, and Stiffness: Interactions in Perceptual Discrimination," *IEEE Transactions on Haptics*, vol. 4, no. 3, pp. 221–228, 2011.

[149] M. Rank, T. Schauß, A. Peer, S. Hirche, and R. L. Klatzky, "Masking Effects for Damping JND," in *Proceedings of the EuroHaptics Conference 2012, Lecture Notes in Computer Science*, vol. 7283, pp. 145–150, 2012.

[150] I. Witten, E. Frank, and M. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, 2011.

[151] P. Mitra and G. Niemeyer, "Model-Mediated Telemanipulation," *The International Journal of Robotics Research*, vol. 27, no. 2, pp. 253–262, 2008.

[152] W. Yoon, T. Goshozono, H. Kawabe, M. Kinami, Y. Tsumaki, M. Uchiyama, M. Oda, *et al.*, "Model-Based Space Robot Teleoperation of ETS-VII Manipulator," *IEEE Transactions on Robotics and Automation*, vol. 20, no. 3, pp. 602–612, 2004.

[153] G. Hirzinger, K. Landzettel, D. Reintsema, C. Preusche, A. A. Schäffer, B. Rebele, and M. Turk, "ROKVISS-Robotics Component Verification on ISS," in *Proceedings of the 8th International Symposium on Artificial Intelligence, Robotics and Automation in Space*, 2005.

[154] J. Y. C. Chen, E. C. Haas, and M. J. Barnes, "Human Performance Issues and User Interface Design for Teleoperated robots," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 37, no. 6, pp. 1231–1245, 2007.

[155] K. Stanney, R. Mourant, and R. Kennedy, "Human Factors Issues in Virtual Environments: A Review of the Literature," *Presence: Teleoperators and Virtual Environments*, vol. 7, no. 4, pp. 327–351, 1998.

[156] T. Beigbeder, R. Coughlan, C. Lusher, J. Plunkett, E. Agu, and M. Claypool, "The Effects of Loss and Latency on User Performance in Unreal Tournament 2003," in *Proceedings of the 3rd ACM SIGCOMM Workshop on Network and System Support for Games*, pp. 144–151, 2004.

[157] R. Ellis, O. Ismaeil, and M. Lipsett, "Design and Evaluation of a High-Performance Haptic Interface," *Robotica*, vol. 14, no. 3, pp. 321–328, 1996.

[158] V. Hayward and O. Astley, "Performance Measures for Haptic Interfaces," in *Proceedings of the 7th International Symposium on Robotics Research*, pp. 195–207, 1996.

[159] A. Erramilli, O. Narayan, and W. Willinger, "Experimental Queueing Analysis With Long-Range Dependent Packet traffic," *IEEE/ACM Transactions on Networking*, vol. 4, no. 2, pp. 209–223, 1996.

[160] J. L. Adams, "An Investigation of the Effects of the Time Lag due to Long Transmission Distances Upon Remote Control," Tech. Rep. D-1211, NASA, 1961.

[161] T. B. Sheridan and W. R. Ferrell, "Remote Manipulative Control With Transmission Delay," *IEEE Transactions on Human Factors in Electronics*, vol. HFE-4, no. 1, pp. 25–29, 1963.

[162] Z. Shi, H. Zou, M. Rank, L. Chen, S. Hirche, and H. J. Müller, "Effects of Packet Loss and Latency on the Temporal Discrimination of Visual-Haptic Events," *IEEE Transactions on Haptics*, vol. 3, pp. 28–36, 2010.

[163] J. C. Mateo, R. H. Gilkey, and J. L. Cowgill, "Effect of Variable Visual-Feedback Delay on Movement Time," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 51/19, pp. 1373–1377, 2007.

[164] R. Chaudhari, C. Schuwerk, V. Nitsch, E. Steinbach, and B. Färber, "Opening the Haptic Loop: Network Degradation Limits for Haptic Task performance," in *Proceedings of the 2011 IEEE International Workshop on Haptic Audio Visual Environments and Games (HAVE)*, pp. 56–61, 2011.

[165] L. B. Rosenberg, "The Use of Virtual Fixtures to Enhance Operator Performance in Time Delayed Teleoperation," tech. rep., Armstrong Laboratory, Biodynamics and Biocommunications Division, Human Systems Center, Air Force Materiel Command, 1993.

[166] C. Aurrecoechea, A. T. Campbell, and L. Hauw, "A Survey of QoS Architectures," *Multimedia Systems*, vol. 6, pp. 138–151, 1998.

[167] X. Xiao and L. Ni, "Internet QoS: a big Picture," *IEEE Network*, vol. 13, no. 2, pp. 8–18, 1999.

[168] Q. Ni, L. Romdhani, and T. Turletti, "A Survey of QoS Enhancements for IEEE 802.11 Wireless LAN," *Wireless Communications and Mobile Computing*, vol. 4, no. 5, pp. 547–566, 2004.

[169] K. Åström and B. Wittenmark, *Adaptive Control*. Addison-Wesley Longman Publishing Co., Inc., 1994.

[170] R. Marler and J. Arora, "Survey of Multi-Objective Optimization Methods for Engineering," *Structural and Multidisciplinary Optimization*, vol. 26, no. 6, pp. 369–395, 2004.

[171] E. Camacho and C. Bordons, *Model Predictive Control*. Springer Verlag, 1999.

[172] M. Althoff, O. Stursberg, and M. Buss, "Stochastic Reachable Sets of Interacting Traffic Participants," in *Proceedings of the 2008 IEEE Intelligent Vehicles Symposium*, pp. 1086–1092, 2008.

[173] M. Althoff, O. Stursberg, and M. Buss, "Model-Based Probabilistic Collision Detection in Autonomous Driving," *IEEE Transactions on Intelligent Transportation Systems*, vol. 10, no. 2, pp. 299–310, 2009.

[174] E. N. Gilbert, "Capacity of a Burst-Noise Channel," *Bell System Technical Journal*, vol. 39, pp. 1253–1265, 1960.

[175] E. O. Elliot, "A Model of the Switched Telephone Network for Data Communications," *Bell System Technical Journal*, vol. 44, pp. 89–109, 1963.

[176] H. Sanneck, *Packet Loss Recovery and Control for Voice Transmission Over the Internet*. PhD thesis, Technische Universität Berlin, 2000.

[177] L. Pantel and L. C. Wolf, "On the Impact of Delay on Real-Time Multiplayer Games," in *Proceedings of the 12th International Workshop on Network and Operating Systems Support for Digital Audio and Video*, (New York, NY, USA), pp. 23–29, ACM, 2002.

[178] U. Nuding, S. Ono, M. Mustari, U. Büttner, and S. Glasauer, "A Theory of the Dual Pathways for Smooth Pursuit Based on Dynamic gain Control," *Journal of neurophysiology*, vol. 99, no. 6, pp. 2798–2808, 2008.

[179] D. A. Braun, A. J. Nagengast, and D. Wolpert, "Risk-Sensitivity in Sensorimotor Control," *Frontiers in Human Neuroscience*, vol. 5, pp. 1–10, 2011.

[180] A. Orda and A. Sprintson, "A Scalable Approach to the Partition of QoS Requirements in Unicast and Multicast," *IEEE/ACM Transactions on Networking (TON)*, vol. 13, no. 5, pp. 1146–1159, 2005.

[181] S. Shakkottai and R. Srikant, "Economics of Network Pricing With Multiple ISPs," in *Proceedings of the 24th Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 1, pp. 184–194, 2005.

[182] K. Tee, E. Burdet, C. Chew, and T. Milner, "A Model of Force and Impedance in Human arm Movements," *Biological Cybernetics*, vol. 90, no. 5, pp. 368–375, 2004.

[183] https://www.rtai.org/.

[184] https://rt.wiki.kernel.org/.

[185] G. Lafferriere, G. J. Pappas, and S. Yovine, "Symbolic Reachability Computation for Families of Linear Vector Fields," *Journal of Symbolic Computation*, vol. 32, no. 3, pp. 231–253, 2001.