uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Nizar Sakr
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

Ph.D. (Electrical and Computer Engineering)
GRADE / DEGREE

School of Information Technology and Engineering
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Haptic Date Reduction and Protection in Multimodal Virtual Environments

TITRE DE LA THÈSE / TITLE OF THESIS

Jying Zhao
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

The late Nicolas Georganas
CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

Abdulmotaleb El Saddik

Peter Liu

Ling Guan
Ryerson University

Shervin Shirmohammadi

Gary W. Slater
Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Haptic Data Reduction and Protection in Multimodal Virtual Environments

by

Nizar Sakr

A thesis

submitted to the Faculty of Graduate and Postdoctoral Studies

in partial fulfillment of the requirements for the degree of

Doctorate of Philosophy in Electrical and Computer Engineering

Ottawa-Carleton Institute for Electrical and Computer Engineering

School of Information Technology and Engineering

University of Ottawa

Your file Votre référence
ISBN: 978-0-494-74245-7
Our file Notre référence
ISBN: 978-0-494-74245-7

Abstract

The emerging field of haptics, which enables the sensing and manipulation of virtual environments through touch, is recognized as an important element in the area of human computer interaction. Many of its applications, including medical training, rehabilitation, security, entertainment, etc. involve the transmission, storage, protection, and analysis of haptic information. However, the multi-dimensional and time-varying nature of haptic data along with the fact that high sampling frequencies are required when generating haptic signals, motivate the research of *haptic data reduction*. This thesis discusses two important and related topics: data reduction and protection in multimodal visual-haptic virtual environments.

The focus of the first topic is on the design, evaluation, and analysis of haptic data reduction techniques in order to: (1) Improve packet transmission in telehaptic systems, and (2) reduce the dimensionality of the inherently large haptic datasets. In the former case, robust perception-based data reduction methods in 3- and 6-degree-of-freedom telehaptic systems are derived. The proposed techniques are evaluated in different experimental settings, including a haptic-enabled telesurgery simulation, and demonstrate a significant reduction in haptic data traffic while maintaining a high-quality telehaptic experience. In the latter case, data reduction in haptic information analysis problems is investigated. Specifically, data reduction methods are presented within the context of haptic data mining and knowledge discovery to help facilitate the analysis of the inherently high-dimensional haptic datasets. Experiments using datasets of haptic-based handwritten signatures demonstrate the accuracy of the techniques in reducing the high dimensionality of haptic feature spaces.

The protection of haptic information through digital watermarking is also investigated in this thesis due to its conceptual similarity to haptic data reduction and compression. A foundational study is conducted to examine the role of multisensory feedback in the perception of a watermark embedded in a haptic-enabled 3D virtual surface. In particular, the analysis aims to answer numerous important questions in this brand-new research area, including: Do water-

marks inspected using multimodal feedback (vision + haptic) result in very different detection thresholds from those detected using a single sensory modality; or more importantly, whether multimodal visual-haptic feedback improves the perception of watermarks embedded in 3D meshes.

A Personal Remembrance of Professor Nicolas D. Georganas

You left us all too soon professor. It is painful to realize that I will never be able to meet with you again, to receive an e-mail from you, or to hear you joyfully say your usual "Sooo, how's everyone doing?" as you walk into the lab.

Professor Georganas passed away only two weeks after sending me his corrections of this thesis dissertation. I have the privilege and honor of being one of his last PhD students. He was a truly great professor and human being and I have very fond memories of him. I first met Dr. Georganas in the winter of 2001 when I took his Circuit Theory II course, a second year Electrical Engineering undergraduate course at the University of Ottawa. In his classroom, it

felt like I had a legend as a teacher. His teaching was stimulating, and captivating. He had the rare ability of holding the students' attention completely. I remember visiting him during his office hours for some assistance in the course material; he was always very helpful, patient, and gracious.

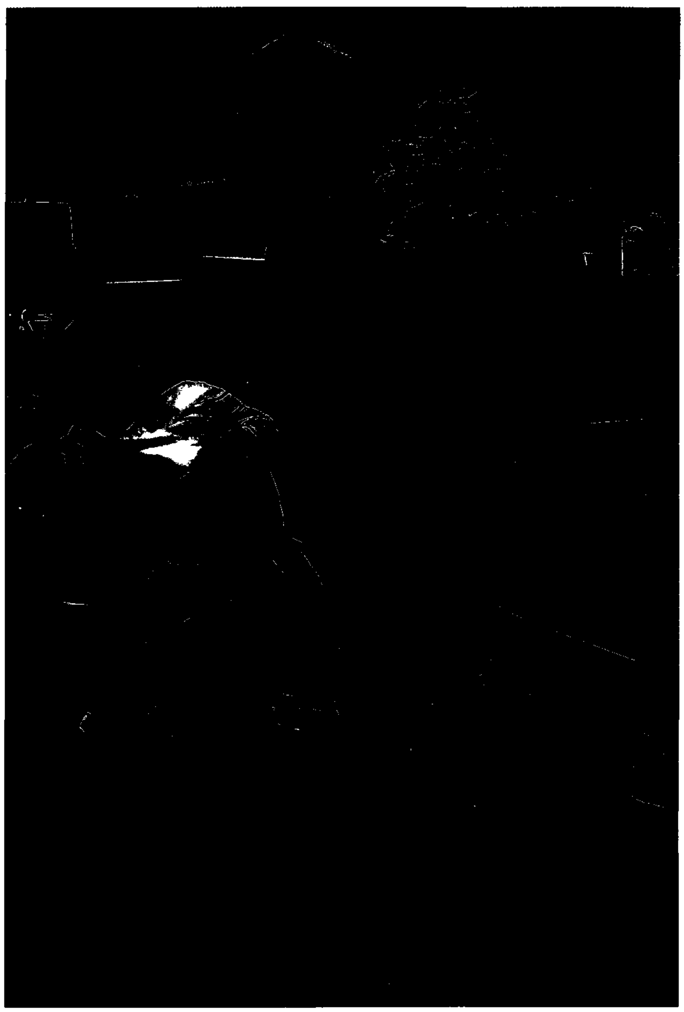
In October of 2005, I remember seeing him at a haptics workshop (IEEE International Workshop on Haptic Audio Visual Environments and their Applications) that was in fact hosted by his research laboratory. At the time, I was already accepted to the PhD program (under his supervision), but I had yet to move to the lab as I was admitted to start my degree in January of 2006. I remember him welcoming me to his lab and proudly introducing me as his new PhD student to some of his colleagues and workshop attendees. It was a very kind and thoughtful gesture that I greatly appreciated, especially coming from such a well-renowned, respected, and accomplished research giant.

Professor Georganas was a brilliant engineer, a wonderful research supervisor and a true gentleman. I will always remember our weekly team meetings with him from which I learned so much not only about scientific research, but also about life. My occasional individual research meetings with Dr. Georganas were always a joy. No matter how bad my thesis and research problems were, I would always leave his office with a smile. He believed in his students and their ability to solve challenging research problems, and succeed.

In addition, professor Georganas was extremely giving of his time, and took an interest in following and supporting the careers of his past students, many of whom I personally know.

A few months ago, I scheduled an early morning meeting with Dr. Georganas over Skype since at the time he was conducting research abroad. That morning, my Skype application was not working properly and despite numerous attempts, I was unable to establish a connection with Dr. Georganas as his status was never showing as "Online". Later that day, we happened to have a team meeting with him, also over Skype, however this time, a voice connection was in fact successfully established. During the meeting, Dr. Georganas was not very pleased with me since he assumed that I had forgotten about our morning one-on-one research meeting. Nonetheless, I explained what had happened and then we simply proceeded with the team meeting. A couple of days later, he sent me an e-mail discussing a research paper and in his message he included the following: “I am sorry that I was bitter at you Wednesday, although you could not see me on Skype! I was on the whole day.” I was humbled by his kind, thoughtful, and considerate gesture.

It never took too long for any newcomer to the team to realize that Dr. Georganas was a genuinely kind and warm-hearted man that very much cared for his students and their well-being.



Professor Georganas also had a very pleasant sense of humor. I remember the first time I gave him a ride in my 1994 white and very rusted Chevrolet Cavalier (that was sometime back in 2007); he looked at it and humorously said, "Ahh! You're taking me in your white limousine!".

I feel proud and privileged for having been part of professor Georganas' life and career, and to have had the fortune to be influenced by his combination of creativity, generosity, dedication, and humanity. Dr. Georganas has not only taught me the ability to dream big, but to also achieve big while enjoying the journey. I will miss him dearly.

His PhD student,
Nizar Sakr

Acknowledgments

This thesis is the result of the support, guidance, and hard work of many wonderful people. I would like to take this opportunity to thank and recognize them.

First, I would like to express my deep appreciation to my supervisor the late Dr. Nicolas D. Georganas. His flexible approach provided me with the freedom to explore different interesting ideas, and enabled me to discover the research path that best suited, and was most appealing to me. I will always remember the many words of wisdom you shared with us during our weekly team meetings, including how to survive in, and make the most out of the engineering research world. I will be forever grateful for the opportunities you have offered me and helped me realize.

I would also like to express my sincere gratitude to my co-supervisor Dr. Jiying Zhao for his guidance, patience, encouragement, and suggestions throughout the preparation of this thesis. Thank you for our regular one-on-one meetings that helped me stay focused and motivated.

My true sincere thanks to Dr. Julio J. Valdés from the NRC-IIT. Chapters 5 and 6 of the thesis would not have been nearly as strong without your invaluable help and supervision. Your passion for science and research is most apparent, admirable, and contagious. I feel blessed for having had the opportunity to work with you and for the wealth of knowledge I have gained from our regular weekly meetings.

My sincere gratitude to Dr. Emil Petriu for being a mentor and a friend to me throughout my graduate years at the University of Ottawa. Thank you for sharing your wisdom and giving me the right advice and support at key times in my PhD experience.

Many thanks to Dr. Abdulmotaleb El Saddik for his kindness and support for the past number of years, and for providing me with the great opportunity to work on the haptic-based biometrics project.

I would also like to thank Dr. Shervin Shirmohammadi and Dr. Claude D'Amours for their great advice and help on numerous occasions during my degree.

I am very grateful to the members of my thesis advisory and examination committees, Dr. Abdulmotaleb El Saddik, Dr. Ling Guan, Dr. Peter Liu, and Dr. Shervin Shirmohammadi, for their constructive comments which helped improve the quality of this thesis.

I owe an immense debt of gratitude to my very dear friend Nicolas Ellaham for his constant insightful, encouraging, and unending support (technically, socially, and spiritually) throughout my university studies. If everyone had a friend like you, the world would be a much better and happier place! I would also like to express my appreciation to Flora Nassrallah, for her substantial help in the statistical analysis of the data presented in this thesis, and for being such a wonderful friend, always enthusiastic, supportive, and strengthening; Many thanks are also due to my very good friend Karen Soueidan, for proofreading many parts of this thesis, and for always being willing to lend a hand regardless of how busy she was.

Many others have contributed to this work through their very valuable friendship, helping to keep me sane during the past four and a half years: Bassam Khouri and Joseph Moubayed, for always being such great friends, and for our many fun and relaxing trips to various places including Cuba, Las Vegas, and New Orleans; Wissam Elhage, Maya Nahas, and Philippe Malouf, for your constant, supportive, and remarkable friendship; Suzan Badel, thank you for our many very enjoyable lunches in the early stages of my degree, and for lending me an ear during one of the most challenging phases of my graduate years; the many others, whom I am afraid to list for fear of forgetting someone, but whose friendship I treasure dearly.

I am thankful to the academic and administrative staff at the School of Information Technology and Engineering, especially Suzanne St-Michel, Annie Trottier, Michèle G. Roy, and

Sylvie Richer. Your help and kindness throughout my graduate years was very much appreciated.

I would like to express my deepest gratitude to Jilin Zhou for his collaboration, help, and for the numerous discussions we shared on *data reduction in telehaptic systems*. A thank you is also due to Fawaz Alsulaiman for his collaboration, assistance, and helpful discussions on the *data reduction in haptic datasets* components of the thesis.

I would like to thank my lab mates, present and past, for the many insightful discussions and laughs we shared. I would like to also express my sincere appreciation to the lab manager, Francois Malric, whose technical expertise and dedication were a major factor in the success of the many experiments conducted in this thesis.

A heartfelt thank you to all my participants who selflessly gave so much of their time and energy to participate in this research. Without your help much of this work would not have been possible.

My thanks to the University of Ottawa, the Government of Ontario, and the Natural Sciences and Engineering Research Council of Canada for their generous financial support throughout my graduate student career.

My special thanks to the Rockcliffe Flying Club, my flight instructors Ryan King and Steven Hyde, and the "Canadian airspace", for providing me a place to escape the sometimes overwhelming realities and responsibilities of life.

Last but certainly not least, to my parents, Violette and Tarek Sakr, thank you for your unwavering love, encouragement, and support. You taught me to not be afraid to take chances, to always aim high and strive for the best. I dedicate this PhD thesis to you!! To my brother Ziad

for inspiring me to do a PhD and for always being the great supportive big bro. that you are, thank you. To my beautiful baby niece Nicole, thank you for reminding me how simple and precious life can really be. To my wonderful sisters Lama and Maysaa, the two most ambitious and hard working young women I know, I thank you for always being there when I needed you most.

To mom and dad, Violette and Tarek Sakr;
without you I would not be where I am today.

Contents

List of Tables

List of Figures

List of Abbreviations

A list of abbreviations employed throughout the thesis is given for reference. The section where each abbreviation is first defined and introduced is also provided.

Symbol	Description	Section
ANOVA	Analysis of Variance	4.1
CFS	Correlation-based Feature Selection	6.4.5
CV	Cross Validation	5.5
DoF	Degree of Freedom	2.1
DISCOVER	Distributed and Collaborative Virtual Environments Research Laboratory	2.1
DPCM	Differential Pulse Code Modulation	2.1
ET	Expression Tree	5.2.4
FR	Fletcher Reeves	5.2.3
GA	Genetic Algorithm	6.1
GEP	Gene Expression Programming	5.2.4
GP	Genetic Programming	5.2.4
HRTC	Haptic Real-Time Controller	3.2
HSI	Human System Interface	4.1
JND	Just Noticeable Difference	2.1
k-NN	k-Nearest Neighbors	5.1
LAN	Local Area Network	3.4.1
LDR	Lossy Data Reduction	2.1
LS	Least Squares	3.1
MF	Median Filter	3.1
OP	Operator	4.1
PCA	Principal Component Analysis	1.5
PDM	Position Deadband-based Method	3.3.4
PL	Packet Loss	3.5.2
SF	Selected Features	6.4.2

(Continued)

Symbol	Description	Section
SVM	Support Vector Machine	5.1
TD	Time Delay	3.5.2
TOP	Teleoperator	4.1
VDM	Velocity Deaband-based Method	3.3.4
VR	Virtual Reality	1.1

Part I

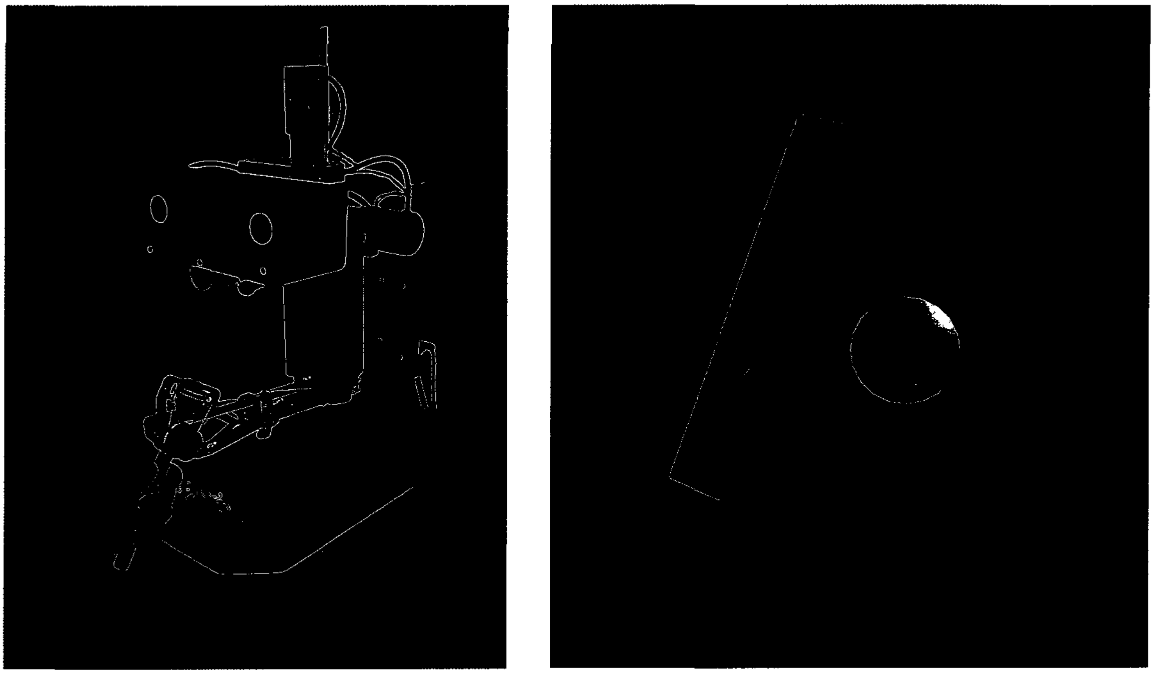
Setting

Chapter 1

Introduction

1.1 Introduction

The integration of haptics into immersive virtual environments has been an active research area over the past decade. Immersive digital environments consist of computer-created scenes within which users can immerse themselves and interact with other users or various objects through a virtual reality experience. Conversely, haptic systems enable physical interactions with remote or virtual objects through the sense of touch (see Fig. 1.1), and are therefore expected to become the next dimension of human-computer interaction. Haptic-based applications are wide, and span many areas, including medicine, rehabilitation, manufacturing, education, and entertainment. Many of the current and future applications will involve the transmission, storage, and analysis of haptic information. Haptic data depict trajectory, cutaneous, and kinesthetic information (which essentially consist of position, velocity, orientation, torque, and force information) that can be directly acquired from a haptic interface upon a user's interaction with a predefined environment. However, the multi-dimensional and time-varying nature of haptic data, along with the fact that relatively high sampling frequencies are required when generating haptic signals, motivates the research of *haptic data reduction*. There are two primary applications in which the use of haptic data reduction may be exceedingly attractive:



(a) Haptic device (b) Virtual environment

Figure 1.1: An example of a haptic-enabled virtual environment (stick-on-ball): A haptic device provides the user with realistic force-feedback sensations when interacting with a virtual 3D ball.

1. Packet reduction for improved transmission in networked haptics.

2. Size and dimensionality reduction of haptic datasets to:

    - reduce storage and network bandwidth requirements;
    - provide compact visual representations for a more rapid and intuitive understanding of the data;
    - facilitate the analysis and possibly reveal knowledge about the data under consideration.

A more detailed description of the important role haptic data reduction plays in each of the two broad applications is presented in Sections 1.2 and 1.3.

In addition, protection of haptic data through digital watermarking is also investigated in this work due to its conceptual similarity to haptic data reduction and compression. In particular, both techniques attempt to intelligently distort signals with the purpose of achieving greater system reliability, robustness, and effectiveness, while maintaining high quality haptic interaction and user experience. A thorough introduction to haptic digital watermarking and its applications is presented in Section 1.4.

1.2 Packet Reduction for Improved Transmission in Networked Haptics

1.2.1 Motivation

Networked haptics, or *telehaptics* refers to the transmission of the sense of touch over network connections. Example applications include tele-medicine, tele-education, tele-maintenance and tele-entertainment. A telehaptic system can take many different physical configurations depending on the particular application. Haptic-enabled telepresence systems, which enable users to operate in a remote environment while incorporating the sense of touch, are one common example. In advanced teleoperation systems of this kind, multimodal sensory (haptic, audio, video, etc.) and control information is bilaterally transmitted over a communication channel, enabling the human operator to be immersed into the remote environment. As it is illustrated in Fig. 1.2, a multimodal haptic-enabled telepresence system consists of three primary modules: the human system interface (HSI), the teleoperator (TOP), and a communication network that serves as a link between the HSI and the TOP. Essentially, a human operator (OP) commands the TOP through the HSI. Multimodal sensor data acquired at the teleoperator are then transmitted back and displayed to the OP. Haptic telepresence systems have various potential applications including, aerospace, military, security, medicine, etc., in which humans are required to be immersed in real (or virtual), but inaccessible environments. A promising application of haptic-enabled telepresence is in the area of explosive ordnance disposal,

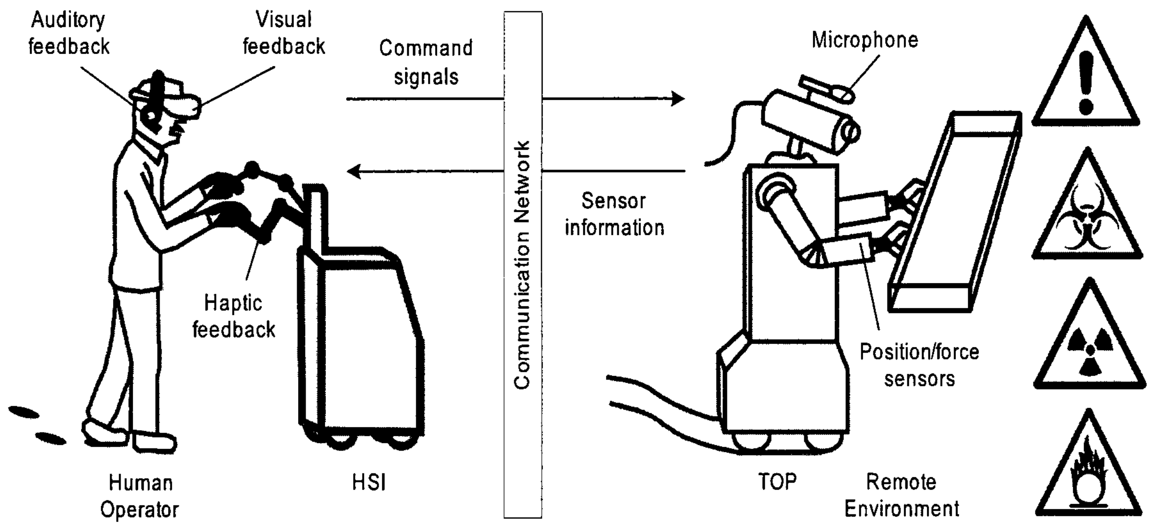


Figure 1.2: An example of a haptic-enabled telepresence system (adapted from [1]).

which involves the detection, identification, on-site evaluation and de-arming of unexploded explosive ordnance such as land mines [7]. Consequently, a haptic telepresence system would enable a human expert to perform explosive ordnance disposal from a remote control station located in a safe environment while relying on a haptic display and a communication medium to command the teleoperator. The haptic feedback in addition to visual and auditory feedback enhances the operator's situation awareness, increasing the efficiency of teleoperation and improving safety.

Other types of telehaptic systems with a distinct physical configuration are also exploited in telementoring applications. Telementoring can be defined as remote mentoring via a communication network. A haptic telementoring system generally encompasses a mentor site and a mentee site linked over a packet switched communication network, where each is equipped with a haptic device and an audio-visual interface. The applications of haptic telementoring in tele-education are numerous. In particular, it has garnered significant interest in recent years in tele-surgical training. This is primarily due to the fact that procedures in medical applications are not easily illustrated using solely audio and visual cues. For example (refer to Fig. 1.3),

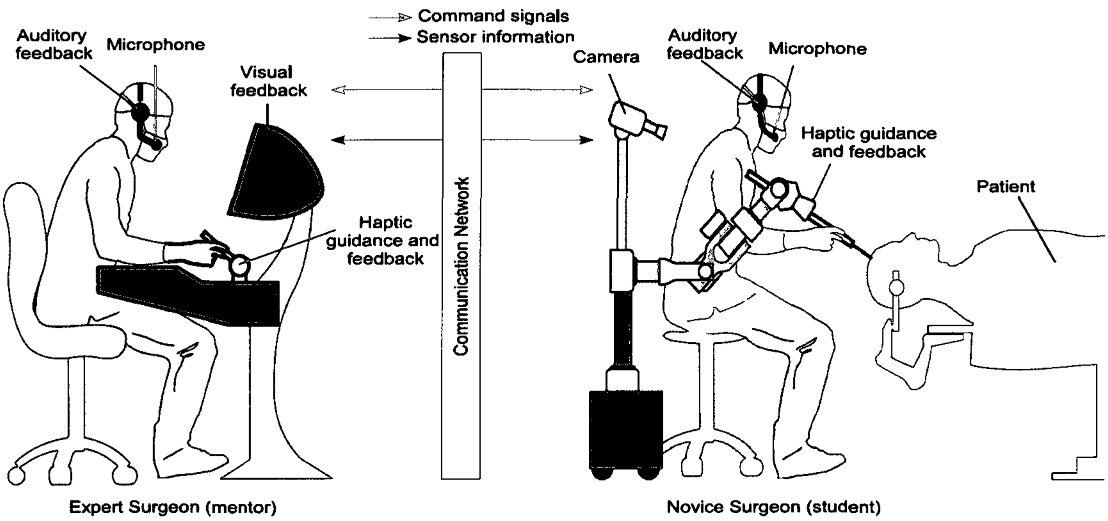


Figure 1.3: An example of a haptic-enabled telementoring system (inspired from [2]).

haptic guidance permits a novice surgeon to perform laparoscopic procedures with the assistance of a remote expert via telementoring [8]. This is of particular importance since expert surgeons are generally primarily concentrated in major cities, and are therefore often not readily available at remote locations especially near battlefield sites.

However, in bilateral telehaptic systems, regardless of the addressed application, haptic data is typically sampled at a constant rate in the range 500-1000 Hz in order to ensure that the stability of the control loops at the local and remote sites (e.g., the HSI and the TOP in a telepresence system, respectively) is not compromised. In order to limit packetization and transmission delays which could adversely affect the stability of the system, every set of sampled data is transmitted in individual packets. This translates to a packet rate that is equal to the sampling rate of the local haptic control loops. However, high packet rates (in the range of 500-1000 packets per second) are difficult to sustain over long distance digital communication networks. Conversely, the payload portion of a haptic data packet is rather small when compared to other data-intensive media such as video. Nevertheless, even though the packet load itself is relatively low, the data transmission requirements are comparable to other bandwidthintensive streaming media. For example, let's consider a haptic-enabled telementoring system where two six Degree-of-Freedom (DoF) haptic devices are coupled over an IP-based communication network. The packet rate is 1000 packets/s as the sampling rate is assumed to be 1000 Hz. In addition, since 6-DoF devices are considered, the following attributes should be bilaterally transmitted (at the least): position, velocity, orientation, and the rate-of-change of orientation data [9]. Therefore, in each packet at least 12 values of type float are sent, which results in a packet payload of 48 bytes, i.e., 4 (attributes) $\times$ 3 (dimensions) $\times$ 4 bytes/sample = 48 bytes. Additionally, assuming that data transmission is performed using RTP (Real-time Transport Protocol) over UDP (User Diagram Protocol) as a communication protocol, the corresponding packet header size is 40 bytes (i.e., 12 bytes for RTP, 8 bytes for UDP, and 20 bytes for IPv4). The total packet size is therefore 88 bytes. Consequently, the minimum required communication bandwidth is 1.41 Mbps, which is comparable to uncompressed audio or compressed video streaming. Moreover, it is important to note that unlike audio and video streams which are considered soft real-time, haptic streams are hard real-time suggesting that minimal delay could render a telehaptic system unstable.

1.2.2 Research Problems

The primary concern in the design of telehaptic systems is to minimize delays and the amount of data transmitted over the network. With this in mind, it is of high interest to develop and exploit haptic data reduction techniques in an attempt to decrease network traffic in telehaptic environments without impairing the fidelity and immersiveness of the system. Therefore, problems that are of interest in this research area consist of the development of human-perception-based haptic data reduction methods in multi-DoF telehaptics and their evaluation in different networked (haptic-enabled) applications. Many factors must, however, be considered when deriving a haptic data reduction technique: It should introduce minimal (if any) delay to the system and real-time compression must be guaranteed; it must be robust against network-induced distortions such as delay and packet loss; it would be desirable if it considers knowledge from human haptic perception to minimize the amount of data transmitted without affecting the quality of the telehaptic experience; additionally, it must ensure that the stability of the system is not compromised.

1.2.3 Goals and Objectives

Novel human-perception-based data reduction methods for haptic communication in multi-DoF (3- and 6-DoF) telehaptic systems are derived. Their application in different telehaptic applications, including haptic-enabled telementoring and telepresence systems is demonstrated. The algorithms rely on knowledge from human haptic perception in order to reduce the number of packets transmitted without compromising transparency. Haptic prediction models are exploited to further reduce the amount of haptic packets transmitted, and to improve the reconstruction of data samples on the receiver side. Several experiments are conducted to provide evidence of the proposed techniques' effectiveness in normal network conditions and in the presence of significant communication delay and packet loss, while preserving the overall quality of the telehaptic environment.

1.3 Size and Dimensionality Reduction of Haptic Datasets

1.3.1 Motivation

Many current and future applications of haptics in rehabilitation, security, training and simulation, medicine, and entertainment (the list is by no means exhaustive), will involve the analysis and interpretation of acquired haptic information in order to possibly reveal patterns and relationships in the data. However, regardless of the intended haptic application, the number of resulting features are significantly large (in the thousands range) as the recorded information consists of multidimensional time-varying data. This is also due to the fact that the data are collected at relatively high sampling rates to capture the inherent frequency components of haptic signals. Consequently, the tremendous size and high dimensionality of haptic datasets creates a need for research, development, and evaluation of haptic data reduction techniques.

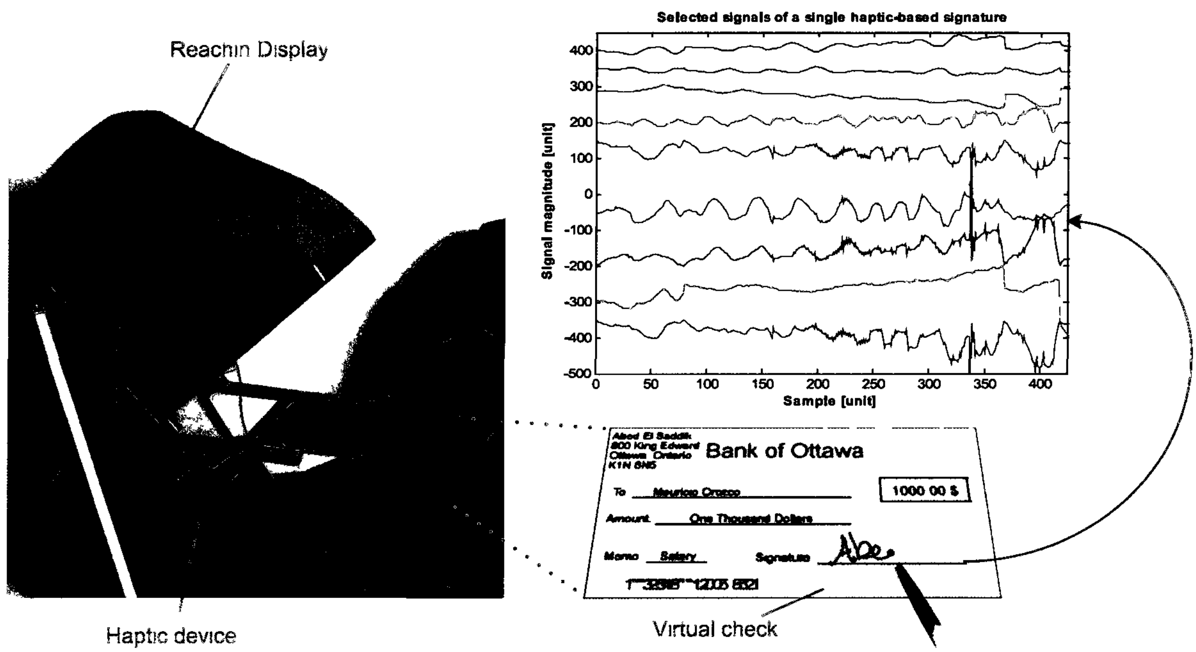


Figure 1.4: A haptic-enabled virtual check application that serves as an advanced biometric identification and verification system.

Hence, in contrast to the methods mentioned in Section 1.2 which are specifically intended for packet reduction within the context of telehaptic systems, data reduction methods briefly introduced in this section and depicted in detail in Chapters 5 and 6, are discussed within the context of haptic data mining and knowledge discovery.

To emphasize the importance of the development and use of haptic data reduction methods consider the haptic-based biometric system illustrated in Fig. 1.4. The possible use of haptic devices in biometric systems has been recently suggested to enable improved user identification/verification performance over more traditional techniques, such as those based on handwritten signatures [10, 11]. The system presented in Fig. 1.4 is a haptic-enabled virtual check application that enables users to perform their haptic-based handwritten signatures. More specifically, users write their haptic signatures on a virtual check using a virtual pen guided by a haptic force-feedback device. The resulting haptic stimuli are force and frictional feedback that attempt to mimic the tactile sensations felt when signing a traditional paper check. This being said, the luxury of such an advanced information-rich handwritten signature comes at the cost of having to deal with a massive array of haptic features. In fact, even when only the fundamental haptic features are considered, i.e., three-dimensional position, force, torque, and orientation data, a user signature can be found to be in the range of thousands of attributes. It should therefore be apparent that extracting and analyzing information from datasets with such a large number of features is a major challenge.

Reducing the size of a large dataset offers numerous well-known advantages. These benefits are especially apparent in a knowledge discovery system in which a classification model is built to distinguish between possible patterns in the data. With this in mind, probably the most important advantage is the fact that feature reduction can help alleviate the "curse of dimensionality" problem. The curse of dimensionality suggests that the size of the classifier's training dataset (number of instances) must increase exponentially with the dimensionality of the feature space. Furthermore, excluding noisy and irrelevant features could potentially improve the accuracy and generalization properties of the classification model. Additionally, a smaller subset of features reduces the space complexity of the classification problem, consequently decreasing the computational overhead and speeding up the learning process. Furthermore, size reduction of haptic datasets can help reduce storage and network bandwidth requirements, which remains an important challenge in various applications despite today's advances in high-capacity storage and high-speed/high-bandwidth communication systems. For example, the use of compression and data reduction techniques is very common in biometric applications to alleviate a system's storage and bandwidth requirements. Depending on the application, users' biometric representations (also referred to as *templates*, e.g., signature, fingerprint, retinal scan, etc.) are generally stored in a central database or recorded on a portable storage device issued to each individual such as a magnetic card or a smart card. However, in the former case, network bandwidth may become a system bottleneck for user identification/authentication. In the latter case, such biometric systems are subject to storage constraints as such portable devices can generally only hold few kilobytes of data.

1.3.2 Research Problems

The development of different methods for data reduction, specifically, dimensionality reduction of large haptic datasets is of great interest. Important problems in this emerging area of research include the derivation and evaluation of different feature selection and generation (also referred to in the literature as feature extraction) methods that best reduce the dimensionality of haptic datasets while preserving the information contained in the data. Feature selection refers to the process of identifying an optimum subset of features from a much larger set of potentially useful features [12]. In other words, the primary goal of feature selection is the identification of irrelevant and redundant features in an attempt to find the best subset satisfying a certain criterion (e.g., classification error rate). Conversely, feature generation (or extraction) is a transformation of a possibly high-dimensional dataset into a new space of lower dimensions using linear or non-linear combinations of existing features. Another important advantage and often a popular application of certain feature reduction techniques (feature generation methods), is the possibility to project the high-dimensional feature space into a space of fewer dimensions to construct a visual representation of the data. This can clearly facilitate the analysis and understanding of any possible patterns that could exist in the data. Advantages and drawbacks of feature selection and feature generation are further discussed in Chapters 5 and 6.

1.3.3 Goals and Objectives

A haptic data preprocessing technique is introduced that accurately maps the time varying attributes of a haptic data instance (e.g., a haptic-based handwritten signature) to a single vector. Furthermore, data reduction through feature generation is initially illustrated to enable visual data mining of high-dimensional haptic datasets. In particular, the suggested techniques make use of linear and non-linear transformations to map the high-dimensional haptic feature space into another space of smaller dimension while minimizing some error measure of information loss. Additionally, the potential use of different feature selection methods in the reduction of high-dimensional haptic datasets is investigated. Moreover, a method is proposed where feature vectors are subsampled prior to the feature selection procedure. This is achieved to further reduce the haptic feature space while maximizing discrimination between instances. In addition, different classifiers are exploited in order to evaluate the potential discrimination performance of the computed low-dimensional sets of attributes. All experiments are conducted while relying on haptic-based handwritten signatures datasets.

1.4 Data Protection in Multimodal Visual-Haptic Virtual Environments

1.4.1 Motivation

With the wide availability and widespread use of 3D models in several fields, such as in entertainment, the medical industry, the military, etc., and the recent advancements of haptic technology, it is expected that in the very near future, the need to protect haptic-enabled virtual scenes from malicious attacks or inadvertent tampering will arise. In recent years, digital watermarking, which deals with the process of embedding information into digital data in an inconspicuous manner, has been proposed as a viable solution to the need of copyright protection and authentication of multimedia information. Example applications of digital watermarking include identifying the origin, owner, use, rights, integrity, or destinations of multimedia content (e.g., digital images, video, audio, and 3D models) [3, 13–17]. Some advantages of watermarking over encryption include, protecting content even after it is decrypted, and monitoring how legitimate customers handle the content after decryption. Moreover, watermarking techniques are different from other intellectual property rights methods in that: Watermarks are imperceptible, they undergo the same transformations as the work, and they are inseparable from the digital content in which they are embedded. In fact, the first requirement of digital watermarking techniques regardless of the addressed media or application is imperceptibility [15, 18, 19]. This refers to the perceptual similarity between the original and watermarked data. Ideally, the perceptual quality of the watermarked media must be identical to the original. Recently, considerable progress has been made in 3D watermarking where the main focus has been on triangle meshes, the most common digital representation of 3D models due to its simplicity and usability. Existing watermarking techniques concerning 3D meshes host the watermark by either modifying the geometry or the connectivity of the surface (spatial domain) or by modifying some kind of spectral-like coefficients (spectral domain) [16]. Accordingly, the watermark's intrusiveness can be evaluated in terms of its visibility in the rendered version of the mesh. The evaluation of 3D watermarking algorithms against the imperceptibility constraint has been thus far exclusively based on the sensitivity of human vision to distortion. Moreover, currently available perceptual metrics generally used to assess the quality of watermarked 3D meshes, have been validated solely through psychovisual experiments [18, 19]. However, the increased popularity of haptic technologies and their integration into immersive virtual 3D environments motivates research in multimodal visual-haptic digital watermarking (see Fig. 1.5).

1.4.2 Research Problems

Research in haptic watermarking conducted in this work is primarily focused on studying the impact of the haptic dimension on the perceptibility of watermarks inserted in 3D virtual objects. In particular, the recent advancements of haptic technology have raised several important questions in the 3D watermarking research community: Is the haptic sensory channel more sensitive than the visual sensory channel in detecting a watermark embedded in a haptic-enabled 3D object? Do watermarks inspected using multimodal feedback (vision + haptic) result in very different detection thresholds from those detected using a single sensory modality (touch-only or vision-only); Or more importantly, does visual feedback, when presented together with haptic feedback, improve the perception of a watermark embedded in a 3D mesh? Additionally, what is the possible impact of different modalities (vision, haptic, and vision + haptic) on

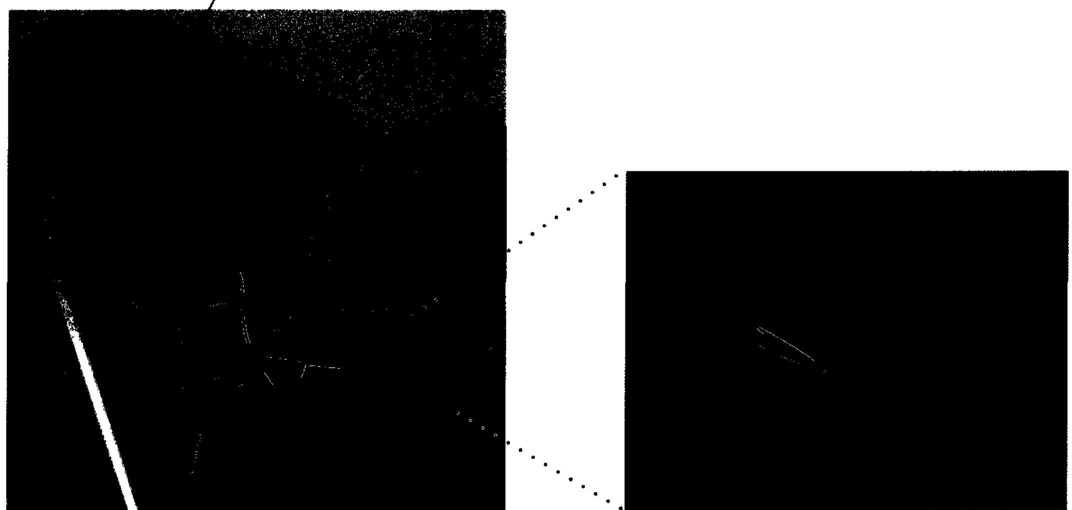


Figure 1.5: A participant relies on a visual-haptic system (the Reachin display and a Sensable PHANTOM Desktop haptic device) to interact with a (watermarked) haptic-enabled virtual environment [3].

the measured watermark detection thresholds if different resolutions of the underlying virtual 3D mesh are considered? It is important to emphasize that while it is intuitive to assume that a multimodal presentation of stimuli should lead to an improvement in performance (e.g., for watermark detection), previous research in human perception suggests otherwise. Specifically, the role and possible advantages of multisensory feedback in roughness perception (the 3D watermarking process can be regarded as a particular form of surface roughness) is quite complex as it can vary across different experimental conditions, including type of surface, haptic device used, surface parameters, etc.

1.4.3 Goals and Objectives

A study is conducted to examine the role of multisensory feedback in the perception of a watermark embedded in a haptic-enabled 3D mesh. Specifically, an emphasis is made on whether the use of a bimodal approach, where a subject is looking at and touching a 3D mesh, could in fact improve the perceptibility of a watermark as opposed to when only one single modality (vision-alone or haptic-alone) is available. The experiments are performed using a visual-haptic interface (also referred to as a hapto-visual or a visuo-haptic interface) that enables users to see and touch virtual objects at the same location in space. This approach enables a superior integration of vision and touch than a conventional 2D screen-based display. Moreover, the advantages of the use of a visual-haptic interface stem from the fact that human perception is generally multimodal, i.e., the senses of touch and vision do not operate independently, but are in fact closely coupled. A statistical analysis to test for significance between modalities is also presented. Research conducted in Chapter 7 has been devoted to this topic with the purpose of providing a foundational study on multimodal vision-haptic digital watermarking.

1.5 Main Contributions of the Thesis

This work contributes to several research fields including: haptics, data reduction and compression, human haptic perception, data mining, and biometrics. The following summarizes the contributions of the thesis:

Data Reduction in Telehaptic Systems

- Novel human-perception-based data reduction methods are derived for haptic communication in multi-DoF (3- and 6-DoF) telehaptic systems.
  - A prediction approach is proposed that relies on the least-squares method and median filtering to reduce the number of packets transmitted and efficiently reconstruct unsuccessfully received data samples.
  - The algorithms rely on the limitations of human haptic perception (i.e., the Just Noticeable Differences) with respect to a user's hand movement (position/orientation) and exerted force (force/torque) in order to reduce the number of packets transmitted without compromising transparency.

  - Several distance metrics are provided to evaluate the acuity of human haptic perception in detecting haptic distortion when data reduction is performed in 3- and 6-DoF settings.
  - An extension of the deadband principle to 6-DoF is derived.

- An application of the perception-based haptic data reduction methods in a haptic-enabled telementoring eye-surgery simulation is presented.

- Experiments are carried out to evaluate human haptic sensitivity to data reduction. Additionally, a thorough statistical analysis is conducted to determine the appropriate multivariate human haptic perception thresholds (force, torque, orientation, etc.) required to minimize the number of packets transmitted while preserving the immersiveness of the 6-DoF telepresence system.

- A performance analysis of the proposed data reduction methods in normal and adverse network conditions is presented. This is conducted for 3- and 6-DoF telehaptic systems.

Data Reduction in Haptic Pattern Analysis

- A haptic data preprocessing method is proposed that accurately maps the multidimensional and time-varying haptic data types (e.g., position, force, torque, etc.) associated with every instance (e.g., a haptic-based handwritten signature) to a single feature vector.

- A novel information-preserving haptic data reduction method is suggested that combines intelligent data subsampling with a feature selection technique based on concepts from computational intelligence, specifically, rough set theory and evolutionary computation.
  - A comparison between the suggested rough set-based method and classical machine learning techniques in the selection of minimal information-preserving subsets of features in high-dimensional haptic datasets, is provided.
  - Small subsets, generated using the proposed feature selection approach, are exploited to give some insight on the possible importance of different haptic data types in a haptic-based handwritten signature identification application.

- A feature generation approach to data reduction is explored where high-dimensional haptic datasets are mapped onto 3D feature spaces through linear and non-linear transformations.
  - An analysis is conducted to explore whether visual data mining is in fact feasible when dealing with inherently high-dimensional haptic datasets that encompass thousands of features.
  - A comparison between linear Principal Component Analysis (PCA), as well as non-linear methods based on genetic programming and classical optimization techniques in the construction of 3D visual spaces using very high-dimensional haptic datasets, is provided.
  - The application of the feature generation techniques to construct 3D visual spaces of a haptic-based handwritten signatures dataset is presented.

Multimodal Vision-Haptic Watermarking

- A study is conducted to investigate the role of multisensory vision-haptic feedback in the perceptibility of 3D watermarks embedded in haptically-enabled virtual surfaces.
  - An analysis is carried out to explore statistically the impact of the considered modalities (vision, haptic and vision-haptic) on measured watermark detection thresholds, across different resolutions of the underlying virtual 3D mesh.
  - Numerous haptic watermarking research questions are answered, including: 1) Whether haptic-alone is superior to vision-alone in detecting a watermark embedded in a 3D virtual object; (2) if multimodal visual-haptic conditions can improve the watermark detectability as opposed to haptic-alone or vision-alone conditions; (3) if the impact of the selected modality (or modalities) on the perceptibility of the 3D watermark is dependent on the chosen virtual surface resolution; and (4) whether the detectability of the watermark increases or decreases with the resolution of the surface virtual mesh when bimodal conditions are considered.

1.6 Publications Arising from this Thesis

The following refereed journal and conference publications have arisen from the work presented in this thesis:

Refereed Journals

1. N. Sakr, J. Zhou, N.D. Georganas, J. Zhao, "Prediction-based Haptic Data Reduction and Transmission in Telementoring Systems," *IEEE Transactions on Instrumentation and Measurement*, vol. 58, no. 5, pp. 1727–1736, May 2009.

2. N. Sakr, N.D. Georganas, J. Zhao, E.M. Petriu, "Multimodal Vision-Haptic Perception of Digital Watermarks Embedded in 3-D Meshes," *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 5, pp. 1047–1055, May 2010.

3. N. Sakr, N.D. Georganas, J. Zhao, "Human Perception-based Data Reduction for Haptic Communication in 6-DoF Telepresence Systems," submitted to the *IEEE Transactions on Instrumentation and Measurement, Special Issue for the IEEE Symposium on Haptic Audio Visual Environments and Games 2010.*

4. N. Sakr, F.A. Alsulaiman, J.J. Valdés, A. El Saddik, N.D. Georganas, "Structure Analysis of High-Dimensional Haptic-based Handwritten Signatures," in preparation.

5. N. Sakr, F.A. Alsulaiman, J.J. Valdés, A. El Saddik, N.D. Georganas, "Feature Selection in High-Dimensional Haptic-based Biometric Data using Rough Set Methods," in preparation.

Refereed Conference Proceedings

1. N. Sakr, N.D. Georganas, J. Zhao, "Exploring Human Perception-based Data Reduction for Haptic Communication in 6-DOF Telepresence Systems," In *Proc. of IEEE International Symposium on Haptic Audio Visual Environments and Games*, pp. 1–6, Phoenix, Arizona, USA, Oct. 2010.

2. N. Sakr, F.A. Alsulaiman, J.J. Valdés, A. El Saddik, N.D. Georganas, "Feature Selection in Haptic-based Handwritten Signatures Using Rough Sets," In *Proc. of IEEE World Congress on Computational Intelligence*, pp. 1–8, Barcelona, Spain, July 2010.

3. N. Sakr, F.A. Alsulaiman, J.J. Valdés, A. El Saddik, N.D. Georganas, "Exploring the Underlying Structure of Haptic-based Handwritten Signatures using Visual Data Mining Techniques," In *Proc. of IEEE Haptics Symposium*, pp. 467–474, Boston, USA, March 2010.

4. N. Sakr, J. Zhou, N.D. Georganas, J. Zhao, E.M Petriu, "Network Traffic Reduction in Six Degree-of-Freedom Haptic Telementoring Systems," In *Proc. of IEEE International Conference on Systems, Man, and Cybernetics*, pp. 2451–2456, San Antonio, TX, USA, Oct. 2009.

5. N. Sakr, F.A. Alsulaiman, J.J. Valdés, A. El Saddik, N.D. Georganas, "Relevant Feature Selection and Generation in High-Dimensional Haptic-based Biometric Data," In *Proc. of International Conference on Data Mining* (DMIN'09), CSREA Press, pp. 71–77, Las Vegas, NV, USA, July 2009.

6. F.A. Alsulaiman, N. Sakr, J.J. Valdés, A. El Saddik, N.D. Georganas, "Feature Selection and Classification in Genetic Programming: Application to Haptic-based Biometric data," In *Proc. of IEEE Symposium on Computational Intelligence for Security and Defence Applications*, pp. 1–7, Ottawa, Ontario, Canada, July 2009.

7. N. Sakr, J. Zhou, N.D. Georganas, J. Zhao, E.M Petriu, "Robust perception-based data reduction and transmission in telehaptic systems," In *Proc. of World Haptics, Third Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pp. 214–219, Salt Lake City, UT, USA, March 2009.

8. N. Sakr, N.D. Georganas, J. Zhao, E.M. Petriu, "Perceptibility of Digital Watermarks Embedded in 3D Meshes using Unimodal and Bimodal Hapto-Visual Interaction," In

Proc. of IEEE International Instrumentation and Measurement Technology Conference, pp. 987–991, Singapore, May 2009.

9. N. Sakr, N.D. Georganas, J. Zhao, "Perceptibility of Watermarked 3D Meshes using a Multimodal Visuo-Haptic Interface," In *Proc. of IEEE International Conference on Virtual Environments, Human-Computer Interfaces, and Measurement Systems*, pp. 20–24, Turkey, Istanbul, July 2008.

10. N. Sakr, J. Zhou, N.D. Georganas, J. Zhao, X. Shen, "Prediction-based Haptic Data Reduction and Compression in Tele-Mentoring Systems," In *Proc. of IEEE International Instrumentation and Measurement Technology Conference*, pp. 1828–1832, Victoria, BC, Canada, May 2008.

11. N. Sakr, N.D. Georganas, J. Zhao, X. Shen, "Motion and Force Prediction in Haptic Media," In *Proc. of IEEE International Conference on Multimedia & Expo*, pp. 2242–2245, Beijing, China, July 2007.

12. N. Sakr, N.D. Georganas, J. Zhao, X. Shen, "Towards an Architecture for the Compression of Haptic Media," In *Proc. of IEEE International Conference on Virtual Environments, Human-Computer Interfaces, and Measurement Systems*, pp. 13–18, Ostuni, Italy, June 2007.

1.7 Outline of the Thesis

The thesis encompasses four different parts. Part I includes Chapters 1 and 2 which aim at providing a concise introduction to the thesis topic and an overview of the related work, respectively. Part II includes Chapters 3 and 4 which discuss data reduction in telehaptic systems. More specifically, Chapter 3 introduces perception-based data reduction techniques for improved haptic communication in 3-DoF telementoring systems. The techniques are evaluated in a basic experimental setting (in normal and adverse network conditions) to validate their performance. Their application in a haptic-enabled telementoring surgery simulation is also demonstrated. Chapter 4 extends the work presented in Chapter 3 and derives a human-perception-based haptic data reduction method to reduce the number of transmitted packets in 6-DoF telepresence systems. The algorithm relies on the limitations of human haptic perception (i.e., the Just Noticeable Differences) with respect to a user's hand movement (position/orientation) and exerted force (force/torque) to reduce the number of packets transmitted without compromising transparency. A validation of the 6-DoF haptic data reduction algorithm is performed under normal network conditions and in the presence of network-induced time delay and packet loss.

Part III encompasses Chapters 5 and 6 that discuss data reduction methods in haptic pattern analysis problems. In particular, data reduction is introduced within the context of haptic data mining and knowledge discovery, where feature selection and generation methods are derived to determine relevant attributes in high-dimensional haptic datasets. In Chapter 5, feature generation methods are presented that make use of linear and non-linear transformations to map the high-dimensional haptic feature space into another space of smaller dimension with different intuitive metrics. The application of the suggested approaches in visual data mining of large haptic datasets is presented. A number of research issues that relate to visual data mining of inherently large haptic datasets are also investigated in this chapter. In Chapter 6, feature selection is explored as an alternative approach to reduce the high dimensionality of haptic datasets. A novel information-preserving haptic data reduction method is suggested that combines intelligent data subsampling with a feature selection technique that is based on concepts from rough set theory and evolutionary computation.

Part IV encompasses Chapter 7 which introduces haptic data protection through digital watermarking. In particular, a study is conducted to explore multimodal visual-haptic perception of digital watermarks embedded in 3D virtual objects. Psychophysical experiments are conducted, and an in-depth analysis is carried out to inspect experimentally and statistically the impact of different modalities (vision-alone, haptic-alone, and vision + haptic) on the measured watermark detection thresholds, across different resolutions of an underlying virtual 3D mesh. This analysis is of great significance to 3D watermarking researchers as the current vision-only-based imperceptibility criteria (in 3D mesh watermarking) will ultimately need to be redefined if a scenario is uncovered in which watermark detection is best achieved using multimodal feedback.

Finally, Part V includes Chapter 8 that summarizes the major contributions and results presented in the thesis and discusses possible future work.

Chapter 2

Literature Review

2.1 Haptic Data Reduction in Telehaptic Systems

Haptic data reduction and compression is a fairly new research area that is progressing at a relatively fast pace. The techniques currently available in the literature can be classified into three distinct categories: *statistical*, *perceptual*, and *hybrid*. Methods in the first category make use of statistical properties of haptic signals to enable efficient data compression (for data storage or transmission). Methods in the second category exploit the limitations of human haptic perception to accurately reduce the packet transmission rate over networked haptic-enabled systems, without compromising the quality of the telehaptic experience. The third category methods combine perceptual and statistical techniques to solely encode perceptually-significant haptic data samples (to be stored or transmitted) using fewer number of bits.

The problem of haptic data reduction and compression was first addressed in [20], where Shahabi *et al.* presented a technique in which a combination of adaptive sampling and adaptive DPCM[1] is performed for the reduction of data samples during storage or transmission of haptic sensory readings. With adaptive sampling, the Nyquist rate (i.e., sampling rate) is computed based on data acquired over the previous sampling window. In turn, a greater compression

[1] Differential Pulse Code Modulation encodes values as differences between the current and the previous value.

performance is achieved with larger sampling windows, but only at the expense of increased delay. In [7], Kron *et al.* discussed a low-delay coding technique that relies on DPCM together with fixed quantization and Huffman coding. The authors incorporated this method in a bimanual haptic telepresence system intended for explosive ordnance disposal, to enable efficient and low-delay haptic data transmission. A similar method was also introduced in [21], however, in this case adaptive DPCM was used, where the quantization step-size is increased or decreased during periods of fast or slow haptic motion respectively. In [22], Borst made use of both lossy uniform and nonuniform quantization of first-order differences (between current and past haptic data values) to investigate the possible data reduction that can be achieved before compression artifacts are haptically sensed. Borst evaluated the compression algorithm using the Rutgers Master II force-feedback glove [23]. The results obtained in [22], although not conclusive, illustrated that nonuniform quantization can slightly outperform uniform quantization.

In more recent work [24, 25], Hirche *et al.* presented the first psychophysically motivated compression approach that relies on a deadband control technique to reduce the amount of haptic data packets transmitted over communication channel (for haptic-enabled telepresence and teleaction), while guarantying the stability of the networked system. In particular, a deadband controller ensures that haptic data are only transmitted if the difference between the current and the most recently transmitted sample exceeds a perceptual threshold; a parameter tuned to match the Just Noticeable Difference (JND) in human haptic perception. On the receiver end, missing data samples are reconstructed using a modified Hold-Last-Sample (or zero order hold) approach that guarantees the stability of the telepresence system. This method is, however, formulated under the assumption that the communication network has no time delay. In [26, 27], Hirche *et al.* further developed their work to enable haptic data reduction in time-delayed communication channels (with unknown, yet constant time-delay), while guaranteeing the stability of the haptic telepresence system. This is performed while exploiting a control method that relies on the scattering (also known as the "wave") transformation [28, 29] which stabilizes a communication subsystem for arbitrarily large constant delays. In addition, the authors introduced a novel energy supervised data reconstruction algorithm which combines two reconstruction methods: a non-conservative, but possibly energy generating standard strategy, such as the Hold-Last-Sample technique, and a strictly passive[2] strategy, such as the aforementioned modified Hold-Last-Sample technique. The results Hirche *et al.* presented in [24–27] demonstrated that considerable network traffic reduction can be achieved without significantly impairing the perception of the remote environment in the case when no delay is present over the network and when the communication channel is time delayed. Their theoretical and experimental derivations were, however, based on the fact that 1-DoF (degree-of-freedom) haptic devices are used. This is a fundamental limitation as 1-DoF devices are less and less common due to recent advances in haptic devices (with 3- and 6-DoF) that are considerably more effective at presenting natural and realistic force-feedback.

This work was further extended by Hinterseer *et al.* in [30–33], where a prediction based on linear extrapolation is introduced in order to further reduce the amount of haptic data packets transmitted over a network. The authors also use a simplified Kalman filter to denoise the sensory readings and improve the prediction performance. An extension to the deadband control principle to devices with 3-DoF is also presented in [30–33]. Their technique is demonstrated to work well in telepresence systems equipped with both, 1-DoF or 3-DoF haptic devices. However, a limitation of their method is that they attempt to predict velocity data using a simple linear extrapolation procedure which relies solely on two previously received sample values. A velocity signal is highly nonlinear, and in consequence, using a basic linear prediction model would generally not yield optimal results. Furthermore, there are important drawbacks in their Kalman filtering procedure. First, the measurement error and the process noise covariance values necessary to initialize the filter must be derived off-line. This process

[2] A system is passive if it absorbs more energy than it produces. In particular, the passivity concept provides a sufficient condition for stability of a haptic force-feedback system. Moreover, this concept is commonly used to stabilize haptic presence systems in telepresence architectures.

is often tedious, time consuming, and error-prone. Moreover, the assumption that the measurement error of a haptic signal remains constant is clearly false, as it will be explained in Section 3.3. This work was, however, finalized in [4] where the authors replaced the Kalman filtering technique with a low-pass filter that has a cutoff frequency beyond the human haptic perception range. The particular type of filter is not discussed in any detail. In addition, the stability (passivity) of the haptic telepresence system using the velocity-prediction-based data reduction method is not investigated.

In [34], Kuschel *et al.* introduced a systematic classification framework for lossy data reduction (LDR) algorithms in haptic telepresence systems. The stability conditions of the presented LDR techniques are also derived; in particular, the methods are passive in nature and consequently guarantee the stability of a telepresence system. The authors divide LDR approaches into two categories: frame-based and sample-based algorithms. The former category includes interpolative and frequency-based LDR techniques. The latter category encompasses extrapolative LDR and direct LDR. Interpolative LDR approximates the incoming signal in the time domain and only the resulting function (e.g., spline) parameters are subsequently transmitted. In addition, the authors argue that a frequency-based approach is not suitable for haptic telepresence as it could introduce (possibly large) additional time delay to the system. This is due to the nature of frame-based LDR techniques which generally make use of large data frames in order to efficiently encode the dominant low frequency components of the signal. Conversely, extrapolative LDR algorithms estimate the future signal parameters based on previous data, whereas direct LDR-based techniques perform data reduction on each sample directly; e.g., coarser quantization, and Hold-Last-Sample techniques. In [35], Kuschel *et al.* extended their previous work and further discussed the implementation and experimental evaluation of two of the four aforementioned haptic data compression techniques, namely the interpolative compression strategy, and the extrapolative compression approach. Psychophysical knowledge from human haptic perception is, however, not incorporated in Kuschel *et al.*'s techniques.

In [36], Kammerl *et al.* proposed an offline coding technique for haptic media (primarily intended for haptic playback applications) that relies on the deadband principle. This is a hybrid compression algorithm that combines the concept of the Just Noticeable Differences with predictive and entropy coding in order to encode the required perceptually-significant haptic data samples, using only a few number of bits. Prediction is performed using a basic first-order linear predictor. The prediction errors are initially encoded using a deadband-based differential coder before being further compressed using adaptive arithmetic coding.

In [37], Vittorias *et al.* presented a novel approach for perceptual coding of haptic data in time-delayed teleoperation. Their approach relied on a novel control scheme and a modified version of the deadband method introduced in [25, 31]. Their control scheme, based on the scattering theory [28, 29], enables the transmission and encoding of haptic data while simultaneously preserving the stability of the system under constant, and known time-delays. Psychophysical experiments with human subjects in telepresence and teleaction systems were, however, not demonstrated.

In very recent work [38], Kammerl *et al.* presented a perception-based haptic data reduction method that relies on a novel velocity-adaptive deadband approach. In particular, the authors extended the deadband method by incorporating an additional perceptual parameter, namely the velocity of the operator's hand movement. Psychophysical experiments conducted by the authors revealed improved data reduction performance when comparing their approach to the traditional deadband method. However, Kammerl *et al.*'s velocity-adaptive deadband scheme was strictly applied on the force channel (i.e., from the TOP to the HSI) of a telepresence system; data reduction on the position/velocity channel (i.e., from the HSI to the TOP) was not conducted.

2.2 Data Reduction in Haptic Pattern Analysis

Relevant feature selection and generation in high-dimensional haptic data is very little explored in the literature. In [10], Orozco *et al.* analyzed different datasets acquired using two haptic-enabled applications: solving a virtual maze, and signing a virtual check. The former application requests a user to solve a 2D maze. Force feedback is provided to restrict a user from passing through the maze walls. The latter application enables users to record their haptic-based handwritten signatures on a virtual check. Users write their haptic signatures while relying on a haptic-enabled virtual stylus. The resulting haptic stimuli mimic the tactile sensations felt when signing a traditional paper check. The purpose of their investigation was to demonstrate the feasibility of user authentication and identification in haptic-enabled applications. In their work, feature selection was performed while relying on a relative entropy measure, which the authors interpreted as a pseudo-metric that defines the distance between an individual's features and those of the entire population of users. Specifically, the relative entropy is recognized as an appropriate information theoretic measure for the information content of a biometric feature and is defined as $D(p \parallel q) = \int p(\mathbf{x}) \log_2(p(\mathbf{x})/q(\mathbf{x}))d\mathbf{x}$. $D(p \parallel q)$ is the relative entropy between the intra-(q(**x**)) and the inter-person (p(**x**)) biometric feature distribution, where **x** denotes a feature vector. Consequently, the relative entropy D corresponds to extra information necessary to express a distribution p based on an assumed distribution q. Orozco *et al.* make use of this measure to evaluate the information content of different haptic-based biometric features. For example, in the analysis of the information content of force data, the authors consider feature vectors of the form $\mathbf{x} = (f_x, f_y, f_z)$. Accordingly, in order to reduce the dimensionality of the haptic datasets, the authors performed feature selection using the relative entropy measure where biometric features with the highest average information content (across a sample of users) were selected. In [11], Orozco *et al.* extended their previous work ([10]), and introduced an alternative technique to significantly reduce the dimensionality of haptic-based digital signatures which relied on a combination of principal component analysis (PCA), quantization, and dynamic time warping. PCA was initially used to linearly map a multidimensional haptic dataset onto a single dimension. The corresponding output is subsequently quantized, and similarity scores between generated compressed signatures are computed using dynamic time warping.

In [39], Iglesias *et al.* exploited a feature generation method to reduce high-dimensional haptic datasets to only three dimensions. This was performed to construct a 3D visual representation of the information and enable 3D visual data mining of haptic datasets. The authors made use of datasets acquired using the aforementioned haptic-based biometric system; three different applications were considered: solving a virtual maze, signing a virtual check, and dialing a number on a virtual phone. The first two applications were described above. Conversely, the haptic-based virtual phone enables a user to simulate a call to another individual. To make a call a user simply dials the number (on the virtual phone) using a virtual stylus guided using a haptic force-feedback device. A single dataset was acquired using each of the aforementioned applications, for a total of three different datasets. First, the authors attempted to explore user-behavior-similarity across the three applications (for a single user and using all the acquired features) using 3D virtual reality-based representations of the recorded data; e.g., to visually inspect whether haptic forces exerted when navigating the virtual maze differ from those applied when signing a virtual check. More precisely, the 3D virtual reality spaces of the haptic datasets were constructed such that the structure of the objects, (e.g., haptic signatures) as determined by the original set of features, is preserved. The experiment were then repeated using only subsets of features to assess the impact of different features on user-discrimination; e.g., from the preliminary results the authors revealed that force and torque features provide information that can improve classification between users.

In follow-up work [40, 41], the authors exploited a visual data mining approach similar to that discussed in [39] to explore the within-user variations and between-user variations of haptic-based handwritten signatures (this dataset was collected with the virtual check application). In particular, a nonlinear transformation was performed on the original high-dimensional haptic dataset to generate Euclidean 3D visual spaces that preserve the similarity structure of the original samples. This mapping is constructed while minimizing an error measure of information loss. Specifically, Iglesias *et al.* exploited a commonly used error measure referred to as the *Sammon error*. The authors, however, do not discuss the optimization method used to minimize the mapping error (*Sammon error*) between the high-dimensional feature space and the 3D visual space.

2.3 Haptic Data Protection through Digital Watermarking

Research on haptic data protection through digital watermarking is in its infancy. All previous research works on this topic (as well as the presented work) address digital watermarking of haptic-enabled 3D models. This is in contrast to the suggested haptic data reduction schemes which are mainly concerned with *raw* haptic signals, i.e., the algorithms deal primarily with acquired haptic data values which include position, force, orientation, etc. Nevertheless, the conceptual similarity between haptic watermarking and haptic data reduction methods is quite apparent. As aforementioned, in both research areas, methods are developed to intelligently manipulate and modify haptic signals with the purpose of achieving greater system reliability, robustness, and effectiveness. This, however, must be achieved while ensuring that any distortions introduced by the algorithms are undetectable to the user. This is important in order to maintain high-quality haptic interaction.

Despite its importance, only one other research group has investigated watermarking of haptic-enabled 3D models. Prattichizzo *et al.* ([42]) were the first to introduce the idea of haptic digital watermarking. In [13, 42] the authors proposed a haptic watermarking technique where a sinusoidal watermark is superimposed onto a sinusoidal host surface texture signal. More specifically, Prattichizzo *et al.* examined the perceptibility of haptic watermarks embedded in the textures of virtual 3D objects. Moreover, in order to achieve imperceptible haptic watermarking, they carefully embedded the sinusoidal watermark such that its amplitude does not exceed the human detection threshold; specifically, the threshold depicts the minimum signal strength required to perceive a change in haptic stimuli. Prattichizzo *et al.* also demonstrated how the haptically imperceptible sinusoidal watermark can be detected by means of spectral analysis.

In [13, 43], psychophysical experiments were conducted to measure the perceptibility of a watermark embedded in a 3D mesh. Watermark embedding was achieved by altering the positions of the mesh vertices. The experiments were performed using several 3D meshes with different resolutions in an attempt to establish a relationship between the watermark strength and the size of the triangular mesh components. The watermark perceptibility is inspected through a haptic or visual interface, i.e., the user can only see or touch the virtual object. Their preliminary results revealed that watermark detection thresholds increased as a function of the mesh resolution, suggesting that more noise (or a higher strength watermark) can be embedded in surfaces with coarser representation. The latter observation was consistent for both, haptic- and vision-based watermark detection, i.e., experiments during which watermark detection was performed using strictly visual or haptic feedback. In addition, their results demonstrated that watermarks embedded in 3D meshes are more easily detected through a haptic interface rather than by visual inspection. In [44], Prattichizzo *et al.* extended their previous study and presented another experiment which aimed at estimating watermark detection thresholds using two different haptic interfaces, a PHANTOM Desktop and a PHANTOM Premium 1.5. The former featuring a stylus type end effector, whereas the latter is equipped with a thimble end effector (both devices provide 3-DoF haptic force feedback). Their results demonstrated that the acquired detection thresholds using the two aforementioned haptic devices where approximately of equal value for different resolutions of the watermarked surface mesh. This observation suggested that haptic watermark detection performance could in fact be independent of the exploited haptic device.

Part II

Data Reduction in Telehaptic Systems

Chapter 3

Data Reduction for Haptic Communication in 3-DoF Telementoring Systems

3.1 Introduction

In the next two chapters, novel prediction-based haptic data reduction and transmission techniques are introduced to efficiently reduce the packet rate in bilateral telehaptic environments, while ensuring that the hard-real time constraints of such systems are met. The focus of this chapter is on haptic-enabled telementoring. Telementoring is a branch of bilateral telehaptics; it enables one experienced individual to mentor another individual (i.e., the mentee) over a network connection while incorporating the sense of touch. Moreover, telementoring differs from conventional teleoperation systems in its flexible control architecture that can adapt to different learning scenarios and modalities. In turn, the proposed haptic packet reduction and transmission techniques consider both the bilateral and flexible nature of telementoring systems and the inherent network limitations that already impact the quality of telehaptic systems, such as network impairments due to delay, jitter, and packet loss.

More specifically, in this chapter, two distinct methods are introduced to reduce the packet rate in 3-DoF haptic-enabled telementoring systems. However, as it will be illustrated in Chapter 4, the techniques can in a straightforward manner, be extended and applied to other telehaptic systems (and applications) that encompass haptic devices with multiple-DoF (e.g., 6-DoF). The first data reduction technique relies on the limitations of human haptic perception (i.e. the Just Noticeable Differences [JND]) with respect to a user's hand velocity in order to reduce the haptic data rate. Similarly, the second technique extends the first to reduce the amount of haptic data packets; however, it relies on the JND of human haptic perception with respect to a user's hand position. The least-squares (LS) method is used to obtain reliable velocity estimates [45], required in the two suggested haptic data reduction architectures. Conversely, median filtering (MF) is exploited to eliminate any fine irregularities and outliers remaining in the computed velocity measures while preserving the true signal discontinuities [46]. In addition, an adaptive prediction technique that makes use of the filtered velocity estimates to predict a user's haptic movement and the corresponding haptic force is derived. This prediction technique is used in the two suggested data reduction methods as will be clearly illustrated in Section 3.3.

The rest of the chapter is organized as follows. In Section 3.2, haptic-enabled telementoring is introduced. In Section 3.3 the proposed haptic data reduction and transmission techniques are presented. In Section 3.4 the materials and methods used during the experimental procedure are illustrated. In Section 3.5 the experimental verifications are provided. In Section 3.6, a discussion of the results is presented. Finally, a summary of the chapter is provided in Section 3.7.

3.2 Introduction to Haptic-enabled Telementoring

Telementoring is a particular form of bilateral telehaptics that enables direct coupling of haptic devices over a network. However, it follows a more flexible control architecture, which can dynamically be tuned to improve its effectiveness according to different learning scenarios and modalities. Practically, it consists of a teaching technique that involves real-time guidance of a less experienced trainee through a procedure in which the student has limited experience.

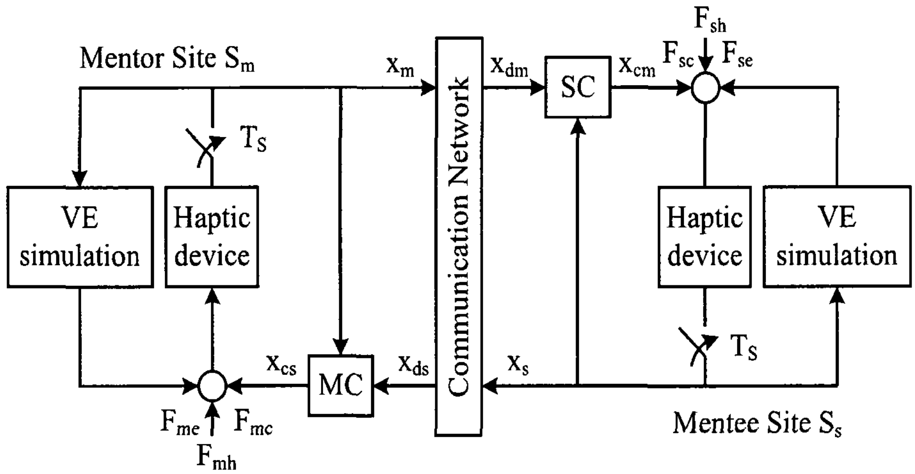


Figure 3.1: General architecture of a haptic-enabled telementoring system.

Hence, a remote expert can guide a student's hand through an intricate procedure over a network connection while incorporating the sense of touch. The general architecture of a haptic-enabled telementoring system is shown in Fig. 3.1; its elements are described as follows:

- Subscripts m and s refer to the master (mentor) and the slave (student), whereas S_m is the mentor site, and S_s is the student site.

- MC and SC represent the haptic real-time controller (HRTC) at sites S_m and S_s. The HRTC serves several purposes, such as compensating for network latency and maintaining the synchronization and stability of the system, and regenerating the output haptic forces applied at the remote site.

- x_m and x_s denote the position of the end effectors of the haptic devices at sites S_m and S_s.

- x_{dm} and x_{ds} correspond to the possibly delayed device position values at sites S_m and S_s.

- x_{cm} and x_{cs} denote the compensated (if required) device position values at sites S_m and S_s.

- The haptic devices' position data are sampled and transmitted at a time interval of length T_s. High sampling rates in the range of 1 kHz are required to guarantee the stability of the control loops at S_m and S_s.

- F_{me} and F_{se} are the output forces computed based on the virtual environment simulation at S_m and S_s.

- F_{mh} and F_{sh} denote the forces exerted by the mentor and the student, respectively.

- F_{mc} and F_{sc} correspond to the output forces generated by the HRTCs at the two sites.

The force values sent to the haptic device motors at S_m and S_s are depicted below in Eqs. (3.1) and (3.2), respectively.

$$\begin{aligned} F &= F_{mc} + F_{me} \\ &= K_{mp}(x_{cs} - x_m) + K_{md}(\dot{x}_{cs} - \dot{x}_m) + F_{me} \end{aligned} \tag{3.1}$$

$$\begin{aligned} F &= F_{sc} + F_{se} \\ &= K_{sp}(x_{cm} - x_s) + K_{sd}(\dot{x}_{cm} - \dot{x}_s) + F_{se}. \end{aligned} \tag{3.2}$$

The first two terms in Eqs. (3.1) and (3.2) form the expression of a classical proportional-derivative controller, where K_{mp} and K_{sp} denote the adjustable proportional gains at S_m and S_s, whereas K_{md} and K_{sd} correspond to the derivative gains. Practically, the mentor can adjust K_{sp} based on the student's performance. If the student is confident of his or her skills, the mentor (or the student for that matter) can lower the value of K_{sp} to enable the trainee to feel more force feedback from the simulation itself.

In the presented work, the general telementoring architecture is extended to include haptic data reduction. This is important as a fundamental requirement of haptic-enabled telementoring and teleinteraction systems is to minimize transmission delay and the amount of haptic

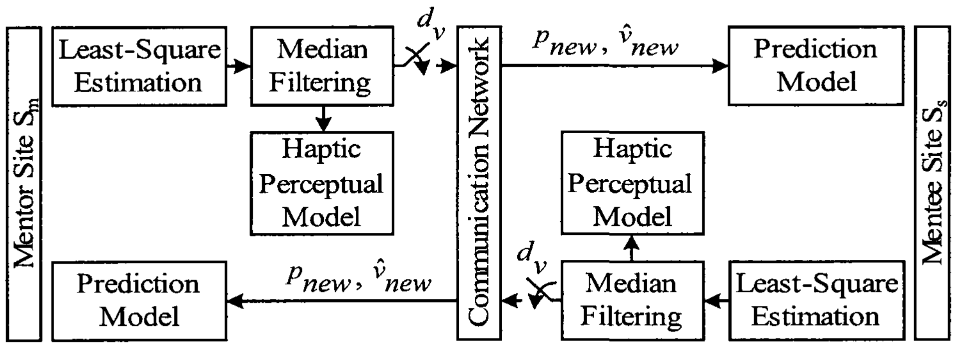


Figure 3.2: Block diagram architecture of the velocity-deadband-based haptic packet reduction technique.

data sent to meet the hard real-time constraints of haptic data streams. The stability of telementoring systems and the implications of additional communication time delay have extensively been examined in other works performed at DISCOVER Laboratory of the University of Ottawa [47, 48] and will not be investigated further in this chapter.

3.3 Perceptual Haptic Data Reduction

In this section, two novel haptic data reduction and transmission methods for reducing the amount of transmitted packets in haptic telementoring systems are presented. The first technique is shown in Fig. 3.2, whereas the second method is shown in Fig. 3.3. The modules that constitute the two haptic data reduction architectures are described in detail in the succeeding sections. The fundamental differences between the two techniques are also highlighted.

3.3.1 Haptic Perceptibility

The suggested transmission techniques rely on the characteristics of haptic perceptibility to reduce the amount of data packets sent over a communication network. Human haptic perception is generally analyzed using Weber's law of JNDs. The JND consists of the minimum amount of change in stimulus intensity, which results in a noticeable variation in sensory experience.

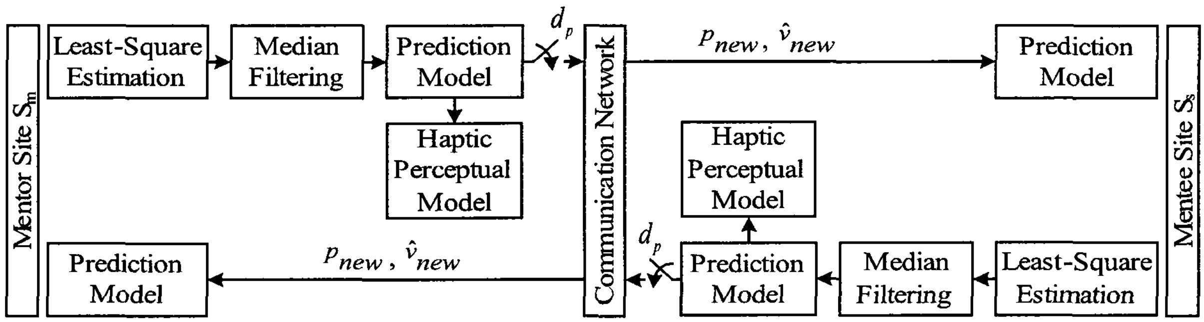


Figure 3.3: Block diagram architecture of the position-deadband-based haptic packet reduction technique.

This relation can be expressed as $\frac{\Delta I}{I} = k$, where ΔI denotes the difference threshold (or the JND) and the constant k expresses the linear relationship between the JND and the initial stimulus I. Numerous psychophysical studies have been conducted in the literature to derive the JND of haptic perception. For kinesthetic force feedback (typically exploited in haptic telementoring systems), several factors impact the overall human haptic perception JND, including force, pressure, velocity, and the position of the joints. Generally, the JND for human haptic perception ranges from 5% to 15% [49–53]. This suggests that, in a telementoring system, if the change in magnitude of the haptic force currently applied by the mentor is less than the JND, the mentee would not perceive a force-feedback variation. This concept is used in the two suggested packet reduction methods, where haptic data are only transmitted if the remote operator is likely to perceive a change in the kinesthetic force-feedback response. Furthermore, the haptic perceptual model included in the two data reduction methods is based on existing psychophysical knowledge of human haptic perception and the deadband principle presented in [4]. The deadband principle essentially states that signal changes do not need to be transmitted over the communication network, unless they exceed a predefined threshold. This can be depicted as follows:

IF $\left(|\mathbf{v}_i - \mathbf{v}_c| > f(\mathbf{v}_i)\right)$ THEN *transmit* $\mathbf{v}_c$ ELSE *do nothing*.

In other words, to determine whether to transmit the currently acquired data vector $\mathbf{v}_c$, the Euclidean distance between $\mathbf{v}_c$ and the initial data vector $\mathbf{v}_i$ must be larger than $f(\mathbf{v}_i)$, which corresponds to the radius of a spherical region (since we are dealing with 3D vectors), referred to as the deadzone. The first suggested method (see Fig. 3.2), relies on a velocity deadband with a deadzone of radius $f(\mathbf{v}) = d_v \cdot \mathbf{v}$, where $\mathbf{v}$ denotes a 3D velocity vector and d_v is the deadband. Conversely, the second method (see Fig. 3.3) relies on a position deadband with a deadzone of radius $f(\mathbf{p}) = d_p$, where $\mathbf{p}$ denotes a 3D position vector and d_p is the deadband. A detailed comparison of the two data reduction methods is provided in Section 3.3.4.

3.3.2 Haptic Prediction Model

To further reduce the amount of haptic packets transmitted and improve the reconstruction of missing data samples, haptic data prediction is exploited in the two suggested techniques, and computed using the position information of the haptic device's end effector as follows:

$$\begin{aligned}\hat{p}(n+1) &= \dot{p}(n)\Delta T + p(n) \\ &= \hat{v}(n)\Delta T + p(n),\end{aligned} \tag{3.3}$$

where $p(n)$ denotes the current position value, $\hat{v}(n)$ is the current estimated velocity, $\hat{p}(n+1)$ corresponds to the predicted position estimate, and ΔT is the sampling period. It can be observed that, in the velocity-deadband-based data reduction technique, prediction is solely used to reconstruct the haptic signal based on a number of received data samples. Similarly, in the position-deadband-based data reduction method, prediction is again used to efficiently reconstruct missing haptic data samples; however, it is also exploited to further reduce the amount of packets transmitted. The reason behind these data reduction strategies will be discussed in detail in Section 3.3.4. Thus, the transmission method exploits the suggested linear prediction technique as follows:

$$p_i = \begin{cases} p_{new} & \text{if packet sent or received} \\ p_{i-1} + \hat{v}_{new} \cdot \Delta T & \text{otherwise} \end{cases} \tag{3.4}$$

where p_{new} and $\hat{v}_{new}$ are the newly transmitted or received position and velocity data samples, respectively; p_i and p_{i-1} are the most recent outputs of the prediction model; and ΔT is the sampling period. Essentially, when the predictor is used on the transmitting end, the predicted value must fall within the predefined deadzone; otherwise a new sample is transmitted over the network. Conversely, when the predictor is exploited on the receiving end of the data reduction techniques, it attempts to reconstruct the missing haptic data samples based on newly transmitted packets. Hence, the prediction model compensates for missing data packets or introduced communication delay to regularly provide the haptic real-time controller (used to calculate forces sent to the haptic device) with data samples at a consistent rate of 1 kHz or better. This is necessary to guarantee the stability of the closed-control loops and, in turn, maintain compelling and stable force feedback at S_m and S_s.

3.3.3 Least-Square Estimation and Median Filtering

It can be observed from Eq. (3.4) that the prediction performance can be positively or negatively affected, depending on the precision of the acquired position and velocity data. Generally, the noise present in position readings is low due to the fact that haptic devices are typically equipped with joint sensors. However, velocity is generally derived from position measurements using the finite-difference method which leads to very noisy velocity estimates at high sampling frequencies [54]. To illustrate this mathematically, consider that $y(n)$ is the true position and that $p(n) = y(n) + e(n)$ is the measurement (as we have so far assumed) with error $e(n)$. The finite-difference method translates to the following formulation:

$$\hat{v}(n) = \frac{p(n)-p(n-1)}{T_s} = \frac{y(n)-y(n-1)}{T_s} + \frac{e(n)-e(n-1)}{T_s}, \tag{3.5}$$

It can be observed that, as the sampling period T_s (which is equivalent to ΔT mentioned earlier) becomes smaller, the variation between successive position samples decrease; however the noisy components increase and are correspondingly amplified. Thus, in order to obtain more precise velocity estimates, the least-squares method is used and can be described as follows:

$$\hat{v}(n) = \sum_{i=0}^{M-1} w(i)p(n-i) \tag{3.6}$$

where $w(i)$ corresponds to the filter coefficients and M denotes the filter length. Essentially, a velocity estimate $\hat{v}(n)$ (i.e., the slope) that would best fit the observation data $p(n)$ in the least-squares sense is derived. The linear equation is then formulated as

$$A\vec{x}(n) = \vec{p}(n) \tag{3.7}$$

where

$$A = \begin{bmatrix} 0 & 1 \\ -T_s & 1 \\ \vdots & \vdots \\ -iT_s & 1 \\ \vdots & \vdots \\ -(M+1)T_s & 1 \end{bmatrix}$$

is the data matrix, $\vec{x}(n) = [v(n) \;\; b(n)]^T$ is the unknown to be solved[1], and $\vec{p}(n) = [p(n)\; p(n-1) \cdots p(n-M+1)]^T$ is the observations. The solution to Eq. (3.7) in the least-squares sense is

$$\hat{\vec{x}}(n) = (A^T A)^{-1}(A^T \vec{p}(n)). \tag{3.8}$$

Solving for $\hat{\vec{x}}(n)$, the velocity estimates of Eq. (3.6) can be derived, where the corresponding filter coefficients can be computed as follows:

$$w(i) = \frac{12\,(\frac{M-1}{2} - i)}{T_s M(M^2-1)}, \quad 0 \le i \le M-1 \tag{3.9}$$

[1] $b(n)$, which denotes the intercept of the linear equation, is not derived, as only the velocity $v(n)$ is required.

where, ideally, the least-squares filter length M should be short when the haptic signal velocity is high to capture the transient behavior while minimizing the computational time and delay, and long when the velocity is low, yielding more precise estimates. Velocity estimation with filter length adaptation was initially attempted using the adaptive-windowing method presented in [54]. However, when comparing this method to the fixed filter alternative, relatively no data reduction improvement was observed, even though the former technique was computationally more expensive. Therefore, the filter length has been selected such that a compromise between reliable velocity estimation and fast computation is achieved. Furthermore, median filtering is subsequently used to eliminate any fine irregularities and outliers remaining in the computed velocity measures while preserving the true signal discontinuities [46]. The filter essentially replaces a signal sample with the median of the samples located within a sliding (filter) window. This classical filtering technique is computationally simple as the window size does not need to be very large to be effective (a typical window size is 5).

Additionally, it should be further clarified that the choice of the proposed first-order prediction approach is primarily based on the numerous advantages of least-squares-based techniques. In particular, least-squares methods have excellent characteristics in terms of stability and efficiency [55]. Moreover, the suggested prediction method is primarily inspired from Janabi-Sharif *et al.*'s aforementioned article [54], which proposed a least-squares-based method for velocity estimation from discrete and quantized (haptic) position samples. The authors demonstrated that a least-squares-based approach, although simple in nature, can potentially outperform more sophisticated methods, including Kalman filter-based velocity estimation.

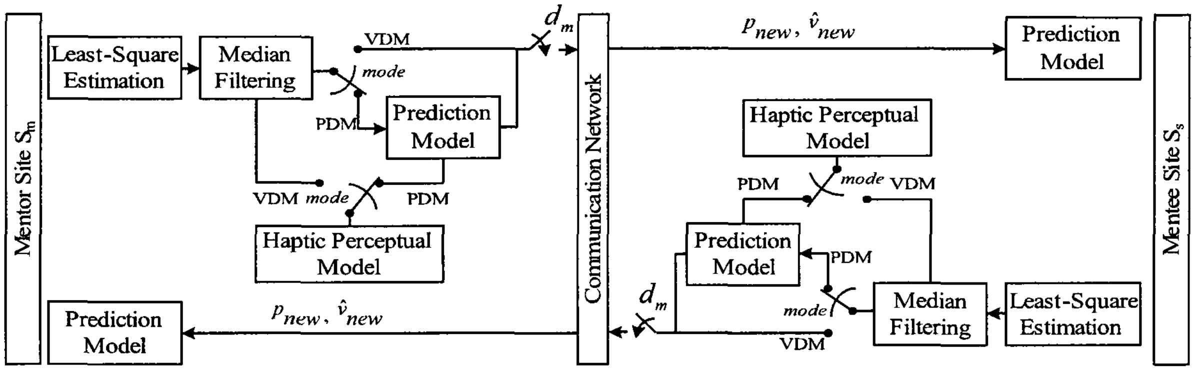


Figure 3.4: Block diagram architecture of the position and velocity-deadband-based haptic data reduction methods.

3.3.4 Comparison between Velocity and Position-Deadband-based Haptic Data Reduction

The fundamental difference between the two suggested data reduction techniques is the fact that the first relies on the JND of haptic force perception with respect to a user's hand velocity, whereas the other is with respect to a user's hand position. The differences and similarities between the position-deadband-based method (PDM) and the velocity-deadband-based method (VDM) is better illustrated in Fig. 3.4 as it integrates both techniques into a single architecture. Furthermore, as it is shown in Figs. 3.2 and 3.3, in contrast to the first method, which only includes a prediction module used to reconstruct the haptic signal based on a number of received data samples, the second method (position-deadband-based) encompasses the latter module and another prediction unit used to further reduce the amount of transmitted data packets. This difference is due to the fact that, contrary to a position signal that is relatively continuous and linear within a certain time interval (considering that we are working with a sampling frequency that is greater than or equal to 1 kHz), a velocity signal is generally highly nonlinear; consequently, using a computationally simple linear prediction model did not result in any improvement in data reduction.

3.4 Materials and Methods

3.4.1 Experimental Setup

The experimental setup consists of two identical haptic devices connected to two separate PCs over a 100 Mbps Ethernet local area network (LAN). The haptic stimuli at both the mentor and mentee (i.e., student) ends are refreshed at a rate of 1000 Hz. Essentially, packets between the two personal computers (PCs) are transmitted at a rate of 1 kHz. The suggested haptic data reduction and transmission techniques are executed on both PCs due to the bilateral nature of telementoring systems. Consequently, haptic position and velocity data are bilaterally transmitted between the mentor and mentee PCs. This is necessary to allow for the direct coupling of the haptic devices over the communication network and, in turn, enable a hand-by-hand teaching experience. Two primary experiments were performed to evaluate the suggested packet reduction methods: The first experiment serves as a performance validation of the two data reduction techniques, where a user is guided to perform a simple predefined task and is only provided with haptic feedback, i.e., no visual feedback is available. Conversely, the second experiment is a direct application of the suggested packet reduction techniques to a haptic-enabled telementoring surgery simulation (see Fig. 3.5), where a participant is guided to perform a relatively complex task and provided with both haptic and visual feedback. Furthermore, the experiments were conducted using two distinct pairs of haptic devices. Initially, results for the two experiments were obtained while using the common Sensable PHANTOM Omni haptic devices. Then, the experiments were repeated while replacing the Omni devices with two high-precision Sensable PHANTOM Desktop haptic devices. The PHANTOM Omni and the PHANTOM Desktop devices are equipped with 6-DoF positional sensing and 3-DoF haptic force feedback. However, of the two devices, the PHANTOM Desktop is recognized for its higher fidelity, stronger forces, and lower friction.

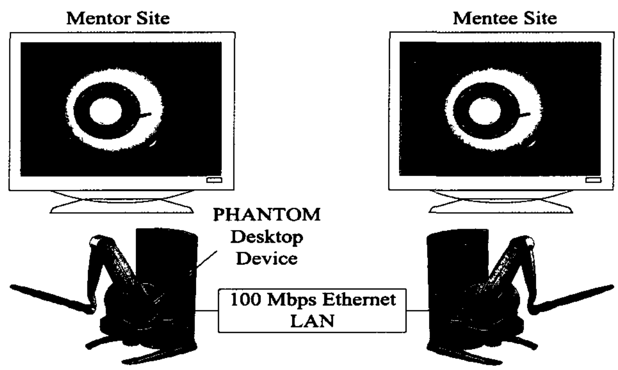


Figure 3.5: Example setup of a haptic-enabled telementoring environment.

3.4.2 Subjects

Altogether, ten subjects (six males and four females, aged between 25 and 35), of which nine are right handed and one is left handed with no known sensorimotor impairments with their hands, took part in the experiments. Their prior experience with PHANTOM devices ranged from novice to expert. All participants were remunerated for their time.

3.4.3 Procedure

The experiments were arranged in three blocks per subject: a practice block and two experimental blocks (i.e., two runs). For each participant, a new experiment is performed (i.e., two experimental blocks) every time a device change is carried out or when the network parameters, such as the simulated communication delay time or the packet loss rate, are modified. Moreover, the practice and the experimental blocks follow the same procedure; however, the practice block is intended to familiarize the subjects with the haptic telementoring system and to give them a feel of the distortion encountered when the suggested haptic packet reduction techniques are enabled. For the velocity-deadband-based data reduction technique, each block consisted of 14 intervals differing in the deadband size. During the 14 intervals of each block, the following velocity deadband (d_v) values were used in a randomly selected order: 0%, 2.5%, 5%, 7.5%, 10%, 12.5%, 15%, 17.5%, 20%, 22.5%, 25%, 30%, 35%, and 40%.

Similarly, for the position-deadband-based data reduction method, each block consisted of 15 intervals, where the position deadband (d_p) values were selected in a random order and are as follows: 0mm, 0.05mm, 0.1mm, 0.2mm, 0.3mm, 0.4mm, 0.5mm, 0.6mm, 0.7mm, 0.8mm, 0.9mm, 1.0mm, 1.2mm, 1.6mm, and 2.0mm. After every interval, the subjects were requested to rate the haptic telementoring experience using the following five-grade impairment scale:

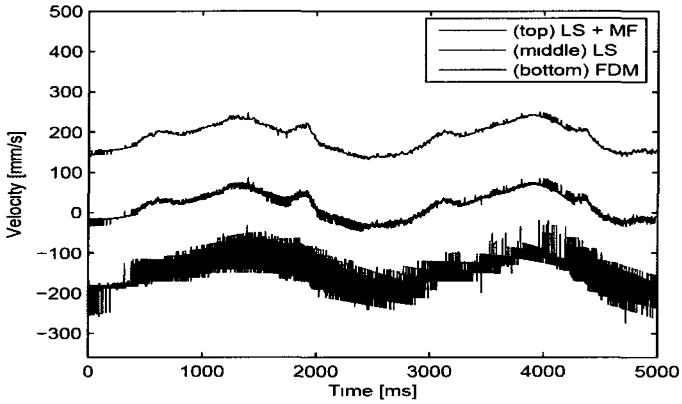
5 – Imperceptible;
4 – Perceptible, but not annoying;
3 – Slightly annoying;
2 – Annoying;
1 – Very annoying.

In addition, the users were also permitted to give in-between values when necessary (i.e., 4.5, 3.5, 2.5, and 1.5) to better assess their opinion.

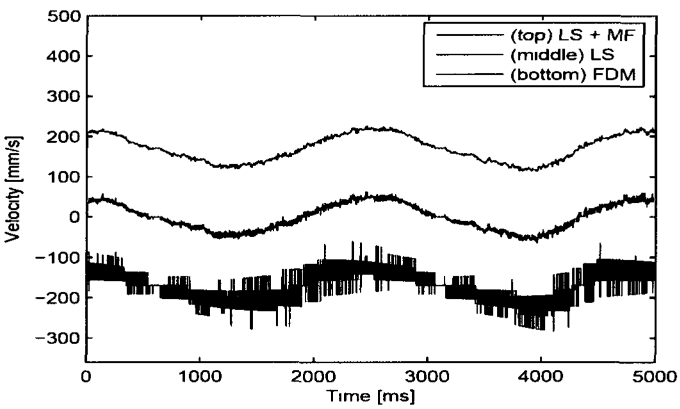
3.5 Experimental Verifications

3.5.1 Velocity Estimation Results using the Least-Squares Method and Median Filtering

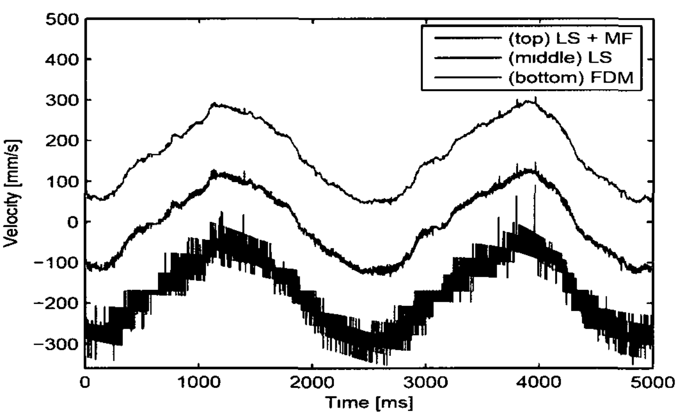
As previously mentioned, velocity estimation has a direct impact on the prediction accuracy, and consequently, on the performance of the packet reduction techniques. In order to emphasize the importance of precise velocity estimation and demonstrate the validity of the suggested technique, a comparison between three different methods which include, the traditional finite-difference method (FDM), least-squares (LS), and least-squares with median filtering (LS + MF), is illustrated in Fig. 3.6. Specifically, Fig. 3.6 (a) – (c) corresponds to the velocity estimates along the x, y, and z axes of a sinusoidal signal (with a period of 5 cm and peak-to-peak amplitude of 10.4 cm) traced offline for a duration of 5 seconds using the PHANTOM



(a)



(b)



(c)

Figure 3.6: Responses of the finite-difference method (FDM), least-squares method (LS), and least-squares with median filtering (LS + MF) along the (a) x-axis, (b) y-axis, and (c) z-axis. (FDM and LS + MF results have been shifted by -170 and 170 mm/s, respectively).

Omni haptic device. It is clear that the finite-difference method is not suitable due to its high-frequency noise. Conversely, although the least-squares method leads to relatively smooth results, velocity estimates computed using least-squares and median filtering are superior.

3.5.2 Experiment I: Validation of the Data Reduction Techniques

In this section, a validation of the two haptic data reduction techniques is performed. The displaying stimuli of the experiment are initially presented, followed by a brief description of the experimental results.

Displaying Stimuli

In this experiment, a simple experimental environment is used to enable an unbiased validation of the suggested methods. Essentially, during each run, users are guided to trace a periodic sinusoidal signal (with a period of 5 cm and peak-to-peak amplitude of 10.4 cm) for a duration of approximately 20 seconds. The mentor repeatedly traces a sketch of the sinusoidal signal while remaining within the boundaries of the 3D workspace of the haptic device. Conversely, the participant (i.e., the mentee) is not provided with a visual description of the signal and is passively following the mentor's movements.

Results

The haptic data reduction techniques are initially examined under normal network conditions using PHANTOM Omni and Desktop devices. This is performed to evaluate the packet reduction methods when a haptic device with reasonable precision and force-feedback quality is used (Omni) as opposed to a higher-end device with higher fidelity, stronger force, and lower friction (Desktop). The average number of transmitted packets (percentage rate) with respect to the applied velocity deadband and the position deadband values are plotted in Fig. 3.7 (a) and (b) respectively. This is performed to evaluate the data reduction techniques under normal network conditions and with network-induced time delays (TD) and packet loss (PL). First, un-

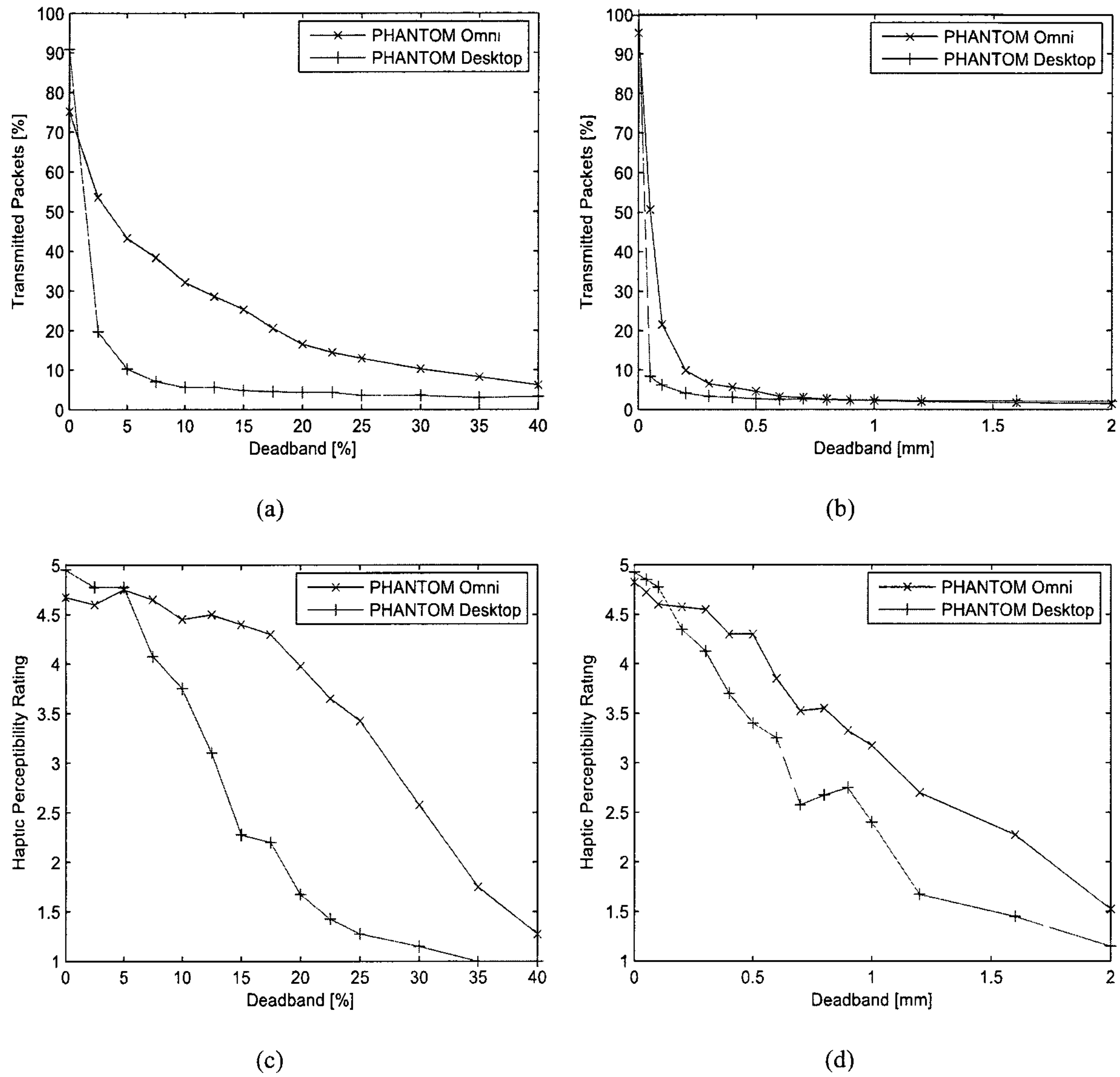


Figure 3.7: Validation results of the haptic packet reduction techniques under normal network conditions.

der normal network conditions and using the VDM, for each of the applied velocity-deadband values, the perceptibility ratings of all ten subjects and for the two experimental blocks were averaged as shown in Fig. 3.7 (c). It can observed from Fig. 3.7 (c) that the velocity deadband can be increased to as much as 20% for the PHANTOM Omni and slightly higher than 7.5% for the PHANTOM Desktop, while still obtaining a perceptibility rating of 4, which

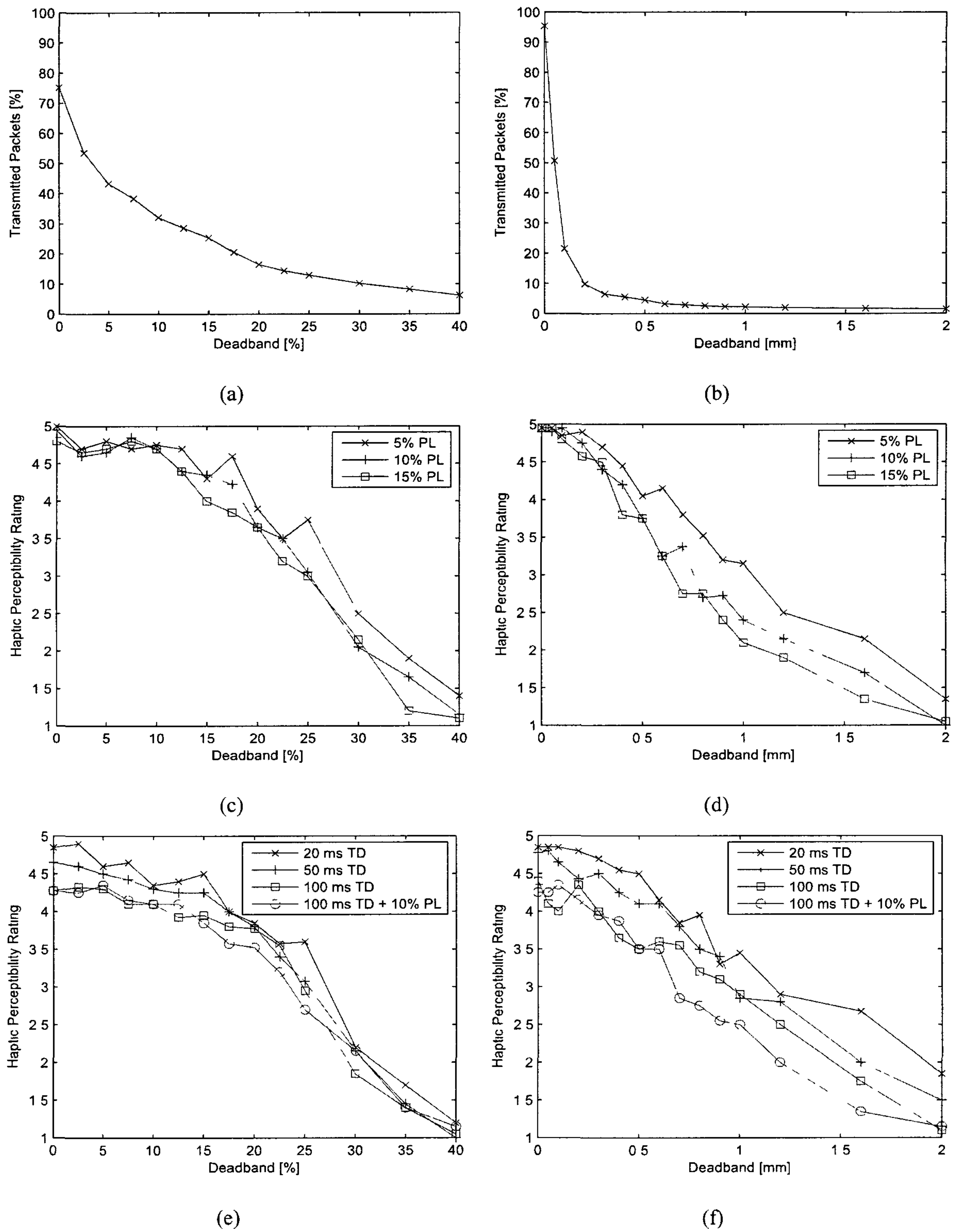


Figure 3.8: Validation results of the haptic packet reduction techniques under network-induced time delays (TD) and packet loss (PL).

most participants described as fairly distortion free. Moreover, as it is shown in Fig. 3.7 (a), with the 20% velocity deadband that can be permitted when using PHANTOM Omni devices, only 16.5% of the packets are transmitted. Conversely, when PHANTOM Desktop devices are utilized, the amount of packets is reduced to less than 7.1% with the barely perceivable 7.5% deadband. Similar to the first technique, to assess the effectiveness of the second method (PDM) under normal network conditions, for each of the applied position-deadband values, the perceptibility ratings of all ten subjects and for the two experimental blocks were averaged as depicted in Fig. 3.7 (d). It can be seen that the position deadband can be increased over 0.5mm for the PHANTOM Omni and slightly higher than 0.3mm for the PHANTOM Desktop while still obtaining a perceptibility rating well above 4. In addition, with the 0.5mm position deadband allowed when using PHANTOM Omni devices, only 4.6% of the packets are transmitted as it is shown in Fig. 3.7 (b). Conversely, when PHANTOM Desktop devices are exploited, the amount of packets is reduced to less than 3.3% with a hardly perceivable deadband of 0.3mm. To avoid redundancy in the presented results, only the more popular and cost effective PHANTOM Omni devices were used to evaluate the robustness of the suggested methods under simulated network-induced delays and packet loss. From Fig. 3.8 (a) and (c), it is shown that with a very high network packet loss rate of 15% only 25.4% of the packets need to be transmitted to obtain a desirable perceptibility rating of 4 when using the VDM. Similarly, from Fig. 3.8 (b) and (d), it is revealed that with a packet loss rate of as much as 15%, a maximum of 6.1% of the packets need to be transmitted to obtain a perceptibility rating of 4 when using the PDM. A quick comparison of Fig. 3.8 (a) with Fig. 3.8 (e) and Fig. 3.8 (b) with Fig. 3.8 (f) reveals that even with poor network conditions, when a time delay (round trip) of 100 ms together with a 10% packet loss rate are induced on the bilateral communication channel, only 27% and 6.6% of the packets need to be transmitted using the VDM and the PDM, respectively.

3.5.3 Experiment II: Application of the Data Reduction Techniques in a Haptic-enabled Surgery Simulation

Surgical simulation is receiving increasing attention throughout the medical community as it can provide high-fidelity training to increase the diffusion of innovative and less-invasive procedures while decreasing the surgeon's learning curve [56]. In this section, the packet reduction techniques were directly applied in a haptic-enabled telementoring eye surgery simulation. The haptic and visual stimuli of the experiment are initially presented, followed by a brief illustration of the experimental results.

Displaying Stimuli

The surgery application exploited here consists of an OpenGL-based hapto-visual eye surgery training simulation developed in collaboration with the University of Ottawa Eye Institute to train novice surgeons in performing eye cataract surgeries [47]. The application encompasses all the (virtual) instruments and tools necessary to effectively perform a cataract surgery. In this experiment, the same visual and haptic stimuli are provided at the mentor and mentee sites (refer to Fig. 3.5). The stimuli consist of the aforementioned eye surgery training simulation environment that is haptically sensed using PHANTOM haptic devices. During each run, the mentor haptically guides the participant to perform a simulated eye surgery for a duration of approximately 30 seconds.

Results

The results obtained in this experiment were evaluated as in Section 3.5.2. However, to avoid redundancy in the reported results, the evaluation of the data reduction architecture against adverse network conditions is not performed in this section as it has already been demonstrated in Section 3.5.2. It can be seen from Fig. 3.9 (c) that the velocity deadband can be increased to as much as 22.5% for the PHANTOM Omni and 12.5% for the PHANTOM Desktop, while still achieving a perceptibility rating of approximately 4. In addition, as it is shown in Fig. 3.9 (a),

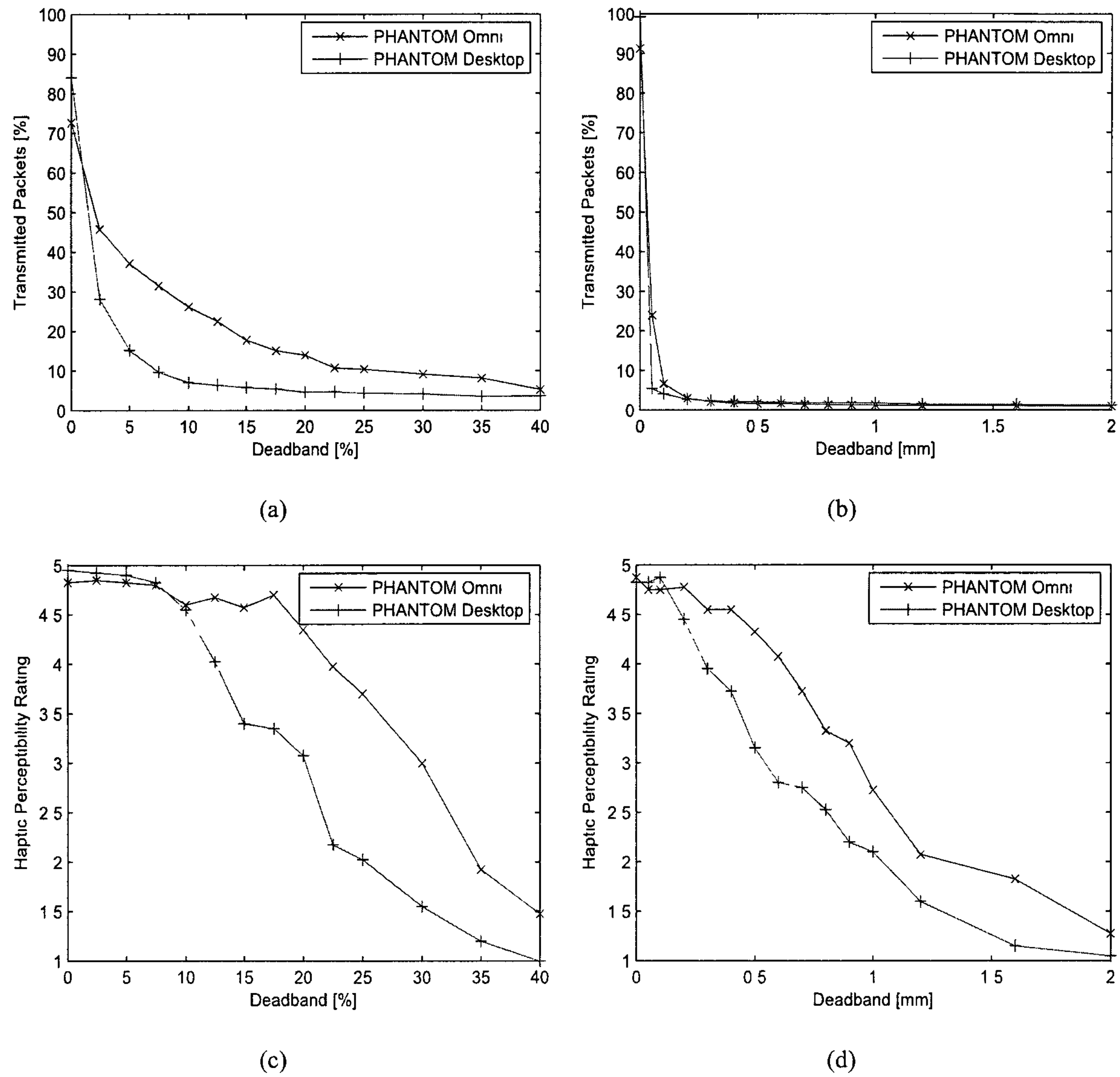


Figure 3.9: Results of the data reduction techniques when applied to a haptic-enabled telementoring eye surgery simulation.

with the 22.5% velocity deadband that is considered acceptable in this case when using PHANTOM Omni devices, only 10.8% of the packets are transmitted. However, when PHANTOM Desktop devices are exploited, the amount of packets is reduced to 6.4% with the barely perceivable 12.5% deadband. Conversely, it can be observed from Fig. 3.9 (d) that the position deadband can be as high as 0.6mm for the PHANTOM Omni and 0.3mm for the PHANTOM

Desktop, and still attain a perceptibility rating of 4. Consequently, with the 0.6mm position deadband tolerated with PHANTOM Omni devices, merely 2% of the packets are transmitted as depicted in Fig. 3.9 (b). Similarly, when PHANTOM Desktop devices are utilized, the amount of packets is reduced to less than 2.4% with the perceptually reasonable 0.3mm deadband.

3.6 Discussion

Many observations can be made from the results presented in Figs. 3.7, 3.8, and 3.9. First, if we consider the validation results shown in Fig. 3.7, it is clear that, of the two suggested packet reduction techniques, the position-deadband-based method is more superior. In addition, as it is demonstrated in Fig. 3.7 (b) and (d), while using the PDM, a rapid drop in the number of transmitted packets can be observed with small perceptually insignificant deadbands. For example, with a minuscule position deadband of 0.05mm, only 8.4% of the packets are transmitted when using the PHANTOM Desktop device, with a corresponding (and impressive) perceptibility rating of 4.85. Furthermore, it can be observed that, at smaller deadband values, greater packet reduction is achieved when using the PHANTOM Desktop device. Moreover, the JNDs of haptic perception with respect to a user's hand velocity and hand position (i.e., velocity and position deadbands) are noticeably lower when using PHANTOM Desktop devices, as opposed to PHANTOM Omnis. This is due to the fact that the PHANTOM Omni has a lower position resolution and higher backdrive friction, which negatively affect a user's haptic perception. Consequently, these device limitations mask certain distortions introduced by the packet reduction methods. Moreover, from Fig. 3.8, it is clearly demonstrated that both, the PDM and the VDM are quite robust against network-induced delays and packet loss. However, in this case as well, the PDM clearly outperforms the VDM by a fairly large margin. Conversely, the results presented in Fig. 3.9, which served to demonstrate the performance of the data reduction methods in a haptic-enabled telementoring surgery application were, for the most part, very comparable to the provided validation results. In fact, it can be seen that the perceptually reasonable deadband values (see Sections 3.5.2 and 3.5.3) in the second experiment were equal to or slightly higher than those obtained in the first experiment. In addition, for different deadband values, better packet reduction results were generally obtained in the second experiment. This can be due to many factors; a possible reason is that, during the telementoring surgery experiments, haptic movements were restricted to a smaller workspace (the virtual eye), making the packet reduction distortion less apparent to the user. The difference in visual stimuli (no visual display was available in the validation experiments) could also have played a role in the obtained results.

A comparison of the results with other related works is difficult due to the fact that we are the first in the literature (to the best of our knowledge) to perform data reduction where two 3-DoF haptic devices are coupled over a network. The other similar perception-based packet reduction architecture presented in [4] is primarily intended for virtual telepresence systems. In addition, the robustness of the authors' prediction-based data reduction technique to network artifacts is not discussed in any detail. Moreover, only the local site is equipped with a haptic device, whereas the remote site is a static virtual environment. Conversely, in our experiments, we have to consider the stability of the device coupling and the dynamics of the virtual environment. Nevertheless, if we disregard these differences, a direct comparison of the results obtained in this work and those reported in [4] suggests that our technique (PDM) outperforms the latter method as it can further reduce the amount of packets (velocity/position) transmitted by over 6%.

3.7 Summary

Two robust perception-based haptic data reduction and transmission techniques are presented to reduce data traffic in telehaptic systems. A prediction approach that relies on the least-squares method and median filtering is exploited to reduce the number of packets transmitted, and efficiently reconstruct unsuccessfully received data samples. Knowledge from human haptic perception is also used and incorporated into the general data reduction architecture. The techniques are initially evaluated in a basic experimental setting in order to validate their performance. Their application in a haptic-enabled (3-DoF) telementoring surgery simulation is also demonstrated. Specifically, the experimental results prove the proposed approach's effectiveness as haptic data packets can be reduced by as much as 96% in normal network conditions and up to 93% in the presence of significant communication delay and packet loss while preserving the overall quality of the telehaptic environment.

Chapter 4

Data Reduction for Haptic Communication in 6-DoF Telepresence Systems

4.1 Introduction

Haptic-enabled telepresence systems enable users to operate in remote (real/virtual) environments while incorporating the sense of touch. Multimodal sensory (haptic, audio, video, etc.) and control information is bilaterally transmitted over a communication channel, enabling a human operator to be immersed into the remote environment. Applications range from toxic/explosive material handling and disposal, to minimal invasive surgery, education, and entertainment [4,25,57]. A classic example of a multimodal telepresence system architecture is shown in Fig. 4.1. It consists of three main modules: the human system interface (HSI), the teleoperator (TOP), and a communication network that serves as a link between the HSI and the TOP. Specifically, a human operator (OP), equipped with a head-mounted display and headphones, manipulates the HSI and thereby commands the position/orientation/velocity of the teleoperator. Multimodal sensor data acquired at the teleoperator through interactions with the remote environment (e.g., video, audio, force/torque feedback) are then transmitted back

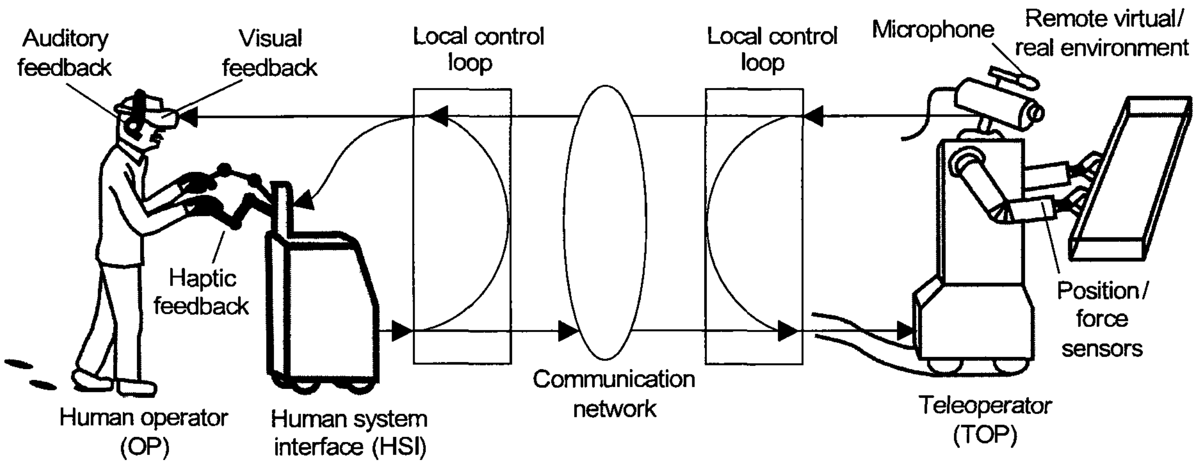


Figure 4.1: Multimodal telepresence system (adapted from [1,4]).

and displayed to the OP. The main focus of this chapter is, however, on haptic data streams; specifically, haptic data reduction in 6-DoF telepresence systems.

As discussed in Section 1.2, haptic data in telepresence systems are sampled at a constant rate in the approximate range of 500-1000 Hz to ensure that the stability of the local control loops at the HSI and the TOP is not compromised. In order to limit packetization and transmission delays which could adversely affect the stability of the system, every set of sampled data is transmitted in individual packets. This translates to a packet rate that is equal to the sampling rate of the local haptic control loops, (i.e., 500-1000 packets per second) which is clearly difficult to sustain over long distance digital communication networks. Consequently, as noted in Chapter 2, over the past few years several techniques have been proposed in the literature for reducing the number of haptic packets transmitted between the HSI and the TOP without affecting the fidelity or immersiveness of the telepresence system. Conversely, all currently available methods and studies in haptic data reduction have been developed and evaluated, strictly in telepresence systems with simpler devices that range from 1- to 3-DoF [4,21,25,27,31,34,35,37,38,57,58]. Teleoperation systems, however, that rely on 3-DoF (or less) haptic devices impose a great restriction on the operator, since manipulation inherently requires six degrees-of-freedom. 6-DoF devices provide the user with a more realistic haptic experience, as full force and torque feedback is available. An increase of the number of DoF, from 3-DoF to 6-DoF evidently renders perceptual haptic data reduction a more challenging problem.

In this chapter, a perception-based haptic data reduction method is derived to reduce the number of transmitted packets in 6-DoF telehaptic systems. In fact, the algorithm is an extension of the (3-DoF) methods presented in Chapter 3. The suggested data reduction algorithm is further generalized while considering haptic-enabled telepresence and teleaction applications, rather than strictly haptic-enabled telementoring. In addition, the algorithm relies on the limitations of human haptic perception (i.e., the Just Noticeable Differences [JND]) with respect to a user's hand movement (position/orientation) and exerted force (force/torque) to reduce the number of packets transmitted without compromising transparency. A haptic prediction model is exploited to further reduce the amount of haptic packets transmitted, and to improve the reconstruction of data samples on the receiver side. Several distance metrics are provided to evaluate the acuity of human haptic perception in detecting haptic distortion when data reduction is performed in 6-DoF settings. An extension of the deadband principle to 6-DoF is also derived. Furthermore, a validation of the 6-DoF haptic data reduction technique is performed under normal network conditions and in the presence of network-induced time delays and packet loss. In the analysis of all conducted experiments, statistical significance tests (using Friedman's nonparametric analysis of variance [ANOVA], and Wilcoxon Signed-Rank Tests) were carried out to determine the appropriate multivariate human haptic perception thresholds (force, torque, orientation, etc.) required to minimize the number of packets transmitted while preserving the immersiveness of the 6-DoF telepresence system.

The rest of the chapter is organized as follows. In Section 4.2, the data reduction algorithm for haptic communication in 6-DoF telepresence systems is illustrated. In Section 4.3, the materials and methods used during the experimental procedure are described. In Section 4.4, the experimental results are provided. In Section 4.5, a discussion of the results is presented. Finally, Section 4.6 presents a summary of the chapter.

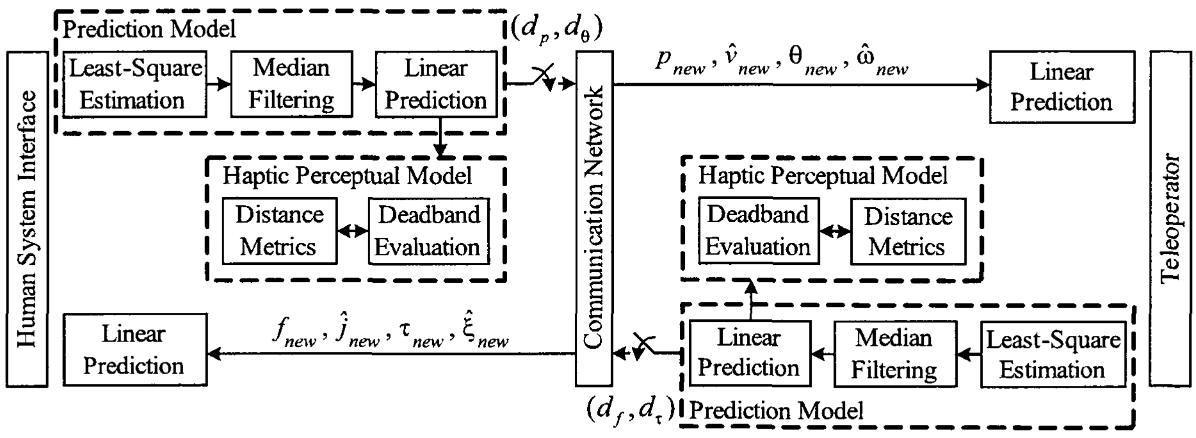


Figure 4.2: Block diagram architecture of the 6-DoF haptic data reduction algorithm.

4.2 Haptic Data Reduction in 6-DoF Telehaptic Systems

In this section, the data reduction algorithm for haptic communication in 6-DoF telehaptic systems is illustrated. The general block diagram of the suggested method is shown in Fig. 4.2. Moreover, the modules that constitute the haptic data reduction architecture are described in detail in succeeding sections.

4.2.1 Haptic Perceptibility

Numerous factors impact human haptic perception when 6-DoF force-feedback devices are considered, including: magnitude of the exerted force/torque, hand movement (i.e., position and orientation) and velocity, stiffness of the environment, etc. Generally, the Just Noticeable Difference of human haptic perception is in the range of 5% to 15% [49–53]. More interestingly, with regard to telehaptic systems, a JND of approximately 10% is reported in [4, 31] for both velocity and force when analyzing the impact of haptic data reduction on human haptic perception in a 3-DoF haptic-enabled telepresence system. This suggests that, in a 3-DoF telepresence system, if the change in magnitude of the haptic force currently applied at the teleoperator side is less than the 10% JND, the operator would not perceive a force-feedback variation. This concept, which is in fact a direct application of Weber's law, is also used in the suggested packet reduction method, where haptic data are only transmitted if the operator is likely to perceive a change in the kinesthetic force-feedback response. However, as the presented algorithm is intended for 6-DoF telehaptic systems, JND thresholds of human haptic perception with respect to force, torque, position, and orientation are derived.

4.2.2 Deadband Principle

This algorithm derives a more generalized formulation of the deadband principle introduced in Chapter 3 which states that signal changes do not need to be transmitted over the communication network unless they exceed a fixed or magnitude-dependent threshold [4,59]. In particular, the reformulation generalizes the deadband principle to multiple multidimensional haptic data types (e.g., force, torque, position, orientation, velocity, etc.) as follows:

$$
\begin{aligned}
&if\Big(D_1(\mathbf{v}_{i_1}, \mathbf{v}_{c_1}) > f(\mathbf{v}_{i_1}) \text{ or } D_2(\mathbf{v}_{i_2}, \mathbf{v}_{c_2}) > f(\mathbf{v}_{i_2}) \text{ or } \ldots\ D_n(\mathbf{v}_{i_n}, \mathbf{v}_{c_n}) > f(\mathbf{v}_{i_n})\Big)\\
&then \text{ transmit } \mathbf{v}_{c_1}, \mathbf{v}_{c_2}, \ldots, \mathbf{v}_{c_n}\\
&else \text{ do not transmit}
\end{aligned}
$$

where $\mathbf{v}_{i_n}$ and $\mathbf{v}_{c_n}$ denote the initial and the currently acquired data vectors of the nth haptic data type. $D_n(\cdot)$ functions consist of different distance metrics used to measure the proximity between data vectors (of type n). $f(\cdot)$ defines the perceptual thresholds for different haptic data types, i.e., the region (referred to as the deadzone) in which transparency is maintained. Therefore, currently acquired data vectors $\mathbf{v}_{c_1}, \mathbf{v}_{c_2}, \ldots, \mathbf{v}_{c_n}$ are not transmitted (from the HSI to the TOP or vice-versa), if it is determined that the distances between $\mathbf{v}_{c_1}$ and $\mathbf{v}_{i_1}$, $\mathbf{v}_{c_2}$ and $\mathbf{v}_{i_2}, \ldots, \mathbf{v}_{c_n}$ and $\mathbf{v}_{i_n}$ are smaller than, or equal to $f(\mathbf{v}_{i_1}), f(\mathbf{v}_{i_2}), \ldots, f(\mathbf{v}_{i_n})$, respectively.

As the proposed method is intended for 6-DoF telepresence applications, data reduction in the forward direction (from the HSI to the TOP) is effectively performed while relying on the limitations of human haptic perception with respect to a user's hand position and orientation. To determine whether it is necessary to transmit newly acquired data samples, perceptual distance metrics are defined such that both, hand position and orientation information is considered. A direct application of the generalized deadband principle to this problem can be depicted as follows:

$if\left(D_{\mathbf{p}}(\mathbf{p}_i,\ \mathbf{p}_c) > f(\mathbf{p}_i) \text{ or } D_\theta(\theta_i,\ \theta_c) > f(\theta_i)\right)$

$then$ transmit $\mathbf{p}_c,\ \theta_c,\ (\hat{v}_{new},\ \hat{w}_{new})$

$else$ do not transmit

where $\mathbf{p}_i$ and $\mathbf{p}_c$ denote the initial and the currently acquired position vectors, θ_i and θ_c correspond to the initial and the currently acquired orientation vectors, whereas $\hat{v}_{new}$ and $\hat{w}_{new}$ denote estimates of the velocity (computed from position data) and the rate-of-change of orientation data respectively. The estimates $\hat{v}_{new}$ and $\hat{w}_{new}$ are exploited in the prediction modules as it will be later discussed in subsequent sections. $D_{\mathbf{p}}(\cdot)$ and $D_\theta(\cdot)$ consist of distance metrics used to measure the proximity between position and orientation vectors respectively, whereas the functions $f(\mathbf{p}_i)$ and $f(\theta_i)$ define the perceptual thresholds for position and orientation data. Therefore, in order to determine whether to transmit the currently acquired data vectors $\mathbf{p}_c$ and θ_c, the distance between $\mathbf{p}_c$ and the initial position vector $\mathbf{p}_i$, or the distance in orientation between θ_c and θ_i must be larger than $f(\mathbf{p}_i)$ or $f(\theta_i)$, respectively. Furthermore, the method relies on a fixed position deadband $d_{\mathbf{p}}$ and a fixed orientation deadband d_θ with deadzones of radius $f(\mathbf{m}) = d_{\mathbf{m}}$, where $\mathbf{m} \in \{\mathbf{p}, \theta\}$.

Conversely, the characteristics of human haptic force and torque perception are exploited in the backward direction (from the TOP to the HSI). To determine whether it is necessary to transmit newly acquired data samples from the TOP back to the HSI, a perceptual distance metric is defined such that both, force and torque information is considered. The application of the generalized deadband principle to this problem can be described as follows:

$if\left(D_{\mathbf{f}}(\mathbf{f}_i, \mathbf{f}_c) > f(\mathbf{f}_i) \text{ or } D_{\boldsymbol{\tau}}(\boldsymbol{\tau}_i, \boldsymbol{\tau}_c) > f(\boldsymbol{\tau}_i)\right)$

$then$ transmit $\mathbf{f}_c, \boldsymbol{\tau}_c, (\hat{j}_{new}, \hat{\xi}_{new})$

$else$ do not transmit

where $\mathbf{f}_i$ and $\mathbf{f}_c$ denote the initial and the currently acquired force vectors, whereas $\boldsymbol{\tau}_i$ and $\boldsymbol{\tau}_c$ correspond to the initial and the currently acquired torque vectors. $\hat{j}_{new}$ and $\hat{\xi}_{new}$ express the rate-of-change of force and torque data respectively. $D_{\mathbf{f}}(\cdot)$ and $D_{\boldsymbol{\tau}}(\cdot)$ consist of distance metrics used to measure the proximity between force and torque vectors respectively, whereas the functions $f(\mathbf{f}_i)$ and $f(\boldsymbol{\tau}_i)$ define the human perceptual thresholds for different force and torque values. More specifically, the method relies on a force deadband $d_{\mathbf{f}}$ and a torque deadband $d_{\boldsymbol{\tau}}$ with deadzones of radius $f(\mathbf{h}_i) = d_{\mathbf{h}} \cdot |\mathbf{h}_i|$, where $\mathbf{h} \in \{\mathbf{f}, \boldsymbol{\tau}\}$.

4.2.3 Distance Metrics for Different Haptic Data Types

In order to measure the proximity of two haptic data vectors $\mathbf{v}_i$ and $\mathbf{v}_c$, different distance metrics are considered. For example, to measure the closeness of two position vectors, the Euclidean distance is exploited as follows:

$$\begin{aligned} D_{\mathbf{p}}(\mathbf{p}_i, \mathbf{p}_c) &= |\mathbf{p}_i - \mathbf{p}_c| \\ &= \sqrt{(\mathbf{p}_{i,x} - \mathbf{p}_{c,x})^2 + (\mathbf{p}_{i,y} - \mathbf{p}_{c,y})^2 + (\mathbf{p}_{i,z} - \mathbf{p}_{c,z})^2}. \end{aligned} \tag{4.1}$$

The Euclidean distance is also used when measuring the proximity of force (between $\mathbf{f}_i$ and $\mathbf{f}_c$) and torque (between $\boldsymbol{\tau}_i$ and $\boldsymbol{\tau}_c$) vectors. Many important reasons motivated the use of the Euclidean distance metric, including its effectiveness, simplicity, computational efficiency, its ability to give us a direct physical meaning of the distance between two points, computed distances are independent of rotations of the coordinate system, and its robustness in the presence of noise [60].

Conversely, measuring the proximity between two orientation vectors is not as easily achieved as with the previously mentioned haptic data types. Specifically, the convention used when representing and parameterizing rotations in three dimensions must be considered when defining a distance metric to express the similarity (or "closeness") between two different orientations. The following two common representations will be discussed: Euler angles and unit quaternions. First, according to Euler's rotation theorem, any orientation can be illustrated by a sequence of three rotations (α, β, γ) about a set of three axes (x_1, x_2, x_3). Using the common roll-pitch-yaw Euler angles convention, the angular distance between two orientations $\boldsymbol{\theta}_i = (\alpha_i, \beta_i, \gamma_i)$ and $\boldsymbol{\theta}_c = (\alpha_c, \beta_c, \gamma_c)$ can be computed as follows:

$$D_\theta(\boldsymbol{\theta}_i, \boldsymbol{\theta}_c) = \sqrt{S(\alpha_{i,x}, \alpha_{c,x})^2 + S(\beta_{i,y}, \beta_{c,y})^2 + S(\gamma_{i,z}, \gamma_{c,z})^2}. \tag{4.2}$$

The $S(\cdot)$ function is defined to ensure that the "shortest path" difference between two angles is considered as follows:

$$S(\phi_1, \phi_2) = \begin{cases} \Delta\phi + 2\pi & \text{if } \Delta\phi < -\pi \\ \Delta\phi - 2\pi & \text{if } \Delta\phi > \pi \\ \Delta\phi & \text{otherwise} \end{cases} \tag{4.3}$$

where $\Delta\phi = \phi_2 - \phi_1$ and $\Delta\phi \in [-\pi, +\pi]$. An important advantage of Euler angles is their intuitive physical interpretation. However, a drawback of this representation, is that its parametrization is not unique; i.e., there are multiple sets of parameter values which can result in the same rotation [61]. An alternative metric to measure rotation distances can be defined using the unit quaternion representation; recognized for its computational simplicity, e.g., the inverse operation is easily achieved. Quaternions are better understood when considering their relationship to axis-angle pairs. In particular, it can be demonstrated that any arbitrary orientation in three dimensions could be achieved by a single rotation θ about a single axis $w = (w_x, w_y, w_z)$. A unit quaternion is defined as:

$$Q = (q_0, q_1, q_2, q_3) = (\cos(\theta/2),\ \mathrm{w}\sin(\theta/2)), \tag{4.4}$$

where q_0 is the scalar component of Q, $\vec{q} = \{q_1, q_2, q_3\}$ is the vector component. Accordingly, given two unit quaternions $Q_i = \{q_{0,i}, q_{1,i}, q_{2,i}, q_{3,i}\}$, and $Q_c = \{q_{0,c}, q_{1,c}, q_{2,c}, q_{3,c}\}$, (where in this case, $\theta_i \Leftrightarrow Q_i$, and $\theta_c \Leftrightarrow Q_c$) the difference in rotation can be computed as follows:

$$D_\theta(Q_i, Q_c) = 2 * \arccos(|Q_i \cdot Q_c|), \tag{4.5}$$

where the inner product between the two quaternions is defined as follows:

$$Q_i \cdot Q_c = q_{0,i}q_{0,c} + q_{1,i}q_{1,c} + q_{2,i}q_{2,c} + q_{3,i}q_{3,c} \tag{4.6}$$

Another approach to measuring the difference in orientation using unit quaternions is to simply compute the absolute inner product between the two vectors. In this case, computed distances are in the range [0,1], where small rotation differences result in distances that are closer to 1.

4.2.4 Haptic Prediction Model

The haptic prediction model derived in Section 3, is further extended and applied in the suggested 6-DoF data reduction algorithm to reduce the amount of haptic packets transmitted in telepresence systems, and to improve the reconstruction of missing data samples. The model encompasses three primary modules: a least-square estimator, a median filter and a linear predictor. It is used to predict multidimensional position, orientation, force, and torque information. Furthermore, at this stage of the data reduction algorithm, orientation is represented using Euler angles as it is more intuitive and convenient to predict angles as opposed to unit vectors (quaternions)[1]. This being said, haptic data prediction is computed as follows[2]:

$$\begin{aligned}\hat{g}(n+1) &= \dot{g}(n)\Delta T + g(n) \\ &= \hat{s}(n)\Delta T + g(n),\end{aligned} \tag{4.7}$$

where $g(n)$, $g \in \{\mathbf{p},\, \theta,\, \mathbf{f},\, \tau\}$ denotes the current haptic data value, $\hat{s}(n)$ is the current estimated first derivative of a haptic data type, i.e., $\hat{s} \in \{\hat{v},\, \hat{w},\, \hat{j},\, \hat{\xi}\}$, $\hat{g}(n+1)$ corresponds to the predicted data estimate, whereas ΔT is the sampling period. In turn, the transmission method exploits the suggested linear prediction technique as follows:

$$g_i = \begin{cases} g_{new} & \text{if packet sent or received} \\ g_{i-1} + \hat{s}_{new} \cdot \Delta T & \text{otherwise} \end{cases}$$

where g_{new} and $\hat{s}_{new}$ are the newly transmitted or received haptic data samples and their corresponding estimated first derivatives (rate-of-change), respectively. g_i and g_{i-1} are the most recent outputs of the prediction model. More specifically, when the predictor is used at the transmitting end (see Fig. 4.2), predicted values must fall within predefined deadzones (defined by the deadband $d_{\mathbf{m}}$, where $\mathbf{m} \in \{\mathbf{p}, \theta\}$ and by the deadband $d_{\mathbf{h}}$, where $\mathbf{h} \in \{\mathbf{f}, \tau\}$ when transmission is performed from the HSI to the TOP, and from the TOP back to the HSI, respectively), otherwise a new sample is transmitted over the network. Conversely, when the predictor is exploited on the receiving end of the data reduction technique, it attempts to reconstruct the missing haptic data samples based on previously received packets. Hence, the prediction model compensates for missing haptic data packets or communication delays to regularly provide the HSI and the TOP with data samples at a consistent rate of approximately 1 kHz or higher. This is necessary in order to guarantee the stability of the global control loop (closed over the communication system), and in turn maintain compelling and stable force feedback at the HSI side.

[1] The extension of the proposed data reduction method to enable orientation data prediction using a unit quaternion representation is left for future work.

[2] The prediction model is illustrated for one-dimensional signals; its extension to multidimensional data is, however, straightforward.

As previously discussed in Section 3.3.2 it can be observed from Eq. (4.7), that the prediction performance can be positively or negatively affected depending on the precision of the acquired haptic data samples. Generally, the noise present in position and orientation readings is low due to the fact that haptic devices are typically equipped with joint sensors. This is also true for force and torque signals as they are computed directly using position/orientation data. However, the first derivative of the signals (whether movement velocity, Euler angles' rate-of-change, or force/torque's rate-of-change) is generally computed using the finite-difference method which leads to very noisy estimates at high sampling frequencies [54]. This being said, in order to obtain more precise estimates, linear filtering is performed as follows:

$$\hat{s}(n) = \sum_{i=0}^{M-1} w(i) g(n-i), \tag{4.8}$$

where $w(i)$ corresponds to the filter coefficients and M denotes the filter length. Essentially, an estimate of the first derivative $\hat{s}(n)$ (i.e., the slope) that would best fit the observation data $g(n)$ in the least-squares sense is derived. Moreover, the corresponding filter coefficients can be computed using the standard least-squares method as follows:

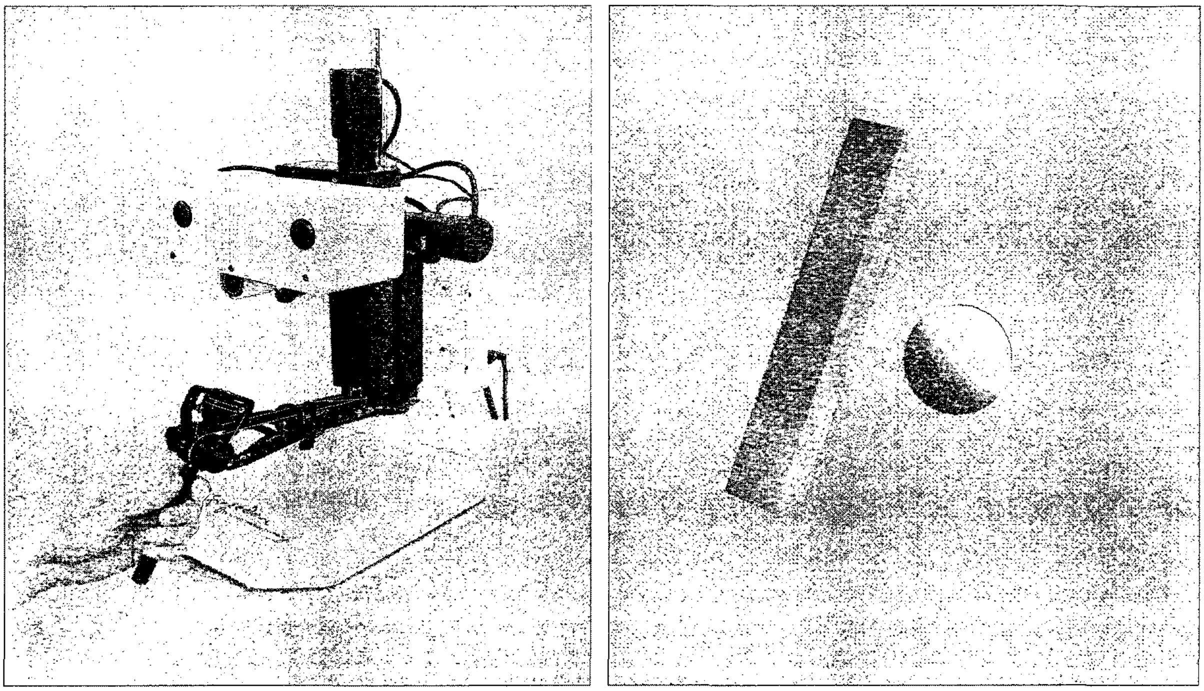
$$w(i) = \frac{12\left(\frac{M-1}{2} - i\right)}{T_s M (M^2 - 1)}, \quad 0 \leq i \leq M-1 \tag{4.9}$$

where T_s is the sampling period (equivalent to ΔT discussed previously). Furthermore, median filtering is also used to eliminate any fine irregularities and outliers remaining in the computed noisy measures, while preserving the true signal discontinuities [46].

4.3 Materials and Methods

4.3.1 Experimental Setup

The experimental setup, as shown in Fig. 4.3, consists of a 6-DoF haptic device (HSI side) used to interact with a remote virtual environment (TOP side) over a 100 Mbps Ethernet local area network. The experiments were conducted using an MPB high fidelity Freedom 6S haptic



(a) (b)

Figure 4.3: The experimental setup which consists of (a) an MPB Freedom 6S haptic device (HSI) used by the OP, to interact with a (b) remote virtual environment (TOP) that is also locally displayed on the HSI side.

device, which provides haptic force feedback in six degrees of freedom. The suggested haptic data reduction and transmission technique is executed on both the HSI and the TOP sides due to the bilateral nature of telepresence and teleaction systems. Packets between the HSI and the TOP sides of the system are transmitted at a rate of 1 kHz to ensure stable and realistic haptic interaction.

Setup of the HSI Side

The experimental setup of the HSI side consists of the Freedom 6S haptic device (see Fig. 4.3 (a)) interfaced with a PC running MS Windows XP. The PC is equipped with an Intel(R) Pentium 4, 3.40 GHz processor and 1 GB of RAM. The haptic device is a six-axis force feedback hand controller that can generate both translational force and rotational torque (twist force).

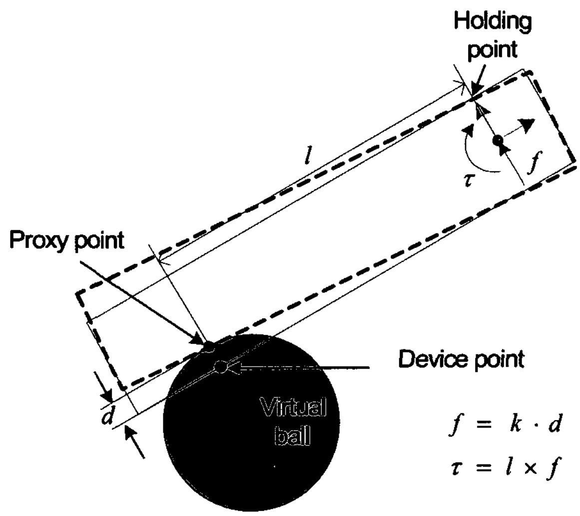


Figure 4.4: Force/torque feedback in 6-DoF haptic rendering (performed using the direct rendering method [5,6]).

Its high spatial resolution, low friction, and accurate force reproduction make it suitable for interaction with virtual worlds and/or as part of a teleoperator system. In addition, a regular 19 inch flat CRT monitor is used for the graphic display of the remote (virtual) environment.

Setup of the TOP Side

The teleoperator side is a simple virtual environment consisting of a stick-on-ball application displayed on a 17 inch LCD monitor, and running on a Windows XP-based PC equipped with an Intel(R) Xeon 1.7 GHz processor and 2 GB of RAM. A simple remote environment is exploited to eliminate the possible introduction of other factors that may considerably influence the analysis of the measured data. Furthermore, the force/torque computation necessary to enable 6-DoF haptic rendering is performed using the direct rendering method [5]. This technique is illustrated in Fig. 4.4, where a virtual 3D stick mimics the motion of the device stylus (device tool) in free space, but is constrained to the surface of the virtual 3D ball when a collision between the two objects is detected. Moreover, when the 3D stick is in contact with the virtual 3D ball, the deepest penetration point (device point) and its corresponding contact point, located at the surface of the ball (proxy point), are first acquired. A feedback force $f \in \mathbb{R}^3$ is then calculated using the Hooke's (linear) law:

$$f = k \cdot d \tag{4.10}$$

where $d \in \mathbb{R}^3$ is the penetration vector between the two points and $k \in \mathbb{R}$ is the stiffness constant. The force f is then used to generate a torque $\tau \in \mathbb{R}^3$ around a predefined holding point on the virtual tool as follows:

$$\tau = l \times f \tag{4.11}$$

where l is the distance from the user holding point to the collision point between the 3D stick and the 3D ball. The resulting contact force/torque pair $F = [f \;\; \tau]^T \in \mathbb{R}^6$ is then directly displayed (haptically) back to the user.

4.3.2 Subjects

Altogether, ten subjects (six males and four females, aged between 24 and 38), of which eight are right handed and two are left handed with no known sensorimotor impairments with their hands, took part in the experiments. Their prior experience with force-feedback haptic devices ranged from novice to expert. All participants were remunerated for their time.

4.3.3 Procedure

Numerous experiments were conducted to evaluate the data reduction performance of the proposed algorithm. Additionally, experiments were carried out to reveal the appropriate multivariate haptic perception threshold (deadband) values required to minimize the number of packets transmitted between the HSI and the TOP while preserving the immersiveness of the 6-DoF haptic-enabled telepresence system. The following experimental scenarios were considered: Assessment of the perceptual thresholds (deadband values) and the corresponding haptic data reduction rates when:

- position and orientation information is transmitted from the HSI to the TOP;
- force and torque information is transmitted from the TOP to the HSI.

Furthermore, for each scenario the following network conditions were examined:

- Normal network conditions;
- network-induced packet loss;
- network-induced time delays and packet loss.

Practice runs were initially conducted to familiarize the subjects with the 6-DoF haptic-enabled telepresence system and to give them a feel of the distortion encountered when the suggested haptic packet reduction technique is enabled. Specifically, throughout the familiarization period and across different practice runs, data reduction distortions introduced in the telepresence system were randomly varied from *no distortion* to *largely noticeable distortion.* Following the familiarization phase, six different test experiments were carried out, each consisting of eighty-one 25 seconds runs, where participants (as in the practice runs) were requested to explore the remotely located virtual environment using the provided 6-DoF haptic telepresence system and assess the quality of the overall haptic experience. The initial three experiments correspond to the first experimental scenario mentioned above where the deadband values for position and orientation data (which are sent from the HSI to the TOP) are assessed in three different network conditions: normal, network-induced packet loss, and network-induced time delays and packet loss. The other three experiments are associated with the second experimental scenario where deadband values for force and torque data (sent from the TOP to the HSI) are also assessed in the three different network conditions.

Experiments where data reduction is applied on packets transmitted from the HSI to the TOP (in the forward direction of the system), consisted of a number of runs that differed in the position and orientation deadband values. Nine different position deadband (d_p) values were used and are as follows: 0mm, 0.025mm, 0.05mm, 0.1mm, 0.2mm, 0.3mm, 0.5mm, 0.7mm, and 0.9mm. Similarly, nine different orientation deadband values were exploited and they consisted of: 0.0rad, 0.0005rad, 0.001rad, 0.002rad, 0.003rad, 0.004rad, 0.006rad, 0.008rad, and 0.01rad. Therefore, in total $9 \times 9 = 81$ different runs per experiment (i.e., per network condition) were performed using different randomly selected combinations of position/orientation deadband values. After every run the subjects were requested to rate the haptic-enabled telepresence experience using the following five-grade impairment scale:

5 – Imperceptible;
4 – Perceptible, but not annoying;
3 – Slightly annoying;
2 – Annoying;
1 – Very annoying.

Users were also permitted to give in-between values when necessary (i.e., 4.5, 3.5, 2.5, and 1.5) to better assess their opinion. Similarly, experiments where data reduction is conducted on packets transmitted from the TOP back to the HSI (backward direction), consisted of a number of runs that differed in the force and torque deadband values. Nine different force deadband (d_f) values were used: 0%, 0.5%, 1%, 1.5%, 3%, 5%, 7%, 10%, and 15%. In addition, nine different torque deadband values were exploited: 0%, 0.5%, 1%, 1.5%, 3%, 5%, 7%, 9%, and 12%. Therefore, in total $9 \times 9 = 81$ distinct runs per experiment were performed using different randomly selected combinations of force/torque deadband values.

4.4 Experimental Verifications

In this section, a validation of the 6-DoF haptic data reduction technique is performed under normal network conditions and in the presence of packet loss and a network-induced time delay. For all experiments, statistical tests were conducted to determine the appropriate haptic perceptual threshold (deadband) values (d_p, d_θ, d_f, and d_τ) required to minimize

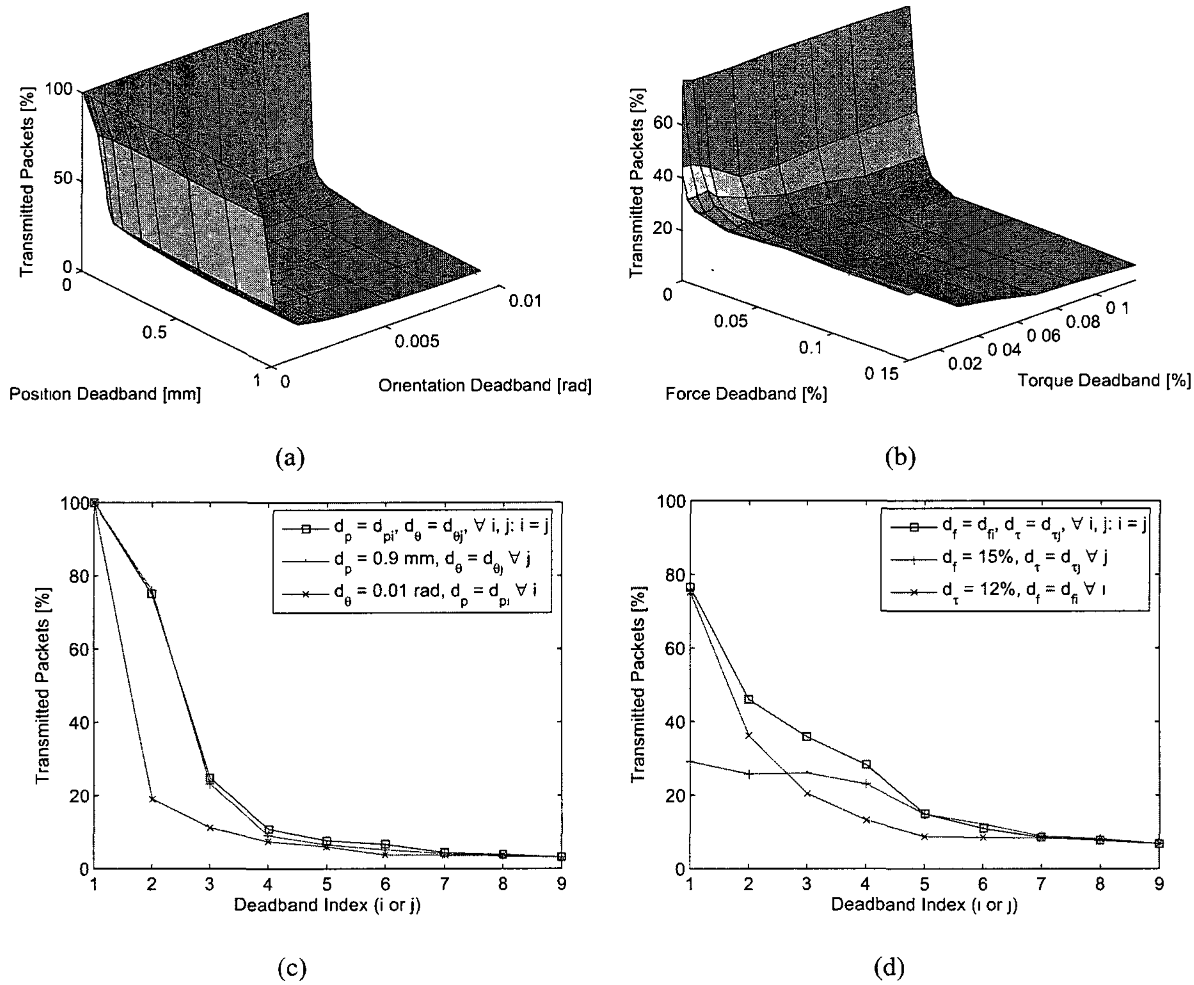
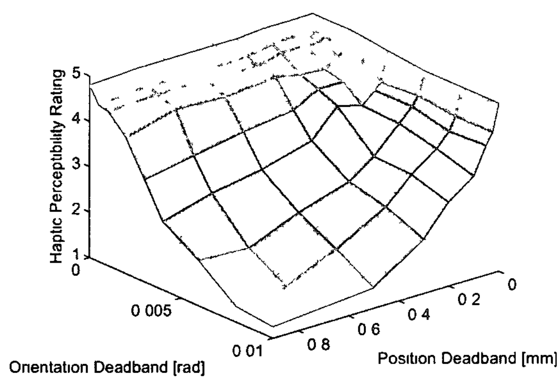
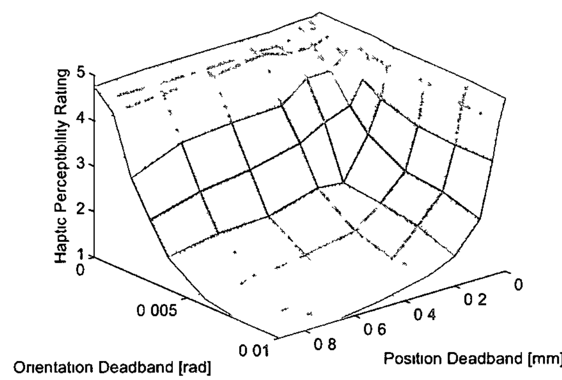


Figure 4.5: Average number of transmitted packets with respect to (a) position and orientation deadband values and, (b) force and torque deadband values. 2D projections of the plots are illustrated in (c) and (d) respectively.

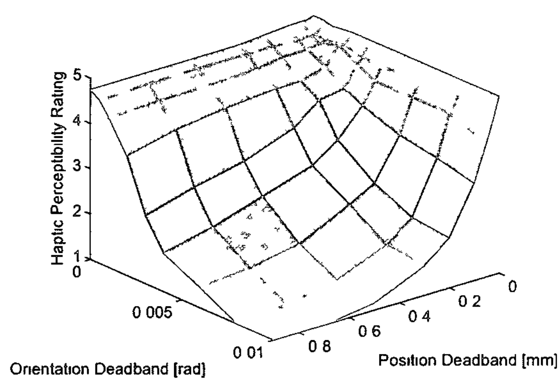
the number of packets transmitted between the HSI and the TOP while preserving the immersiveness of the 6-DoF telepresence system. Specifically, Friedman (nonparametric ANOVA) Tests were conducted to detect whether a statistically significant change in perceptibility ratings is overall apparent for different position/orientation, force/torque deadband values. It should be emphasized that a nonparametric approach is exploited as the data are of ordinal nature (and the assumptions of normality were not met). Furthermore, a post hoc analysis with Wilcoxon Signed-Rank Tests was performed when necessary to determine significant differ-



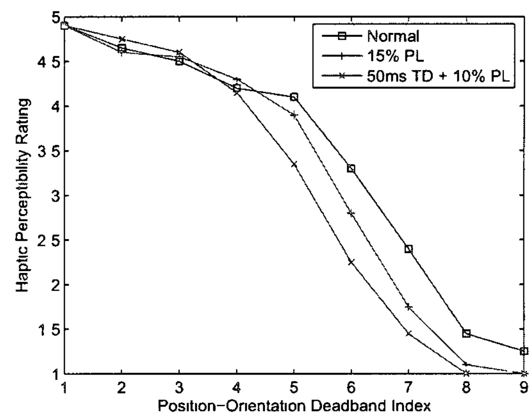
(a) Normal



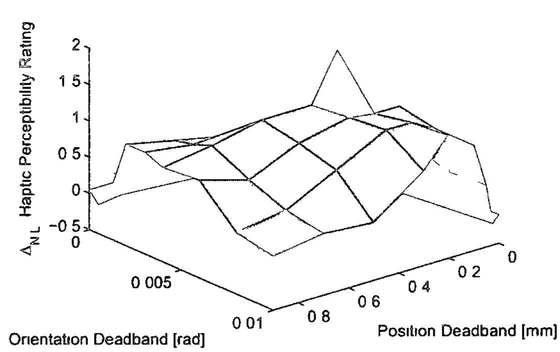
(b) 15% PL



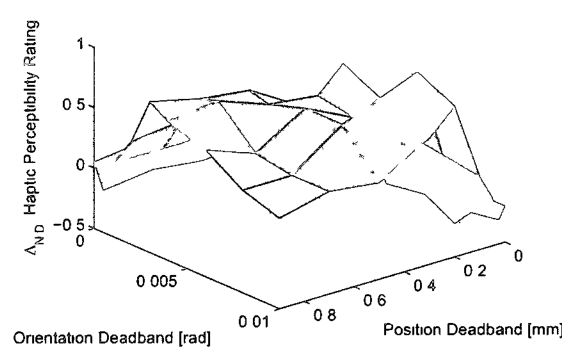
(c) 50ms TD + 10% PL



(d)

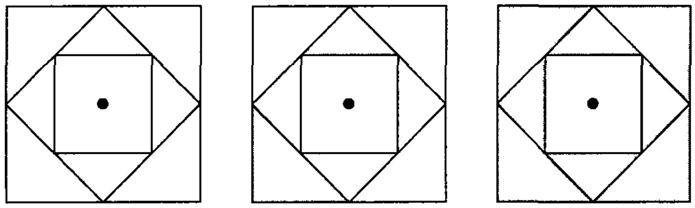


(e)



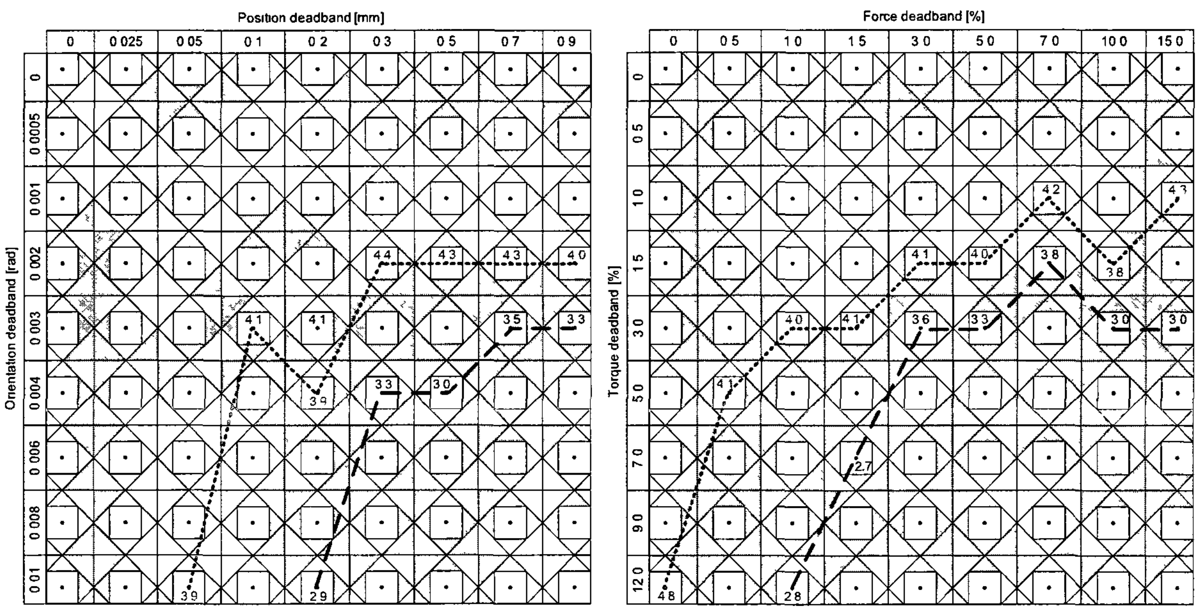
(f)

Figure 4.6: Average perceptibility rating scores in the forward direction (from the HSI to the TOP).



(a) (b) (c)

Figure 4.7: Square tile templates used to visually represent statistically significant results.



(a) (b)

Figure 4.8: A visual representation of post hoc pairwise comparisons of perceptibility scores (using Wilcoxon Signed-Rank Tests method) in normal network conditions, is shown across different (b) position/orientation (forward direction) and (c) force/torque (backward direction) deadband values.

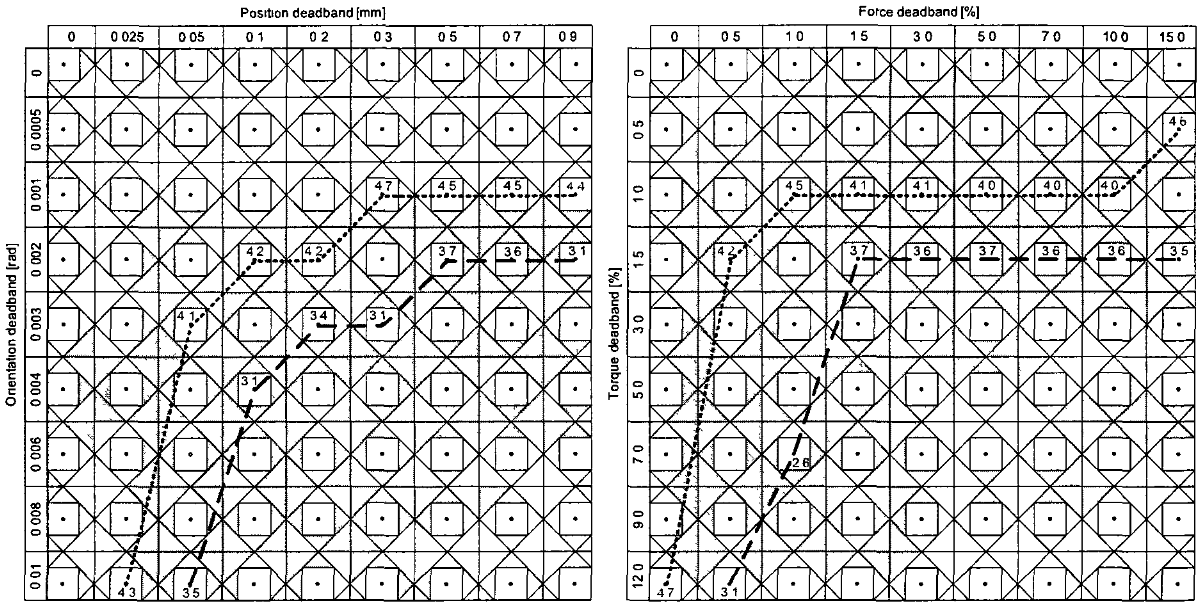
ences in perceptibility ratings between two deadband-value pairs, e.g., $d_{f_1,\tau_1} = [d_{f_1}, d_{\tau_1}]$ and $d_{f_2,\tau_2} = [d_{f_2}, d_{\tau_2}]$. In addition, for all experiments, the significance level adopted was $\alpha = 0.05$.

4.4.1 Experimental Scenario I: Data Reduction of 6-DoF Haptic Movement Information

The suggested 6-DoF haptic data reduction algorithm is initially applied in the forward direction of the system, i.e., strictly on packets transmitted from the HSI to the TOP. The average number of transmitted packets with respect to the applied position and orientation deadband values are plotted in Fig. 4.5 (a). For a better illustration of the results, 2D projections of the 3D plot (Fig. 4.5 (a)) are provided in Fig. 4.5 (c). One projection corresponds to the packet reduction rate along the diagonal line of the 3D plot which extends from $d_{p_1,\theta_1} = [d_{p_1} = 0.0mm,\ d_{\theta_1} = 0.0rad]$ to $d_{p_9,\theta_9} = [d_{p_9} = 0.9mm,\ d_{\theta_9} = 0.01rad]$, i.e., $d_p = d_{p_i}$, $d_\theta = d_{\theta_j}, \forall i,j : i = j$ and $i,j \geq 1,\ i,j \leq 9$, where i and j correspond to the deadband indices for position and orientation data respectively (e.g $d_{p_1} = 0.0mm$, $d_{p_2} = 0.025mm$, ..., $d_{p_9} = 0.9mm$). The second and third 2D projections are subject to the following constraints $d_p = d_{p_9} = 0.9mm,\ d_\theta = d_{\theta_j}\ \forall j$, and $d_\theta = d_{\theta_9} = 0.01rad,\ d_p = d_{p_i}\ \forall i$, respectively. In addition, the corresponding perceptibility ratings for all ten subjects, when packet reduction is performed from the HSI to the TOP in normal network conditions, are averaged as shown in Fig. 4.6 (a). A 2D projection of the perceptibility ratings along the diagonal line of the plot, i.e., $d_p = d_{p_i}$, $d_\theta = d_{\theta_j}, \forall i,j : i = j$ and $i,j \geq 1,\ i,j \leq 9$, is presented in Fig. 4.6 (d) (line with square markers). The results of the Friedman Test indicated that there was a statistically significant difference in perceptibility rating scores across the eighty-one different position/orientation deadband values, $\chi^2(80, n = 10) = 634.48,\ p < 0.0005$.

Wilcoxon Signed-Rank Tests are then conducted to determine significant differences in perceptibility rating scores between (position/orientation) deadband-value pairs. A visual representation of the significance test results are presented in Fig. 4.8 (a). This is achieved using a square tile representation divided into eight different triangles (which point in the direction of all neighboring tiles) as demonstrated in Fig. 4.7. For example, a grayed-out triangle at deadband-value pair d_{p_1,θ_1} that is pointing in the direction of d_{p_2,θ_2}, suggests a statistically significant difference in perceptibility rating scores between d_{p_1,θ_1} and d_{p_2,θ_2}. Accordingly, a deadband-value pair (e.g., $d_{p,\theta} = [d_p = 0.0mm,\ d_\theta = 0.0rad]$) with a tile representation such as that of Fig. 4.7 (a) has perceptibility rating scores that are not statistically significantly different than any of its neighboring deadband-value pairs. In contrast, a deadband-value pair with a tile representation such as that of Fig. 4.7 (b) has perceptibility rating scores that are statistically significantly different when compared to its neighboring deadband-value pairs located at the top and the bottom-left corner. Similarly, the deadband-value pair with a tile representation such as that of Fig. 4.7 (c) has perceptibility rating scores that are statistically significantly different than the scores associated with all neighboring deadband-value pairs. Consequently, in Fig. 4.8 (a), position/orientation perceptual threshold dotted lines are drawn to identify the regions in which average perceptibility rating scores are equal to, or higher than 4/5 (red, short-dashed line) and 3/5 (blue, long-dashed line). This is achieved while taking into consideration statistically significant differences in perceptibility rating scores between distinct (position/orientation) deadband-value pairs. For example, the average perceptibility rating scores at $d_{p_5,\theta_5} = [d_{p_5} = 0.2mm,\ d_{\theta_5} = 0.003rad]$, and $d_{p_5,\theta_6} = [d_{p_5} = 0.2mm,\ d_{\theta_6} = 0.004rad]$ are 4.1 and 3.9 respectively. However, the 4-and-above perceptual threshold line passes through $d_{p_5,\theta_6} = [d_{p_5} = 0.2mm,\ d_{\theta_6} = 0.004rad]$, despite it having a perceptibility score lower than 4, since a Wilcoxon Signed-Rank Test demonstrated that a statistically significant difference does not exist between the two deadband-value pairs, ($z = -1.03$, $p = 0.31$, with a small effect size $r = 0.23$). It should be clarified that for each deadband-value pair (i.e., tile) statistically significant differences with only top and bottom neighbors were considered when drawing the perceptual threshold lines, the remaining six neighbors are provided for illustration purposes.

It can be observed from Figs. 4.5 (a), 4.6, (a) and 4.8 (a) that the results are quite promising. For example, Figs. 4.5 (a) and 4.8 (a) reveal that with a small deadband-value pair

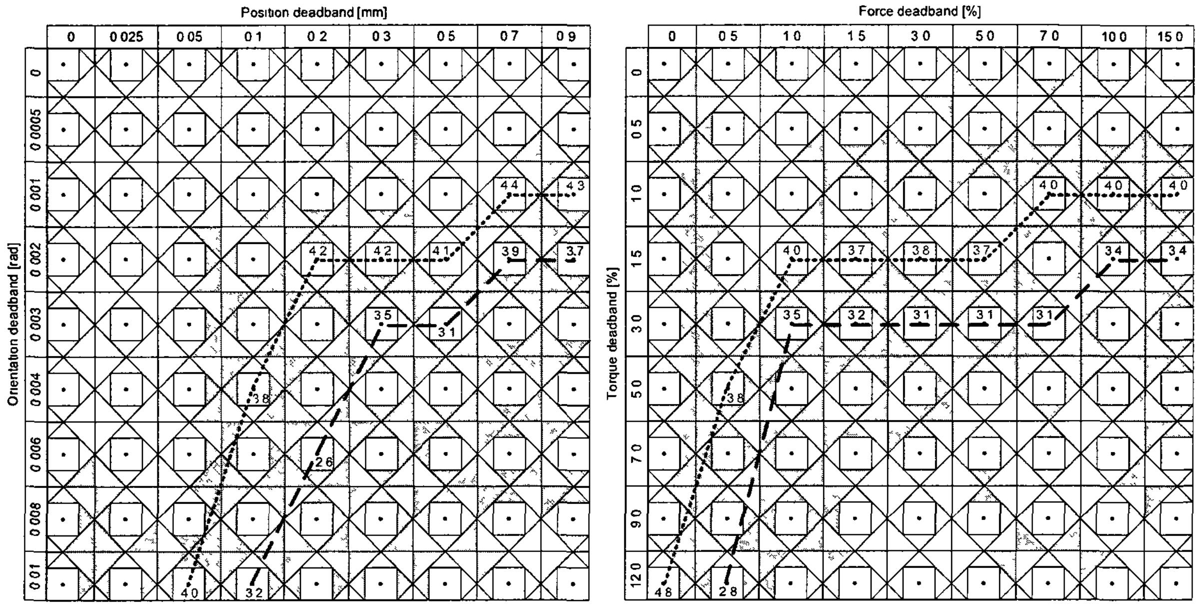


(a) (b)

Figure 4.9: A visual representation of post hoc pairwise comparisons of perceptibility scores in adverse network condition with 15% packet loss, is shown across different (b) position/orientation (forward direction) and (c) force/torque (backward direction) deadband values.

$d_{p_5,\theta_5} = [d_{p_5} = 0.2mm,\ d_{\theta_5} = 0.003rad]$ a desirable average perceptibility rating of 4.1 (which participants described as fairly distortion-free), and a corresponding packet reduction rate of 92.3% is achieved, i.e., only 7.7% of the packets are transmitted from the HSI to the TOP.

In addition, the average perceptibility rating scores when the suggested 6-DoF haptic data reduction algorithm is evaluated in adverse network conditions with 15% packet loss and in conditions with a network-induced delay of 50ms (round-trip) + 10% packet loss, are presented in Fig. 4.6 (b) and (c) respectively. 2D projections of the average scores along the diagonal line of the plots, i.e., $d_p = d_{p_i}$, $d_\theta = d_{\theta_j}, \forall i,j : i = j$ and $i,j \geq 1,\ i,j \leq 9$, are presented in Fig. 4.6 (d). Additionally, to provide a clearer visual depiction of the results presented, the difference between the average perceptibility scores in normal network conditions (Fig. 4.6 (a))



(a) (b)

Figure 4.10: A visual representation of post hoc pairwise comparisons of perceptibility scores in a network-induced delay of 50ms and 10% packet loss, is shown across different (b) position/orientation (forward direction) and (c) force/torque (backward direction) deadband values.

and average scores in conditions with 15% packet loss (Fig. 4.6 (b)) are presented in Fig. 4.6 (e). Furthermore, the difference between the average scores in normal network conditions (Fig. 4.6 (a)) and those obtained in conditions with a 50ms delay + 10% packet loss (Fig. 4.6 (c)) are presented in Fig. 4.6 (f). The difference computation is performed as follows:

$$\Delta_{N,A}(i,j) = score_N(i,j) - score_A(i,j) \tag{4.12}$$

where $score_N(i,j)$ refers to the average perceptibility rating score in normal network conditions at position/orientation deadband indices i, j, whereas $score_A(i,j)$ denotes the average score when network artifacts ($A \in \{15\%$ packet loss, 50ms delay + 10% packet loss$\}$) are considered. The results of the Friedman Test indicated that there was a statistically significant difference in perceptibility rating scores across the eighty-one different position/orientation deadband values when network conditions with 15% packet loss are considered, $\chi^2(80, n =$

$10) = 726.22,\ p < 0.0005$. A similar result was obtained with network-induced time delays and packet loss, $\chi^2(80, n = 10) = 690.67,\ p < 0.0005$. Post hoc pairwise comparisons of perceptibility scores across different position/orientation deadband-value pairs in adverse network conditions were then conducted using the Wilcoxon Signed-Rank Test method. A visual representation of the significance test results with strictly 15% packet loss, and in a network-induced delay of 50ms combined with 10% packet loss are presented in Figs. 4.9 (a) and 4.10 (a), respectively.

4.4.2 Experimental Scenario II: Data Reduction of 6-DoF Haptic Force-Feedback Information

The average number of transmitted packets in the backward direction, i.e., from the TOP to the HSI, with respect to the applied force and torque deadband values are plotted in Fig. 4.5 (b). Moreover, 2D projections of the 3D plot are provided in Fig. 4.5 (d). Similar to the results presented in Section 4.4.1, one projection corresponds to packet reduction rate along the diagonal line of the plot which extends from $d_{f_1,\tau_1} = [d_{f_1} = 0.0\%,\ d_{\tau_1} = 0.0\%]$ to $d_{f_9,\tau_9} = [d_{f_9} = 15.0\%,\ d_{\tau_9} = 12.0\%]$, i.e., $d_f = d_{f_i},\ d_\tau = d_{\tau_j}, \forall i,j : i = j$ and $i,j \geq 1,\ i,j \leq 9$, where i and j correspond to the deadband indices for force and torque data respectively. The second and third 2D projections are subject to the following constraints $d_f = d_{f_9} = 15\%,\ d_\tau = d_{\tau_j}\ \forall j$, and $d_\tau = d_{\tau_9} = 12\%,\ d_f = d_{f_i}\ \forall i$ respectively. In addition, the corresponding average perceptibility rating scores, when all ten participants are considered in normal network conditions, are shown in Fig. 4.11 (a). A 2D projection of the perceptibility ratings along the diagonal line of the plot, i.e., $d_f = d_{f_i},\ d_\tau = d_{\tau_j}, \forall i,j : i = j$ and $i,j \geq 1,\ i,j \leq 9$, is presented in Fig. 4.11 (d) (line with square markers). The results of the Friedman Test indicated that there was a statistically significant difference in perceptibility rating scores across the eighty-one different force/torque deadband values, $\chi^2(80, n = 10) = 692.40,\ p < 0.0005$. Similar to Section 4.4.1, post hoc pairwise comparisons of perceptibility scores across different combinations of force/torque deadband values were then conducted using the Wilcoxon Signed-Rank

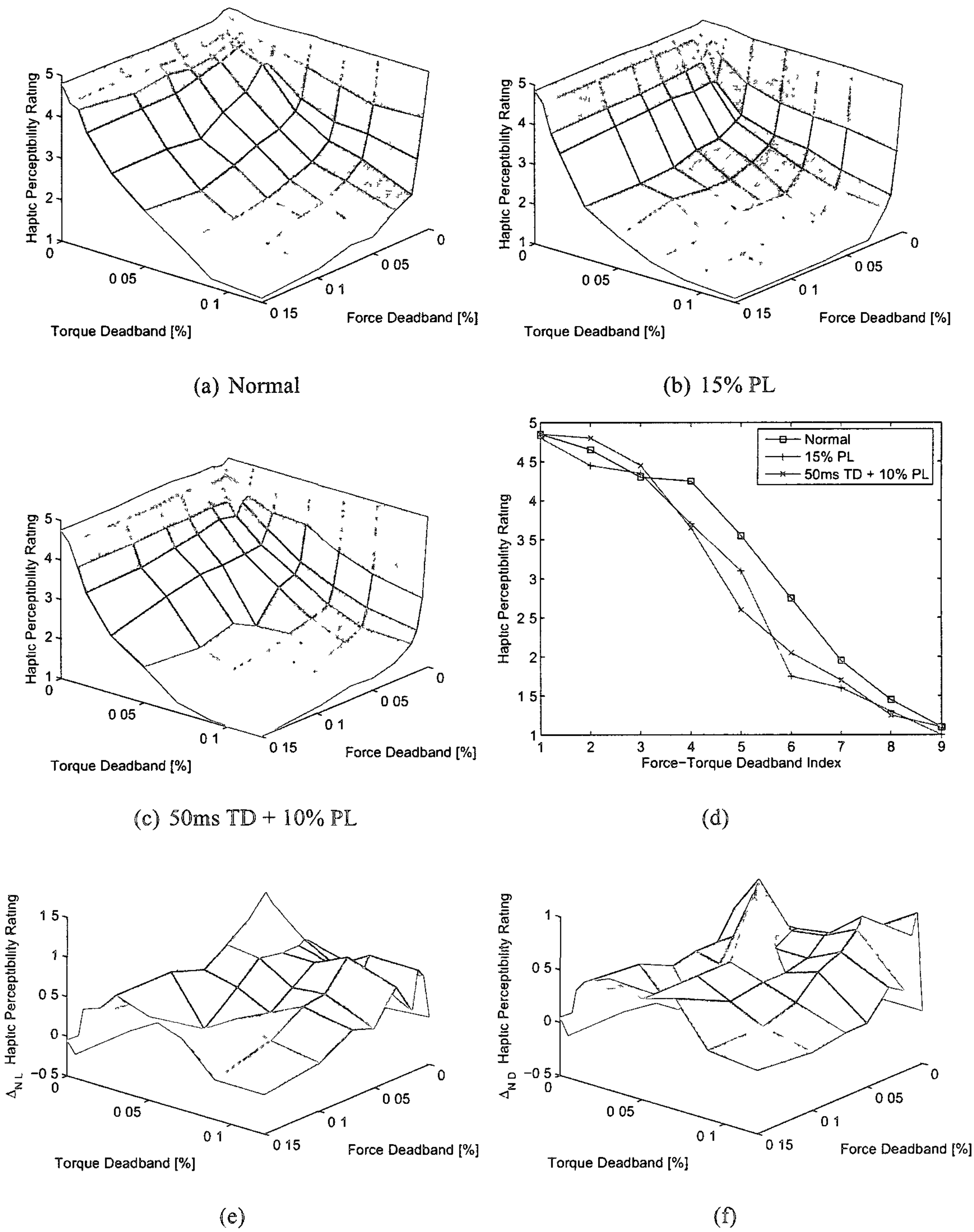


Figure 4.11: Average perceptibility rating scores in the backward direction (from the TOP back to the HSI).

Test method. A visual representation of the significance test results are presented in Fig. 4.8 (b).

A first examination of the plots revealed encouraging results. For example, with a small force/torque deadband-value pair $d_{f_4,\tau_5} = [d_{f_4} = 1.5\%,\ d_{\tau_5} = 3.0\%]$ a desirable average perceptibility rating of 4.1 with a corresponding packet reduction rate of 80.7% is achieved. In addition, from Figs. 4.5 (b) and 4.8 (b) a slightly lower average perceptibility rating of 3.6, with a corresponding data reduction rate of 85.1% can be achieved with the following force/torque deadband-value pair: $d_{f_5,\tau_5} = [d_{f_5} = 3.0\%,\ d_{\tau_5} = 3.0\%]$. It should be mentioned that in Fig. 4.8 (b) the average perceptibility score with a deadband $d_{f_5,\tau_5} = [d_{f_5} = 3.0\%,\ d_{\tau_5} = 3.0\%]$ (score = 3.6) is statistically significantly lower than the average score with $d_{f_5,\tau_4} = [d_{f_5} = 3.0\%,\ d_{\tau_4} = 1.5\%]$ (score = 4.1), $z = -2.25$, $p = 0.024$, with a large effect size ($r = 0.50$). However, a statistically significant difference does not exist between the average perceptibility scores $d_{f_5,\tau_5} = [d_{f_5} = 3.0\%,\ d_{\tau_5} = 3.0\%]$ (score = 3.6) and $d_{f_4,\tau_5} = [d_{f_4} = 1.5\%,\ d_{\tau_5} = 3.0\%]$ (score = 4.1), $z = -1.83$, $p = 0.068$, with a medium effect size ($r = 0.41$). This suggests that the distortion felt with the deadband pair $d_{f_5,\tau_5} = [d_{f_5} = 3.0\%,\ d_{\tau_5} = 3.0\%]$ is potentially significantly low as it is statistically indistinguishable from a deadband pair ($d_{f_4,\tau_5} = [d_{f_4} = 1.5\%,\ d_{\tau_5} = 3.0\%]$) with an average perceptibility rating of 4.1.

Furthermore, the average perceptibility rating scores when the proposed 6-DoF haptic data reduction algorithm is evaluated in the backward direction of the telepresence system, and under simulated network conditions with 15% packet loss, and a network-induced delay of 50ms + 10% packet loss, are shown in Fig. 4.11 (b) and (c) respectively. 2D projections of the average scores along the diagonal line of the plots, i.e., $d_f = d_{f_i},\ d_\tau = d_{\tau_j}, \forall i,j : i = j$ and $i,j \geq 1,\ i,j \leq 9$, are presented in Fig. 4.11 (d). The difference between the average perceptibility scores in normal network conditions (Fig. 4.11 (a)) and the average scores in conditions with 15% packet loss (Fig. 4.11 (b)) are presented in Fig. 4.11 (e). Similarly, the difference between the average scores in normal network conditions (Fig. 4.11 (a)) and the average scores in conditions with a 50ms delay + 10% packet loss (Fig. 4.11 (c)) are presented in Fig. 4.11 (f). The difference computation is achieved using Eq. (4.12), where $score_N(i,j)$ refers to the average perceptibility rating score in normal network conditions at force/torque deadband indices i, j, whereas $score_A(i,j)$ denotes the average score when network artifacts are considered. Furthermore, results of the Friedman Test indicated that there was a statistically significant difference in perceptibility rating scores across the eighty-one different force/torque deadband values when network conditions with 15% packet loss are considered, $\chi^2(80, n = 10) = 753.05,\ p < 0.0005$. A similar result was obtained in a network-induced time delay of 50ms and 10% packet loss, $\chi^2(80, n = 10) = 730.52,\ p < 0.0005$. Post hoc pairwise comparisons of perceptibility scores across different force/torque deadband-value pairs in adverse network conditions were then conducted using the Wilcoxon Signed-Rank Test method. A visual representation of the statistical significance test results with strictly 15% packet loss, and in network-induced delays combined with packet loss are presented in Figs. 4.9 (b) and 4.10 (b), respectively.

4.5 Discussion

Many observations can be made from the results presented in Figs. 4.5 – 4.10. First, in the forward direction (reduction of packets transmitted from the HSI to the TOP), it can be observed from Fig. 4.6 (e) and (f) that, as expected, the average perceptibility scores in normal network conditions are generally higher than in adverse network conditions (computed difference values are mostly positive; see Eq. (4.12)). More interestingly, it can be observed from Figs. 4.8 (a), 4.9 (a), and 4.10 (a) that in adverse network conditions, the position/orientation perceptual threshold dotted lines (identifying deadband-pairs with which perceptibility scores equal or higher than a 4/5 or 3/5 can be achieved) tend to shift to the upper-left corner of the perceptibility rating tables. Moreover, fewer deadband-pairs (with lower corresponding data reduction rates) fall within their bounded region. Nevertheless, the data reduction results remain almost equally as impressive in moderately degraded network conditions as in normal conditions. For example, a close inspection of Figs. 4.5 (a) and 4.9 (a) reveals that with a very high network packet loss rate of 15%, using the deadband-pair $d_{p_5,\theta_4} = [d_{p_5} = 0.2mm,\ d_{\theta_4} = 0.002rad]$, a haptic packet reduction rate of 90.0% can be achieved while maintaining an average perceptibility score of 4.2. However, in contrast to perceptibility scores when normal network traffic is considered, in conditions with 15% packet loss there is a significant leap in average perceptibility rating (at the perceptual threshold lines) when comparing the average scores obtained with the deadband-pair $d_{p_5,\theta_4} = [d_{p_5} = 0.2mm,\ d_{\theta_4} = 0.002rad]$ (score = 4.2) and the deadband-pair $d_{p_5,\theta_5} = [d_{p_5} = 0.2mm,\ d_{\theta_5} = 0.003rad]$ (score = 3.4), where a statistically significant difference between the two results clearly exists ($z = -2.345$, $p = 0.019$, with a large effect size $r = 0.52$). Similarly, a careful examination of Figs. 4.5 (a) and 4.10 (a) shows that using the suggested algorithm in a moderate network-induced delay of 50ms + 10% packet loss, haptic packet reduction rate of 91.4% can be achieved with a corresponding average perceptibility score of 4.1 when a deadband-pair $d_{p_7,\theta_4} = [d_{p_7} = 0.5mm,\ d_{\theta_4} = 0.002rad]$ is selected.

Similar to the forward direction, it can be observed from Fig. 4.11 (e) and (f) that average perceptibility scores collected when data reduction is performed on packets transmitted from the TOP back to the HSI (backward direction), in normal network conditions, are generally higher than scores acquired when applying the algorithm in adverse network conditions. Accordingly, it can be seen from Figs. 4.8 (b), 4.9 (b), and 4.10 (b) that in degraded network conditions, the force/torque perceptual threshold dotted lines shifted toward the upper-left corner of the perceptibility rating tables. Nonetheless, good data reduction results are still achieved despite undesirable network conditions, while simultaneously preserving the overall quality of the telepresence experience. For example, Figs. 4.5 (b) and 4.9 (b) reveal that with a high network packet loss rate of 15%, using the deadband-pair $d_{f_8,\tau_4} = [d_{f_8} = 10.0\%,\ d_{\tau_4} = 1.5\%]$, a force/torque packet reduction rate of 77.7% can be achieved while maintaining a good average perceptibility score of 3.6. Similarly, a closer inspection of Figs. 4.5 (b) and 4.10 (b) shows that using the proposed algorithm in a network-induced delay of 50ms + 10% packet loss, a haptic force/torque packet reduction rate of 74.9% can be achieved with a corresponding average perceptibility score of 4.0 when a deadband-pair $d_{f_7,\tau_3} = [d_{f_7} = 7.0\%,\ d_{\tau_3} = 1.0\%]$ is selected.

A comparison of the results with other published studies is difficult due to the fact that (to the best of the authors' knowledge) this is the first work in the literature that addresses data reduction in 6-DoF haptic-enabled telepresence systems. As discussed in Section 4.1, a similar perception-based packet reduction approach is presented in [4], however, it is intended for 3-DoF telepresence systems. In addition, the robustness of the authors' algorithm to network artifacts is not discussed in any detail. Furthermore, an analysis using advanced statistical methods to test for significant differences between perceptibility rating scores across different deadband values was not provided. Nevertheless, the authors state that in normal network conditions, packet reduction rates of 91.3% and 92.6% can be achieved in the forward (position packets) and backward (force packets) directions respectively when applying their 3-DoF algorithm, with barely noticeable influence on immersiveness. In contrast, in the suggested 6-DoF algorithm, packet reduction rates of 92.3% and 85.1% can be achieved in the forward (position/orientation packets) and backward (force/torque packets) directions respectively, with little or no influence on the quality of haptic-enabled telepresence interaction. Therefore, despite the significantly added realism provided with a 6-DoF force/torque haptic feedback telepresence system, the proposed algorithm achieved results that are comparable to those obtained using a simpler telepresence environment with strictly 3-DoF (point-contact) haptic force feedback.

4.6 Summary

In this chapter a human-perception-based data reduction method is suggested to reduce the number of packets transmitted in six-degrees-of-freedom telehaptic systems; specifically in haptic-enabled telepresence. In fact, to the best of the authors' knowledge, this is the first work in the literature to introduce and analyze haptic data reduction in 6-DoF telehaptic systems. The algorithm relies on knowledge from human haptic perception to reduce the number of packets transmitted without compromising transparency. Several distance metrics are also discussed to best examine the acuity of human perception in detecting haptic distortion when data reduction is performed in 6-DoF settings. A validation of the proposed haptic data reduction technique is performed under normal network conditions and in the presence of network-induced time delays and packet loss. Statistical significance tests (using Friedman's nonparametric ANOVA, and Wilcoxon Signed-Rank Tests) were carried out to determine the appropriate human haptic perceptual thresholds (force, torque, orientation, etc.) required to minimize the number of packets transmitted while preserving the immersiveness of the 6-DoF telehaptic environment. It was observed that the suggested algorithm can significantly reduce haptic data traffic in normal and adverse network conditions, with little or no influence on the quality of haptic-enabled telepresence interaction.

Part III

Data Reduction in Haptic Data Analysis

Chapter 5

Data Reduction based on Feature Generation

5.1 Introduction

The recent surge of interest in the development of sophisticated haptic displays for various virtual environment and remote sensing applications is expected to stimulate significant research in the field of haptic data analysis and processing. This research direction is of great interest as it enables a better understanding of the effects and significance of haptics in different systems and applications. For example, the analysis and interpretation of haptic information acquired in haptic-enabled applications such as, rehabilitation, security, training and simulation, medicine, entertainment, etc. is highly relevant in order to possibly reveal patterns and relationships in the data. However, regardless of the intended haptic application the amount of recorded data is significantly large as the acquired information consists of multidimensional time-varying signals. Equally important, the large volume of information is also due to the fact that the data are collected at relatively high sampling rates to capture the inherent frequency components of haptic signals. Therefore, in contrast to Chapters 3 and 4 which introduced different data reduction techniques intended for packet reduction in networked haptic environments, the next two chapters discuss data reduction methods within the context of haptic data mining and knowledge discovery. In particular, data reduction is performed while relying on feature generation and selection methods to determine relevant attributes in high-dimensional haptic datasets. This in turn minimizes the size of the large haptic datasets while retaining only information that contains potentially useful knowledge about the data. The importance and advantages of data reduction methods in reducing the size of large datasets have been briefly discussed in Section 1.3 and will be further emphasized throughout the next two chapters.

In this chapter, data reduction through feature generation is illustrated to enable visual data mining of high-dimensional haptic datasets. Visual data mining techniques have proven to be remarkably effective in exploratory data analysis. Generally, visual data mining algorithms attempt to present datasets in some visual form, enabling a user to gain insight into the data to identify potentially useful and ultimately understandable patterns, rapidly and with ease. Moreover, there are several advantages of visual data exploration over automatic data analysis techniques. First, visual data mining can handle highly inhomogeneous and noisy data with relative ease. Moreover an understanding of complex mathematical or statistical methods is not required. Furthermore, visual data exploration is intuitive and provides users with direct interaction with the data, resulting in a greater degree of confidence in the interpretation of the findings [62–64]. Conversely, a possible drawback of visual data mining lies mainly in the subjectivity of the interpretation performed by different individuals; in particular, the interpretation of a visual space may vary significantly across different users depending on their background and expertise. Nevertheless, this chapter aims to investigate the possibility of constructing visual spaces of the inherently high-dimensional haptic datasets while relying on data mining techniques. The dataset in this study is a collection of haptic-based handwritten signatures acquired using a haptic-enabled biometric application developed at DISCOVER laboratory of the University of Ottawa. Biometric systems provide a solution to ensure that protected services are solely accessible by a legitimate user. This is achieved while relying on users' behavioral and/or physiological characteristics. In recent years, the possible use of haptic devices in biometric systems has been suggested to enable improved user identification/verification performance over more traditional techniques, such as those based on handwritten signatures [10, 11, 39–41].

In all of the related works in the literature [10, 11, 39–41] (which also relied on the same datasets exploited in this chapter, refer to Chapter 2 for more details.), when preprocessing the dataset, the authors distributed the high-dimensional attributes of each signature across different instances. For example, position, velocity, orientation, torque, and force data acquired at time t_1 are assigned to instance $Inst_1$, data acquired at time t_2 are assigned to instance $Inst_2$, etc., yielding a number of instances per user signature with only few attributes per instance. An alternative (and perhaps more intuitive) approach, would be to assign all the generated haptic data attributes per signature to a single instance, i.e., each instance contains the entire (haptic-based) signature for a single user. This approach can lead to better analysis and interpretation of the haptic dataset, and improved discrimination between users. This, however, comes at the expense of having to deal with instances that consist of thousands of attributes (as opposed to only 18 attributes as it was performed in [10, 11, 39–41]).

Accordingly, in this chapter, a haptic data preprocessing technique is introduced that accurately maps the time-varying attributes of a haptic data instance (e.g., a haptic-based handwritten signature) to a single multidimensional vector. Subsequently, a high-dimensional haptic dataset, specifically, a set of haptic-based handwritten signatures, is analyzed within a visual data mining paradigm while relying on unsupervised construction of virtual reality spaces using classical optimization and genetic programming. In particular, the suggested approaches essentially consist of feature generation techniques that make use of nonlinear transformations to map a high-dimensional feature space into another space of smaller dimension while minimizing some error measure of information loss. A number of research issues that relate to visual data mining of inherently large haptic datasets are investigated in this work. The first is whether visual data mining is in fact feasible when dealing with such very high-dimensional haptic datasets that encompass thousands of features. A comparison between the well-known linear PCA, as well as nonlinear methods based on genetic programming and classical optimization techniques in the construction of visual spaces using very high-dimensional haptic datasets, is provided. In addition, when investigating the nonlinear mapping methodologies, different distance functions are examined to explore whether the choice of measure affects the representation accuracy of the computed visual spaces. Furthermore, different classifiers (Support Vector Machines [SVM], k-nearest neighbors [k-NN], Naïve Bayes, and Random Forest) are exploited in order to evaluate the potential discrimination performance of the generated low-dimensional sets of attributes.

The rest of the chapter is organized as follows. In Section 5.2 the space mapping and visualization methodology is presented. In Section 5.3 the classification techniques used to explore the potential discrimination power of the generated attributes (reduced datasets) are illustrated. In Section 5.4 the experimental settings including the haptic data preprocessing methodology are described. In Section 5.5 the results are provided. In Section 5.6 a discussion of the experimental results is presented. Finally, a summary of the chapter is provided in Section 5.7.

5.2 Space Mapping and Visualization Methodology

There are numerous approaches to presenting visual spaces constructed using data mining techniques. In this work, visual exploration of large haptic datasets is performed through the construction of Virtual Reality (VR) spaces. Several reasons make VR desirable in visual data mining, including: its flexible-nature (it permits the choice of different representation models), 3D-interactivity, and immersion. In this section a taxonomy of the construction of VR spaces is first presented. A detailed description of the methods used to linearly and nonlinearly map the original (high-dimensional) set of haptic attributes into another space of fewer dimensions (typically 2-3 to enable visualization) is also presented.

5.2.1 Taxonomy of the Construction of Virtual Reality Spaces

The construction of a VR space is performed via a transformation that maps the original set of objects under study $\mathcal{O}$ into another space of small dimension $\hat{\mathcal{O}}$, with an intuitive metric. A feature generation approach of this kind generally implies information loss as it usually involves different (linear or nonlinear) transformations ($\varphi : \mathcal{O} \rightarrow \hat{\mathcal{O}}$). There are essentially three kinds of spaces generally sought [65]: *i)* spaces preserving the structure of the objects as determined by the original set of attributes or other properties, *ii)* spaces preserving the distribution of an existing class or partition defined over the set of objects, and *iii)* hybrid spaces. Depending on the properties that the constructed objects must satisfy, different strategies should be considered:

- Unsupervised: The location of the objects in the constructed space must preserve some structural property of the data. This is achieved while relying only on the set of descriptor attributes, i.e., class information is ignored.
- Supervised: The objective is to generate a space where the objects are maximally discriminated with respect to a class distribution. Preservation of structural properties is considered. In addition, the space is distorted as much as necessary such that class discrimination is maximized.
- Mixed: A space is constructed in such a manner that a compromise between the two goals is achieved.

Conversely, with regard to the mathematical nature of the space construction algorithms, the mapping can be:

- Implicit: Transformed objects are computed directly; i.e., the algorithm does not provide a function representation of the mapping between the original space and the smaller VR space.

- Explicit: This type of algorithm can be exploited to determine a mapping function. Objects of the reduced space can be determined by simply applying this function. Two sub-types are as follows:
  - Analytical functions;
  - general function approximators: neural networks, fuzzy systems, etc.

In this work, unsupervised spaces are constructed. This is because data structure is a very important element to consider when the location and adjacency relationships between different objects in the newly generated space could give an indication of the *similarity relationships* [66, 67] between the objects defined by the original set of attributes [68]. In the following sections, three different approaches are introduced to map (i.e., reduce or compress) a high-dimensional dataset to a lower dimensional space while minimizing the potential loss of information. The first two (principal component analysis, and a classical optimization-based method) define an implicit mapping strategy, whereas the third (genetic programming-based approach) enables the generation of explicit mapping functions.

5.2.2 Implicit Mapping using Principal Component Analysis

Principal component analysis or, equivalently, Karhunen-Loève transformation [69,70] is concerned with explaining the variance-covariance structure of the data through a few linear combinations of the original variables. Its general objectives are: (1) Data reduction and, (2) interpretation. This is achieved by transforming a possibly high-dimensional dataset to a new set of uncorrelated features, referred to as the *principal components* (PCs). Suppose $\mathbf{x}$ denotes a p-dimensional dataset. The first principal component corresponds to the linear combination of the elements of $\mathbf{x}$ that depict the greatest amount of variability. Specifically, a linear function $\mathbf{y}_1 = \mathbf{e}_1^T \mathbf{x}$ is sought such that $\mathbf{y}_1$ has the largest variance possible, where $\mathbf{e}_1$ (the superscript T denotes the transpose) is a vector of p constants $e_{11}, e_{12}, \ldots, e_{1p}$, so that:

$$\mathbf{y}_1 = \mathbf{e}_1^T \mathbf{x} = e_{11}x_1 + e_{12}x_2 + \cdots + e_{1p}x_p = \sum_{j=1}^{p} e_{1j}x_j. \tag{5.1}$$

Therefore, $\mathbf{e}_1$ defines the direction of maximum variance, whereas $\mathbf{y}_1$ represents the projection of each observation vector onto $\mathbf{e}_1$. In particular, selected principal axes are chosen so that they are of length 1 and are orthogonal; algebraically, this indicates that:

$$\mathbf{e}_i^T \mathbf{e}_{i'} = \begin{cases} 1 & \text{if } i = i', \\ 0 & \text{if } i \neq i'. \end{cases}$$

The second PC, $\mathbf{y}_2$ is a linear function $\mathbf{y}_2 = \mathbf{e}_2^T\mathbf{x}$ which is uncorrelated with $\mathbf{e}_1^T\mathbf{x}$ and has maximum variance. Accordingly, at the kth stage a linear function $\mathbf{e}_k^T\mathbf{x}$ is derived such that it has maximum variance and is uncorrelated with $\mathbf{e}_1^T\mathbf{x}, \mathbf{e}_2^T\mathbf{x}, \ldots, \mathbf{e}_{k-1}^T\mathbf{x}$. A total of p PCs can be computed. However, if the variables in $\mathbf{x}$ are correlated, one often finds that most of the variation is accounted for by only q PCs, where $q \ll p$ (for visual data mining of high-dimensional data, $q \leq 3$ PCs are generally selected).

Having defined principal components, let's proceed and discuss how PCs are in fact computed. Assuming that the covariance matrix Σ of $\mathbf{x}$ is known[1], it is possible to show that the direction of maximum variance is parallel to the eigenvector associated with the largest eigenvalue of Σ. In fact, it can be shown that for $k = 1, 2, \ldots, p$, the kth PC is given by $\mathbf{y}_k = \mathbf{e}_k^T\mathbf{x}$ where $\mathbf{e}_k$ is an eigenvector of the kth largest eigenvalue (λ_k) of Σ. In addition, if $\mathbf{e}_k$ is selected to have unit length as mentioned above, then var($\mathbf{y}_k$) = λ_k, where var($\mathbf{y}_k$) corresponds to the variance of $\mathbf{y}_k$.

5.2.3 Implicit Mapping using Classical Optimization

The second feature generation method defines a mapping φ that is constructed while minimizing some error measure of information loss [71]. This is achieved as follows: if δ_{ij} is a

[1] In practice, Σ is not known, instead, the sample covariance S is computed as follows: $S_{(c \times c)} = \frac{1}{m-1}\sum_{j=1}^{m}(\mathbf{z}_j - \overline{\mathbf{z}})(\mathbf{z}_j - \overline{\mathbf{z}})^T$, where $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_m$, correspond to a set of c-dimensional vector observations, whereas $\overline{\mathbf{z}} = \frac{1}{m}\sum_{j=1}^{m}\mathbf{z}_j$.

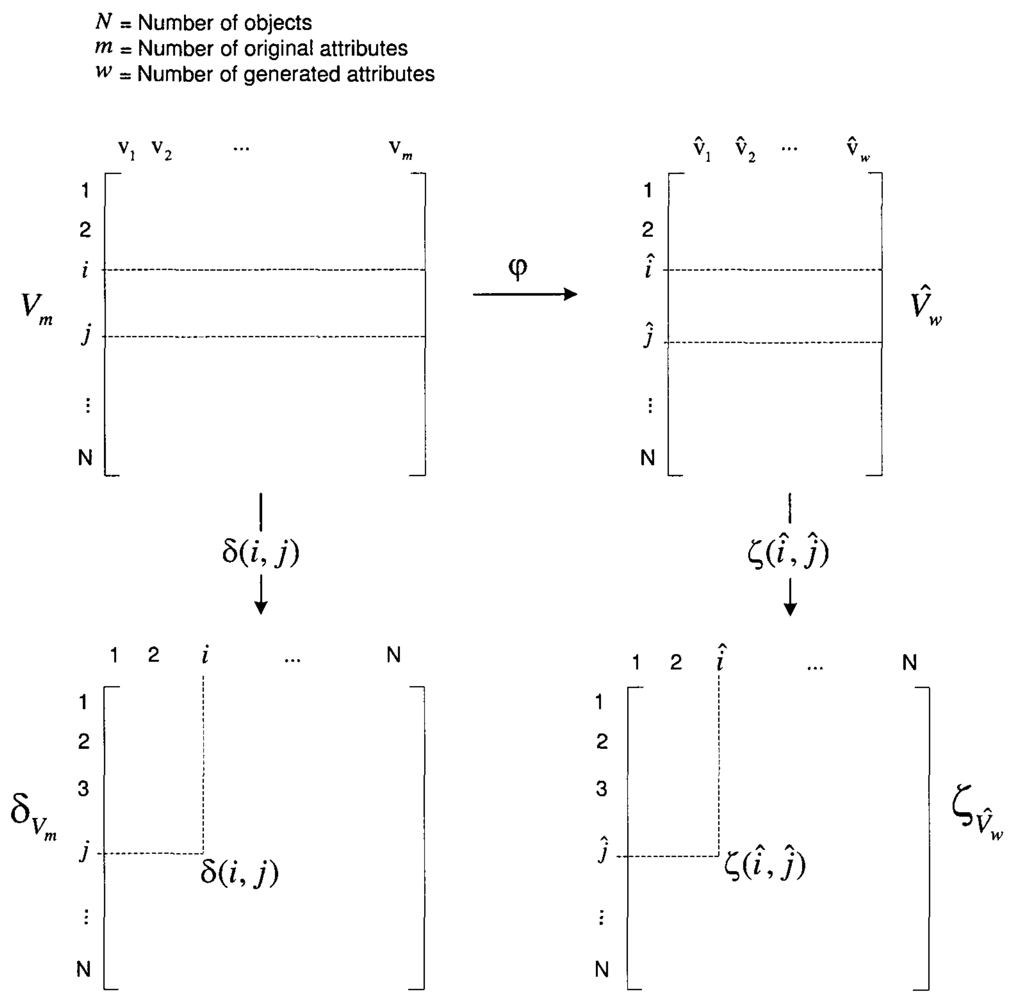


Figure 5.1: 3D nonlinear mapping procedure.

dissimilarity measure between any two objects $i, j \in \mathcal{O}$, and $\zeta_{\hat{i}\hat{j}}$ is another dissimilarity measure defined on objects $\hat{i}, \hat{j} \in \hat{\mathcal{O}}$ $(\hat{i} = \varphi(i), \hat{j} = \varphi(j))$, a frequently used error measure associated to the mapping φ is:

$$\text{Sammon error} \;=\; \frac{1}{\sum_{i<j} \delta_{ij}} \frac{\sum_{i<j} (\delta_{ij} - \zeta_{\hat{i}\hat{j}})^2}{\delta_{ij}}. \tag{5.2}$$

An illustration of the aforementioned procedure is presented in Fig. 5.1, where a mapping from an original dataset (V_m) of m dimensions (attributes) and N objects, to another set (V_w) with the same number of objects and only few dimensions (w), i.e., $w \ll m$, is performed. It can

Table 5.1: Distance functions

Metric Name	Metric Formulation
Bray Curtis	$d(x,w) = \sum_i \frac{1}{\sum_i(\lvert x_i\rvert + \lvert w_i\rvert)} \lvert x_i - w_i \rvert$
Normalized Canberra	$d(x,w) = \frac{1}{N}\sum_i \frac{1}{(\lvert x_i\rvert + \lvert w_i\rvert)} \lvert x_i - w_i \rvert$
Normalized Euclidean	$d(x,w) = \sqrt{\frac{1}{N}\sum_i (x_i - w_i)^2}$
Normalized City Block	$d(x,w) = \frac{1}{N}\sum_i \lvert x_i - w_i \rvert$

be observed that in order to perform the mapping φ, two different dissimilarity matrices are generated δ_{V_m} and $\zeta_{\hat{V}_w}$, which are associated with the original and generated sets respectively. Values in δ_{V_m} and $\zeta_{\hat{V}_w}$ are determined as follows: the computed dissimilarity δ_{ij} between any two objects $i, j \in V_m$ is stored in the (i,j)'s index of δ_{V_m}. Similarly, the obtained dissimilarity $\zeta_{\hat{i}\hat{j}}$ between any two objects $\hat{i}, \hat{j} \in \hat{V}_w$ is stored in the $(\hat{i},\hat{j})$'s index of $\zeta_{\hat{V}_w}$. In this study, ζ_{ij} is given by the Euclidean distance, whereas δ_{ij} is computed using several distance functions (for comparison purposes) presented in Table 5.1.

Consequently, an information preserving-mapping φ is initially attempted by optimizing the Sammon error measure (Eq. 5.2) using a classical optimization algorithm. Classical methods such as Steepest Decent, Fletcher-Reeves, Powell, Levenberg-Marquardt, etc. have been commonly used to optimize such measures. In this study, Fletcher-Reeves' (FR) algorithm, a well known technique used in deterministic optimization [72], is exploited. It assumes that the function f to be minimized (where in this case, f corresponds to the Sammon error function) is roughly approximated as a quadratic form in the neighborhood of an N dimensional point $\mathbf{P}$. Mathematically,

$$f(\vec{x}) \approx c - \vec{b}\cdot\vec{x} + \frac{1}{2}\vec{x}\cdot\mathbf{A}\cdot\vec{x}, \tag{5.3}$$

where

$$c \equiv f(\mathbf{P})\ ,\ b \equiv -\nabla f|_{\mathbf{P}} \text{ and } [\mathbf{A}]_{ij} \equiv \frac{\partial^2 f}{\partial_{xi}\partial_{xj}}|_{\mathbf{P}}. \tag{5.4}$$

The matrix **A** whose components are the second partial derivatives of the function is called the Hessian matrix of the function at **P**. Starting with an arbitrary initial vector $\vec{g}_0$ and letting $\vec{h}_0 = \vec{g}_0$, the conjugate gradient method constructs two sequences of vectors from the recurrence:

$$\vec{g}_{i+1} = \vec{g}_i - \lambda_i\, \mathbf{A} \cdot \vec{h}_i, \quad \vec{h}_{i+1} = \vec{g}_{i+1} + \gamma_i \vec{h}_i, \text{ where } i = 0, 1, 2, \ldots \tag{5.5}$$

The vectors satisfy the orthogonality and conjugacy conditions:

$$\vec{g}_i \cdot \vec{g}_j = 0, \quad \vec{h}_i \cdot \mathbf{A} \cdot \vec{h}_j = 0, \quad \vec{g}_i \cdot \vec{h}_j = 0, \quad j < i \tag{5.6}$$

where λ_i, and γ_i are given by

$$\lambda_i = \frac{\vec{g}_i \cdot \vec{g}_j}{\vec{h}_i \cdot \mathbf{A} \cdot \vec{h}_i}, \quad \gamma_i = \frac{\vec{g}_{i+1} \cdot \vec{g}_{i+1}}{\vec{g}_i \cdot \vec{g}_i}. \tag{5.7}$$

It can be proven [72] that if $\vec{h}_i$ is the direction from point $\mathbf{P}_i$ to the minimum of f located at $\mathbf{P}_{i+1}$, then $\vec{g}_{i+1} = -\nabla f(\mathbf{P}_{i+1})$, therefore, not requiring the Hessian matrix. Accordingly, Fletcher-Reeves' algorithm can be depicted as follows:

Select a starting point $\mathbf{P}_0$
$\vec{g}_0 \leftarrow -\nabla f(\mathbf{P}_0)$
$\vec{h}_0 \leftarrow \vec{g}_0$
for $i \leftarrow 1$ **to ... do**
  Minimize $f(\mathbf{P})$ along $\vec{h}_i$, and let the minimum point be $\mathbf{P}_{i+1}$.
  $\vec{g}_{i+1} \leftarrow -\nabla f(\mathbf{P}_{i+1})$.
  if $|\vec{g}_{i+1}| < \varepsilon$ or $|f(\mathbf{P}_{i+1}) - f(\mathbf{P}_i)| < \varepsilon$ exit loop, (convergence criterion has been met).
  $\gamma \leftarrow \frac{\vec{g}_{i+1} \cdot \vec{g}_{i+1}}{\vec{g}_i \cdot \vec{g}_i}$.
  $\vec{h}_{i+1} \leftarrow \vec{g}_{i+1} + \gamma \cdot \vec{h}_i$.
end

The φ mappings obtained using an approach of this kind are implicit, as the images of the transformed objects are computed directly and the algorithm does not provide a function representation. Hence, the accuracy of the mapping depends on the final error obtained in the optimization process. Explicit mappings can, however, be obtained from these solutions using neural networks, genetic programming, and other techniques.

5.2.4 Explicit Mapping using Genetic Programming

Analytic functions are among the most important building blocks for modeling, and consist of a classical form of expressing knowledge. In data mining, however, direct discovery of general analytic functions poses significant challenges due to the (in principle) infinite size of the search space. Within computational intelligence, Genetic Programming (GP) techniques are a promising approach to overcome this problem, as they aim at evolving computer programs which, ultimately, are functions. There are many variants of GP algorithms in the literature; the one exploited in this work is the so-called Gene Expression Programming (GEP) [73, 74]. It is essentially an evolutionary algorithm which uses populations of individuals, selects them according to fitness, and introduces genetic variation using one or more operators.

GEP individuals are nonlinear entities of different sizes and shapes (expression trees) encoded as strings of fixed length. For the interplay of the GEP chromosomes and the expression trees (ET), GEP uses an unambiguous translation system to transfer the language of chromosomes into the language of expression trees and vise-versa. The structural organization of GEP chromosomes allows a functional genotype/phenotype relationship, as any modification made in the genome always results in a syntactically correct ET or program. The set of genetic operators applied to GEP chromosomes always produces valid ETs.

Chromosomes in GEP itself are composed of genes structurally organized in a head and a tail [73]. The head contains symbols that represent both functions (elements from a function set F) and terminals (elements from a terminal set T), whereas the tail contains only terminals. Therefore, two different alphabets occur at different regions within a gene. For each problem, the length of the head h is chosen, whereas the length of the tail t is a function of h and the number of arguments of the function with the largest arity.

As an example, consider a gene composed of the function set F=$\{Q, +, -, *, /\}$, where Q represents the square root function, and the terminal set T=$\{a, b\}$. Such a gene looks like (the tail is shown in **bold**): `*Q-b++a/-b`**`baabaaabaab`**, and encodes the ET which corresponds to the mathematical equation $f(a, b) = \sqrt{b} \cdot \left(\left(a + \frac{b}{a}\right) - ((a - b) + b)\right)$ simplified as $f(a, b) = \frac{b \cdot \sqrt{b}}{a}$.

Moreover, GEP chromosomes are usually composed of more than one gene of equal length. For each problem the number of genes as well as the length of the head have to be chosen. Each gene encodes a sub-ET and the sub-ETs interact with one another forming more complex multi-subunit ETs through a connection function. As an evolutionary algorithm GEP defines its own set of crossover, mutation, and other operators [74]. Furthermore, to evaluate GEP chromosomes, different fitness functions can be used.

For the research described in this chapter, GEP is exploited in an attempt to generate analytic functions that nonlinearly map the original highly multidimensional haptic objects into a 3D visual space. Consequently, in contrast to the Fletcher Reeves-based mapping technique which is implicit, space mapping using GEP is achieved in an explicit fashion.

Furthermore, GEP-generated analytic functions are typically modeled as $y = f(v_1, \ldots, v_m)$, where $(v_1, \ldots, v_m)$ is the set of independent or predictor variables (attributes), and y the dependent or predicted variable (decision classes), so that $v_1, \ldots, v_m, y \in \mathbb{R}$, where $\mathbb{R}$ are the reals. In general terms, the model describing the program is given by $y = f(\vec{v})$, where $y \in \mathbb{R}$ and $\vec{v} \in \mathbb{R}^m$. However, the case presented in this work involves vector functions, where the model associated to the evolved programs is in effect $\vec{y} = f(\vec{v})$. This is due to the fact that the mapping is performed between vectors of two spaces of different dimensions (m and w). Specifically, the transformation $\varphi : \mathbb{R}^m \rightarrow \mathbb{R}^w$ which maps vectors $\vec{v} \in \mathbb{R}^m$ to vectors $\vec{y} \in \mathbb{R}^w$ result in a reformulation of Eq. (5.2) (and the other associated measures) to:

$$\text{Sammon error} \quad = \quad \frac{1}{\sum_{i<j} \delta_{ij}} \frac{\sum_{i<j} (\delta_{ij} - d(\vec{y}_i, \vec{y}_j))^2}{\delta_{ij}} \tag{5.8}$$

where $\vec{y}_i = \varphi(\vec{x}_i)$ and $\vec{y}_j = \varphi(\vec{x}_j)$. $\vec{x}_i$ and $\vec{x}_j$ correspond to objects at indices i and j respectively.

The implication of this modification within the context of genetic programming, is that instead of evolving expressions trees where a one-to-one correspondence exists between an expression tree and a fitness function (as in the classical case), the evolution has to consider populations of *forests*. Specifically, the evolution of the fitness function depends on a set of trees within a forest. The cardinality of the forest is equal to the dimension of the desired visual space (w).

5.3 Classification

In this section, different classification techniques, including Support Vector Machines, k-nearest neighbors (k-NN), Naïve Bayes, and Random Forest classifiers are introduced. The classifiers are later exploited in Section 5.5 to evaluate the discrimination power of constructed visual spaces. It should, however, be emphasized that this is only done for the purpose of data and results analysis. Since, as previously mentioned, unsupervised mapping algorithms such as those investigated in this work, have as a goal to preserve the structural properties of the data and not (specifically) to maximize or maintain the discrimination capabilities of a dataset.

5.3.1 Support Vector Machine

Support Vector Machines (SVM) were introduced in [75, 76] to solve pattern classification and nonlinear regression estimation problems. The aim of support vector classification is to derive a computationally efficient approach of learning accurate separating hyperplanes in a high-dimensional feature space. Without loss of generality, given the set of training vectors belonging to two separate classes,

$$\mathcal{D} = \left\{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_l, y_l)\right\}, \quad \mathbf{x} \in \mathbb{R}^n, \mathbf{y} \in \{-1, 1\}, \tag{5.9}$$

the SVM requires the solution of the following optimization problem:

$$\begin{aligned}
&\min_{\mathbf{w},b,\xi} \ \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{i=1}^{l}\xi_i \\
&\text{subject to} \ \ y_i(\mathbf{w}^T\phi(\mathbf{x}_i)+b) \geq 1-\xi_i, \\
&\qquad\qquad\quad \xi_i \geq 0.
\end{aligned} \tag{5.10}$$

where the vector $\mathbf{w}$ is known as the *weight vector* and the scalar b is called the bias. ξ_i is a slack variable that generalizes the SVM approach to allow some classification errors (hence, ξ_i is a measure of the misclassification errors), whereas $C > 0$ is a penalty parameter of the error term that must be chosen to reflect the knowledge of the noise present in the data. Moreover, in Eq. (5.10), the training vectors $\mathbf{x}_i$ are mapped into a higher dimensional space by the function ϕ. In the linearly separable case (i.e., a mapping function is not required), SVMs attempt to find the hyperplane (class boundary) with the maximum distance from the nearest training patterns. In the nonlinear case, where the data is not linearly separable, a transformation to a higher dimensional space can be performed. With an appropriate nonlinear mapping $\phi(\cdot)$ to an adequate high dimension, data from two categories can always be separated by a hyperplane. In order for this mapping to be computationally feasible, a kernel function $K(x_i, x_j) \equiv \phi(\mathbf{x}_i)^T\phi(\mathbf{x}_j)$ is typically exploited. Many commonly used functions can be used to perform this mapping, including polynomial, radial basis functions, and certain sigmoid functions.

5.3.2 *k*-Nearest Neighbor

The k-Nearest Neighbor algorithm is a simple nonparametric classifier that is widely used in many pattern classification problems. Nonparametric methods differ from their parametric counterparts which assume a certain model whose parameters are estimated from a given sample of data (e.g., using maximum likelihood estimation). Nonparametric approaches are commonly used in problems where an appropriate parametric model is not known or parameter estimation may be too costly. Moreover, as nonparametric techniques make fewer assumptions, they tend to be more robust, and their applicability is much broader than the corresponding parametric methods. In the k-NN rule, a test instance $\mathbf{x}$ is classified by assigning it to the pattern class that appears most frequently amongst the k-nearest mining samples. If no single class appears with the greatest frequency, then the class with the minimum average distance from $\mathbf{x}$ is selected. The distance metric that is most commonly used in nearest-neighbor methods is the Euclidean distance. For example, the Euclidean distance between two attribute value vectors $\mathbf{x}_1 = (x_{11}, x_{12}, \ldots, x_{1n})$ and $\mathbf{x}_2 = (x_{21}, x_{22}, \ldots, x_{2n})$ is computed as follows:

$$d(\mathbf{x}_1, \mathbf{x}_2) = \sqrt{\sum_{j=1}^{n} a_j (x_{1j} - x_{2j})^2}, \tag{5.11}$$

where a_j is a scaling factor for dimension j. It is generally used to scale the attributes in order for the spread of feature values along each dimension is approximately identical. The selection of the parameter k is essential in the design of the k-NN model as it can strongly influence the classifier's performance. A small value of k will lead to a large variance in classification accuracy. In contrast, a large value of k could decrease the chance of the classification decision being influenced by noisy data samples close to $\mathbf{x}$. However, setting k to a large value may potentially lead to a significant model bias. Nevertheless, it is clear that the denser the training samples are around $\mathbf{x}$ and the larger the value of k, the better the prediction estimate. Conversely, it has been established that it is only in the limit as the number of training samples goes to infinity that the nearly optimal behavior of the k-nearest neighbor rule can be assured [77].

5.3.3 Naïve Bayes

The Naïve Bayesian classifier is a straightforward and widely used method designed for use in supervised induction problems, where the purpose is to accurately predict the class membership of test instances, given a set of labeled training instances (which contains class information). The classifier is referred to as *naïve*, since it relies on two important simplifying assumptions: (1) All attributes are assumed to be independent from each other given the class, and (2) hidden or latent attributes have no affect on the prediction performance [78]. Accordingly, based on Bayes theorem, the Naïve Bayes classifier is best described using the following formulation [78]:

$$p(C = c \mid \mathbf{X} = \mathbf{x}) = \frac{p(C = c)p(\mathbf{X} = \mathbf{x} \mid C = c)}{p(\mathbf{X} = \mathbf{x})} \tag{5.12}$$

where

$$p(\mathbf{X} = \mathbf{x} \mid C = c) = \prod_i p(X_i = x_i \mid C = c) \tag{5.13}$$

Here C is a random variable that denotes the class of an instance, $\mathbf{X}$ corresponds to a vector of random variables of the observed attribute values. c denotes a particular class label and $\mathbf{x}$ represents a specific observed attribute value vector. The classifier is therefore entirely defined by the conditional probabilities of each attribute given the class. The Naïve Bayes classifier handles discrete and numeric attributes quite differently. In the former case, each discrete attribute is represented by a real number in the range [0,1], which essentially denotes the probability $p(X = x \mid C = c)$, i.e., the probability that the attribute X will equal the value x, given that the class is c. In the latter case, each numeric attribute is modeled using a certain probability distribution, which is generally assumed to be Gaussian. It is therefore apparent that the conditional independence property significantly simplifies the model's learning procedure. Despite the simplification assumption that underlie its design, the Naïve Bayesian classifier is well reputed for its robustness, achieving classification accuracies comparable to other much more sophisticated methods in practical settings, even when the independence assumption does not hold. Due to its simplicity and effectiveness, the Naïve Bayes classifier is often considered as a reference technique in pattern classification research and studies.

5.3.4 Random Forest

A Random Forest classifier is a classification algorithm that encompasses a collection of tree-like classifiers [79]. The method relies on a large number of individual decision trees, all of which are trained to tackle the same classification problem. Each tree is built on a randomly selected subset of the training data, where final classification of a test sample is based on a majority vote determined from all the trees in the forest. Specifically, there are three factors that contribute to the individuality of each tree:

1. Each tree is built using a bootstrapped sample of the training data, i.e., each tree is trained with a distinct set of n instances which are drawn independently with replacement from the original training set of n instances.

2. If the data set comprises instances of M features, then, at each node of each tree a split is determined by searching through $m \ll M$ randomly chosen features.

3. Each tree is grown to the maximum size and kept with no pruning.

Furthermore, the forest error rate depends on two important factors that are directly related to the parameter m (the size of the random subset of features):

1. Increasing the value of m increases the correlation between any two trees in the forest. However, an increase in the correlation leads to an increase in the error rate of the forest.

2. Increasing the value of m increases the strength (i.e., improves the classification accuracy) of each individual tree and results in a decrease in the forest error rate.

The above reasons make m the only adjustable parameter to which random forests is quite sensitive. The optimal range of m, which is generally quite wide, is selected such that the best compromise between strength and correlation is achieved. This is performed while relying on the out-of-bag (OOB) error rate. More specifically, the training sets of individual trees are constructed by bootstrap replication, where roughly one-third are left out: they are called OOB samples. As each tree is built, its OOB samples are used to evaluate its classification accuracy. The OOB error estimate is therefore the number of misclassified OOB samples for each tree, averaged over the total number of trees.

Random forests have proven to be very effective when dealing with high-dimensional data sets. Other important advantages of Random forests include: computational effectiveness, impressive classification accuracy, ability to overcome the problem of overfitting, and low sensitivity to noisy data when compared to other similar techniques, such as boosting.

5.4 Experimental Settings

In this section, the haptic-enabled virtual check application as well as the exploited haptic dataset are initially described. The proposed haptic data preprocessing method used to map a multidimensional haptic data instance to a one-dimensional vector is introduced. In addition, the parameter settings of the implicit Fletcher-Reeves-based and the explicit GEP-based mapping algorithms will be illustrated.

5.4.1 Haptic-enabled Virtual Environment

The experiments are performed using the Reachin Display [80], which integrates a haptic device with stereo graphics for an immersive and high-quality 3D experience. The Reachin hapto-visual interface enables users to see and touch virtual objects at the same location in space. This approach enables a superior integration of vision and touch than a conventional 2D screen-based display. The haptic stimulus is sensed using the Sensable PHANTOM Desktop force-feedback device, which is equipped with an encoder stylus that provides six-degree-of-freedom single contact point interaction and positional sensing. In the case presented here, the visual stimuli consist of a virtual pen and a virtual check on which users can record their handwritten signature. The latter haptic-enabled virtual environment has been selected since handwritten signatures have been widely accepted as a mean to prove authenticity and authorship of a document. Conversely, the haptic stimuli are force and frictional feedback that attempt to mimic the tactile sensations felt when signing a traditional paper check. More specifically, the check is built on an elastic membrane surface with particular texture features, providing the users with a user-friendly and realistic feel of the virtual object. Moreover, similar to con-

John

Sample	Position_X	Postion_Y	Position_Z
1	121.66	35.43	-89.09
2	122.23	35.92	-90.22
3	117.09	37.46	-91.64
...			
N	113.75	45.35	-113.50

Alice

Sample	Position_X	Postion_Y	Position_Z
1	119.40	33.48	-93.15
2	119.86	34.01	-93.32
3	120.10	34.42	-93.65
...			
N	101.70	46.44	-110.88

Class	A_1	A_2	A_3	A_4	A_5	A_{3N}
John	121.66	35.43	-89.09	122.23	35.92 ...	-113.50
Alice	119.40	33.48	-93.15	119.86	34.01 ...	-110.88

Figure 5.2: Mapping the time-varying haptic-based signatures to single multi-feature vectors.

ventional dynamic signature verification technologies, the virtual check application records a wide array of attributes that depict a user's physical and behavioral traits.

5.4.2 Haptic Dataset and Preprocessing

The haptic-based handwritten signatures were diligently obtained from 13 different participants, where 10 signatures were collected per individual. A number of haptic data types are considered that characterize the instantaneous state of the haptic system, including, three-dimensional position, force (pressure exerted on the virtual check), torque, and angular orientation. Moreover, the multi-feature and multidimensional haptic data were recorded at 100 Hz. As the data is time-varying, the resulting number of attributes per signature is in fact the number of haptic data types considered (position, force, torque, ...) times the number of samples recorded per data type during each signature acquisition (it is clear that the number of samples across different haptic data types is the same since samples of all data types are acquired at the same frequency). This evidently leads to significantly large feature vectors that encompass thousands of haptic-based attributes. An example of the mapping performed between the acquired signature data and the resulting multi-attribute vector used in subsequent steps of the biometric system, is demonstrated in Fig. 5.2. However, in this case, for the sake of simplicity, a single haptic signature per user is shown, and only position data types (along

x, y, and z) are provided. It can be observed that the position values associated with each haptic data sample (along each row) are contiguously linked to form a single vector per object. Moreover, in the example of Fig 5.2, the number of data samples N is assumed to be equal for both, John, and Alice. However, it is evident that the latter assumption is seldom true, suggesting that in practice the number of attributes of different signatures will almost always differ. Length-normalization of the haptic-based signatures is therefore necessary in order to guarantee the same number of attributes across all recorded signatures. This is performed by upsampling/downsampling the data types (e.g., the column vectors $position_x$, $position_y$, and $position_z$ in Fig. 5.2) to a predetermined number of samples prior to mapping the signatures to a 1D vector. This can be performed using a basic nearest-neighbor interpolation approach as follows:

$$\hat{v}(k) = v(round[(k-1)\cdot d/u]+1) \tag{5.14}$$

where v corresponds to the original data sequence, whereas d/u denotes the reciprocal of the sampling factor (e.g., $d/u = 1/2$ suggests that the resulting vector $\hat{v}$ will be 2 times longer than the original). An advantage of this approach is its great simplicity. However, an important drawback of this method is that the obtained re-sampled vector is generally not smooth as staircase-like artifacts are often produced. This can be best observed when upsampling is done, where a nearest-neighbor replication of the samples is performed. This limitation can be overcome by computing the best estimate of each sample using a weighted average of its neighbors. This can be performed as follows:

$$\hat{v}(k) = \frac{\sum_{j=1}^{knn} \frac{B_j}{(d_j)^p}}{\sum_{j=1}^{knn} \frac{1}{(d_j)^p}} \tag{5.15}$$

where knn denotes the number of nearest neighbors (B) of a sample at index k, d_j is the distance between the sample at index k and its jth neighbor, whereas p is a constant with a suggested value that is ≥ 1.

Accordingly, in order to ensure accurate discrimination between the signatures, the obtained feature vectors were normalized to a common length of 10000. Essentially, the acquired haptic data types are re-sampled (upsampled/downsampled) when necessary to ensure a common feature vector length across all instances. Moreover, the feature vector length was selected in such a manner to minimize the information loss that is most apparent when downsampling is performed. Consequently, the computed preprocessed dataset contains 130 instances, where each consists of 10000 features. However, it should be mentioned that the suggested size normalization procedure removes information describing the length of the acquired signatures. In the future work, the original length of recorded signatures (or user-specific information, depending on the application) must be considered and added as a separate feature.

5.4.3 Implicit and Explicit Mapping Algorithms Settings

The implicit classical mapping method corresponds to Fletcher-Reeves' algorithm. This technique attempts to solve Eq. (5.2) to optimize the transformation $\varphi : \mathbb{R}^m \rightarrow \mathbb{R}^w$. Moreover, in order to reduce the chance of getting trapped in a local minima ten runs of the algorithm were performed and the best result was then selected. This procedure was performed during each experiment (for each distance measure used of Table 5.1). The solutions were then used to construct 3D visual spaces of the data.

As for the GEP experiments, the primary experimental settings used are listed in Table 5.2. It can be seen that the mathematical expressions are composed of three chromosomes per individual and each chromosome is made from 3 gene expressions, all linked by the addition operator. The gene expressions can be formed using constants, the independent variables (features) associated with each problem, and any of the functions listed. The weightings for the functions allow the user to give a preference to some operators over others. In this case, the addition operator is favored over the others (subtraction, multiplication, power, and sine). It is beyond the scope of this work to discuss the GEP parameters in detail. For more information regarding the GEP algorithm settings please refer to [73, 74]. For each of the presented experiments a set of 100 runs was performed with random seeds. It is important to mention that during each run, there are in fact 1000 different functions that are obtained (since the process is initiated with a population of size 1000), however, only the GEP model with the best compromise between mapping error (Sammon error) and discrimination power is selected.

5.5 Results

Table 5.3 presents the results of the best (nonlinearly mapped) models found using the classical optimization-based approach (FR) and the GEP method. In addition, the experiments are repeated with different distance measures (δ_{ij} of Eqs. (5.2) and (5.8)) which include the

Table 5.2: Experimental settings of the GEP algorithm.

GEP Parameters	Experimental Values
No. generations	100000
Population size	1000
No. chromosomes / individual	3
No. genes / chromosome	3
Gene head size	16
Linking function	Addition
No. constants / gene	2
Bounded range of Constants	[0,10]
Inversion rate	0.1
Mutation rate	0.044
is-transposition rate	0.1
ris-transposition rate	0.1
One-point recombination rate	0.3
Two-point recombination rate	0.3
Gene recombination rate	0.1
Gene transposition rate	0.1
rnc-mutation rate	0.01
dc-mutation rate	0.044
dc-inversion rate	0.1
Function set (e.g. $\{\text{function}_1(\text{weight}), \text{function}_2(\text{weight}), \ldots\}$)	add(2), sub(1), mult(1), pow(1), sin(1)

Table 5.3: Nonlinear mapping and discrimination results of constructed haptic feature spaces.

Distance Metric	*Fletcher-Reeves*					*Genetic Expression Programming*				
	Sammon Error	*SVM*	*k-NN*	*Naïve Bayes*	*Random Forest*	*Sammon Error*	*SVM*	*k-NN*	*Naïve Bayes*	*Random Forest*
Normalized City Block	0.03126	79.2	81.5	80.7	66.9	0.07178	74.6	70.8	65.4	65.4
Normalized Euclidean	0.03958	79.2	76.9	70.0	67.7	0.07084	68.5	59.2	61.5	54.6
Bray Curtis	0.04926	77.7	77.6	75.3	69.2	0.09281	63.8	60.0	61.5	54.6
Normalized Canberra	0.06014	73.8	73.0	69.2	69.2	0.11192	50.0	45.5	42.3	40.8

Table 5.4: Discrimination results of the original, and linearly constructed feature spaces.

Datasets	*Dimensions*	*Classifiers*			
		SVM	*k-NN*	*Naïve Bayes*	*Random Forest*
Original	10000	92.3	80.0	85.3	88.4
PCA	3	77.7	72.3	74.6	66.9

Normalized Euclidean, Normalized Clark, Bray Curtis, and Normalized Canberra as depicted in Table 5.1. The results encompass mapping errors (Sammon errors) and classification rates using four different classifiers, SVM, k-NN, Naïve Bayes, and Random Forest. It is important to mention that the presented classification rates are in fact cross-validation (CV) results, specifically, 5-fold cross-validations of the generated 3D spaces (recall that a mapping from 10000 features to 3 features is performed to enable visualization of the underlying structure of haptic-based signatures). In k-fold CV, a dataset is divided into k subsets of equal size. Sequentially one subset is tested using the classifier trained on the remaining $k - 1$ subsets. Thus, each instance of the whole dataset is predicted once. The k results from the folds are then averaged to produce a single estimation. The purpose of cross-validation is therefore to estimate the generalization capability of a model with respect to a data set that is independent of the data used during training. Consequently, for our purpose, it provides us with an estimate of the discrimination power preserved in the constructed space. It should, however, be emphasized that cross-validation is only done for the purpose of data and results analysis. Since, as previously mentioned, unsupervised mapping algorithms such as the ones investigated in this work, have as a goal to preserve the structural properties of the data and not (specifically) to maximize or maintain the discrimination capabilities of a dataset.

Furthermore, for the sake of comparison with linear-based feature generation methods, discrimination (5-fold CV) performance in a PCA-mapped 3D haptic feature space is illustrated in Table 5.4. Specifically, the PCA-generated 3D dataset is obtained while only considering the first three principal components. Basically, it is assumed that the structural information for the dataset is stored in at most three principal components, so that dimensionality reduction from 10000 dimensions to only 3 dimensions can be achieved, with no important loss of information. In addition, for illustration purposes, 5-fold cross-validation results computed using the entire original preprocessed dataset (130 objects, 10000 attributes) are also presented in Table 5.4.

Snapshots of the VR 3D spaces constructed using the FR and GEP mapping algorithms as well as PCA, are presented in Figs. 5.3 – 5.6 (a), Figs. 5.3 – 5.6 (b), and Fig. 5.7, respectively. The figures clearly demonstrate the similarity in structure between the original classes (Users 1 to 13). Moreover, Figs. 5.3 – 5.6 are constructed while relying on the Normalized City Block distance, Euclidean distance, Bray Curtis distance, and the Normalized Canberra distance measures, respectively. Furthermore, in order to better distinguish between users, all classes have been wrapped with transparent membranes (Convex Hulls).

5.6 Discussion

Numerous observations can be made from the results presented in Tables 5.3 and 5.4, as well as Figs. 5.3 – 5.6. First, from Table 5.3, with regard to the Sammon error, it can be seen that the FR method has consistently outperformed GEP by approximately a factor of 2. Similarly, the discrimination power of the spaces constructed using FR is higher than that of the spaces obtained using GEP. However, looking at the spaces constructed using the Normalized

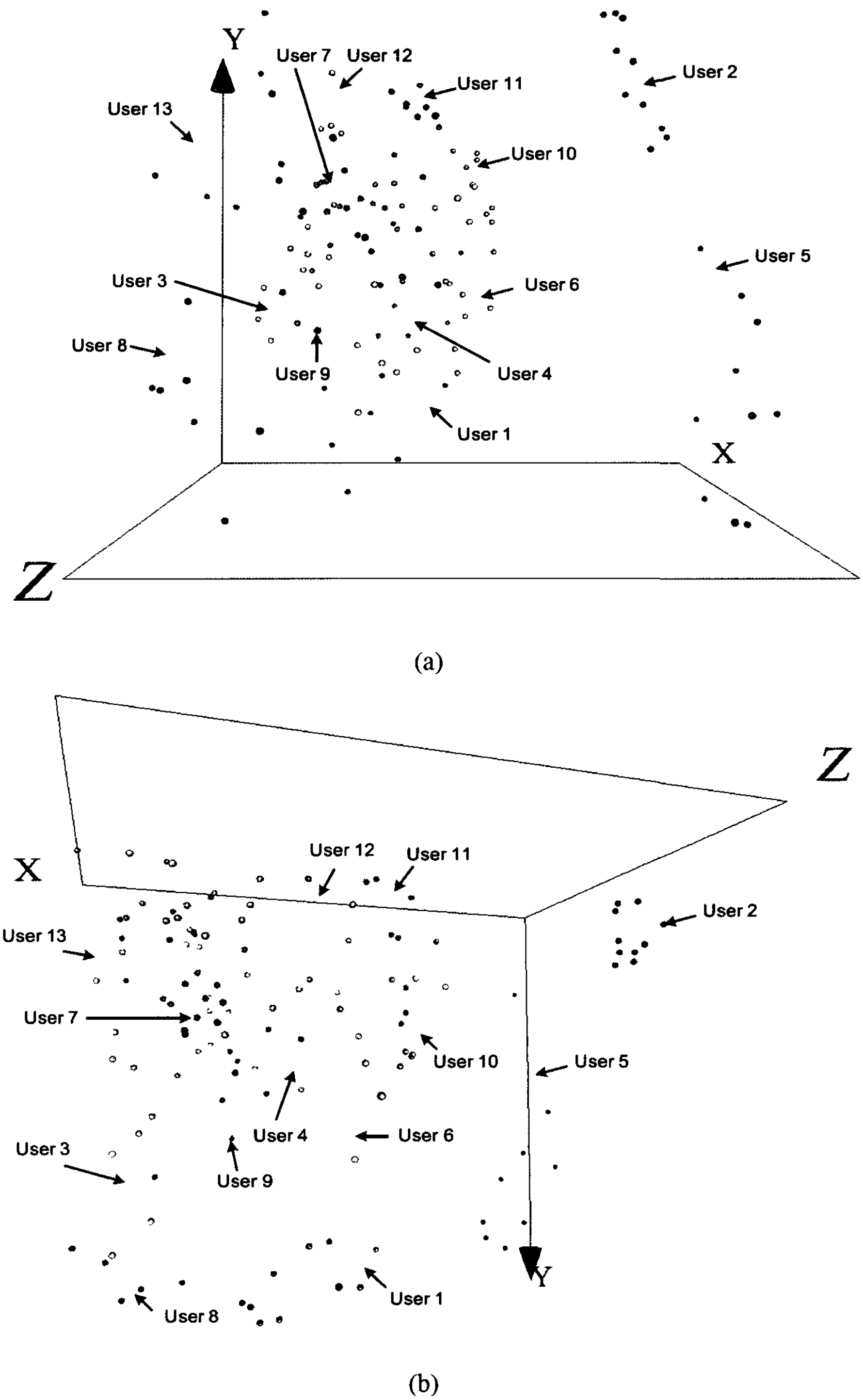
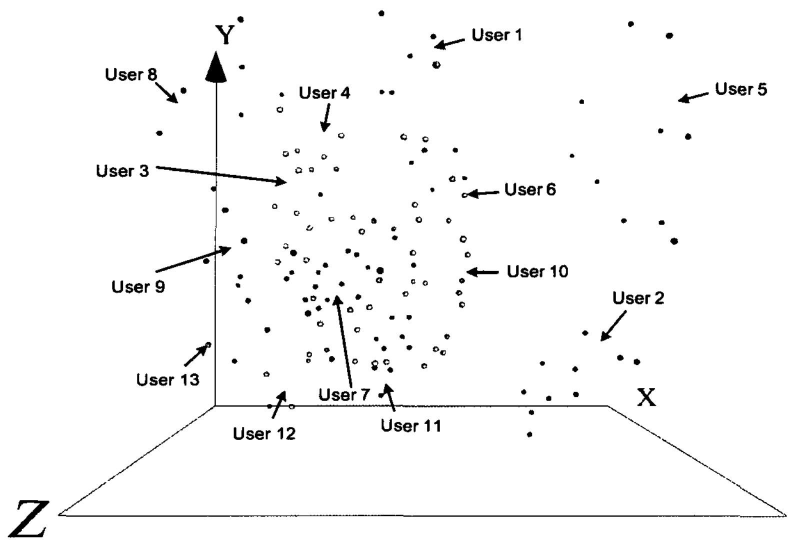
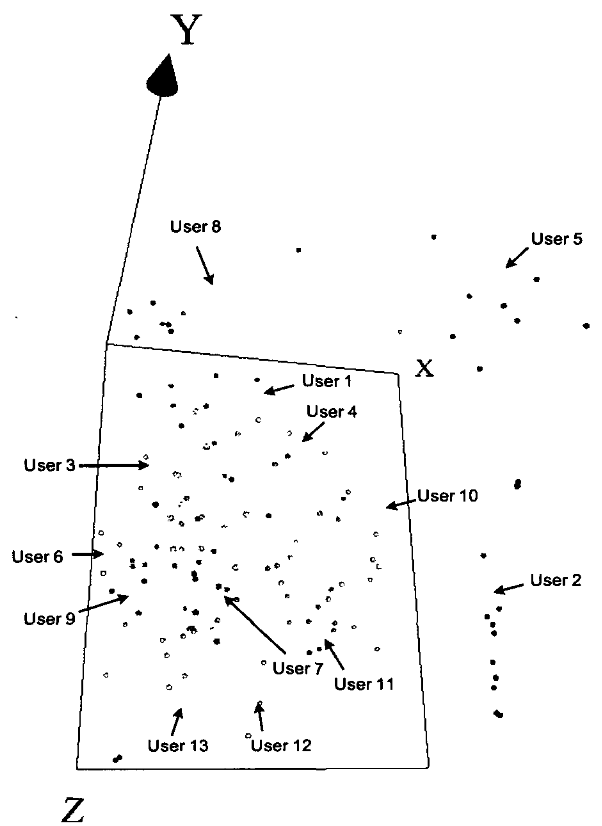


Figure 5.3: 3D VR spaces of the haptic-based handwritten signatures dataset constructed using (a) FR , (b) GEP, and the Normalized City Block distance measure.

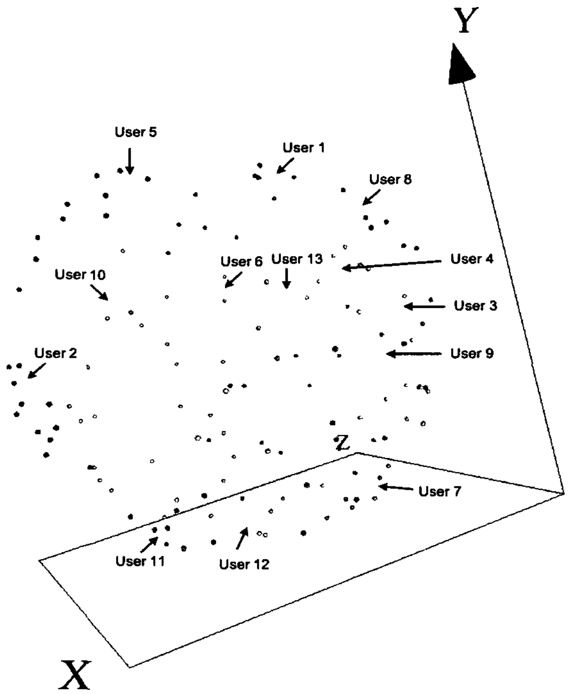


(a)

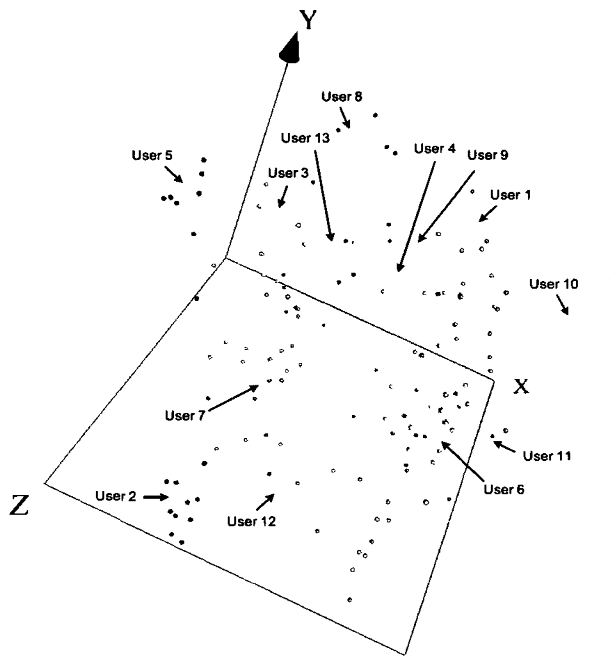


(b)

Figure 5.4: 3D VR spaces of the haptic-based handwritten signatures dataset constructed using (a) FR , (b) GEP, and the Normalized Euclidean distance measure.

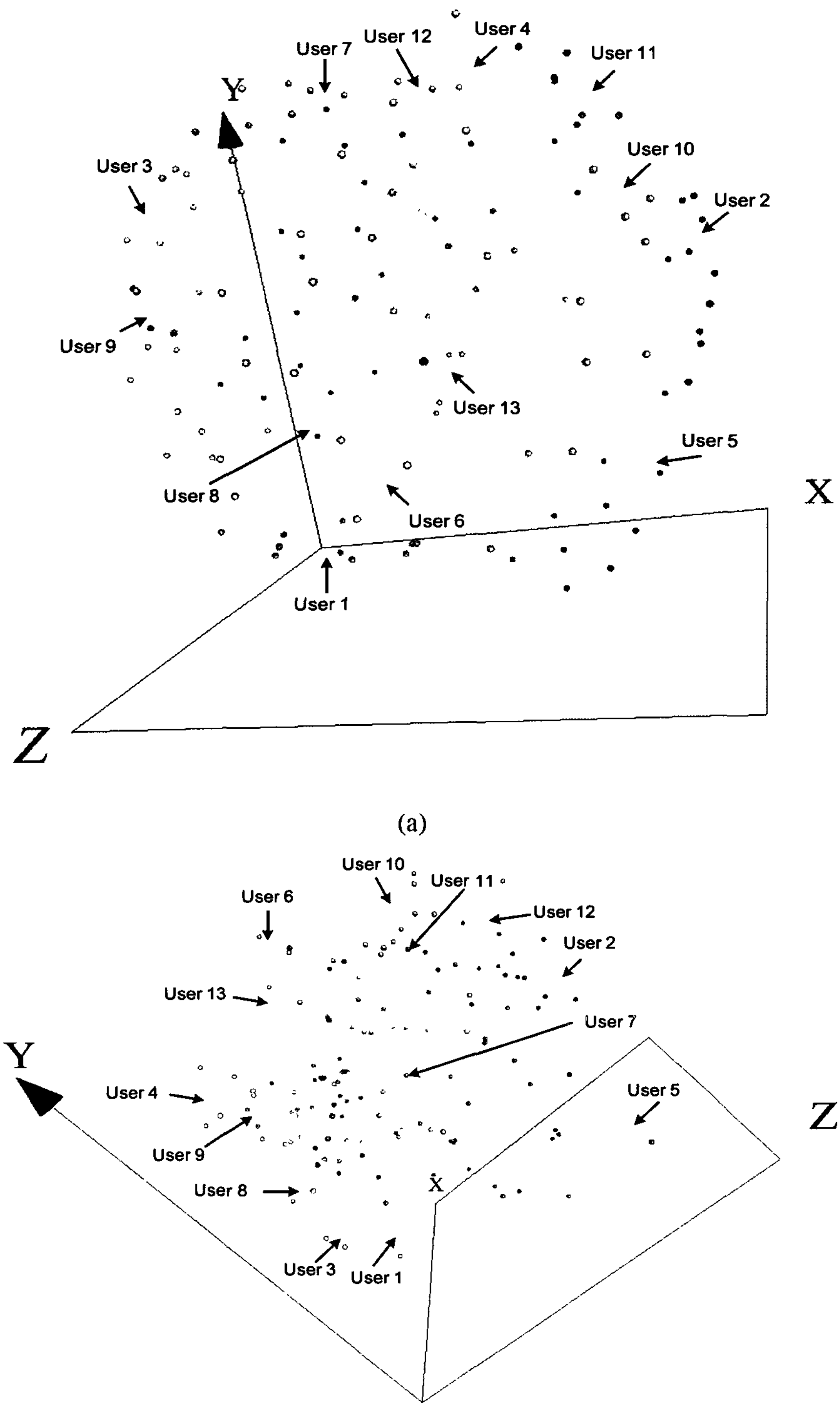


(a)



(b)

Figure 5.5: 3D VR spaces of the haptic-based handwritten signatures dataset constructed using (a) FR , (b) GEP, and the Bray Curtis distance measure.



(a)

(b)

Figure 5.6: 3D VR spaces of the haptic-based handwritten signatures dataset constructed using (a) FR , (b) GEP, and the Normalized Canberra distance measure.

City Block distance measure, it can be observed that although FR outperforms GEP by a factor of 2.3 (with respect to the Sammon error), their discrimination power are quite similar. Moreover, comparing the Normalized City Block-based spaces of FR and GEP presented in Fig. 5.3 (a) and (b) respectively, it can be seen that a similar class structure is exhibited by the GEP method despite an important difference in the mapping error. The same can be said regarding the spaces constructed while relying on the Normalized Euclidean distance measure (see Fig. 5.4). Conversely, a quick look at Fig. 5.6 (which depicts the obtained spaces when the Normalized Canberra distance is exploited), reveals heavy intersection among the classes when GEP is explored. The discrimination power of the space constructed in this scenario can therefore be expected to be quite low. This assumption is confirmed when inspecting the results presented in Table 5.3, where a poor maximum-discrimination-power of 50% was obtained for this experiment.

Nevertheless, the GEP solution will generally always be preferred when the space constructed is similar in structure to that constructed using FR (e.g., this is the case with the haptic-based signature dataset when the Normalized City Block distance is used, and somewhat so when the Normalized Euclidean and Bray Curtis distance measures are exploited). This is due to the fact that GEP provides explicit solutions. For example, the explicit GEP solution when Normalized Euclidean distances are considered is presented in Eq. (5.16).

$$
\begin{aligned}
\varphi_X &= 7.992 + sin(v_{4853}) + 0.67264 \\
\varphi_Y &= sin(sin(sin(sin(sin(sin(v_{2773})))))) + \\
&\quad sin(sin(v_{8588})) + 1.05797 \\
\varphi_Z &= 7.69287 + sin(sin(v_{1166})) + \\
&\quad sin((((((((v_{7680} + v_{7340}) + (v_{1938} + v_{2339})) - \\
&\quad (v_{6191} - v_{202})) + ((v_{2227} + v_{8997}) + v_{7954})) + \\
&\quad (v_{6667} + v_{8342})) * 0.084373)).
\end{aligned}
\tag{5.16}
$$

Similarly, the explicit GEP solution when Normalized City Block distances are explored is presented in Eq. (5.17).

$$\begin{aligned} \varphi_X &= sin(v_{6061}) + 15.68325 \\ \varphi_Y &= 14.69131 + sin(sin((v_{1646} * sin(sin((v_{3353} - \\ &\quad v_{441}))))) \\ \varphi_Z &= 0.91044 + sin(v_{1541}) + sin(v_{443}) \end{aligned} \tag{5.17}$$

The example mapping functions presented in Eqs. (5.16) and (5.17) reveal that many of the original 10000 variables (e.g., v_{4853} denotes the $4853th$ feature of the original 10000) are disregarded suggesting that this technique may help in identifying very relevant attributes (the arguments of the space transformation equations). Specifically, Eqs. (5.16) and (5.17) define different solutions for φ, the parameter which corresponds to a mapping from the original space V_m to the constructed 3D visual space $\hat{V}_m$ as previously illustrated in Fig. 5.1. In particular, $\varphi = (\varphi_x, \varphi_y, \varphi_z)$, where φ_x, φ_y, φ_z, corresponds to the explicit equations that map the high dimensional space V_m to the x, y, and z dimensions of the 3D visual space $\hat{V}_w$, respectively.

Additionally, it is interesting to observe from Table 5.4 and Fig. 5.7 that despite it's linear nature, PCA can in certain cases outperform (nonlinear) FR and GEP-based feature generation methods. Specifically, a comparison of Fig. 5.3 (a) and (b) (which correspond to the best mapping results obtained when FR and GEP are exploited, and while using the Normalized City Block distance measure) with Fig. 5.7 shows that a similar structure is in fact observed when nonlinear FR/GEP-based mapping and linear PCA are considered respectively. Nevertheless, from Tables 5.3 and 5.4 it can be seen that nonlinear feature generation can in fact outperform linear PCA when FR-based optimization (and the Normalized City Block distance measure) is considered. This is demonstrated by comparing the 5-fold CV results of the generated features

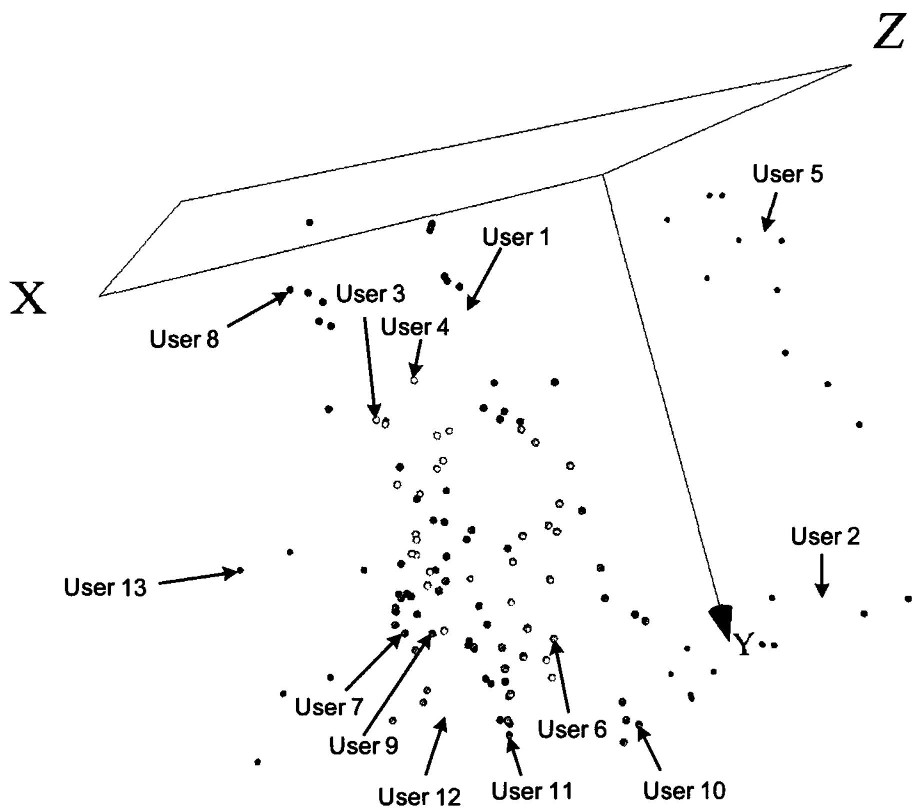
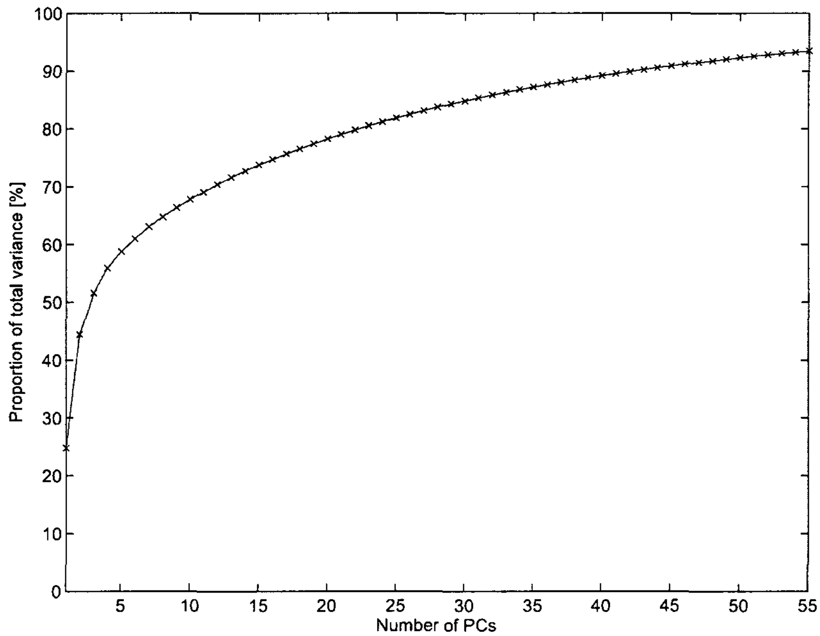


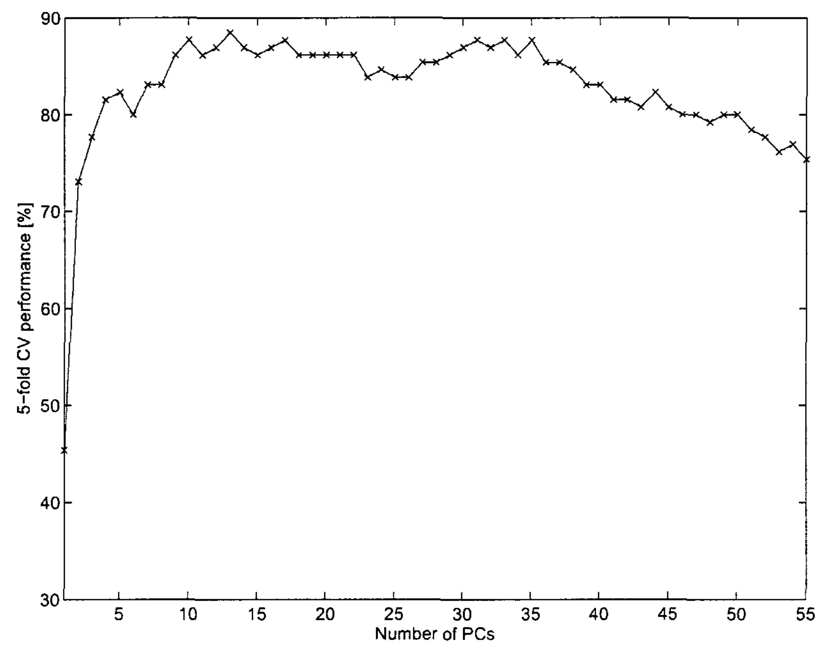
Figure 5.7: 3D VR spaces of the haptic-based handwritten signatures dataset constructed using PCA.

obtained using the corresponding methods. In order to have a rough estimate of the information loss incurred by any of the suggested dimensionality reduction procedures, a comparison between the 5-fold CV result achieved with the original set (10000 features), and that obtained with the generated feature set (3 features) can be performed. For example, from Table 5.4, it can be observed that when a PCA-generated dataset (3 features) is considered, a 5-fold CV of 77.7% is achieved using an SVM classifier, whereas a 92.3% 5-fold CV performance is obtained with the entire original dataset (10000 features).

More interestingly, it was observed that the first three PCs account for 51.6% of the total variance in the data. Two plots are shown in Fig. 5.8 that better illustrate the relationship between the proportion of the total variance and the 5-fold CV performance (in this case an SVM classifier is used) with an increase of the number of PCs. The proportion of total variance is



(a)



(b)

Figure 5.8: Plots showing the relationship between (a) the proportion of the total variance in the data, and (b) the 5-fold CV performance (in this case an SVM classifier is used), with respect to the number of PCs selected.

a simple and approximate criterion for estimating the number of nontrivial principal components. Naturally, the higher the number of PCs selected, the greater the proportion of the total variance that would be accounted for in the data. However, the selection of a greater number of PCs (and account for a higher proportion of the total variance) will not necessarily result in features that better represent and discriminate patterns in the data. This can be seen in Fig. 5.8 where a CV performance clearly decreases when the number of selected PCs is greater than 35 (despite an increase of the proportion of the total variance). Moreover, it can be observed that with the considered haptic dataset, a peak CV of 88.4% is achieved when the first 13 PCs are selected, and which comprise 71.6% of the total variance in the data.

Another interesting observation from the results presented in Table 5.4 is that the discrimination power of the original 10000 features using k-NN is 80.0%; however, when considering only the three features constructed using the nonlinear FR feature generation approach (and the normalized city block distance measure) the resulting discrimination rate is higher (81.5%). This is possibly due to a well-known drawback of nearest neighbors methods: their sensitivity to irrelevant features [81]. Thus, in the high-dimensional haptic dataset of 10000 features, there are potentially many irrelevant features that contributed to the significant degradation in performance of the k-NN classifier.

In addition, with regard to human haptic behavior and the underlying structure of different signatures, several conclusions can be drawn. For example, from Figs. 5.3 – 5.7, it can be seen that convex hulls of users 2 and 5 are consistently well separated from all other envelopes, with relatively no intersection. This suggests that users 2 and 5 have very distinct haptic features and their signature should therefore be easily distinguished from that of all other users. Furthermore, from Figs. 5.3 – 5.5 and 5.7, it can be observed that user 8 has a convex hull with objects that are quite dispersed, suggesting that this participant's haptic signature varies quite significantly between trials. Without going into further details, it should now be apparent that the generated 3D spaces (using FR or GEP) can provide an accurate visual representation of the underlying structure of haptic signatures that can be analyzed with great ease.

Overall, for all suggested feature generation methods, whether linear PCA or a nonlinear FR/GEP-based approach, feature reduction of very high-dimensional haptic datasets to only three dimensions proved to be indeed feasible. Despite some apparent loss of information (seen from the cross-validation and the Sammon error results) introduced in the mapping procedure, all three methods generated relatively accurate 3D spaces that can enable easy visual analysis and interpretation of haptic datasets.

5.7 Summary

In this chapter, data reduction through feature generation is explored to enable visual exploration of patterns in very high-dimensional haptic datasets. In particular, the considered datasets consist of haptic-based handwritten signatures. This was achieved through the unsupervised construction of virtual reality spaces using three very different approaches: linear feature generation using PCA, and nonlinear feature generation using classical optimization and genetic programming techniques. Additionally, different distance functions were evaluated (within the nonlinear mapping procedures) to investigate whether the choice of measure affects the visual representation and discrimination power of the constructed spaces. It was observed that with respect to the mapping error, classical optimization outperforms a genetic programming approach by approximately a factor of 2. Similarly, the potential discrimination power of spaces constructed using a classical optimizer is higher than that of spaces obtained using genetic programming. Nevertheless, it was revealed that in most cases, the relationship between the haptic data objects and their classes can be visually appreciated in all of the computed spaces regardless of the mapping error. Finally, it can be safely stated that a genetic programming-based solution will generally always be preferred when the space constructed is visually similar (in structure) to that constructed using PCA or classical optimization. This is due to the fact that GP provides explicit solutions that offer several advantages, such as, the ability to explore the source and role of newly constructed haptic data attributes by examining the corresponding mappings functions, or more specifically, to help identify relevant attributes in haptic datasets such as in haptic-based handwritten signatures.

Chapter 6

Data Reduction based on Feature Selection

6.1 Introduction

This chapter builds on the previous chapter and introduces alternative methods that can be exploited to reduce the high dimensionality of haptic datasets. Specifically, the focus of this chapter is primarily on data reduction in haptic datasets through *feature selection*. *Feature selection* refers to the process of identifying an optimum subset of features from a much larger set of potentially useful features [12]. Reducing the size of a large dataset offers numerous well-known advantages. These benefits are especially apparent in a knowledge discovery system in which a classification model is built to distinguish between possible patterns in the data. With this in mind, probably the most important advantage is the fact that feature reduction can help alleviate the "curse of dimensionality" problem. Furthermore, excluding noisy, redundant, and irrelevant features could potentially improve the accuracy and generalization properties of the classification model. In addition, a smaller subset of features reduces the space complexity of the classification problem, consequently decreasing the computational overhead, and speeding up the learning process.

Feature selection techniques can be divided into three different categories: Filters, wrappers, and embedded methods [82, 83]. Filters select subsets of features as a preprocessing step, independently of the learning algorithm. Wrappers, rely on a learning algorithm to identify and select relevant subsets of features according to their predictive power. Embedded methods incorporate feature selection as part of the model fitting/training procedure and, therefore, are generally specific to the classification algorithm.

Feature selection algorithms of the wrapper and embedded models can usually select features that result in higher learning performance for a particular learning algorithm. However, wrapper methods are computationally very expensive. Furthermore, an important limitation of embedded methods stems from their definition as they are designed to only suit a particular classification algorithm. An advantage of the filter method lies in the fact that it is independent of the learning algorithm, therefore, no bias is introduced with any classification models. Another advantage of the filter model is that it permits the algorithms to have a very simple structure. In turn, the algorithms are easy to design, and are usually very computationally efficient. This chapter discusses primarily feature selection methods that adapt the filter model.

Feature selection in high-dimensional haptic data has been little explored in the literature. As discussed in Chapter 2, the most closely related works are presented by Orozco *et al.* in [10, 11] where feature selection is performed in a haptic-based signature verification and identification problem using two different approaches: (1) relying on a relative entropy measure, which can be interpreted as a pseudo-metric that defines the distance between an individual's features and those of the entire population of users, and (2) using a Principal Component Analysis (PCA)-based approach to reduce the dimensionality of the data matrix while preserving relevant information in the haptic datasets. The authors, however, preprocess the haptic datasets in such a manner that possibly reduces the effectiveness of the pattern classification procedure which can result in less than expected (or potentially achievable) user classification (verification and recognition) rates. In particular, the time-varying features of each signature are distributed across different objects, i.e., velocity, orientation, torque, and force data acquired at time t_1 are assigned to object $Inst_1$, data acquired at time t_2 are assigned to object $Inst_2$, etc., yielding a number of objects per user signature with only few features per instance. For example, a general feature vector $\mathbf{r}$, which the authors refer to as a state vector, is defined as $\mathbf{r} = (v_x, v_y, v_z, f_x, f_y, f_z, t_x, t_y, t_z, \theta)$, where v, f, t, and θ denote velocity, force, torque, and orientation data respectively. This leads to a data matrix that's relatively low-dimensional, however, massive in terms of the number of objects. An alternative, and possibly a more adequate approach to defining a haptic feature vector would be to assign all the generated haptic data features per signature to a single object, i.e., each object contains the entire (haptic-based) signature for a single user. This approach can possibly lead to better analysis and interpretation of a haptic dataset, and potentially close-to-optimum user verification and identification rates (that may be achievable with haptic information). This, however, comes at the expense of having to work with and analyze data matrices that contain a fewer number, yet very high-dimensional objects that consist of thousands, and possibly (depending on the haptic application) tens of thousands of features.

This chapter investigates the potential use of feature selection methods in the reduction of high-dimensional haptic datasets. In particular, it explores the use of rough set theory for feature selection in large haptic datasets, where the data in this case (as in Chapter 5) consist of a collection of haptic-based handwritten signatures. Two rough set-based methods for feature selection are analyzed, the first is a greedy approach while the second relies on genetic algorithms (GA) to find minimal subsets of features. Several other contributions have been incorporated into this work. A method is proposed where feature vectors are subsampled prior to the feature selection procedure. This is performed to further reduce the haptic feature space while maximizing discrimination between users. Furthermore, a comparison between rough set-based methods and classical machine learning techniques in the selection of small information-preserving subsets of features in high-dimensional haptic datasets, is provided. The criteria for comparison are the length of the selected subsets of features and their corresponding discrimination power. In addition, minimal subsets generated using rough set-based approaches are also exploited to give some insight on the importance of different haptic data types (e.g., force, position, torque, and orientation) in a haptic-based handwritten signature identification application, or more precisely, in discriminating between different haptic users.

The rest of the chapter is organized as follows. In Section 6.2 an introduction to the basic concepts of the rough set theory is provided, including data reduct computation and discretization algorithms. In Section 6.3 the data preprocessing steps are described. In Section 6.4 the results are presented and analyzed. Finally, a summary of the chapter is provided in Section 6.5.

6.2 Concepts and Methods

In this section basic concepts of the rough set theory are introduced, followed with a concise description of several discretization and reduct computation algorithms, which together, are used in the feature selection procedure.

6.2.1 Rough Set Theory

The rough set theory originally introduced by Pawlak [84] is a mathematical approach to deal with vagueness and uncertainty in data analysis. The theory is based on the assumption that some information is associated with every object of the universe of discourse. In contrast to crisp sets, rough sets cannot be uniquely defined in terms of their elements. Objects characterized by the same information are considered indiscernible. The *indiscernibility* relation constructed in this manner is the mathematical foundation of rough set theory. The purpose is to approximate a *rough* concept using a pair of crisp sets, determined through the *indiscernibility* relation. The crisp sets are called the lower and upper approximations of the *rough* concept. The lower approximation consists of all objects that certainly belong to the set and the upper approximation contains all objects that possibly belong to the set. The *boundary region* is defined as the difference between the upper and the lower approximation, and consists of objects that cannot decisively be labeled as members or non-members of the set. A formal definition of the above concepts and others required for a better understanding of this chapter are given below.

Information System An information system is a pair $\mathbf{A} = (U, A)$ where U is a non-empty finite set of objects and A is a non-empty finite set of attributes such that $a : U \rightarrow V_a$ for every $a \in A$, where V_a corresponds to the value set of a. An important subclass of an information system is called a *decision table*. It consists of any information system of the form $\mathbf{A} = (U, A \cup \{d\})$, where $d \in A$ is the decision attribute and the elements of A are the condition attributes. In particular, the decision attribute determines the partition of U into k disjoint classes X_1, $X_2, \ldots, X_k$ called decision classes, where k denotes the number of different values assigned to attribute d.

Discernibility Matrices A *discernibility matrix* M_A is defined as an $|U| \times |U|$ matrix of the information system $\mathbf{A}$. Each entry $M_A(x, y) \subseteq A$ consists of the set of attributes that can be exploited to discern between objects $x, y \in U$ as in Eqs. (6.1) and (6.2).

$$M_A(x, y) = \{a \in A : discerns(a, x, y)\} \tag{6.1}$$

$$discerns(a, x, y) \Leftrightarrow a(x) \neq a(y) \tag{6.2}$$

where, $discerns(a, x, y)$ may be tailored to the application of interest.

Indiscernibility Relations A discernibility matrix M_A defines a binary relation $R_A \subseteq U^2$. The relation R_A is referred to as an *indiscernibility relation* [85] (see Eq. (6.3)) with respect to A, and expresses which pairs of objects cannot be distinguished from each other.

$$xR_Ay \Leftrightarrow M_A(x, y) = \emptyset \tag{6.3}$$

Implicants It has been described [85] that an m-variable function $f : \mathbf{B}^m \rightarrow \mathbf{B}$, where $\mathbf{B} = \{0, 1\}$, is called a *Boolean function* if and only if it can be expressed by a Boolean formula. An *implicant* of a Boolean function f is a product term p that is included in the function f (i.e., $p \leqslant f$, where $\leqslant$ is a partial order called the inclusion relation). For example, both xy' and xyz are implicants of $xy' + yz$. A *prime implicant* is an implicant of f that ceases to be so if any of its literals are omitted. Consequently, it can be seen that if a term $p \leqslant f$ is not prime, then it is possible to obtain another implicant of f by simply removing one of the literals of p.

Discernibility Functions A discernibility function [85] defines how an object (or a set of objects) can be discerned from a subset of the full universe of objects. Specifically, discernibility functions can be defined relative to a particular object, or constructed such that all objects can be discerned from each other. In the former case, the function is constructed relative to an object $x \in U$ from a discernibility matrix M_A according to Eq. (6.4).

$$f_A(x) = \prod_{y \in U} \left\{ \sum a^* : a \in M_A(x,y) \text{ and } M_A(x,y) \neq \emptyset \right\} \tag{6.4}$$

The function $f_A(x)$ contains $|A|$ Boolean variables, where variable a^* corresponds to the attribute a. Each conjunction of $f_A(x)$ stems from an object $y \in U$ from which x can be discerned and each term within that conjunction represents an attribute that discerns between those objects. The prime implicants of $f_A(x)$ provide the minimal subsets of A that are required to discern object x from the objects in U that are not associated with $R_A(x)$. Conversely, in order to discern all objects in U from each other the full discernibility function $g_A(U)$ is defined as follows:

$$g_A(U) = \prod_{x \in U} f_A(x). \tag{6.5}$$

Reducts If an attribute subset $B \subseteq A$ preserves the indiscernibility relation R_A then the attributes $A \backslash B$ are said to be *dispensable*. An information system can have several such attribute subsets B. All such subsets that do not contain any dispensable attributes (i.e., minimal subsets), are called reducts. However, the problem of computing all reducts (or a minimum-length reduct) is NP-hard, and several heuristics have been proposed [86]. This issue is discussed further in subsequent sections.

6.2.2 Discretization

Discretization is a preprocessing step that most symbolically based data mining methods require in order to provide adequate performance. It essentially corresponds to defining a coarser view of the world by making the attributes value sets smaller. The transformation is performed in such a manner that ranges or intervals are used instead of the exact data themselves. The observations can then be treated as qualitative rather than quantitative information.

- Naïve Scaler

  The naive algorithm is a basic heuristic that may result in a greater number of cuts (that determine the intervals) than are necessary. Its description makes the assumption that all condition attributes A are numerical. For every condition attribute a, its value set V_a can be sorted to obtain the following ordering:

$$v_a^1 < \ldots < v_a^i < \ldots < v_a^{|V_a|} \tag{6.6}$$

  Then let C_a denote the set of all determined cuts for attribute a, defined as depicted in Eqs. (6.7) – (6.9).

$$X_a^i = \{x \in U \mid a(x) = v_a^i\} \tag{6.7}$$

$$\Delta_a^i = \{v \in V_d \mid \exists x \in X_a^i \text{ such that } d(x) = v\} \tag{6.8}$$

$$C_a = \left\{ \frac{v_a^i + v_a^{i+1}}{2} \mid \ |\Delta_a^i| > 1 \text{ or } |\Delta_a^{i+1}| > 1 \text{ or } \Delta_a^i \neq \Delta_a^{i+1} \right\} \tag{6.9}$$

The set C_a corresponds to all cuts midway between two observed values, with the exception of those that are not needed because the corresponding objects have the same decision value. The implementation of this discretization technique leaves an attribute unprocessed in the case where no cuts are found [85]. Additionally, missing values are ignored during the search for cuts.

- Boolean Reasoning-based Orthogonal Scaler

This algorithm discussed in [85, 87], is based on the combination of the aforementioned Naïve Scaler algorithm and a Boolean reasoning procedure for discarding all but a small subset of the generated cuts. The algorithm is initiated by first creating a Boolean function f (using Eq. (6.10)) from the set of candidate cuts computed using Eq. (6.9).

$$f = \prod_{(x,y)} \sum_{a} \Big\{ \sum c^* \mid c \in C_a \text{ and } a(x) < c < a(y) \text{ and } \partial_{A(x)} \neq \partial_{A(y)} \Big\}. \tag{6.10}$$

where ∂_A is a *generalized decision attribute*, defined as a function $\partial_A : U \to V_d$ (V_d is assumed to be a set of integers $\{1, \ldots, r(d)\}$, where $r(d)$ is said to be the rank of the decision attribute d) that, when applied to object x, expresses the set of values that the objects in the indiscernibility set $R_A(x)$ take on for the decision attribute d. This can be formulated mathematically as follows:

$$\partial_A(x) = \{\text{v} \in V_d \mid \exists\, y \in R_A(x) \text{ such that } d(y) = \text{v}\} \tag{6.11}$$

The prime implicant of f is then computed using Johnson's greedy algorithm [88], described in Section 6.2.3. However, in certain cases, this discretization technique may result in no cuts being deemed necessary for some attributes. This occurs when such attributes are not needed to preserve discernibility. The implementation (of this algorithm) exploited in this work leaves such attributes untouched [85].

- Equal Frequency Scaler

  In order to discretize attributes that may have been left untouched when the Boolean Reasoning-based Orthogonal Scaler algorithm is used, equal frequency binning can be performed, which is a simple unsupervised and univariate discretization technique. The procedure consists of fixing a number of intervals n and subsequently examining the histogram of each attribute. As a result, $n - 1$ cuts are determined so that approximately the same number of objects fall into each of the n intervals.

6.2.3 Reduct Computation

Two different algorithms to compute reducts or reduct approximations are exploited in this work. Reduct approximations generate attribute subsets that in a sense "almost" preserve the indiscernibility relation. As minimum-length reduct computation (which is of great interest in feature selection) is NP-hard, reduct approximation can therefore be found to be necessary especially when the number of attributes $|A|$ is very high.

- Johnson Reducer

  Johnson's algorithm is a polynomial-time heuristic approach that invokes a variation of a simple greedy algorithm to compute a single reduct [85,88]. The algorithm has a natural tendency towards finding a single approximate prime implicant of minimal length. A step by step illustration of the algorithm is depicted below, where B corresponds to the computed reduct, and $\mathcal{S}$ denotes the set of sets corresponding to the discernibility function. In addition, $w(S)$ denotes a weight for set S in $\mathcal{S}$ that is automatically computed from the data [85].

  1. Let $B = \emptyset$

  2. Let a denote the attribute that maximizes $\sum w(S)$, where the sum is taken over all sets S in $\mathcal{S}$ that contain a.

3. Add a to B.
4. Remove all sets S from $\mathcal{S}$ that contain a.
5. If $\mathcal{S} = \emptyset$ return B. Otherwise, goto step 2.

Computation of approximate reducts is performed by aborting the loop when "enough" sets have been removed from $\mathcal{S}$. This approach is computationally more efficient than exiting the loop only when $\mathcal{S}$ is entirely emptied.

- Genetic Reducer

This method relies on a genetic algorithm to compute minimal hitting sets (reducts), as depicted in [89, 90]. The technique has support for both cost information and approximate solutions. Moreover, the algorithm's fitness function f is defined as in Eqs. (6.12) and (6.13):

$$f(B) = (1-\alpha)\times\frac{cost(A)-cost(B)}{cost(A)}+\alpha\times\min\{\varepsilon, h(B)\}, \tag{6.12}$$

$$h(B) = \frac{|\,[S \text{ in } \mathcal{S} \mid S \cap B \neq \emptyset]\,|}{|\,\mathcal{S}\,|} \tag{6.13}$$

The hitting fraction $h(B)$ is defined as the ratio of the number of nonempty intersection of the subset B with $\mathcal{S}$ over the multiset's ($\mathcal{S}$) cardinality, where $\mathcal{S}$ is the set of sets that correspond to the discernibility function. The parameter α defines a weighting between subset cost and hitting fraction $h(B)$, while ε controls approximate solutions as it represents a minimal value for the hitting fraction. The function $cost(\cdot)$ specifies the cost of an attribute subset. Moreover, subsets B of A are found through the evolutionary search driven by the fitness function, where sets that have a hitting fraction of at least ε, are collected in a "keep list". The genetic operators crossover, mutation, and inversion are used in the algorithm, whereas selection is done by universal stochastic sampling.

6.3 Haptic Dataset and Preprocessing Steps

The exploited high-dimensional haptic dataset selected is the same as that of Chapter 5. It essentially consists of haptic-based handwritten signatures recorded from 13 different participants, where 10 signatures were collected per individual. The data preprocessing steps were performed as illustrated in Section 5.4.2, where all computed one-dimensional feature vectors are normalized to a common length of 10000. This feature vector length was selected in such a manner to minimize the information loss that can occur during the size normalization process, which is most apparent when downsampling is performed. Consequently, the computed preprocessed dataset contains 130 instances, where each consists of 10000 features. Feature selection is then performed using the discretization/reduct computation strategy discussed in Sections 6.2.2 and 6.2.3, where the preprocessed dataset is initially discretized using the *Boolean Reasoning-based Orthogonal Scaler* and the *Equal Frequency Scaler* algorithms. As mentioned in Section 6.2.2, *Equal Frequency Scaler* is only used to discretize attributes that may have been left untouched by the *Boolean Reasoning-based Orthogonal Scaler* algorithm. Reduct computation is then performed using two different algorithms: *Johnson Reducer* and the *Genetic Reducer*.

6.4 Results

The presented results aim to analyze the performance of rough set-based feature selection in high-dimensional haptic signatures (used for user identification). This is performed while comparing the results obtained using rough set-based feature selection algorithms with those acquired using two other methodologies: (1) Feature reduction through haptic data subsampling, and (2) feature selection using classical machine learning-based techniques. In the former case, the time-varying and low-frequency nature of hand movements (when signing a check) is exploited to reduce the number of features. Specifically, the normalized haptic signatures discussed in Section 6.3 (10000 features/object) are further subsampled to reduce the number of features while possibly retaining the discrimination power in the data, despite some informa-

Table 6.1: Feature reduction based on data subsampling.

No. of features per dataset	*5-fold CV on all objects*	*Train 60%, Test 40%*		*Train 80%, Test 20%*	
		5-fold CV (%)	*Classif. (%)*	*5-fold CV (%)*	*Classif. (%)*
10000	90.8	87.2	88.5	88.5	92.3
5000	91.5	87.2	88.5	88.5	96.2
2500	92.3	87.2	88.5	88.5	96.2
2000	91.5	87.2	88.5	88.5	96.2
1000	90.8	87.2	88.5	88.5	96.2
500	90.8	87.2	86.5	88.5	96.2
250	90.8	85.9	**88.5**	87.5	96.2
200	88.5	88.5	84.6	85.6	92.3
150	86.2	85.9	82.7	83.7	92.3
100	86.2	83.3	80.8	84.6	88.5
80	86.9	73.1	80.8	83.7	88.5
50	80.0	75.6	84.6	75.9	92.3
20	74.6	73.1	**67.3**	72.1	80.8
10	67.7	67.9	61.5	72.1	69.2

tion loss. In the latter case, two different machine learning-based feature selection techniques are evaluated: *Best First*, and *Subset Size Forward Feature Selection*. A brief description of each method is later depicted in Section 6.4.5. In all cases, SVM classifiers are exploited to evaluate the accuracy of the selected minimal feature vectors.

6.4.1 Feature Reduction based on Haptic Data Subsampling

The results obtained using the *haptic data subsampling* feature reduction strategy are illustrated in Table 6.1. Essentially, the original feature vectors are reduced in length using different subsampling factors (generated subsets range from 10000 features to as few as 10 features per object). Then, using an SVM classifier, a 5-fold cross-validation (CV) is initially performed on each of the generated subsets of features (on all objects). The purpose of cross-validation in this case is to provide us with an estimate of the discrimination power preserved in the reduced sets.

Table 6.2: Classification analysis of approximate reducts generated using Johnson's algorithm.

No. of Features	*Train 60%, Test 40%*			*Train 80%, Test 20%*		
	No. of SF	*5-fold CV (%)*	*Classif. (%)*	*No. of SF*	*5-fold CV (%)*	*Classif. (%)*
10000	5	78.2	76.9	5	66.3	65.4
5000	5	82.1	69.2	5	67.3	65.4
2500	**5**	**78.2**	**78.8**	5	68.3	80.8
2000	5	61.5	57.7	5	70.2	73.1
1000	5	61.5	55.8	5	62.5	53.8

In addition, for each subset, two distinct datasets are generated: one that divides the objects (haptic signatures) into 60% training and 40% test sets, whereas the other divides the objects into 80% training and 20% test sets. The 5-fold CV accuracy on the training data and the classification accuracy of test data is then computed as reported in Table 6.1. One interesting observation from the provided results (while inspecting the cross-validation and classification rates) is that subsets with as few as 1000 features seem to preserve the discrimination power achieved with the original haptic dataset (10000 features per object). Consequently, feature selection methods discussed in subsequent sections will be analyzed using the original dataset and generated subsets with a number of features ≥ 1000. This should give us some insights on the possibility of combining feature vector subsampling with feature selection methods in order to further reduce the haptic feature space while maximizing discrimination between users.

6.4.2 Feature Selection in Haptic Data using Johnson's Algorithm

The feature selection results achieved using the rough set theory-based Johnson algorithm are shown in Table 6.2 (columns labeled *No. of SF* describe the number of selected features). Reduct computation is performed on the original dataset (10000 features per object), as well as on undersampled versions of the data which include subsets containing 5000, 2500, 2000, and 1000 features per object. In addition, for each subset, two distinct datasets are generated:

Table 6.3: Classification analysis of approximate reducts generated using the Genetic Reducer algorithm (using default parameters).

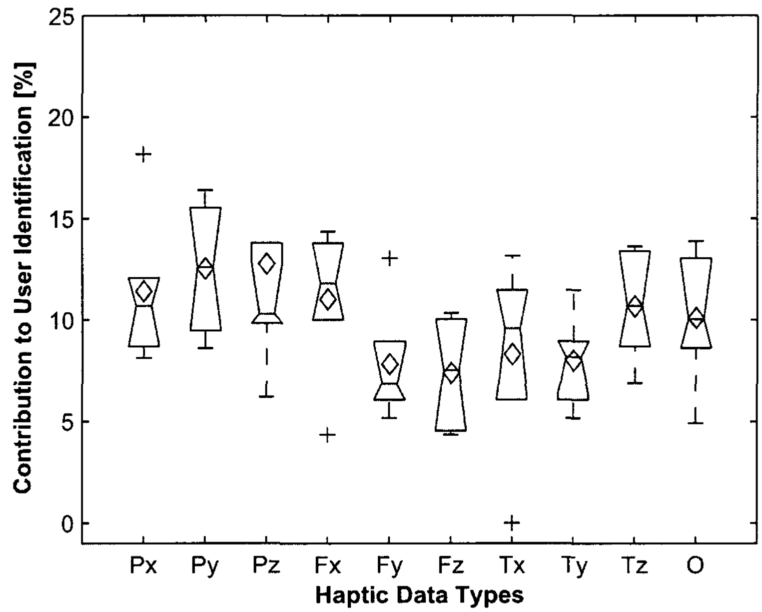
	Train 60%, Test 40%			*Train 80%, Test 20%*		
No. of Features	*No. of SF*	*5-fold CV (%)*	*Classif. (%)*	*No. of SF*	*5-fold CV (%)*	*Classif. (%)*
10000	209	84.6	86.5	294	89.4	92.3
	190	85.9	84.6	371	89.4	92.3
5000	**61**	**83.3**	**88.5**	71	87.5	88.5
	66	80.8	86.5	**136**	**90.4**	**96.2**
2500	23	83.3	76.9	28	80.8	88.5
	58	83.3	86.5	84	86.5	88.5
2000	35	78.2	86.5	30	80.8	88.5
	28	76.9	75.0	44	81.7	88.5
1000	16	78.2	65.4	9	76.9	76.9
	24	76.9	71.2	20	78.8	76.9

One that divides the objects (haptic signatures) into 60% training and 40% test sets, whereas the other divides the objects into 80% training and 20% test sets. The 5-fold CV accuracy on the training data and the classification accuracy of test data is then computed as reported in Table 6.2. Overall, approximate reducts (minimal subsets) obtained using Johnson's greedy algorithm did not demonstrate significant improvement (when inspecting the length of the reducts and the corresponding classification rates) over subsets generated when relying solely on feature vector subsampling. Nevertheless, in certain cases some improvements were apparent. For example, considering the case when reduct computation is performed on the 60% training set with 2500 features, a minimal subset of length 5 is obtained, with a classification rate of 78.8%. This suggests a feature reduction improvement by a factor of at least 4 in comparison to subsampling where even with a feature vector of length 20, the classification rate is a mere 67.3 (see Table 6.1 when 60% training data are considered).

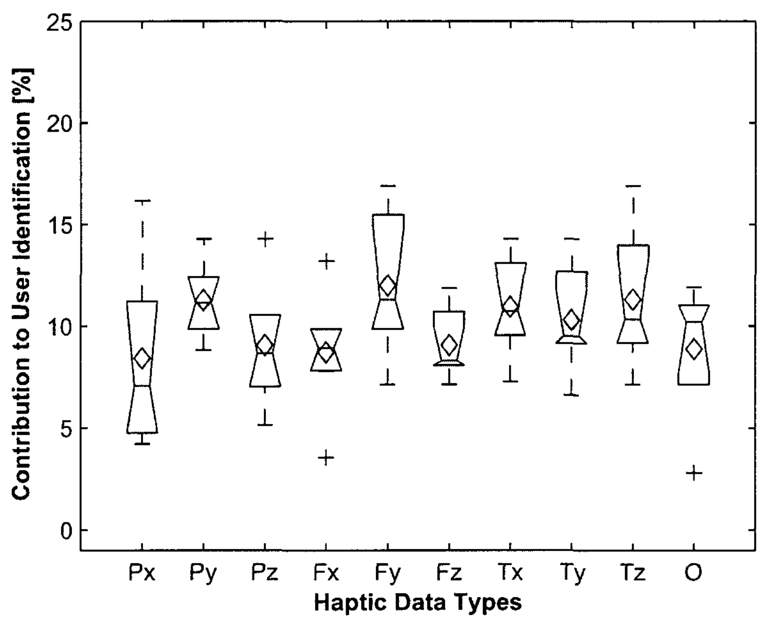
6.4.3 Feature Selection in Haptic Data using the Genetic Reducer Algorithm

Comparison between the Genetic Reducer Algorithm and Data Subsampling

The feature selection results achieved using the rough set theory-based Genetic Reducer technique are presented in Table 6.3. Minimal hitting sets (reducts) are computed using the algorithm's default parameters, i.e., $mutation\ probability = 0.05$, $crossover\ probability = 0.3$, $population\ size = 70$. In order to compute reducts within reasonable computation time, approximate solutions were sought with a $hitting\ fraction = 0.95$. As in Section 6.4.2, reduct computation is performed on the original dataset (10000 features per object), and on subsampled subsets ranging from 5000 to as few as 1000 features per object. Additionally, two versions of each subset is created, one that divides the objects into 60% training and 40% test sets, whereas the second divides the objects into 80% training and 20% test sets. For each version of every subset, 20 independent runs of the GA-based reduct computation procedure was conducted and the top 10% of the reducts obtained are selected (i.e., in this case, the top 2 reducts). Reducts are selected based on the best compromise between the following two criteria: (1) Maximum 5-fold CV accuracy on the training data and, (2) hitting set (reduct) is of minimal length. The top two reducts computed for each subset are reported in Table 6.3. A comparison between the results presented in Tables 6.1 and 6.3 reveals that approximate reducts computed using the Genetic Reducer algorithm led to significant feature reduction improvements over subsets generated using strictly feature vector subsampling. For example, considering the case when reduct computation is performed on the 60% training set with 5000 features, a minimal subset of length 61 is obtained, with a classification rate of 88.5%. This is a feature reduction improvement by a factor of approximately 4 in comparison to subsampling alone, where a classification rate of 88.5% is only obtained with subsampled feature vectors of length ≥ 250 (see Table 6.1 when 60% training data are considered).



(a)



(b)

Figure 6.1: Box plots illustrating the frequency of haptic data types found in subsets generated using the Genetic Reducer when (a) 60% training data, (b) 80% training data, are considered.

Table 6.4: Classification analysis of approximate reducts generated using the Genetic Reducer algorithm while varying the mutation and crossover probability parameters.

GA Parameters		*Train 60%, Test 40%*			*Train 80%, Test 20%*		
Crossover Probability	*Mutation Probability*	*No. of SF*	*5-fold CV (%)*	*Classif. (%)*	*No. of SF*	*5-fold CV (%)*	*Classif. (%)*
0.3	0.01	77	83.3	82.7	52	84.6	88.5
		65	82.1	84.6	97	89.4	92.3
0.3	0.03	35	82.1	76.9	38	86.5	88.5
		49	82.1	84.6	96	85.6	92.3
0.3	0.05	23	83.3	76.9	28	80.8	88.5
		58	83.3	86.5	84	86.5	88.5
0.5	0.01	50	83.3	82.7	**37**	**86.5**	**96.2**
		90	85.9	88.5	64	87.5	88.5
0.5	0.03	19	74.4	75.0	17	79.8	88.5
		24	80.8	75.0	23	81.7	76.9
0.5	0.05	17	82.1	75.0	16	75.0	84.6
		22	76.9	76.9	31	84.6	80.8
0.7	0.01	27	78.2	75.0	46	81.7	92.3
		44	79.5	80.8	51	83.7	92.3
0.7	0.03	18	78.2	76.9	13	80.8	84.6
		29	79.5	76.9	16	76.0	76.9
0.7	0.05	9	75.6	71.2	18	77.9	80.8
		28	76.9	75.0	31	83.7	88.5

Relevance of Haptic Data Types in User Identification

Feature selection using a rough set-based approach is also exploited to gain some insight on the importance of different haptic data types in discriminating between different users. As aforementioned, haptic data types include three-dimensional force, position, torque, and orientation (exceptionally, orientation is one-dimensional and consists of the haptic device's tracker angle). In Fig. 6.1, box plots are provided illustrating the frequency of different haptic data types in subsets generated using the Genetic Reducer algorithm when 60% training data (Fig. 6.1 (a)), and 80% training data (Fig. 6.1 (b)), are considered. Each box has lines at the lower quartile, the median, and the upper quartile values (mean values are identified with the diamond-shaped marker symbol). Lines extending from each end of the box (whiskers) are provided to show the extent of the frequency of different haptic data types found in the obtained subsets. Outliers are represented using the '+' symbol and consist of frequency values that go beyond the ends of the whiskers. The reducts considered are those presented in Table 6.3, and that are associated with subsets ranging from 10000 to 2500 features. Subsets with 2000 and 1000 features per object were not considered due to their low CV performance on the training data. It can be observed that overall, haptic data types are approximately of equal importance when exploited for user identification. Specifically, position information (P_x, P_y, and P_z) appear approximately as frequently as haptic-specific data types, i.e., force, torque, and orientation (F_x, F_y, F_z, T_x, T_y, T_z, and O) in the generated reducts. This is an interesting result as the apparent importance of haptic-specific data types suggests that haptic-based handwritten signatures can potentially outperform traditional handwritten signatures (in user identification applications) as the latter rely solely on position information.

6.4.4 An In-Depth Analysis of the Genetic Reducer Algorithm

In order to gain better insight into the reduct computation capabilities of the Genetic Reducer method, the experiments presented in Section 6.4.3 were repeated while varying the most fundamental genetic parameters of the algorithm. Specifically, the *crossover probability* and *mutation probability* parameters were assigned values ranging from 0.3 to 0.7 and 0.01 to 0.05 respectively as shown in Table 6.4. Moreover, to avoid much redundancy in the presented results, the analysis was performed solely on the (subsampled) subset with 2500 features per object. In fact, in Table 6.1, it is observed that subsets with 2500 not only preserve the discrimination power of the original datasets, but also achieve a better CV result when all the objects are considered (this may be due to the fact that some possibly redundant and noisy features were omitted during the subsampling procedure). Two versions of the subset were created, one that divides the objects into 60% training and 40% test sets, whereas the second

Table 6.5: Classification analysis of feature subsets generated using the Subset Size Forward Selection algorithm.

Original Attributes	***Subset Size Forward Selection***					
	Train 60%, Test 40%			***Train 80%, Test 20%***		
	No. of SF	***5-fold CV (%)***	***Classif. (%)***	***No. of SF***	***5-fold CV (%)***	***Classif. (%)***
10000	88	91.0	86.5	80	91.3	92.3
5000	79	89.7	84.6	79	92.3	92.3
2500	67	93.6	78.8	**63**	**93.3**	**96.2**
2000	58	93.6	82.7	63	91.3	96.2
1000	47	93.6	80.8	50	88.5	80.8

divides the objects into 80% training and 20% test sets. Similar to Section 6.4.3, for every *crossover probability* and *mutation probability* value assignment, 20 independent runs of the GA-based reduct computation procedure was conducted and only the top two reducts were selected (see Table 6.4). A first inspection of the results reveals that reduct computation with higher crossover probabilities resulted in hitting sets that are shorter in length, but only at the expense of classification accuracy. This observation is more pronounced when 60% training sets are considered. Furthermore, it can be observed that generally, a higher classification accuracy is found with reducts generated using a low mutation probability of 0.01; this is, however, at the expense of obtaining approximate hitting sets that are longer in length (greater number of selected attributes). This observation, however, is more pronounced when 80% training sets are considered.

6.4.5 Feature Selection using the Rough Set Theory versus Classical Machine Learning-based Algorithms

In this section a comparison between rough set-based methods and classical machine learning techniques in the selection of minimal information-preserving subsets of features in high-dimensional haptic datasets is presented. The machine learning methods were evaluated using

Table 6.6: Classification analysis of feature subsets generated using the Best First algorithm.

Original Attributes	***Best First***					
	Train 60%, Test 40%			***Train 80%, Test 20%***		
	No. of SF	***5-fold CV (%)***	***Classif. (%)***	***No. of SF***	***5-F CV (%)***	***Classif. (%)***
10000	153	93.6	82.7	158	92.3	84.6
5000	114	92.3	86.5	130	93.3	84.6
2500	**103**	**92.3**	**88.5**	103	90.4	88.5
2000	91	91.0	82.7	95	92.3	92.3
1000	71	91.0	84.6	72	93.3	92.3

the Weka environment [91], a public data mining library in Java for feature selection and classification. Weka contains several methods for searching through the space of feature subsets, as well as evaluation measures for features and feature subsets. Different search and evaluation methods can be combined as desired by the user. In this work, two search methods capable of handling high-dimensional datasets within reasonable computation time are considered: *Best First, and Subset Size Forward Selection*. *Best First* searches the space of attribute subsets by greedy hill-climbing augmented with a backtracking facility. Setting the number of consecutive non-improving nodes allowed controls the level of backtracking done. Conversely, *Subset Size Forward Selection* [92] is a recent algorithm that modifies the standard search method known as *forward selection* to yield a computationally efficient wrapper-based feature selection technique for high-dimensional data. Explicit subset size determination in forward selection is also achieved to combat overfitting, where the search is forced to stop at a pre-computed optimal subset size. In addition, *Correlation-based Feature Selection* (*CFS*) [93] was selected as the feature evaluation method (combined with the Best First and Subset Size Forward Selection search methods). *CFS* evaluates the worth of a subset of attributes by considering the individual predictive ability of each feature along with the degree of redundancy between them. Subsets of features that are highly correlated with the class while having low intercorrelation are preferred.

Similar to Sections 6.4.2 and 6.4.3, feature selection using the machine learning techniques is performed on the original dataset (10000 features per object), and on subsampled subsets ranging from 5000 to as few as 1000 features per object (see Tables 6.5 and 6.6). In addition, two versions of each subset is created, one that divides the objects into 60% training and 40% test sets, whereas the second divides the objects into 80% training and 20% test sets. First, considering the case when feature selection is performed on the 60% training sets, it can be observed that between the two techniques, *Best First* generated the subset with the highest classification rate. The minimal subset is of length 103, and achieves a classification rate of 88.5% on the test data. Conversely, the subset with the highest classification accuracy obtained with feature selection using the Genetic Reducer algorithm is a minimal approximate reduct of length 61, with a classification rate of 88.5% (see Table 6.3). This is a feature reduction improvement by a factor of approximately 1.7 (in comparison to the machine learning techniques). Furthermore, considering the case when feature selection is performed on the 80% training sets, it can be observed that between the two techniques, *Subset Size Forward Selection* generated the subset with the highest classification rate. Specifically, the minimal subset is of length 63, and achieves an impressive classification rate of 96.2% on the test data. Conversely, the subset with the highest classification accuracy obtained with feature selection using the Genetic Reducer algorithm is a minimal approximate reduct of length 136, with a classification rate of 96.2%. However, if the algorithm's genetic parameters are selected such that $crossover\ probability = 0.5$ and $mutation\ probability = 0.01$, approximate reducts with as few as 37 features, and a corresponding classification rate of 96.2% can be achieved (see Table 6.4). The latter comparison is, however, not entirely fair as results using machine learning techniques with default parameters are compared to a rough set-approach with adapted parameters. Nevertheless, the comparison is only intended to illustrate that rough-set based feature selection techniques can, in most cases, (possibly) potentially overcome machine learning techniques.

6.5 Summary

This chapter explores feature selection as a possible approach for reducing the high dimensionality of haptic datasets. Specifically, the use of rough set theory for feature selection in high-dimensional haptic data is investigated and applied to a haptic-based handwritten signatures dataset (exploited for user identification). Two rough set-based methods for feature selection are analyzed, the first is a greedy approach while the second relies on a genetic algorithm to find minimal subsets of attributes. In addition, to further reduce the haptic feature space while maximizing user identification accuracy, a method is proposed where feature vectors are subsampled prior to the feature selection procedure. Rough set-generated minimal subsets are initially exploited to gain some insight on the importance of different haptic data types (e.g., force, position, torque, and orientation) in discriminating between different users. In addition, a comparison between rough set-based methods and classical machine learning techniques in the selection of minimal information-preserving subsets of features in high-dimensional haptic datasets, is provided. The criteria for comparison are the length of the selected subsets of features and their corresponding discrimination power. Support vector machine classifiers are used to evaluate the accuracy of the selected minimal feature vectors. The results demonstrated that the combination of rough set and genetic algorithm techniques can potentially outperform well-established machine learning methods in the selection of minimal subsets of features present in high-dimensional haptic datasets, specifically, in haptic-based handwritten signatures.

Part IV

Data Protection in Multimodal Visual-Haptic Environments

Chapter 7

Multimodal Vision-Haptic Watermarking

7.1 Introduction

The increased processing capabilities of today's modern computers have contributed to the widespread use of 3D models in several fields, such as in entertainment (video games, movies), the medical industry (3D visualization of medical data, education, and training) and the Internet (3D virtual worlds, e-commerce). In turn, the protection of 3D models from theft or tampering has received much attention in the literature [16]. An effective measure that overcomes the limitations of traditional encryption is digital watermarking; a research field that deals with the process of embedding information into digital data in an inconspicuous manner to identify the origin, owner, use, rights, integrity, or destinations of multimedia content (e.g., digital images, video, audio and 3D models) [3, 13, 15–17]. The first requirement of digital watermarking techniques regardless of the addressed media or application is imperceptibility [15, 18, 19]. Imperceptibility refers to the perceptual similarity between the original and watermarked data. Ideally, the perceptual quality of the watermarked media must be identical to the original. Recently, considerable progress has been made in 3D watermarking, where the main focus has been on triangle meshes, which is the most common digital representation of 3D models due to its simplicity and usability. Existing watermarking techniques concerning 3D meshes host the watermark either by modifying the geometry or the connectivity of the surface (spatial domain) or by modifying some kind of spectral-like coefficients (spectral domain) [16]. Accordingly, the watermark's intrusiveness can be evaluated in terms of its visibility in the rendered version of the mesh. The evaluation of 3D watermarking algorithms against the imperceptibility constraint has been thus far exclusively based on the sensitivity of human vision to distortion. Moreover, currently available perceptual metrics generally used to assess the quality of watermarked 3D meshes have been validated solely through psychovisual experiments [18,19].

In recent years, the integration of haptics into immersive virtual environments to enable sensing and manipulation of virtual 3D objects through touch has become increasingly popular due to its many potential applications, including medical training, physical rehabilitation and entertainment [47,94,95]. In consequence, the recent advancements of haptic technology have raised several important questions in the 3D watermarking research community: Is the haptic sensory channel more sensitive than the visual sensory channel in detecting a watermark embedded in a haptic-enabled 3D object? Do watermarks inspected using multimodal feedback (vision + haptic) result in very different detection thresholds from those detected using a single sensory modality (touch-only or vision-only); Or more importantly, does visual feedback, when presented together with haptic feedback, improve the perception of a watermark embedded in a 3D mesh? If it is revealed that the lowest watermark detection thresholds are obtained under haptic-alone and/or haptic + vision conditions, as compared with vision-alone, then 3D watermarking researchers must carefully design (or redesign) their algorithms to ensure watermark imperceptibility when both visual and haptic feedback are considered.

In this chapter, a study is conducted to examine the role of multisensory feedback in the perception of a watermark embedded in a haptic-enabled 3D mesh. Specifically, we focus on whether the use of a bimodal approach, where a subject is looking at and touching a 3D mesh could in fact improve the perceptibility of a watermark as opposed to when only a single modality (vision-alone or haptic-alone) is available. The experiments are performed using a hapto-visual interface that enables users to see and touch virtual objects at the same location in space. This approach enables superior integration of vision and touch than a conventional 2D screen-based display. Moreover, the advantages of the use of a hapto-visual interface stem from the fact that human perception is generally multimodal, i.e., the senses of touch and vision do not independently operate, but are in fact closely coupled. A statistical analysis to test for significance between modalities is also presented. An analysis of variance (ANOVA) is conducted to statistically explore the impact of the considered modalities (vision, haptic, and vision-haptic) on the measured watermark detection thresholds across different resolutions of the underlying virtual 3D mesh.

The rest of this chapter is organized as follows: In Section 7.2 the related work is presented. In Section 7.3, the importance of multimodal vision-haptic watermarking is emphasized. In Section 7.4, the materials and methods of the conducted experiments are described in detail. In Section 7.5, the experimental results are presented and discussed. Finally, a summary of the chapter is provided in Section 7.6.

7.2 Related Work

In this section, previous works closely related to the present study are described in two parts. First, an overview is presented of similar experimental studies that attempt to assess whether multimodal conditions (haptic + vision) can produce very distinct percepts from those produced when a single sensory modality is available. The focus of this discussion is, however, limited to studies that explore the impact of multisensory vision-haptic feedback on the perception of surface roughness. This is due to the fact that surface distortions produced from the 3D watermarking process can be regarded as a particular form of surface roughness. The second part of this section discusses existing work in the literature that have investigated the area of haptic digital watermarking (this is a brief summary of a more detailed review presented in Section 2.3).

7.2.1 Multimodal Vision-Haptic Perception

A fair amount of research has been devoted to understanding how observers combine multisensory (vision-haptic) inputs to attain an integrated percept. For example, whether multisensory vision + haptic conditions can produce different percepts from those produced when a single sensory modality is present, and whether the observers' behavior is consistent across different perceptual tasks, e.g., visual-haptic stiffness [96, 97], size [97], and shape [98]. Of more particular interest, a number of studies have also investigated the role of multisensory feedback on the perception of surface roughness [99–105].

Earlier studies of roughness perception with real surfaces have shown that multimodal conditions can lead to very different percepts than when only single modalities are considered. For example, in Heller [101], participants were requested to judge smoothness of real surfaces while relying on vision-only, touch-only, and vision + touch sensing conditions. It was observed that the roughness discrimination task was better performed using touch-only than vision-only, particularly for very smooth surfaces. More importantly, bimodal vision-touch perception resulted in better discrimination results than any of the single modalities alone. In [102], Lederman and Abbott suggested that the extent to which one modality can dominate another varies depending on the structure of the sensing task. It is mentioned that when participants were requested to judge surface roughness in discrepant modalities, approximately equal weighting was assigned to the visual and haptic sensory inputs. In follow-up work [103], the authors modified the instructions of the experimental task given to participants and determined that visual feedback dominated haptic when observers were asked to judge spatial density; however, haptic feedback dominated vision when observers were asked to judge roughness. In a recent paper, however [104], the authors reinvestigated this topic using virtual reality techniques and a haptic force-feedback device. It was found that in contrast to [103], performance for roughness and density judgments did not differ significantly, i.e., the roughness estimates could not be discriminated from those of spatial density. It was also observed that for certain textures, there was evidence of vision-haptic intermodal integration that resulted in improved texture discrimination performance.

Conversely, in [105], Guest and Spence examined the role of multisensory integration in vision-tactile perception of texture. Their experiments revealed that vision and touch act as independent sources of roughness information. Consequently, there was no evidence of multimodal vision-haptic integration as roughness discrimination performance did not improve in bimodal conditions.

In [99], Poling *et al.* conducted a study to investigate whether roughness discrimination of virtual surfaces is different under bimodal visual + haptic conditions, as compared with unisensory conditions where only vision or haptic feedback is available. The virtual surfaces were sinusoidal gratings that varied in amplitude and spatial period. It was determined that when the grating amplitude was very small, the bimodal (vision + haptic) condition was highly dominated by haptic feedback (i.e., performance in the vision + haptic condition is identical to that in the haptic-only condition). Conversely, when the grating amplitude was large, the bimodal (vision + haptic) condition was highly dominated by visual feedback. For intermediate grating amplitudes, however, the bimodal condition proved to be better than haptic-alone and vision-alone.

7.2.2 Haptic Watermarking

Despite its importance, only one other research group has investigated watermarking of haptic-enabled 3D models. In [13,42], Prattichizzo *et al.* introduced a haptic watermarking technique where a sinusoidal watermark is superimposed onto a sinusoidal host surface texture signal. The authors achieved imperceptible haptic watermarking while relying on past experimental research on human detection thresholds for sinusoidal stimuli, which depict the minimum signal strength required for producing a sensation. Prattichizzo *et al.* also demonstrated how the haptically imperceptible sinusoidal watermark can be detected by means of spectral analysis.

In [13, 44], psychophysical experiments are conducted to measure the perceptibility of a watermark embedded in a 3D mesh. The watermark embedding is achieved by altering the positions of the mesh vertices. The experiments were performed using several 3D meshes with different resolutions in an attempt to establish a relationship between the watermark strength and the size of the triangular mesh components. The watermark perceptibility is inspected through a visual or a haptic interface, i.e., the user can only see or feel the virtual object. Experiments in which the subject can inspect the watermark perceptibility while simultaneously looking at and touching the virtual surface (bimodal conditions) were never demonstrated. An analysis using advanced statistical methods to test for significance between modalities (and possibly surface resolutions) was never provided. Furthermore, the experiments performed in the study relied on a small sample of participants (five subjects). This is an important limitation as it makes it more difficult to test hypotheses of differences among factors or to make any statistical generalizations.

7.3 Why Multimodal Vision-Haptic Watermarking?

It is intuitive to assume that a multimodal presentation of stimuli (e.g., for watermark detection) should lead to an improvement in performance when compared with the unimodal alternative. However, studies discussed in Section 7.2.1 have revealed that the role of multisensory feedback in roughness perception is not well understood (i.e., evidence for multisensory integration is mixed in this area) and certainly not straightforward as it can vary across different experimental conditions, including type of surface (real or virtual), haptic device used (tactile or force feedback), surface parameters (spatial density, roughness), etc. Specifically, in certain cases, it has experimentally been illustrated that the addition of a second modality could divide attentional resources and consequently reduce the information available from each of the individual modalities [105]. It is also suggested that this lack of multisensory integration can be due to the fact that information available from distinct modalities is possibly redundant. This finding has not only been observed in roughness perception but also in other studies investigating the role of multimodal integration, such as vision-haptic length perception [106]. Other studies [97, 99] support the notion that, in certain scenarios, an observer attends to cues from the dominant modality and discards information from the less dominant modality. Conversely, other findings in vision-haptic roughness perception [101, 104] suggest that inputs from the two modalities are efficiently integrated to produce a percept that is better than that provided by either modality alone.

Based on the various conclusions previously depicted regarding the integration of multisensory feedback in roughness perception (and assuming that 3D haptic watermark detection is a subcategory of virtual surface roughness perception), it is impossible to predict with certainty the role of bimodal visual-haptic feedback in the detection of a 3D watermark. To further emphasize the significance of analyzing multimodal vision-haptic watermarking, it is important to realize that the current vision-only-based imperceptibility criteria (in mesh watermarking) will ultimately be redefined if a scenario is uncovered in which watermark detection is best achieved using bimodal feedback. Consequently, 3D watermarking researchers will be required to carefully design (or redesign) their algorithms to ensure watermark imperceptibility in multimodal vision-haptic conditions. Conversely, if it is revealed that watermark detection performance in the haptic + vision condition is identical to (or worse than) that in the haptic or vision-alone condition, then it would be confirmed that the evaluation of a watermarking algorithm in a bimodal setup is not necessary for it to be categorized as *imperceptible*.

In this chapter, a study is conducted to investigate the role of bimodal vision-haptic feedback in the perception of a watermark embedded in a haptic-enabled 3D virtual surface. To the best of the authors' knowledge, this is the first work in the literature that examines multimodal vision-haptic watermarking. The experiments are performed using a hapto-visual interface that enables users to see and touch virtual objects at the same location in space. Hence, user interactions are more natural, and thus provide the ideal environment for high-precision visual-haptic experiments. A two-way within-subjects ANOVA is conducted to statistically explore the impact of the considered modalities on the measured watermark detection thresholds across different resolutions of the underlying virtual 3D mesh. The contributions of this work are as follows: Experimentally and statistically analyze the following: (1) Whether haptic-alone is superior to vision-alone in detecting a watermark embedded in a 3D mesh; (2) more impor-

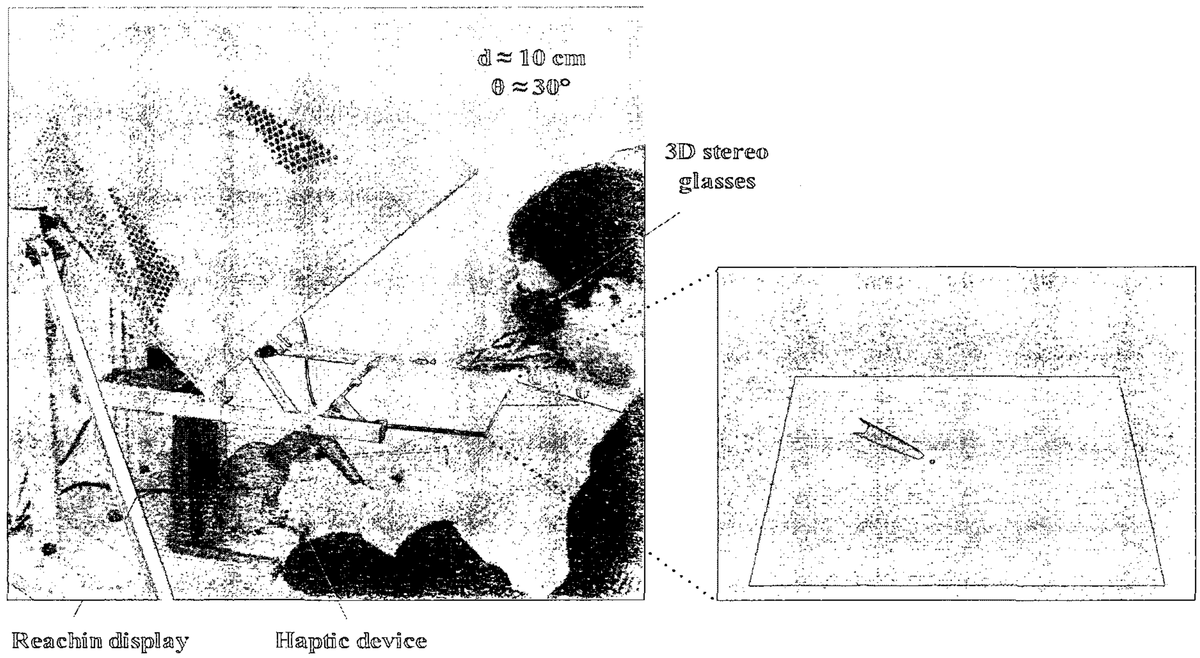


Figure 7.1: Participant's position and orientation with respect to the Reachin Display (left), as well as an example of the (watermarked) haptic-enabled virtual environment presented to the user (right) [3].

tantly, if bimodal vision-haptic conditions can improve the watermark detectability as opposed to either modality alone; (3) if the impact of the selected modality (or modalities) on the perceptibility of the 3D watermark is dependent on the chosen surface resolution; and (4) whether in bimodal conditions the detectability of the watermark increases or decreases with the resolution of the surface mesh.

7.4 Materials and Methods

As previously mentioned, the objective of the undertaken experiments is to measure the perceptibility of a watermark embedded in a 3D mesh while looking at, touching, or simultaneously looking at and touching the virtual surface. To simplify the analysis of the experiments, the virtual object is selected as in [13] and it consists of a flat plane surface represented using a triangle mesh. All three experiments, i.e., visual, haptic, and visual-haptic, are conducted using the Reachin Display (see Fig. 7.1) [80], a hapto-visual system that integrates a haptic device with stereo graphics for an immersive and high quality 3D experience.

The importance of using a colocated hapto-visual setup with stereoscopic information as opposed to a non-collocated display (i.e., the workspace of the haptic device is located outside the visual area) has been discussed in several research studies [107–111]. In [107, 111], the authors emphasized the relevance of vision-haptic collocation in improving a user's perception of depth when interacting with virtual environments. Similarly, in [108, 109], several experiments were conducted to demonstrate that colocation of the visual and haptic sensory conditions does in fact lead to improved task performance and enhanced sense of presence within a virtual environment. More importantly, in [108], the authors revealed that collocation is of great benefit to tasks requiring accurate positioning (which is of crucial significance when conducting haptic watermarking experiments). In addition, in [110], Congedo *et al.* observed that for spatial tasks, delocation of visual and haptic percepts prevented a balance in the integration of the two modalities by favoring the dominant sense. For the foregoing reasons, and to obtain reliable and high-precision visual-haptic results, all our experiments were conducted in collocated conditions using the Reachin Display. The Reachin Display's high-accuracy and collocated hapto-visual setup has made it very popular in various precision-demanding applications, specifically in medical training, where surgical simulations of complex procedures are typically required.

The haptic stimulus was sensed using the Sensable PHANTOM Desktop force-feedback device, which is equipped with an encoder stylus that provides six-degree-of-freedom single-contact-point interaction. The PHANTOM Desktop device is reputed for its low backdrive friction (< 0.23 oz), high-fidelity force feedback output, and high-resolution positional sensing (nominal position resolution of 0.02 mm), making it ideal for high-precision haptic experiments. Furthermore, to guarantee a "real" multimodal visual-haptic sensing of watermarked 3D objects, the Reachin Display (including the haptic device) was periodically recalibrated to ensure a near-perfect collocation between the virtual environment projected on the mirror and the actual physical position of the haptic device (the collocation error is sufficiently small for it not to be perceived even by human subjects with acute senses).

7.4.1 Subjects

Altogether, 12 participants (six males and six females, aged 23–28), 11 right handed and one left handed with no known sensorimotor impairments with their hands took part in the experiments. Their prior experience with PHANTOM haptic devices ranged from novice to expert. All participants were remunerated for their time.

7.4.2 Haptic Watermark Embedding and Detection

The watermark is embedded in the haptic-enabled virtual surface by perturbing the coordinates of the mesh vertices. This is achieved using the following method:

$$H_w(i) = H(i) + \beta \cdot w(i) \cdot n(i), \tag{7.1}$$

where

$$H(i) = [\, x \quad y \quad z \,]^T \ and, \ n(i) = [\, 0 \quad 0 \quad 1 \,]^T.$$

More precisely, $H(i)$ corresponds to the ith vertex of the host plane, whereas $H_w(i)$ denotes the ith vertex of the watermarked plane. $w(i)$ is an independent and identically distributed watermark embedded along the surface normal $n(i)$, and consists of random numbers that follow a Gaussian distribution with zero mean and unit variance. Finally, β is a parameter that controls the watermark embedding strength. Conversely, the watermark detection is performed by (1) looking at the plane; (2) touching the plane; and (3) looking at and touching the plane.

7.4.3 Conditions

There were three experimental conditions in the analysis that depended on the type of sensory feedback presented to each participant during each experiment. The following sensory conditions were used for this study:

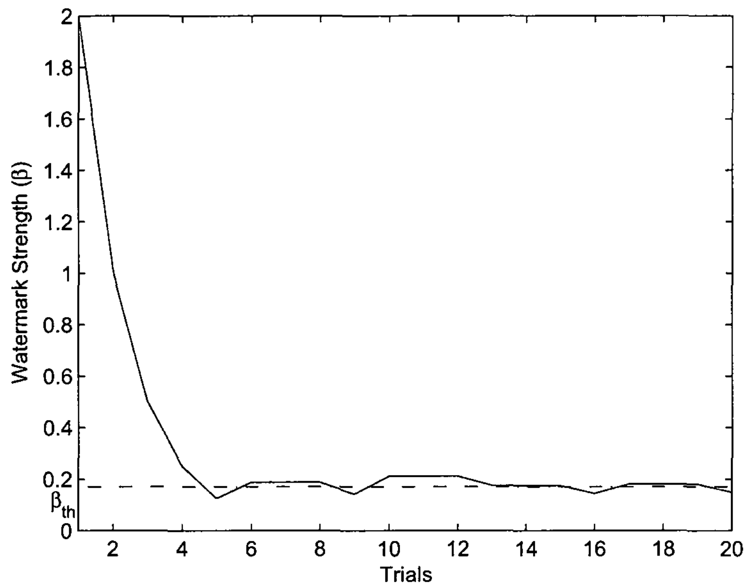


Figure 7.2: A typical experimental staircase procedure for one run [3]. The dashed line represents the resulting detection threshold (β_{th}).

(1) only visual feedback;

(2) only haptic feedback;

(3) visual and haptic feedback.

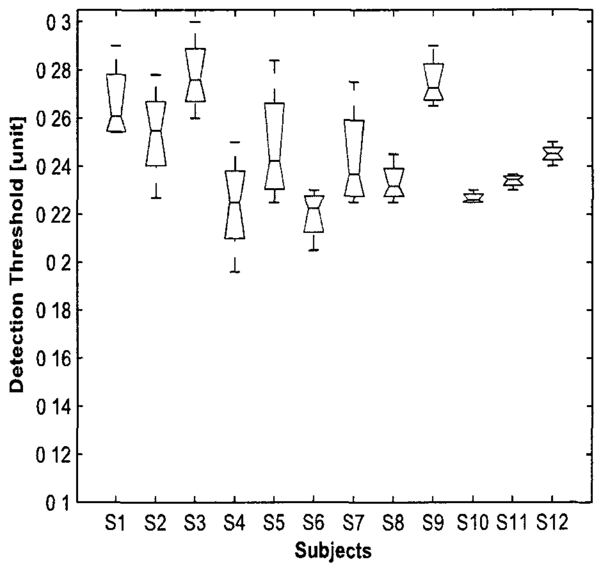
Under all three conditions, for the sake of consistency, the participants were requested to comfortably sit in front of the Reachin Display at a distance $d \approx 10\ cm$ and an orientation $\theta \approx 30°$ from the device's half-mirror (used to align the visual and haptic stimuli and prevent a participant from seeing his or her hand when manipulating the virtual object) as depicted in Fig. 7.1 (left). An example of the virtual environment presented to the user during a visual-haptic experiment is shown in Fig. 7.1 (right). It encompasses a (clearly watermarked) virtual plane and the virtual stylus used to interact with the surface mesh. Moreover, the color of the surface used in the experiments is in fact blue. A light-gray color is used in Fig. 7.1 (right) for better quality black and white printouts of the chapter. Furthermore, the experiments were conducted in a dark room to avoid any effects of ambient lighting.

7.4.4 Procedure

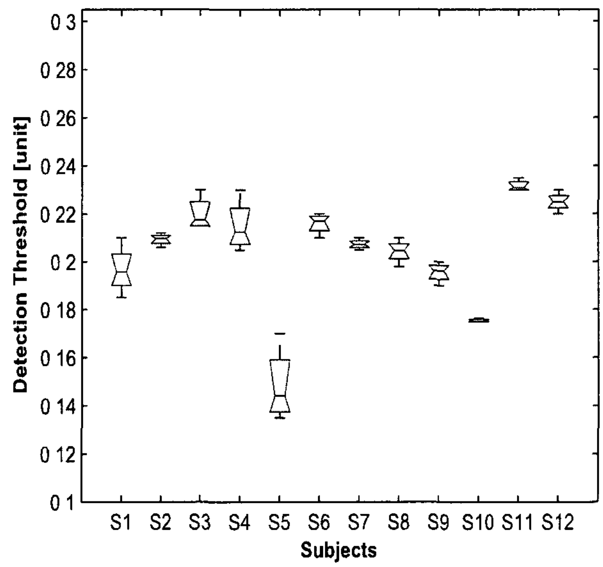
The watermark detection thresholds for each subject, and for the three experiments, were determined using the 2-interval forced choice (2IFC) paradigm along with the adaptive staircase (AS) method [112]. In each trial, subjects viewed two stimuli in subsequent intervals and in random order. One of the stimuli was the host plane; the other was the watermarked plane. The subjects' task was to report which of the two stimuli contained the watermark. Trial-by-trial correct-answer feedback was never provided during the experiment. To determine the detection threshold β_{th}, a "lead-in" one-up one-down AS rule was initially used to speed up the approach to the β region of interest. During this phase, β was halved after each correct response, until the first error occurred. The staircase then followed a one-up three-down rule (β is only decreased after three consecutive correct answers and raised again following a single wrong answer). In this second phase, β was increased or decreased by multiplying it by a factor $k = 1.5$ or 0.75, respectively. After three reversals are obtained, a third phase is initiated (also follows the one-up three-down AS method) where β is increased or decreased by multiplying it by a factor $k = 1.25$ or 0.825, respectively. Trials continued until a total of three reversals were obtained in the third phase of the staircase. Reversals being defined as when the watermark strength changes from increasing to decreasing or vice versa. The detection threshold β_{th} is computed while taking the average over the last phase of reversals. As an example, a plot of one of the experiments using the AS procedure is presented in Fig. 7.2.

7.5 Results

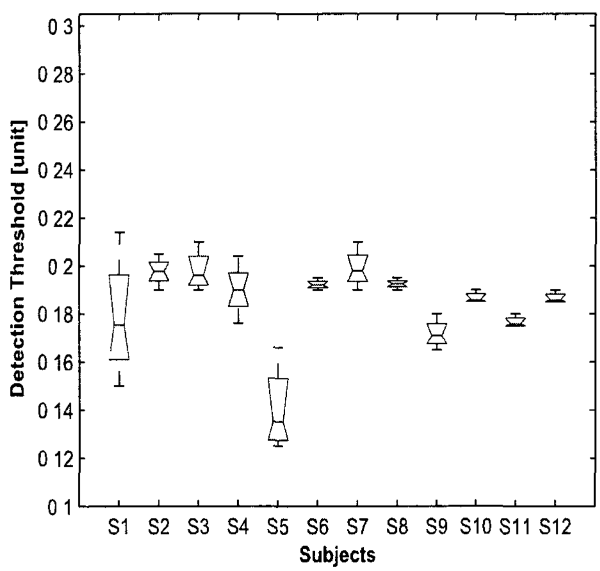
In this section, detection thresholds resulting from each of the three experiments (i.e., visual, haptic, and visual-haptic perceptibility of the watermark) for every participant, and repeated for different resolutions of the virtual surface, are presented. General observations of the provided results are initially discussed, followed with a detailed statistical analysis of the data.



(a)



(b)



(c)

Figure 7.3: Results obtained in experiments I, II, and III, where (a) visual feedback, (b) haptic feedback and (c) visual and haptic feedback were provided. The surface used in these cases is a low-resolution rectangular mesh of size 24 $cm \times 18$ cm represented using 672 triangles.

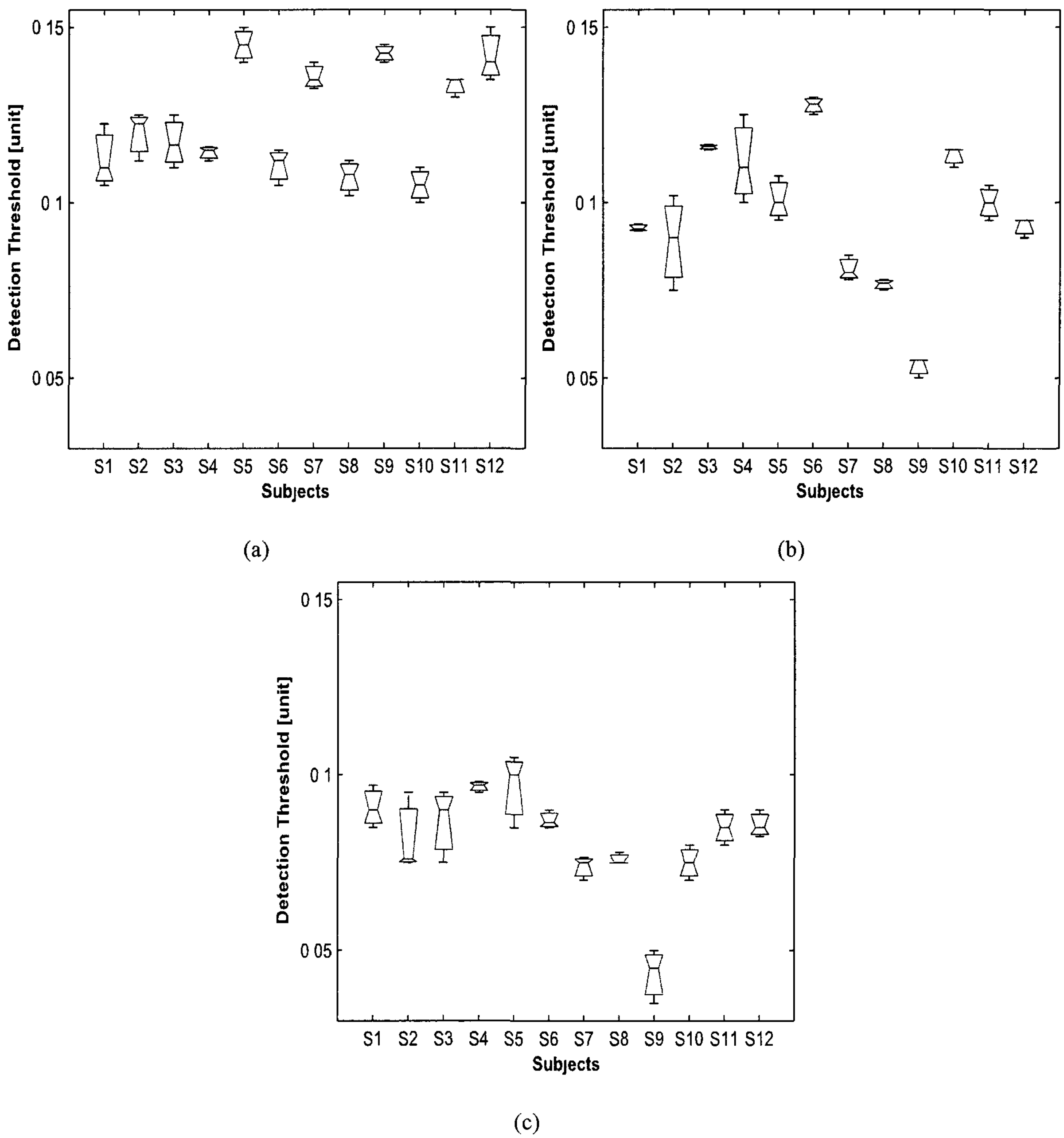


Figure 7.4: Results obtained in experiments I, II, and III, where (a) visual feedback, (b) haptic feedback and (c) visual and haptic feedback were provided. The surface used in these cases is a high-resolution rectangular mesh of size 24 $cm \times 18$ cm represented using 2048 triangles.

7.5.1 Detection Thresholds

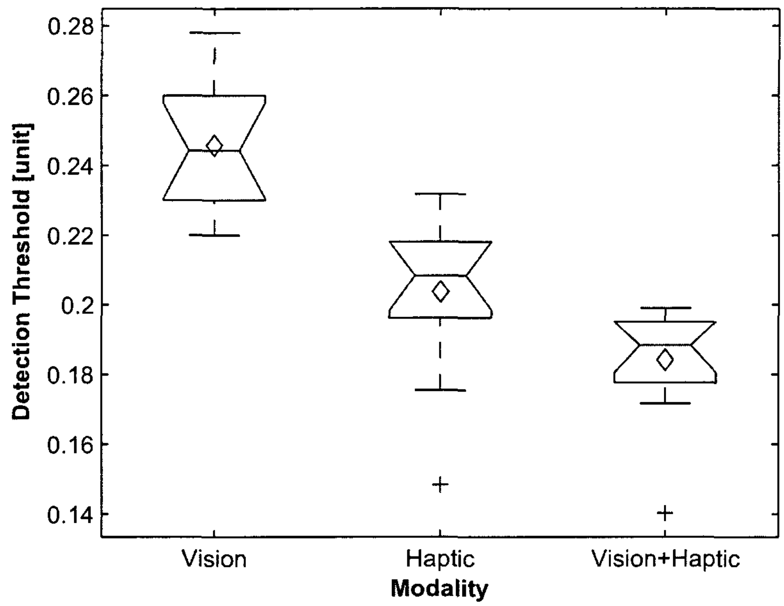
For each of the three experiments, the participants were presented with a lower and a higher resolution version of the virtual surface where the former is represented using 672 triangles, whereas the latter is constituted of 2048 triangles. The plane mesh displayed in all of the undertaken experiments is a rectangular surface of size $24\ cm \times 18\ cm$. Furthermore, haptic rendering (i.e., the synthetically generated haptic stimuli) is performed via the common virtual proxy method [113]. Effort was made to make sure that the subjects are well rested prior to the experiments, as lack of rest could impact the subject's sense of vision and touch, as well as their ability to focus. Also, participants took breaks as needed during each experiment.

Experiment I – Visual Perceptibility

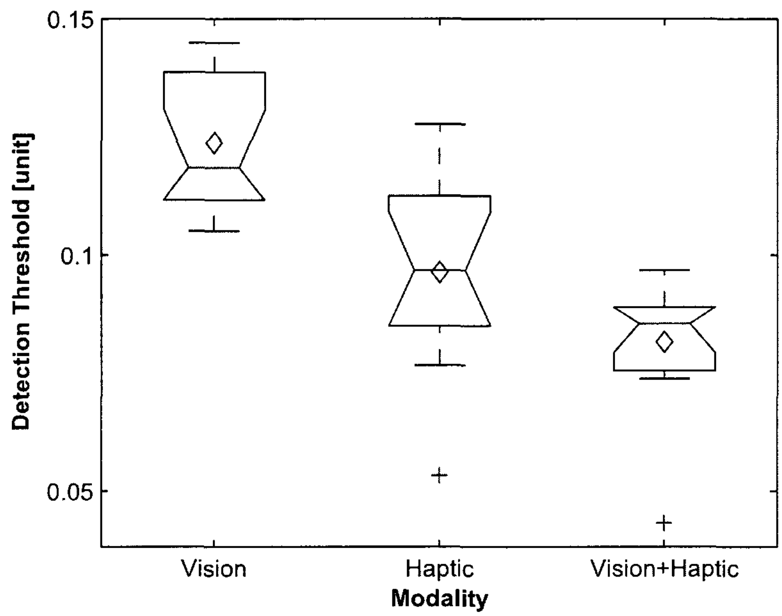
In this experiment, the participants were requested to detect the watermarked plane while solely relying on visual feedback. Three runs of the experiment were conducted for each of the two available resolutions of the surface, and the corresponding detection thresholds were obtained using the aforementioned AS technique. A box plot of the resulting vision-based detection thresholds of all 12 participants (three detection thresholds are acquired per participant) for the lower and the higher resolution meshes are illustrated in Figs. 7.3 (a) and 7.4 (a) respectively.

Experiment II – Haptic Perceptibility

In the second experiment, the participants were requested to detect the watermarked plane while solely relying on haptic feedback. Three runs of the experiment were carried out for each of the two available resolutions of the surface, and the corresponding detection thresholds were obtained using the AS technique. A box plot of the resulting haptic-based detection thresholds of all participants, for the lower and higher resolution meshes, is illustrated in Figs. 7.3 (b) and 7.4 (b), respectively.



(a)



(b)

Figure 7.5: Average detection threshold values of all the participants for each run of every experiment and the corresponding standard deviations. The results for the low- and high-resolution meshes are presented in (a) and (b), respectively.

Experiment III – Visual and Haptic Perceptibility

As for the visual-haptic case, the participants were requested to detect the watermarked plane while relying on both, visual and haptic feedback. Similar to the first two experiments, three runs of the experiment were performed for each of the two available resolutions of the surface, and the corresponding detection thresholds were obtained using the AS method. A box plot of the resulting bimodal vision-haptic-based detection thresholds of all participants for the lower and higher resolution meshes is illustrated in Figs. 7.3 (c) and 7.4 (c) respectively.

7.5.2 Initial Observations

To better visualize and analyze the results considering that the obtained detection thresholds across the different modalities are fairly similar, the threshold values of all the participants for each run of every experiment were averaged and plotted for the lower and higher resolution meshes as shown in Fig. 7.5 (a) and (b), respectively. Several observations can be made from the results presented in Figs. 7.3 – 7.5. First, for all 12 subjects, the average detection thresholds obtained in the haptic-alone experiment are for the most part below the detection thresholds determined when the participants solely relied on visual feedback. This suggests that the watermark is possibly more easily detected when touching a surface mesh as opposed to performing a visual assessment. This assumption is in agreement with the results already obtained in [44], where the perceptibility of a watermark embedded in a 3D mesh is also inspected using vision-alone and haptic-alone (bimodal visual-haptic experiments were not conducted, and as aforementioned, a statistical analysis was not performed to test for significance between modalities). However, in [44], the results illustrated that the average detection thresholds of the haptic-only experiments are in most cases less than half the threshold values obtained with vision-only. Conversely, in the results presented here, it can be observed from Fig. 7.5 that the average detection thresholds (mean values are identified with the diamond-shaped marker symbol) of the haptic-only conditions are only approximately 20% less than the threshold values obtained with vision-only. This variation in the results is possibly due to several reasons.

First, in [44], a fewer number of participants took part in the experiments, and a smaller number of runs per experiment were performed. These factors can evidently lead to less precise results. Moreover, the haptic-only experiments performed in [44] make use of the same high-precision haptic device used in the experimental setup exploited in this work. However, their vision-only experiments were performed using a primitive computer display (cathode ray tube [CRT] monitor). Conversely, the experimental setup presented here relies on stereo graphics for an immersive and high quality 3D experience, which in turn enabled us to perform high-precision vision-based experiments. Thus, the large margin between the detection threshold values of the haptic-only and vision-only experiments presented in [44] is probably also due to the considerable difference in quality between their haptic and visual displays.

Another observation that can be deduced from the results illustrated in Fig. 7.5 is that the detectability of the watermark increases with the resolution of the surface mesh in all three sensory conditions (vision, haptic, and vision + haptic). Moreover, from Fig. 7.5, it is shown that the between-subject variability (the standard deviations) of the detection thresholds for both experiments and the two presented resolutions is quite high. Conversely, from Figs. 7.3 and 7.4 it can be seen that the within-subject variability is relatively low, suggesting that our experimental procedure and setup are accurate. The large between-subject variability however, is probably due to the fact that when performing a perceptual judgment, participants exploit in different manner the information available through their distinct sensory channels. In addition, it is known that PHANTOM haptic devices induce low but perceivable backdrive friction, which can negatively affect a user's haptic perception [114]. In turn, during the haptic and bimodal visual-haptic experiments, it is believed that certain observers were not able to isolate the backdrive distortion as well as others when attempting to feel the barely noticeable watermark. This could also have contributed to the relatively large between-subject variability of the results.

Furthermore, in the bimodal case, it seems that visual-haptic feedback is better than any of the single modalities when detecting a watermark embedded in a 3D mesh. This can best be observed from the results depicted in Fig. 7.5, where the lowest detection threshold values were obtained in the visual-haptic conditions. This suggests that there is possibly a higher level of discrimination performance resulting from the combined modality condition.

7.5.3 Statistical Data Analysis

To further confirm the observations discussed in Section 7.5.2, a two-way within-subjects ANOVA is conducted to statistically explore the impact of the considered modalities (visual, haptic, and visual-haptic) and the virtual mesh resolutions (low and high) on the acquired watermark detection thresholds. The two-way ANOVA enables us to simultaneously test for the effect of each of the independent variables (modality, resolution) on the dependent variable (watermark detection threshold) and also identifies any interaction effect. For each participant, the detection thresholds used in the analysis are in fact the average values of the runs carried out for every (modality, resolution) experimental condition, e.g., the detection threshold for Subject 1 (S1) under the experimental condition [Modality = Haptic, Resolution = Low] is equal to the average of the three threshold values obtained during the three corresponding runs.

Throughout the analysis, when violations of the assumption of sphericity for the repeated measures ANOVA were observed, Greenhouse Geisser adjustments were used. In addition, pairwise comparisons and Bonferroni adjustments for multiple comparisons were conducted as necessary [115]. For all experimental conditions, the significance level adopted was $\alpha = 0.05$. A statistically significant main effect for *Modality* was observed $F(1.225, 13.480) = 41.4$; $p < 0.0005$ with a large effect size (partial eta squared) $\eta_p^2 = 0.79$ [116]. The effect size is specified to indicate the proportion of variance of the acquired detection thresholds that is explained by the main effect (Modality), independent of the number of samples. A post-test pairwise comparison with Bonferroni adjustment indicated a significant difference between the vision and haptic modalities ($p = 0.002$), between the vision and vision + haptic modalities ($p < 0.0005$), as well as between haptic and vision + haptic modalities ($p < 0.0005$).

In addition, as alleged in Fig. 7.5, the factor *Resolution* also has a statistically significant main effect, where $F(1, 11) = 445.2$; $p < 0.0005$ with a large effect size $\eta_p^2 = 0.976$. However, it was observed that the *Modality* × *Resolution* interaction effect, which is the combined effect of the *Modality* and *Resolution* factors on the detection thresholds, was not statistically significant as $F(2, 22) = 3.097;\ p = 0.065$ with a small effect size $\eta_p^2 = 0.220$.

Overall, the results indicate that the lowest watermark detection thresholds were obtained in visual + haptic conditions, as compared with unisensory conditions where only vision or haptic feedback is available. This finding suggests that the addition of a second modality provides information, some maybe redundant and some non redundant, that improves the detectability of watermarks embedded in 3D models. Furthermore, if we are to compare the unisensory conditions, we can state with a fair degree of certainty that haptic-alone conditions dominate vision-alone when detecting a watermark embedded in 3D mesh. In addition, although it is very apparent that the detectability of the watermark increases with the resolution of the surface mesh, the low F value for *Modality*×*Resolution* interaction confirms that the *Resolution* of the surface has absolutely no effect on the exploited *Modality*, i.e., surface resolution and modality conditions are entirely independent. Consequently, the impact of the selected modality on the perceptibility of the 3D watermark is independent of the chosen surface resolution.

7.6 Summary

In this chapter, a study is conducted to investigate the role of multisensory feedback in the perception of a watermark embedded in a haptic-enabled 3D virtual surface. Specifically, the focus is on determining whether the use of a bimodal approach, where a subject is looking at and touching a 3D mesh could in fact improve the perceptibility of a watermark as opposed to when only one single modality (vision-alone or haptic-alone) is available. The experiments are performed using a hapto-visual interface that enables users to see and touch virtual objects at the same location in space. This approach enables a superior integration of vision and touch than a conventional 2D screen-based display. An analysis of variance is also conducted to statistically explore the impact of the considered modalities (vision, haptic, and vision +

haptic) on the measured watermark detection thresholds, across different resolutions of the underlying virtual 3D mesh. Overall, the results suggest that relying on bimodal visual-haptic feedback is better than any of the single modalities when detecting a watermark embedded in a 3D model. In addition, it has been assessed that the impact of the selected modality on the perceptibility of the 3D watermark is independent of the chosen surface resolution.

Part V

Synopsis

Chapter 8

Conclusion and Future Work

8.1 Concluding Remarks

This work focused on the design, evaluation, and analysis of haptic data reduction techniques in order to: (1) Improve packet transmission in haptic-enabled teleoperation systems, and (2) reduce the high dimensionality of the inherently large haptic datasets. In addition, an important part of this work was dedicated to multimodal haptic watermarking due to its conceptual similarity to haptic data reduction and compression.

Data reduction algorithms for improved haptic communication in multi-DoF telehaptic systems were discussed in Chapters 3 and 4. In Chapter 3, two human-haptic-perception-based data reduction methods for telehaptic systems were introduced with a focus on haptic-based telementoring applications. A haptic prediction model that relies on least-squares estimation and median filtering was exploited to reduce the amount of haptic packets transmitted, and to improve the reconstruction of data samples on the receiver side. The suggested techniques proved to considerably reduce haptic data traffic in normal network conditions and in the presence of communication time delays and packet loss, while guarantying the hard real-time constraints of haptic data streams. The experimental results demonstrated that haptic data packets can be reduced by as much as 96% in normal network conditions and up to 93% in the presence of significant communication delays and packet loss, while preserving the overall quality of the telehaptic experience. Conversely, Chapter 4 extended Chapter 3 and introduced an algorithm that relies on knowledge from human haptic perception to reduce the number of packets transmitted in 6-DoF haptic-enabled telepresence systems without compromising user immersiveness and transparency. Distance metrics for different haptic data types were also derived to evaluate the acuity of human haptic perception in detecting haptic distortions when data reduction is performed in 6-DoF telepresence systems. Experimental results proved that the suggested 6-DoF algorithm can considerably reduce haptic data traffic in normal network conditions (92.3% and 85.1% in the forward (position/orientation packets) and backward (force/torque packets) directions respectively) and in the presence of communication time delays and packet loss, with little or no influence on the quality of telepresence interaction.

The significance of data (feature) reduction in haptic information analysis problems was highlighted in Chapters 5 and 6. Data reduction methods were presented within the context of haptic data mining and knowledge discovery to primarily help facilitate the analysis of the inherently high-dimensional haptic datasets. More precisely, Chapter 5 explored a feature generation approach to data reduction where high-dimensional haptic datasets are mapped onto 3D VR-based visual spaces through linear and nonlinear transformations; this ultimately enables visual data mining of high-dimensional datasets. More interestingly, this allows a user to visually explore patterns in very high-dimensional (and potentially noisy) haptic datasets. The considered datasets consisted of haptic-based handwritten signatures. Construction of 3D visual spaces was performed in an unsupervised manner using three different methods: linear feature generation using PCA, and nonlinear feature generation using classical optimization and genetic programming techniques. Experimental results illustrated that feature reduction of very high-dimensional haptic datasets to only three dimensions and with minimal loss of information, is feasible. In fact, despite some apparent loss of information introduced in the mapping procedure, all three explored methods generated relatively accurate 3D spaces that enable easy visual analysis and interpretation of haptic datasets. Alternatively, in Chapter 6, feature selection is explored as another possible approach to reducing the high dimensionality of haptic datasets. A novel information-preserving haptic data reduction method is suggested that combines intelligent data subsampling with a feature selection technique that is based on concepts from rough set theory and evolutionary computation (genetic algorithms). The obtained results demonstrated that the proposed data reduction technique can outperform some well-established machine learning methods in the selection of minimal subsets of features present in high-dimensional haptic datasets.

Lastly, in Chapter 7, haptic data protection through digital watermarking was investigated. In particular, a study was presented that explored the role of multisensory feedback in the perception of a watermark embedded in a haptic-enabled 3D virtual object. The study aimed at answering the following questions: Is the haptic sensory channel more sensitive than the visual sensory channel in detecting a watermark embedded in a haptic-enabled 3D object? Do watermarks inspected using multimodal feedback (vision + haptic) result in very different detection thresholds from those detected using a single sensory modality (touch-only or vision-only); or more importantly, does visual feedback, when presented together with haptic feedback, improve the perception of a watermark embedded in a 3D mesh? The obtained experimental results suggested that haptic-alone conditions dominate vision-alone when detecting a watermark embedded in 3D mesh. In addition, relying on multimodal vision-haptic feedback is better than any of the single modalities when detecting a watermark. Another interesting finding that was observed from the acquired results is that the impact of the selected modality on the perceptibility of a 3D watermark is independent of the chosen surface resolution. It is expected that these findings will stimulate the reevaluation of existing, and the development of new, mesh watermarking algorithms to achieve multimodal vision-haptic imperceptibility, and will serve as a basis for further studies in haptic digital watermarking.

8.2 Directions for Possible Future Work

Data Reduction in Telehaptic Systems: The suggested data reduction algorithms for haptic communication in 3- and 6-DoF telehaptic systems proved to be effective in both, normal and undesirable network conditions, while simultaneously preserving realistic haptic perception of remote environments. However, an extension of the data reduction algorithms to ensure passive (and stable) haptic-enabled teleoperation under a wide variety of network conditions would be desirable. More specifically, it would be interesting to analyze the passivity of the suggested data algorithms and their impact on human haptic perception when applied in telepresence and teleaction systems.

Data Reduction in Haptic Pattern Analysis Problems: The suggested methods that enable data reduction through feature generation and feature selection should be further analyzed with a larger collection of datasets acquired using different haptic-enabled virtual reality applications, e.g., haptic-based rehabilitation, gaming, medical training simulations, etc. This can further assess and confirm the validity of the suggested data reduction techniques, including the haptic data preprocessing methodologies. In addition, with regard to data reduction through feature selection, it would be of interest to investigate the possible haptic feature selection improvements that could result if a multi-objective genetic algorithm is exploited in the reduct computation procedure. A multi-objective approach can better overcome the conflicting properties of reducts, of having small cardinality and the ability to discern among all objects. This approach can ultimately yield a more performing information-preserving haptic data (feature) reduction method.

Multimodal Vision-Haptic Digital Watermarking: The results obtained in Chapter 7 are very interesting and encouraging as they clearly demonstrate the importance of multimodal vision-haptic feedback in the perceptibility of 3D digital watermarks. Many issues are, however, left open for further investigation, including (some of which stem from the discussion presented in [43]):

- Investigating the possible impact of different haptic rendering techniques.
- Different graphic rendering conditions should be analyzed, including: surface color, viewpoint, shading, degree of opacity or transparency, lighting, etc.
- Other more complex shapes of the original 3D object should be considered.
- Evaluating the perceptual impact of bimodal auditory-haptic feedback and comparing with visual-haptic conditions. Or of more importance, investigating the role of multisensory visual-haptic-auditory feedback in the perceptibility of a 3D watermark embedded in haptically-enabled meshes.

Bibliography

[1] M. Buss, K.K. Lee, N. Nitzsche, A. Peer, B. Stanczyk, and U. Unterhinninghofen. *Advanced telerobotics: Dual-handed and mobile remote manipulation*. Advances in Telerobotics: Human System Interfaces, Control, and Applications, Star Series, Springer, 2007.

[2] N. Nathoo, M.C. Cavusoglu, M.A. Vogelbaum, and G.H. Barnett. In touch with robotics:neurosurgery for the future. *Neurosurgery*, 56(3):421–430, Mar. 2005.

[3] N. Sakr, N.D. Georganas, J. Zhao, and E.M. Petriu. Multimodal vision-haptic perception of digital watermarks embedded in 3-D meshes. *IEEE Transactions on Instrumentation and Measurement*, 59(5):1047–1055, May 2010.

[4] 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*, 56(2):588–597, Feb. 2008.

[5] M.A. Otaduy and M.C. Lin. *High fidelity haptic rendering*. San Rafael, CA: Morgan & Claypool Publishers, 2006.

[6] J. Zhou, F. Malric, E.M. Petriu, and N.D. Georganas. Uniform hardness perception in 6-dof haptic rendering. *IEEE Transactions on Instrumentation and Measurement*, (99):1–12, 2010.

[7] A. Kron, G. Schmidt, B. Petzold, M.F. Zah, P. Hinterseer, and E. Steinbach. Disposal of explosive ordnances by use of a bimanual haptic telepresence system. In *Proc. of IEEE International Conference Robotics and Automation*, pages 1968–1973, Apr. 2004.

[8] V. Nistor, B. Allen, G.P. Carman, P. Faloutsos, and E. Dutson. Haptic guided telementoring and videoconferencing system for laparoscopic surgery. In *Proc. of SPIE International Symposium Smart Structures and Materials & Nondestructive Evaluation and Health Monitoring*, Mar. 2007.

[9] N. Sakr, J. Zhou, N.D. Georganas, J. Zhao, and E.M Petriu. Network traffic reduction in six degree-of-freedom haptic telementoring systems. In *Proc. of IEEE International Conference on Systems, Man, and Cybernetics*, pages 2451–2456, Oct. 2009.

[10] A. El Saddik, M. Orozco, Y. Asfaw, S. Shirmohammadi, and A. Adler. A novel biometric system for identification and verification of haptic users. *IEEE Transactions on Instrumentation and Measurement*, 56(3):895–906, June 2007.

[11] M. Orozco, M. Graydon, S. Shirmohammadi, and A. El Saddik. Experiments in haptic-based authentication of humans. *Journal of Multimedia Tools and Applications*, 37(1):73–92, 2008.

[12] I. Czarnowski and P. Jedrzejowicz. Data reduction algorithm for machine learning and data mining. *New Frontiers in Applied Artificial Intelligence, Lecture Notes in Computer Science, Springer-Verlag, Heidelberg*, 5027:276–285, 2008.

[13] D. Prattichizzo, M. Barni, G. Menegaz, A. Formaglio, H.Z. Tan, and S. Choi. Perceptual issues in haptic digital watermarking. *IEEE Multimedia*, 143(3):84–91, 2007.

[14] N. Sakr, N.D. Georganas, J. Zhao, and E.M. Petriu. Perceptibility of digital watermarks embedded in 3D meshes using unimodal and bimodal hapto-visual interaction. In *Proc. of IEEE International Instrumentation and Measurement Technology Conference*, pages 987–991, 2009.

[15] I.J. Cox, M.L. Miller, J.A. Bloom, J. Fridrich, and T. Kalker. *Digital watermarking and steganography*. 2nd edn. Morgan Kaufmann Publishers, 2008.

[16] M. Corsini, M. Barni, F. Bartolini, R. Caldelli, V. Cappelini, and A. Piva. Towards 3D watermarking technology. In *Proc. of IEEE Region 8 EUROCON Conference*, pages 393–396, 2003.

[17] S. Biswas, S.R. Das, and E.M. Petriu. An adaptive compressed mpeg-2 video watermarking scheme. *IEEE Transactions on Instrumentation and Measurement*, 54(5):1853–1861, 2005.

[18] M. Corsini, E.D. Gelasca, and T. Ebrahimi. A multi-scale roughness metric for 3D watermarking quality assessment. In *Proc. of Workshop on Image Analysis for Multimedia Interactive Services*, 2005.

[19] P.R. Alface and B. Macqi. Shape quality measurement for 3D watermarking schemes. In *Proc. of SPIE-IS&T Electronic Imaging, Security and Watermarking of Multimedia Contents*, pages 622–634, 2006.

[20] C. Shahabi, A. Ortega, and M.R. Kolahdouzan. A comparison of different haptic compression techniques. In *Proc. of IEEE International Conference on Multimedia and Expo*, pages 657–660, Aug. 2002.

[21] M.L. McLaughlin, J.P. Hespanha, and G.S. Skhatme. *Lossy compression of haptic data, in Touch in virtual environment: haptics and the design of interactive systems*. M. McLaughlin, Ed., IMSC Press, Prentice Hall, 2002.

[22] C. W. Borst. Predictive coding for efficient host-device communication in a pneumatic force-feedback display. In *Proc. of Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pages 596–599, Mar. 2005.

[23] M. Bouzit, G. Burdea, G. Popescu, and R. Boian. The rutgers master II-new design force-feedback glove. *IEEE/ASME Transactions on Mechatronics*, 7:256–263, 2002.

[24] S. Hirche, P. Hinterseer, E. Steinbach, and M. Buss. Towards deadband control in networked teleoperation systems. In *Proc. of IFAC World Congress, International Federation of Automatic Control*, 2005.

[25] S. Hirche, P. Hinterseer, E. Steinbach, and Martin Buss. Transparent data reduction in networked telepresence and teleaction systems. part I: Communication without time delay. *Presence: Teleoperators and Virtual Environments*, 16(5):523–531, Oct. 2007.

[26] S. Hirche, P. Hinterseer, E. Steinbach, and M. Buss. Network traffic reduction in haptic telepresence systems by deadband control. In *Proc. of IFAC World Congress, International Federation of Automatic Control*, 2005.

[27] S. Hirche and Martin Buss. Transparent data reduction in networked telepresence and teleaction systems. part II: Time delayed communication. *Presence: Teleoperators and Virtual Environments*, 16(5):531–542, Oct. 2007.

[28] R. Anderson and M. Spong. Bilateral control of teleoperators with time delay. *IEEE Transactions on Automatic Control*, 34:494–501, 1989.

[29] G. Niemeyer and J.-J.E. Slotine. Stable adaptive teleoperation. *IEEE Journal of Oceanic Engineering*, 16(1):152–162, Jan. 1991.

[30] P. Hinterseer, E. Steibach, and S. Chaudhuri. Model based data compression for 3D virtual haptic teleinteraction. In *Proc. of IEEE International Conference on Consumer Electronics*, pages 23–24, Jan. 2006.

[31] P. Hinterseer, E. Steinbach, and S. Chaudhuri. Perception-based compression of haptic data streams using kalman filters. In *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 14–19, May 2006.

[32] P. Hinterseer and E. Steinbach. A psychophysically motivated compression approach for 3D haptic data. In *Proc. of Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pages 35–41, Mar. 2006.

[33] P. Hinterseer, S. Hirche, E. Steinbach, and M. Buss. A novel, psychophysically motivated transmission approach for haptic data streams in telepresence and teleaction systems. In *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 1097–1100, 2005.

[34] M. Kuschel, P. Kremer, S. Hirche, and M. Buss. Lossy data reduction methods for haptic telepresence systems. In *Proc. of IEEE International Conference on Robotics and Automation*, pages 2933–2938, May 2006.

[35] M. Kuschel, P. Kremer, and M. Buss. Passive haptic data-compression methods with perceptual coding for bilateral presence systems. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 39(6):1142–1151, Nov. 2009.

[36] J. Kammerl and E. Steinbach. Deadband-based offline-coding of haptic media. In *Proc. of ACM international conference on Multimedia*, pages 549–558, 2008.

[37] I. Vittorias, J. Kammerl, S. Hirche, and E. Steinbach. Perceptual coding of haptic data in time-delayed teleoperation. In *Proc. of World Haptics*, pages 208–213, Mar. 2009.

[38] J. Kammerl, I. Vittorias, V. Nitsch, B. Faerber, E. Steinbach, and S. Hirche. Perception-based data reduction for haptic force-feedback signals using velocity-adaptive deadbands. *PRESENCE - Teleoperators and Virtual Environments*, 19(5), 2010.

[39] R. Iglesias, M. Orozco, J.J. Valdés, and A. El Saddik. Behavioral features for different haptic-based biometric tasks. In *Proc. of IEEE International Workshop on Haptic, Audio and Visual Environments and Games*, pages 102–106, 2007.

[40] R. Iglesias, M. Orozco, F.A. Alsulaiman, J.J. Valdés, and A. El Saddik. Characterizing biometric behavior through haptics and virtual reality. In *Proc. of IEEE International Carnahan Conference*, pages 174–179, 2008.

[41] R. Iglesias, M. Orozco, J.J. Valdés, and A. El Saddik. Visualizing human behavioral features based on signature haptic data. In *Proc. of International Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pages 451–456, 2008.

[42] D. Prattichizzo, M. Barni, H.Z. Tan, and S. Choi. Perceptibility of haptic digital watermarking of virtual textures. In *Proc. of World Haptics, First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pages 50–55, 2005.

[43] S. Belloni, A. Formaglio, G. Menegaz, H.Z. Tan, D. Prattichizzo, and M. Barni. Is haptic watermarking worth it? In *Proc. of IS&T/SPIE International Symposium on Electronic Imaging*, pages 15–19, 2006.

[44] A. Formaglio, S. Belloni, G. Menegaz, H.Z. Tan, D. Prattichizzo, and M. Barni. Perceptibility of digital watermarking in haptically enabled 3D meshes. In *Proc. of Euro-Haptics Conference*, pages 407–412, 2006.

[45] N. Sakr, J. Zhou, N. D. Georganas, J. Zhao, and X. Shen. Prediction-based haptic data reduction and compression in tele-mentoring systems. In *Proc. of IEEE Instrumentation and Measurement Technology Conference*, pages 1828–1832, May 2008.

[46] L. Rabiner, M. Sambur, and C. Schmidt. Applications of a nonlinear smoothing algorithm to speech processing. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 23(6):552–557, Dec. 1975.

[47] X. Shen, J. Zhou, A. Hamam, S. Nourian, N.R. El-Far, F. Malric, and N.D. Georganas. Haptic-enabled telementoring surgery simulation. *IEEE Multimedia*, 15(1):64–76, 2008.

[48] J. Zhou, X. Shen, A. El Saddik, and N.D. Georganas. Design and implementation of haptic tele-mentoring over the internet. In *Proc. of IEEE International Symposium Distributed Simulation and Real-Time Applications*, pages 201–208, Oct. 2007.

[49] L. A. Jones and I. W. Hunter. Human operator perception of mechanical variables and their effects on tracking performance. *ASME Advances in Robotics*, 42:49–53, 1992.

[50] G.C. Burdea. *Force and touch feedback for virtual reality*. New York: Wiley, 1996.

[51] Mehrdad Zadeh, David Wang, and Eric Kubica. Perception-based lossy haptic compression considerations for velocity-based interactions. *Multimedia Systems*, 13(4):275–282, Nov. 2008.

[52] V. Nitsch, J. Kammerl, B. Faerber, and E.Steinbach. On the impact of haptic data reduction and feedback modality on quality and task performance in a telepresence and teleaction system. In *Haptics: Generating and Perceiving Tangible Sensations*, volume 6191, pages 169–176. Springer Berlin / Heidelberg, Cambridge, 2010.

[53] H. Pongrac, B. Farber, P. Hinterseer, J. Kammerl, and E. Steinbach. Limitations of human 3D force discrimination. In *Proc. of Human-Centered Robotics Systems*, pages 109–114, 2006.

[54] F. Janabi-Sharif, V. Hayward, and C.J. Chen. Discrete-time adaptive windowing for velocity estimation. *IEEE Transactions on Control Systems Technology*, 8(6):1003–1009, Nov. 2000.

[55] J.A. Rod Blais. Least squares for practitioners. *Mathematical Problems in Engineering*, 2010:1–19, 2010.

[56] H. Delingette and N. Ayache. Medical image modeling tools and applications: hepatic surgery simulation. *ACM Communications*, 48(2):31–36, Feb. 2005.

[57] N. Sakr, J. Zhou, N.D. Georganas, and J. Zhao. Prediction-based haptic data reduction and transmission in telementoring systems. *IEEE Transactions on Instrumentation and Measurement*, 58(5):1727–1736, May 2009.

[58] N. Sakr, J. Zhou, N.D. Georganas, J. Zhao, and E.M Petriu. Robust perception-based data reduction and transmission in telehaptic systems. In *Proc. of World Haptics, Third Joint EuroHaptics conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pages 214–219, Mar. 2009.

[59] P.G. Otanez, J.R. Moyne, and D.M. Tilbury. Using deadbands to reduce communication in networked control systems. In *Proc. of American Control Conference*, pages 3015–3020, May 2002.

[60] A. Galka. *Topics in nonlinear time series analysis: With implications for EEG analysis*. River Edge, NJ, Singapore, 2000.

[61] J. Kuffner. Effective sampling and distance metrics for 3D rigid body path planning. In *Proc. of IEEE International Conference on Robotics and Automation*, pages 3993–3998, May 2004.

[62] D.A. Keim. Information visualization and visual data mining. *IEEE Transactions on Visualization and Computer Graphics*, 8(1):1–8, 2002.

[63] M. Berthold and D.J. Hand. *Intelligent data analysis: An introduction*. New York: Springer-Verlag, 1999.

[64] M. Kreuseler and H. Schumann. A flexible approach for visual data mining. *IEEE Transactions on Visualization and Computer Graphics*, 8(1):39–51, 2002.

[65] J. Valdés and A. Barton. Virtual reality spaces for visual data mining with multiobjective evolutionary optimization: Implicit and explicit function representations mixing unsupervised and supervised properties. In *Proc. of IEEE Congress of Evolutionary Computation*, pages 5592–5598, 2006.

[66] J.L. Chandon and S. Pinson. *Analyse typologique. Théorie et applications*. Masson, Paris, 1981.

[67] I. Borg and J. Lingoes. *Multidimensional similarity structure analysis*. Springer-Verlag, 1987.

[68] J. Valdés. Virtual reality representation of information systems and decision rules: an exploratory technique for understanding data knowledge structure. In *Lecture Notes in Artificial Intelligence, ser. LNAI, vol. 2639. Springer-Verlag*, pages 615–618, 2003.

[69] K. Karhunen. Uber lineare methoden in der wahrscheinlichkeitsrechnung. *Annales Academiae Sciientiarum Fennicae, Series AI:Mathematica-Physica*, 37:3–79, 1946.

[70] M.M. Loève. *Probability Theory*. Princeton, NJ: Van Nostrand, 1955.

[71] J.W. Sammon. A non-linear mapping for data structure analysis. *IEEE Transactions on Computers*, C18:401–408, 1969.

[72] W. Pres, B. Flannery, S. Teukolsky, and W. Vetterling. *Numeric Recipes in C*. Cambridge University Press, 1992.

[73] C. Ferreira. Gene expression programming: A new adaptive algorithm for problem solving. *Journal of Complex Systems*, 13(2):87–129, 2001.

[74] C. Ferreira. *Gene Expression Programming: Mathematical Modeling by an Artificial Intelligence*. Springer Verlag, Germany, 2006.

[75] B. E. Boser, I. M. Guyon, and V. N. Vapnik. A training algorithm for optimal margin classifiers. In *Proc. of ACM Workshop on Computational Learning Theory*, pages 144–152, 1992.

[76] V. Vapnik. *Statistical Learning Theory*. Wiley, New York, 1998.

[77] R.O. Duda, P.E. Hart, and D.G. Stork. *Pattern Classification*. second ed. Wiley-Interscience, 2000.

[78] G. John and P. Langley. Estimating continuous distributions in bayesian classifiers. In *Proc. of Conference on Uncertainty in Artificial Intelligence*, pages 338–345, 1995.

[79] Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.

[80] Reachin Technologies ab. Reachin Display. http://www.reachin. se/products/.

[81] A.L. Blum and P. Langley. Selection of relevant features and examples in machine learning. *Artificial Intelligence*, 97(1-2):245–271, 1997.

[82] H. Liu, H. Motoda, R. Setiono, and Z. Zhao. Feature selection: An ever evolving frontier in data mining. *Journal of Machine Learning Research, Workshop and Conference Proceedings*, 10(10):4–13, 2010.

[83] I. Guyon and A. Elisseeff. An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3:1157–1182, Mar. 2003.

[84] Z. Pawlak. *Rough sets: Theoretical aspects of reasoning about data*. Kluwer Academic Publishers, Netherlands, 1991.

[85] A. Ohrn. *Discernibility and rough sets in medicine: tools and applications*. Ph.D. Thesis. Department of Computer and Information Science, Norwegian University of Science and Technology, Trondheim, Norway, 1999.

[86] J. Wróblewski. Ensembles of classifiers based on approximate reducts. *Fundamenta Informaticae*, 47(3–4):351–360, 2001.

[87] H.S. Nguyen and A. Skowron. Quantization of real-valued attributes. In *Proc. of International Joint Conference on Information Sciences*, pages 34–37, Sept. 1995.

[88] D. S. Johnson. Approximation algorithms for combinatorial problems. In *Proc. of ACM Symposium on Theory of Computing*, pages 38–49, 1974.

[89] S. Vinterbo and A. Ohrn. Minimal approximate hitting sets and rule templates. In *Predictive Models in Medicine: Some Methods for Construction and Adaptation, Department of Computer and Information Science, NTNU report*, Dec. 1999.

[90] S. Vinterbo and A. Ohrn. Minimal approximate hitting sets and rule templates. *International Journal of Approximate Reasoning*, 25(2):123–143, 2000.

[91] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and Ian H. Witten. The weka data mining software: an update. *SIGKDD Explorations*, 11(1):10–18, 2009.

[92] M. Gutlein, E. Frank, M. Hall, and A. Karwath. Large-scale attribute selection using wrappers. In *Proc. of IEEE Symposium on Computational Intelligence and Data Mining*, pages 332–339, Mar. 2009.

[93] M.A. Hall. *Correlation-based feature selection for machine learning*. Ph.D. dissertation, Dept. Comput. Sci., Waikato Univ., New Zealand, 1999.

[94] M. Paterson. *The senses of touch: Haptics, affects and technologies*. Berg Publishers, 2007.

[95] A. Alamri, M. Eid, R. Iglesias, S. Shirmohammadi, and A. El Saddik. Haptic virtual rehabilitation exercises for poststroke diagnosis. *IEEE Transactions on Instrumentation and Measurement*, 57(9):1876–1884, 2008.

[96] M.A. Srinivasan, G.L. Beauregard, and D.L. Brock. The impact of visual information on haptic perception of stiffness in virtual environments. In *Proc. of ASME Dynamic Systems and Control Division*, pages 555–559, 1996.

[97] W. Wu, C. Basdogan, and M.A. Srinivasan. Visual, haptic, and bimodal perception of size and stiffness in virtual environments. In *Proc. of ASME Dynamic Systems and Control*, pages 19–26, 1999.

[98] J.F. Norman, H.F. Norman, A.M. Clayton, J. Lianekhammy, and G. Zielke. The visual and haptic perception of natural object shape. *Perception & Psychophysics*, 66(2):342–351, 2004.

[99] G.L. Poling, J.M. Weisenberger, and T. Kerwin. The role of multisensory feedback in haptic surface perception. In *Proc. of International Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pages 187–194, 2003.

[100] J.M. Weisenberger and G.L. Poling. Multisensory roughness perception of virtual surfaces: Effects of correlated cues. In *Proc. of International Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pages 161–168, 2004.

[101] M.A. Heller. Visual and tactual texture perception: Intersensory cooperation. *Perception & Psychophysics*, 31(4):339–344, 1982.

[102] S.J. Lederman and S.G. Abbott. Texture perception: Studies of intersensory organization using a discrepancy paradigm, and visual versus tactual psychophysics. *Journal of Experimental Psychology:Human Perception & Performance*, 7(4):902–915, 1981.

[103] S.J. Lederman, G. Thorne, and B. Jones. Perception of texture by vision and touch: Multidimensionality and intersensory integration. *Journal of Experimental Psychology: Human Perception and Performance*, 12(2):169–180, 1986.

[104] K. Drewing, M.O. Ernst, S.J. Lederman, and R. Klatzky. Roughness and spatial density judgments on visual and haptic textures using virtual reality. In *Proc. of EuroHaptics Conference*, pages 203–206, 2004.

[105] S. Guest and C. Spence. What role does multisensory integration play in the visuotactile perception of texture? *International Journal of Psychophysiology*, 50(1–2):63–80, 2003.

[106] A. Kumazaki, K. Terada, and A. Ito. Role of vision on haptic length perception. In *Proc. of World Haptics, Second Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pages 336–341, 2007.

[107] L. Bouguila, M. Ishii, and M. Sato. Effect of coupling haptics and stereopsis on depth perception in virtual environments. In *Proc. of Workshop on Haptic Human-Computer Interaction*, pages 54–62, 2000.

[108] D. Swapp, V. Pawar, and C. Loscos. Interaction with co-located haptic feedback in virtual reality. *Virtual Reality Journal, Springer*, 10(1):24–30, 2006.

[109] S.A Wall, K. Paynter, A.M. Shillito, M. Wright, and S. Scali. The effect of haptic feedback and stereo graphics in a 3D target acquisition task. In *Proc. of EuroHaptics Conference*, pages 23–29, 2002.

[110] M. Congedo, A. Lécuyer, and E. Gentaz. The influence of spatial de-location on perceptual integration of vision and touch. *Presence: Teleoperators and Virtual Environments*, 15(3):353–357, 2006.

[111] G. Jansson and M. Ostrom. The effects of co-location of visual and haptic space on judgments of form. In *Proc. of EuroHaptics Conference*, pages 516–519, 2004.

[112] H. Levitt. Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, 49(2):467–477, 1971.

[113] D.C. Ruspini, K. Kolarov, and O. Khatib. The haptic display of complex graphical environments. In *Proc. of Annual Conference on Computer Graphics and Interactive Techniques*, pages 345–352, 1997.

[114] T.M. Massie and J.K. Salisbury. The phantom haptic interface: A device for probing virtual objects. In *Proc. of ASME Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pages 295–301, 1994.

[115] S.E. Maxwell and H.D. Delaney. *Designing experiments and analyzing data: A model comparison perspective*. Belmont, CA:Wadsworth, 1990.

[116] B.G. Tabachnick and L.S. Fidell. *Using multivariate statistics*. Fifth Edition. Boston:Pearson Education, 2007.