Recognition of Haptic Interaction Patterns in Dyadic Joint Object Manipulation

by

Çığıl Ece Madan

A Thesis Submitted to the
Graduate School of Sciences and Engineering
in Partial Fulfillment of the Requirements for
the Degree of

Master of Science

in

Mechanical Engineering

Koç University

April, 2014

Koç University

Graduate School of Sciences and Engineering

This is to certify that I have examined this copy of a master's thesis by

Çığıl Ece Madan

and have found that it is complete and satisfactory in all respects,

and that any and all revisions required by the final

examining committee have been made.

Committee Members:

Prof. Çağatay Başdoğan

Asst. Prof. Tevfik Metin Sezgin

Assoc. Prof. Erhan Öztop

Date:

to my family

ABSTRACT

This thesis aims to present a perspective to build a robot which can identify the interaction between its human partners and assist accordingly for physical human-robot cooperation. The development of robots that can physically cooperate with humans has attained interest in the last decades. Obviously, this effort requires a deep understanding of the intrinsic properties of interaction. Up to now, many researchers have focused on inferring human intents in terms of intermediate or terminal goals in physical tasks. On the other hand, working side by side with people, an autonomous robot additionally needs to come up with in-depth information about underlying haptic interaction patterns that are typically encountered during human-human cooperation. However, to our knowledge, no study has yet focused on characterizing such detailed information. In this sense, this work is novel as an effort to gain deeper understanding of interaction patterns involving two or more humans in a physical task. We present an expert-labeled human-human interaction dataset, which captures the interaction of two humans, who collaboratively transport an object in an haptics-enabled virtual environment. In the light of information gained by studying this dataset, we propose that the actions of cooperating partners can be examined under three interaction types: In any cooperative task, the interacting humans either 1) work in harmony, 2) cope with conflicts, or 3) remain passive during interaction. In line with this conception, we present a taxonomy of human interaction patterns; then propose five different feature sets, comprising force-, velocity- and power-related information, for the classification of these patterns. Our evaluation shows that using a multi-class support vector machine (SVM) classifier, we can accomplish a correct classification rate of 86 percent for the identification of interaction patterns, an accuracy obtained by fusing the selected features by Minimum Redundancy Maximum Relevance (mRMR) feature selection method.

ÖZETÇE

Bu tezin amacı, insanlar arasındaki etkileşimi anlamlandırabilen ve bu etkileşime göre onlara yardım eden bir robot tasarlanmasını sağlayacak bakış açısı geliştirmektir. Fiziksel olarak insanlarla birlikte çalışan robotların geliştirilmesi son yıllarda büyük ilgi çekmektedir. Şüphesiz ki, bu çaba insanlar arasındaki dokunma içeren fiziksel etkileşim özelliklerini anlamayı gerektirmektedir. Bu amaçla, bu zamana kadar, birçok araştırmacı insan niyetlerinin araştırılmasına ve bu niyetlerin robota aktarılmasına odaklanmıştır. Halbuki, insanlarla birlikte çalışan otonom robotların geliştirilmesi için fiziksel etkileşim örüntülerinin de incelenmesi gerekir. Etkileşim örüntüleri, sürekli ve tekrarlanan etkileşim nitelikleridir. Bu zamana kadar bu konu detaylı olarak araştırılmamıştır. Bu bağlamda, bu çalışma dokunma içeren fiziksel insan-insan etkileşimini anlamlandırmaya çabalayan ilk örnektir. Ayrıca bu çalışma vasıtasıyla, uzmanlar tarafından etiketlenmiş bir etkileşim veri seti de sunulmuştur. Bu veri seti, haptik (dokunsal) özellikli sanal bir dünyada obje taşıyan iki insanın etkileşimini içermektedir. Bu veri setinden elde edilen bilgiler ışığında, insan etkileşimi 3 etkileşim örüntü tipine ayırılmıştır. Herhangi bir işbirliği içinde, etkileşimde bulunan insanlar: 1) uyum içinde çalışırlar, 2) çelişkinin üstesinden gelmeye çalışırlar veya 3) etkileşim sırasında edilgen kalırlar. Bu fikir doğrultusunda, insan etkileşim örüntülerinin taksonimisini yapan ve etkileşim örüntülerini sınıflandıran bir makine öğrenme yöntemi geliştirilmiştir. Makine öğrenme yöntemini eğitmek için beş farklı öznitelik kümesi önerilmiştir. Bu öznitelik kümeleri kuvvet, hız ve güç bazlı öznitelikler olarak belirlenmiştir. Çok sınıflı sınıflayıcı (support vector machine SVM) kullanarak, insan etkileşim örüntülerinin tanınmasında %86 başarı oranı elde edilmiştir. Bu başarı oranı mRMR(Minimum Redundancy & Maximum Relavance) yöntemiyle füzyon edilen öznitelik kümesi kullanılmasıyla sağlanmıştır.

ACKNOWLEDGMENTS

Though only my name appears on the cover of this thesis, a great many people have contributed to its production. I owe my gratitude to all those people who have made this thesis possible and because of whom my graduate experience has been one that I will cherish forever.

First and foremost, I would like to express the deepest appreciation to my advisor Çağatay Başdoğan for his guidance throughout my studies at Koc University. I owe my deepest gratitude to him for his encouragement and enthusiasm in conducting great research. His wisdom, knowledge and commitment to the highest standards have inspired and motivated me. I appreciate all he has done for me, and those I will never forget. I will always recall him with his unique phrases which aim to motivate his students. From beginning of my undergraduate education, I have learned a great deal about professionalism, and academia, from him. I would like to thank Dr. Metin Sezgin for his guidance, understanding, and most importantly his patience during my graduate studies. Thanks for spending his time for discussing my work in detail and for developing my background in machine learning. I would like to thank Dr. Erhan Öztop for agreeing to be part of my thesis committee.

I am most grateful to Ayşe Küçükyılmaz whose advices and insight was invaluable to me, for her continuous help, guidance in all stages of this thesis and co-authoring me. I would also like to thank her for her friendship and her support .

I would like to express the deepest appreciation to my family, for always believing in me, for their continuous love and their supports in my decisions. Without whom I could not have made it here. In particular, the understanding and patience shown by them during my graduate study is greatly appreciated. I am deeply grateful to my mother who has been supporting me and encouraging me with her best wishes despite of my pessimistic perspective.

I would also like to thank Güven Yılmaz for his support, encouragement, caring, and constant patience. He has been always there cheering me up and stood by me through the good times and bad. He can manage to motivate me and enlighten my life and soul.

Finally, I would like to thank my friends for supporting me in many different ways. Thanks to Buket Baylan for being a close friend and the greatest housemate ever . It would have been a lonely lab without her. Thanks to Nazlı Noyan, Atakan Arasan, Yasemin Vardar, Senem Sancar, Senem Ezgi Emgin, Burak Özen, Banuçiçek Gürcüoğlu, Çağla Çığ and İpek Mumcu for their friendship and support. I am lucky to have such good friends. I would like to thank my colleagues: Berkay Yarpuzlu, Soner Cinoğlu, Enes Selman Ege, Ömer Şirin, Yusuf Aydın, Ozan Çaldıran, Nasser Arghavani, Mohammad Ansarin, Gökhan Nadar, Mustafa Yılmaz, Özem Kalay and Erelcan Yanık. Also I would like to thank people who spent their time to participate my user study .

This thesis is funded by TÜBİTAK-BİDEB 221O - National Scholarship Programme for MSc Students.

TABLE OF CONTENTS

List of Tables xi

List of Figures xii

Nomenclature xiii

Chapter 1: Introduction 1

Chapter 2: Related Work 5

Chapter 3: Experiment 9

3.1 Experimental Environment 9

3.2 Physics-Based Engine 11

3.3 Scenarios 13

3.4 Procedure and Participants 15

Chapter 4: A Taxonomy of Haptic Interaction Patterns 16

Chapter 5: Recognition of Haptic Interaction Patterns 21

5.1 Annotation of Interaction Pattens 22

5.2 Identification of Meaningful Features 22

5.2.1 Mean Magnitude of the Individual Forces Applied by the Agents . . 23

5.2.2 Mean Magnitude of the Net Force Applied by the Agents 24

5.2.3 Mean Magnitude of the Interactive Force Acting on the Object . . . 24

5.2.4 Mean Magnitude of the Linear Velocity of the Object 24

5.2.5 Mean Magnitude of the Angular Velocity of the Object 25

5.2.6 Mean Normalized Power Transferred by the Agents to the Object . 25
5.3 Dataset Generation and Feature Extraction 26
5.4 Classifier Design 28
5.5 Evaluation 28
5.5.1 Normalized Confusion Matrix: 28
5.5.2 Correct Classification Rate (Accuracy): 28
5.5.3 Balanced Error Rate (BER): 28

Chapter 6: Results and Discussion **30**

6.1 Classification Results Individual Feature Sets 30
6.2 Classification Results with the Combined Feature Set 32
6.3 Selection of the Optimal Feature Set 32

Chapter 7: Conclusions **37**

Chapter 8: Contributions and Future Directions **39**

Bibliography **41**

Appendix A: Parameters Used in the Experiment **45**

Appendix B: Physical Model of the Experiment **47**

Appendix C: Instructions Used in the Experiment **49**

LIST OF TABLES

5.1 Feature sets 27

A.1 Object and board information for both scenes(Section 3.1) 46
A.2 Spring and damper coefficients 46
A.3 Static and kinetic friction coefficient 46

LIST OF FIGURES

1.1 Daily collaboration scenario 2

3.1 Interaction between two humans in a virtual environment 10
3.2 The straight scene 11
3.3 The bifurcated scene 11
3.4 Physical model of the interaction between two humans 12
3.5 Scenarios 14

4.1 Taxonomy of interaction patterns in dyadic object manipulation 17
4.2 Force signals of agents for pattern classes C1, C2 and C3. 19
4.3 Force signals of agents for pattern classes C4, C5 and C6. 20

5.1 Stages of classifier learning. 22
5.2 Percentage of instances per interaction patterns 23
5.3 Mean values of variables for each pattern class. 25
5.4 Regions of supports 27

6.1 Classification results of individual feature sets. 30
6.2 Confusion matrices of classifiers trained with feature sets. 31
6.3 Classification accuracies for the feature sets 33
6.4 Confusion matrix of classifier trained with the optimal feature set. 34
6.5 Percentage of features from each region 34
6.6 Number and percentage of features 35

B.1 Detailed physical model of the interaction 48

NOMENCLATURE

ANOVA	**:**	Analysis of variance
BER	**:**	Balanced error rate
HIP	**:**	Haptic interface point
HMM	**:**	Hidden Markov Model
mRMR	**:**	Minimum Redundancy and Maximum Relevance
SVM	**:**	Support Vector Machine

Chapter 1

INTRODUCTION

The research presented throughout this thesis aims at recognizing the properties of interaction between two humans in order to provide insight into building intelligent robotic agents that can proactively and intuitively collaborate with humans under joint action in virtual environment. In particular, we discuss intrinsic properties of interaction during physical task to enhance human-robot collaboration. In this thesis, we will present extensive summary of the experimental study we have done to investigate interaction patterns, which remain latent under the interaction of two or more humans in a physical task.

With the emergence of the idea of autonomy in the robotics domain, a significant amount of research has shifted towards discovering how to make robots act in a more human-like manner in terms of their social, cognitive, and motor abilities. Significant attention is now directed towards building interactive and proactive robotic systems, which are capable of cooperating with humans in everyday situations instead of assisting with specific and possibly industrial tasks. In order to build cooperative robotic systems that allow natural and intuitive interaction, an understanding of human behavior and intentions, as well as a capability for communication and coordination is required. In this paper, we follow a human-centric experimental approach to discover human behavior characteristics in everyday physical tasks. We believe that the information extracted from the operation of two humans will be invaluable for developing a robotic *partner* that can effectively cooperate with humans.

Humans cooperate through numerous physical activities during their daily routines. These activities cover a wide range of tasks, such as jointly moving objects, assembling machine parts, hand shaking, and dancing. In its broader sense, cooperation addresses inter-

Figure 1.1: Daily collaboration scenario: two humans jointly carry a table.

action characteristics that provide mutual benefit to the partners. Thus we expect partners to work in harmony or at least without inhibiting the natural course of a given task. However, from time to time, the continuous and complex nature of physical tasks may necessitate partners to adopt some non-cooperative behaviors (i.e. conflicts). Imagine a couple, which has trouble in synchronizing their movements while dancing waltz. The conflict they face can be solved as soon as they manage to move along with the music simultaneously. Such conflicts -unintentional as they are- may be due to differences in partners' intentions or discrepancies in reaction times to each other's actions. Determining how and when interaction behaviors change is a key issue in understanding human collaboration.

A robot, which can comprehend how humans interact, would be able to either mimic the behaviors of one of the partners, or complement the interaction of humans as an assistant.

As a motivating example, think of a robotic system that aids two people with the installation of a rooftop car rack. The humans stand on both sides of the car and try to place the rack in the correct pose while the robotic system helps them with carrying the heavy load. In this case, humans do not act as dyads just because they need help from one another, but because dyadic interaction becomes the medium of communication. In this example, assume that the robot is not fitted with tools to determine where the rack should be installed, but is only capable of lifting the rack up or down as well as monitoring the interaction between humans. In this scenario, the task needs to be led by the humans. However the robotic system can effectively help in completing the task by speeding up the operation in the right direction when it recognizes harmony between the partners and stabilizing the rack when it infers a conflict between them. In other words, the robotic system recognizes the interaction patterns of the humans partners, and assists them as needed only.

This study is an effort to investigate interaction patterns in human-object-human scenarios, where two humans cooperate to move an object (see Fig. 1.1). We focus on dyadic joint object manipulation tasks to identify human interaction patterns when partners collaborate in the existence of conflicts[1]. In this sense, this study is a first step towards exploring how the partners' intentions change over the interaction in a physical task. In order to observe the interaction patterns of the partners, we designed four different dyadic object manipulation scenarios in a haptics-enabled virtual environment. Two of these scenarios were designed to promote collaboration between the partners without imposing any conflict on them, while the other two artificially invoke conflicts between the partners. Real human-human interaction data is collected through a controlled user study with 20 dyads. Through offline examination of this data, we observed that partners exhibit specific interaction patterns during joint operation. Specifically, we first defined three possible interaction types (harmonious, conflicting, and neutral) and then identified six interaction patterns based on the intentions of the dyad on the object. An expert observed the video recordings of the trials, which were captured during the experiment, and manually annotated the data. Through this process, we identified the meaningful parts of the collected data and labeled them with the aforementioned interaction patterns to form a labeled set of data for supervised learning.

We conducted a set of statistical analyses on the data in order to find descriptive variables that are used to recognize the interaction patterns. These descriptive variables are: 1) forces applied by individual agents on the manipulated object, 2) net force applied by the partners on the manipulated object, 3) interactive force among the partners, 4) velocity of the manipulated object, and 5) power transferred to the manipulated object by the partners. We have formed five different feature sets, four of which are composed of haptic information, by extracting features through taking the means, medians, standard deviations, and interquartile ranges of from these descriptive variables. For the recognition of interaction patterns, we used multi-class support vector machine (SVM) classifiers with these 5 feature sets. Inspecting the classification results, we observed that each individual feature set was successful in recognizing at least 4 of the 6 interaction patterns.

Even though the individual feature sets fail to recognize all interaction patterns, when they are used altogether to form a combined feature set, we achieve a correct classification rate of 84.2%. However, this set contains redundant and irrelevant information. In order to eliminate superfluous features, mRMR (Minimum Redundancy Maximum Relevance) feature selection method is used to determine the optimal feature set, which, in our case, consists of only 243 features out of 576. With this optimal feature set, we achieve an even higher recognition rate of 86%.

This paper is organized as follows: Chapter 2 presents the related work. The experimental setup used for data collection is described under Chapter 3. The interaction patterns observed in dyadic joint object manipulation and the proposed taxonomy are discussed in Chapter 4. The machine learning method that used for the classication of interaction patterns is explained in Chapter 5. The results and the discussion are presented in Chapter 6, followed by conclusions in Chapter 7, finally contributions and future directions are revisited in Chapter 8.

Chapter Notes

[1)]Note that even though we focus on dyadic interaction in this paper, the ideas we present can be easily extended to multiple human scenarios.

Chapter 2

RELATED WORK

Developing robots that can collaborate with human partners during physical interaction requires the robots to display proactive behavior. So far, the widespread approach to realize proactive behavior has been to improve the control schemes of the robots based on an estimation of human intentions.

In an early study, Rahman et al. programmed the robot to replay task-specific trajectories recorded in human-human experiments to generate human-like velocity trajectories in human-robot cooperation [21]. Later, Tsumugiwa et al. estimated human arm stiffness through the observation of measured position and forces, and adapted the admittance parameters accordingly [26]. Similarly, Duchaine and Gosselin implemented variable admittance based on the velocity and force derivative information obtained from the human [6]. Corteville et al. developed a human-inspired robotic assistant, which assumed that the humans follow a minimum jerk trajectory [9] during motion, and estimated the intended motions of the human partner based on his/her position and velocity profile [5]. The robot then adjusted its velocity profile to fit along with the intended velocity.

Alternatively, some other investigators have focused on role allocation and sharing in human-robot interaction. Evrard and Kheddar defined two distinct extreme behaviors (leader and follower) for partners and switched between the behaviors via two distinct and independently-varying functions [8]. Later, Kheddar illustrated the use of this mechanism during collaboration with a humanoid robot [13]. Similarly, Bussy et al. proposed a control law for physical interaction with a humanoid robot in an object transportation task [3]. Their control law enabled the robot to proactively switch between standalone (i.e. performing the task alone) and collaborative (i.e. leader or follower) roles depending on the intentions of its human partner. Oguz et al. [19] and Kucukyilmaz et al. [14, 15] proposed a method to infer the intentions of the human during a joint object manipulation task. They implemented a dynamic role exchange model, where the robot inferred human's intentions based on the forces applied by him/her. Depending on the inferred intentions, the robot chose between leader or follower roles. Later, Moertl et al. presented a similar dynamic role exchange mechanism for a joint object manipulation task, in which a man-sized mobile robot sensed the human partner's intentions through the evaluation of an agreement criterion based on the human's force input, and helped accordingly [18]. These studies are effective in enhancing human-robot interaction via generating more natural trajectories. However, the rule based nature of the control laws utilized in these studies makes it difficult to generalize them for different tasks. Furthermore, even though the robots are capable of adapting to their human partners, they lack the ability to comprehend how human behaviors change during interaction, and what these changes signify.

A widely accepted perspective advocates the investigation of human-human interaction to learn from the behavioral mechanisms utilized by humans. Based on the insight gained from human-human interaction data, Reed and Peshkin illustrated that two opposing intentions, to accelerate or to decelerate, exist in a dyadic target acquisition task [22]. Similarly, Stefanov et al. specified conductor and executor roles, which bear information about how two humans cooperate in a joint manipulation task [24]. They presented a model based on velocities and the interaction forces applied through haptic devices to define the roles. Groten et al. focused on the consistency of dominance behavior during a tracking task where two humans collaborated with each other [10]. They demonstrated that the participants' interaction can be represented with a personal dominance distribution. Later they investigated how partners communicate through intentions, and suggested that in order to achieve a joint goal, partners need to integrate their individual action plans in both collaborative and conflicting situations [11]. Even though these studies adopt a similar approach to that of ours, in a sense that they examine human-human interaction data, they are inherently different. All these studies focus on presenting the existence of different patterns in human behaviors, however, none of them attempt a systematic classification of these behaviors using machine learning techniques. Additionally, they mainly define individual labels for human intentions, but do not focus on how partners work with each other over time.

In order to address this shortcoming, some researchers have used statistical learning models to infer about the intentions of the human partner. Evrard et al. implemented a learning-by-demonstration technique [2] to differentiate between leader and follower roles [7]. Their system was able to capture the role switching moments using Gaussian Mixture Models. Takeda et al. [25] and Wang et al. [27] proposed HMM based algorithms to estimate human intentions in physical dyadic tasks, where a robot collaborated proactively with its human partner. Schrempf et al. presented a new approach that allows a robot to plan its actions even if the human intention estimation was uncertain [23]. In their system, the robot computed a confidence for possible actions and executed the task by selecting actions proactively. Even though these studies present task-independent solutions to intention recognition, they fall short in interpreting the meaning of the intentions and the interaction patterns.

Characterization of interaction patterns is an emerging topic in human-human and human-robot interaction domains. As the name implies, interaction patterns describe the interaction between agents, not the behavior of an individual. In this sense, it provides a different perspective to the same problem. There are a few studies in literature, which mainly focus only on identifying a taxonomy of interaction patterns and performing task-dependent classification. Jarrasse et al. has introduced a taxonomy of interaction patterns in physical tasks recently [12]. They described a general taxonomy of human-robot interaction patterns and defined controllers for each pattern. The proposed framework provided a description of interaction patterns of a dyad executing a joint task, along with an interpretation of the patterns. Even though the utility of this taxonomy was demonstrated by simulated interactions of two humans, it lacks the identification of patterns in real data. Melendez-Calderon et al. defined five human interaction patterns in a tracking task where two humans worked together [17]. The patterns were defined as templates, which indicate the action of each partner, such as one agent accelerating the movement while the other is braking. They proposed a rule-based classification system using the interaction torques and EMG recordings of partners' activities to identify these patterns. However, their technique is highly task-dependent. Besides, it requires manual construction of templates and a lot of fine tuning when the task dynamics changes. Furthermore, the system is not robust against the addition of new interaction strategies. On the contrary, the classification method proposed in this paper aims at discovering the descriptive features of interaction, hence, given training data, our technique can be applied to a diverse set of tasks.

Even though the aforementioned studies provide valuable knowledge about human interaction patterns, to our knowledge, no effort has yet been put into defining a systematic way of defining and recognizing these patterns. In this sense, our work is a first to both present a taxonomy and propose a recognition framework for real human-human interaction data.

Chapter 3

EXPERIMENT

We conducted an experimental study to generate data that can be used to identify human-human haptic interaction patterns and learn models for capturing salient characteristics of dyadic interactions. This section presents the experimental design and the scenarios used in this study, as well as the physics-based engine underlying the virtual environment and the procedures.

3.1 Experimental Environment

In order to identify human interaction patterns, we have developed an application where two human subjects interact in a virtual environment through the haptic channel. Our setup requires the subjects to be situated in different rooms, so they only interact through haptic devices.

The application requires the subjects to coordinate their actions in order to move the rectangular object together in a 2D maze-like scene (see Figs. 3.1(a) and 3.1(b)). Due to the selection of friction coefficients, the object rotates easily within the environment, resembling the motion of a table moving on four caster wheels. The goal of the task is to move the object toward a target parking configuration and stay there for a predetermined period of 5 seconds. However, subjects may have the same or different targets but they are not informed about their partner's target (see Appendix C).

During the experiment, the subjects are presented with two different scenes to observe interaction patterns in both translational and rotational motion. The first scene, which will be called the *straight scene* from now on, depicts a horizontal path, whereas the second scene, called the *bifurcated scene*, presents a fork-shaped path for the users to follow. Obviously, the straight scene involves translation along a straight line, while the bifurcated

(a) Screens of Agent 1 and 2

(b) Agent 1 and 2

Figure 3.1: Two humans interact through haptic devices in order to jointly move an object in a virtual environment.

scene entails both translation and rotation. Screenshots of the application for each scene can be seen in Figs. 3.2 and 3.3.

As seen in Figs. 3.2 and 3.3, the jointly manipulated object is depicted as a pink rectangular block. The grasping points of agents are represented as blue and green spheres attached to the short edges of the object. The target is visually represented with a green rectangle that resembles the object and clearly conveys the desired orientation for parking. Once the object reaches the target configuration, the target turns blue and a counter appears in the middle of the screen to alert the user. If the user succeeds in staying on the target until the end of the countdown, a new target appears somewhere else in the scene. In both

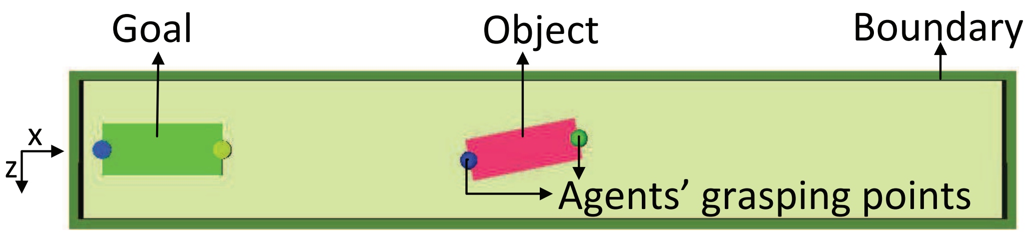


Figure 3.2: The straight scene

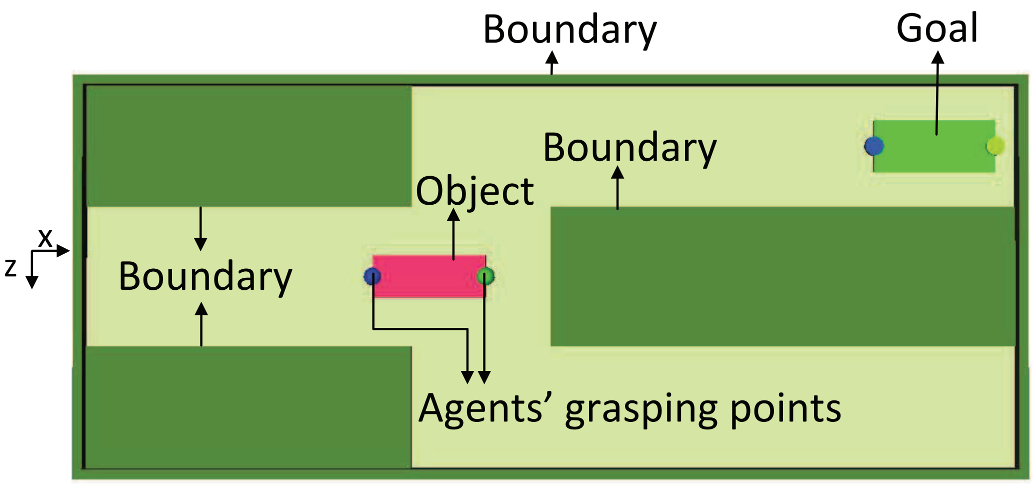


Figure 3.3: The bifurcated scene

scenes, boundaries constrain the movement of the user. Hitting the boundaries during the task is considered an error. In order to signal such errors to the users, the boundaries turn red on object collision.

3.2 Physics-Based Engine

This section details the physics-based engine underlying the virtual environment. Note that bold-face symbols are used to denote vectors throughout the section.

The manipulated object is modeled as a rigid body that moves in 2D, in a way similar to the movement of a table moving on four caster wheels. The physics based engine conveys the dynamic nature of the joint manipulation task to the agents both visually and through haptics. The agents interact with the environment via haptic devices. The end-effector

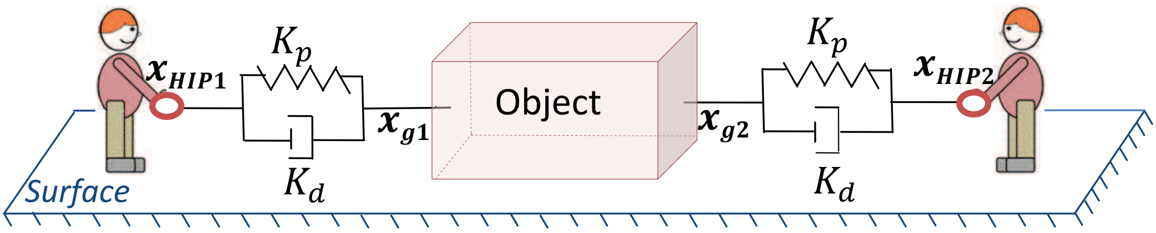


Figure 3.4: HIPs are connected to the object with spring/damper systems. K_p and K_d are spring and damper coefficients and $\dot{\boldsymbol{x}}_{obj}$ is velocity of the object.

positions of haptic styli along x- and z-axes map to the positions of the individual *haptic interface points* (HIPs). A spring and damper model is used between each agent's HIP and the grasping point on the object as shown in Fig. 3.4. The model is used to calculate the individual forces applied by the agents on the object:

$$\boldsymbol{F}_{HIP_1} = K_p(\boldsymbol{x}_{HIP_1} - \boldsymbol{x}_{g_1}) + K_d(\dot{\boldsymbol{x}}_{HIP_1} - \dot{\boldsymbol{x}}_{g_1}) \tag{3.1}$$

$$\boldsymbol{F}_{HIP_2} = K_p(\boldsymbol{x}_{HIP_2} - \boldsymbol{x}_{g_2}) + K_d(\dot{\boldsymbol{x}}_{HIP_2} - \dot{\boldsymbol{x}}_{g_2}) \tag{3.2}$$

where K_p and K_d are spring and damper coefficients (see Appendix A), $\boldsymbol{x}_{HIP_2}$, $\dot{\boldsymbol{x}}_{HIP_1}$, $\dot{\boldsymbol{x}}_{HIP_2}$ are the positions and velocities of HIPs, and $\boldsymbol{x}_{g_1}, \boldsymbol{x}_{g_2}, \dot{\boldsymbol{x}}_{g_1}, \dot{\boldsymbol{x}}_{g_2}$ are the positions and velocities of the grasping points of the agents.

Reciprocally, the agents are fed back with forces $-\boldsymbol{F}_{HIP_1}$ and $-\boldsymbol{F}_{HIP_2}$ through the haptic devices, so that they can feel the dynamics of the object[a)].

In addition to the applied forces, in case the object collides with the boundaries, an impact force, $\boldsymbol{F_I}$ is applied on the object to prevent penetration of the object into the boundaries. Furthermore, since the object acts as a rigid body, forces acting on it generate moments. The moments about the y-axis due to the forces applied by the agents and the impact force are respectively calculated by:

$$\begin{aligned} M_{HIP_u} &= \boldsymbol{l_u} \times \boldsymbol{F}_{HIP_u} \quad u = 1,2 & (3.3)\\ M_I &= \boldsymbol{d} \times \boldsymbol{F}_I \quad , & (3.4)\end{aligned}$$

where l_u, $u = 1,2$ denote the distance between the agent's grasping points and the center of mass of the object and d is the distance between the collision point at the boundary and the center of mass of the object. The object is also affected by frictional forces due to its contact with the surface. Translational and rotational friction ($\boldsymbol{F}_{\boldsymbol{f}}$ and M_f) are calculated using the Coulomb friction model[b)]. Thus, the net force and moment acting on the object becomes:

$$\boldsymbol{F}_{net} = \boldsymbol{F}_{HIP_1} + \boldsymbol{F}_{HIP_2} + \boldsymbol{F}_I + \boldsymbol{F}_f \tag{3.5}$$

$$M_{net} = M_{HIP_1} + M_{HIP_2} + M_I + M_f \quad . \tag{3.6}$$

The state of the object at each time step ($\boldsymbol{x}_{obj}$, $\dot{\boldsymbol{x}}_{obj}$, Θ_{obj}, $\dot{\Theta}_{obj}$) is calculated from M_{net} and $\boldsymbol{F}_{net}$ using Euler integration.

3.3 Scenarios

In order to elicit different interaction patterns, we presented the subjects with different manipulation scenarios, in which conflicts between partners are artificially invoked by providing each agent with different visual information about the location of the target configuration. Apart from the target locations, both subjects observe the motion of the object and view the same path. The subjects are not aware of the whereabouts of their partner's target, but they are informed that it can be different from that of their own, or either they or the other agent might not be given a target at all (see Appendix C).

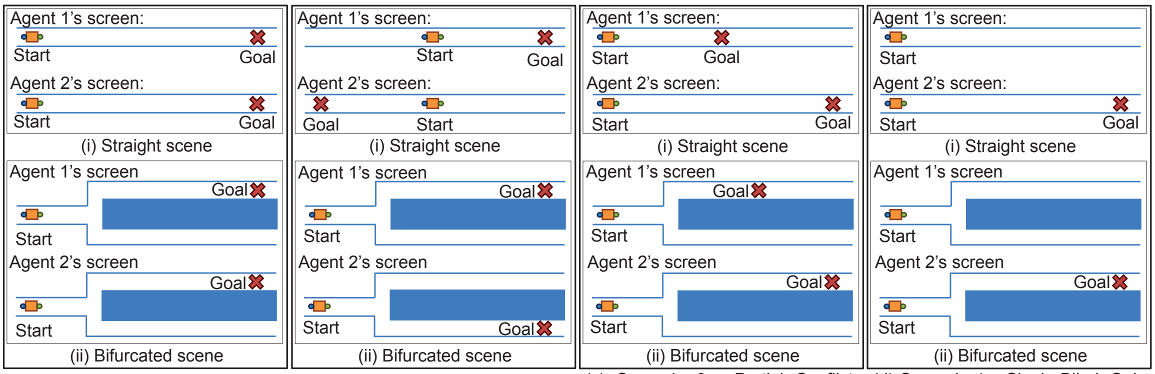
The following manipulation scenarios are considered in the experimental study:

Scenario 1: Harmony

In this scenario, both subjects are given the same target. Hence, we expect no conflict in terms of final goals. Fig. 3.5(a) represent the screen visual shown to each subject for both straight and bifurcated scenes.

Scenario 2: Full Conflict

The subjects are presented with conflicting goals in this scenario. The target configurations are arranged so that only one of them can be achieved at the end of the task. As a result,



(a) Scenario 1 - Harmony: The agents are provided with the same goals, which lie at the farthest end of the path.

(b) Scenario 2 – Full Conflict: The agents have conflicting goals, which lie (i) at opposing ends of the corridor for the straight path and (ii) at the end of different branches of the bifurcated path.

(c) Scenario 3 - Partial Conflict: Both agents' goals are on the same path, however one of the agent's goal is closer.

(d) Scenario 4 – Single Blind: Only one agent is provided with a goal at the farthest end of the path, whereas the other agent does not see any goal on screen.

Figure 3.5: Four scenarios in straight and bifurcated scenes.

one of the subjects needs to yield to the authority of the other in order to accomplish the task. Fig. 3.5(b) shows the screen visual shown to each subject for both scenes.

Scenario 3: Partial Conflict

Similar to the previous scenario, conflicting targets are given to subjects. The achievement of both tasks is not possible, yet the conflict manifests itself later during the trial and the amount of conflict is expected to be less than that of Scenario 2. Fig. 3.5(c) represent the screen visual shown to the subjects for both scenes.

Scenario 4: Single Blind

In this scenario, only one subject is assigned a goal. The other subject (i.e. the blinded subject) is informed that s(he) needs to observe and follow the actions of his/her partner. It is possible to accomplish the task, but the blinded subject is expected to get confused. Fig. 3.5(d) represents the screen visual shown to the subjects for both scenes. Note that in this figure, the blinded subject is agent 1, however a dual scenario, where agent 2 acts as the blinded subject, is also considered in the experiments.

3.4 Procedure and Participants

40 subjects (6 female and 34 male), aged between 21 and 29, participated in our study. The subjects were separated in two different rooms, so that they could not see or hear each other. They interacted with the object and each other through Geomagic® (formerly Sensable®) Phantom® Premium™ haptic devices using a stylus attachment. The haptic devices were connected to separate PCs and communicated through a UDP connection over the local area network.

At the beginning of the experiments, each participant was presented with the same goals (i.e. Scenario 1) for two practice trials in order to familiarize him/her with the system. During the experiments, each manipulation scenario was presented twice, hence, there were a total of 10 trials[c)] to be analyzed. In order to balance the learning effects, the order of the scenarios were permuted using a Latin square design. The subjects were not given detailed descriptions of the scenarios or the interaction patterns, but they were informed that their partners may have conflicting goals or no goal at all (see Appendix C).

Chapter Notes

[a)] Due to mechanical constraints, the forces fed back to the humans are thresholded at 2.8 N.

[b)] During the experiments the values of the static and kinetic friction coefficients for the translational and rotational cases are given in Appendix B.

[c)] Note that Scenario 4 was presented in a twofold fashion so that each agent gets to act as the blinded user within the experiment.

Chapter 4

A TAXONOMY OF HAPTIC INTERACTION PATTERNS

Based on our interpretations of user interactions after the experiments, we have identified a set of interaction patterns that were observed frequently in dyadic object manipulation task. They constitute our taxonomy of human interaction patterns as illustrated in Fig. 4.1. The interaction patterns, which are commonly encountered in dyadic joint object manipulation can be classified under three main types of interaction:

1. Harmonious Interaction:

The partners move the object while agreeing on the direction of the movement. In other words, the intention of both agents are the same; thus, no conflict exists between the agents. We examine this interaction type in two subclasses:

a) **Common intention to start/continue motion:** The acceleration of the manipulated object is greater than or equal to zero.

   i) **Harmonious translation (C1):** The partners agree on translating the object. In other words, both agents apply forces in the same direction to translate the object. Force signals of agents in the object frame during interaction pattern C1 segment are presented in Figs. 4.2(a) and 4.2(d). Fig. 4.2(a) demonstrates the forces of agents in the direction along the motion. It is seen that forces of agents in the same direction.

   ii) **Harmonious rotation with translation (C2):** The partners voluntarily rotate the object by agreeing on moving it along an arc or about its center. Force signals of agents in the object frame during interaction pattern C2 segment are presented in Figs. 4.2(b) and 4.2(e). Fig. 4.2(e) demonstrates the forces of

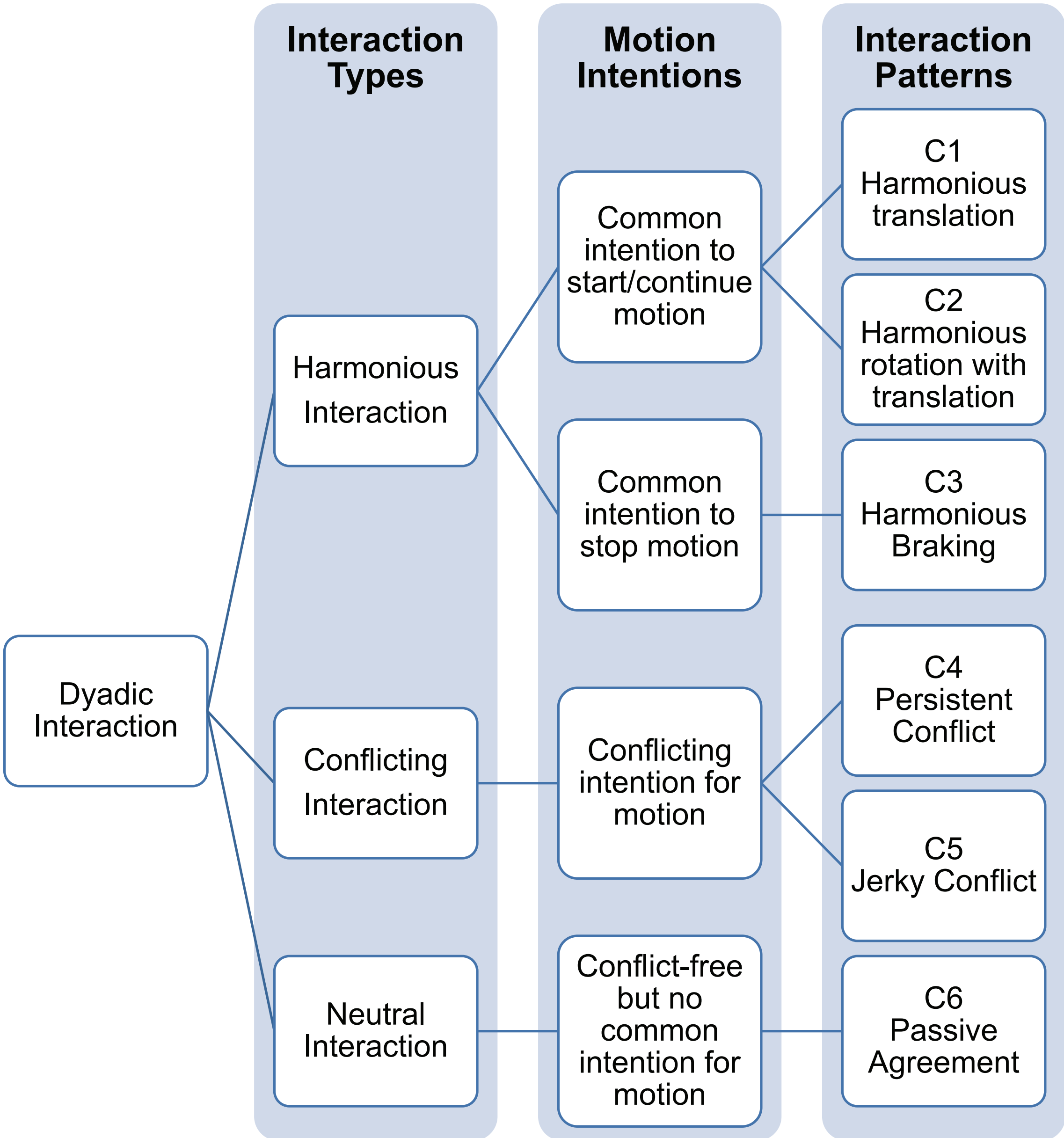


Figure 4.1: Taxonomy of interaction patterns in dyadic object manipulation

agents in the direction perpendicular to the motion. According to this figure, agents rotate the object to the same direction by applying opposing forces perpendicular to the motion.

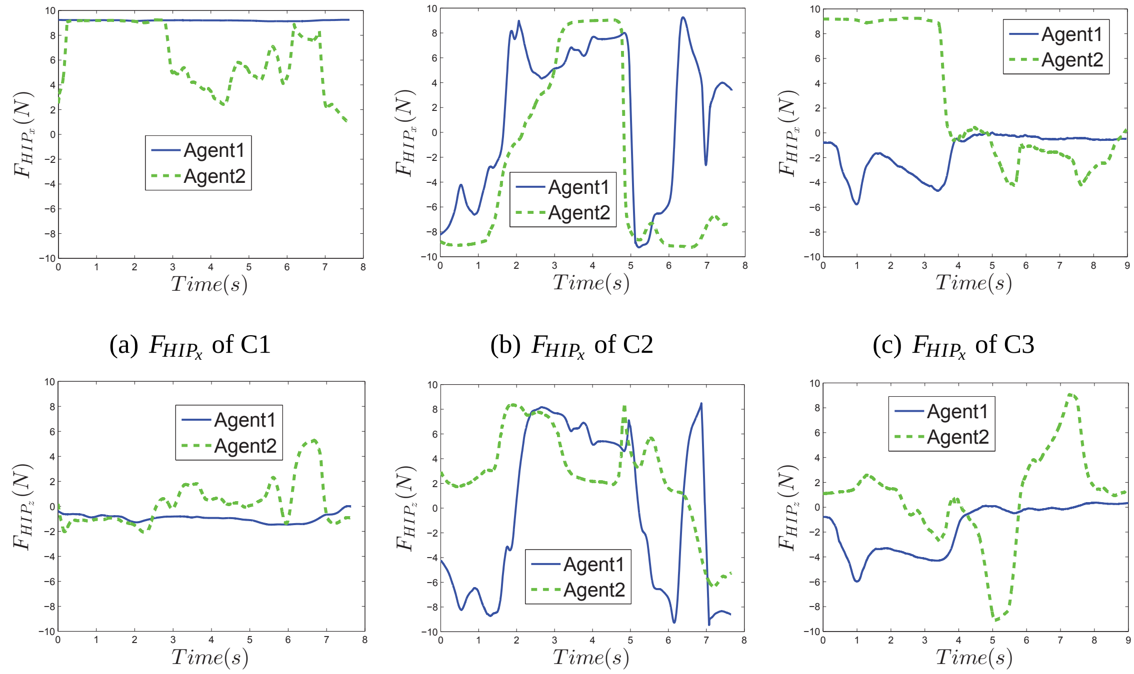
b) **Harmonious braking (C3):** The acceleration of the object is negative. In this case, one or both partners voluntarily decelerate the object with the purpose of stopping the motion. In practice, at least one agent starts applying force in the direction opposite to the movement until the object is stationary. Force signals of agents in the object frame during interaction pattern C3 segment are presented in Figs. 4.2(c) and 4.2(f). Fig. 4.2(c) demonstrates the forces of agents in the direction along the motion. It is seen that agents agree to decelerate the object by applying force in the opposite direction of movement.

2. Conflicting Interaction:

The interaction is dominated by some form of conflict between the agents. In other words, the partners have no common intention for motion. In this type of interaction, we expect that the partners can neither move the object smoothly nor achieve their goal. Two patterns can be defined for this interaction type:

i) **Persistent conflict (C4):** The partners insists on moving the object in opposite directions and hence the object does not move much. Force signals of agents in the object frame during interaction pattern C4 segment are presented in Figs. 4.3(a) and 4.3(d). Fig. 4.3(a) demonstrates the forces of agents in the direction along the motion. It is seen that agents apply opposing forces with approximately same magnitude.

ii) **Jerky conflict (C5):** The users disagree on the movement of the object, but not in a persistent fashion. This typically causes the object to rotate involuntarily or follow undesired trajectories, possibly ending with collisions with the environment. In more general terms, this pattern can be thought of any apparent conflict between agents that is not persistent. Force signals of agents in the object frame during interaction



(a) F_{HIP_x} of C1 (b) F_{HIP_x} of C2 (c) F_{HIP_x} of C3

(d) F_{HIP_z} of C1 (e) F_{HIP_z} of C2 (f) F_{HIP_z} of C3

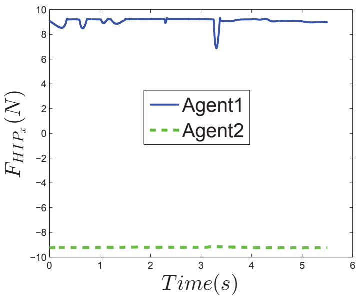
Figure 4.2: Agent force signals for pattern classes C1, C2 and C3 are presented in the object frame. The x direction presents the direction along the motion, whereas z direction presents the direction perpendicular to the motion. C1: Harmonious translation, C2: Harmonious rotation with translation, C3: Harmonious braking.

pattern C5 segment are presented in Figs. 4.3(b) and 4.3(e). It is seen from these figures that there is rapid changes in the forces of agents.

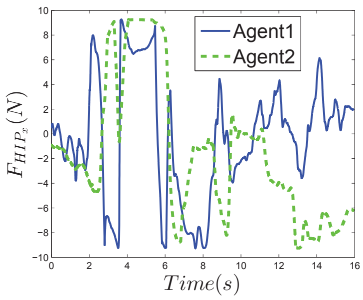
3. Neutral Interaction:

This interaction type implies no conflict between the partners. However, the agents share no common intention for the motion. This interaction type is mainly governed by an agent being passive, and defines a single interaction pattern:

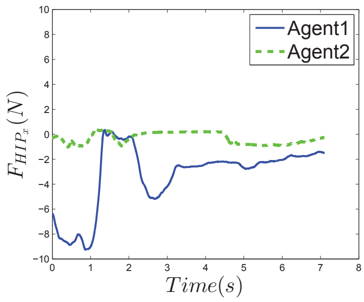
i) **Passive agreement (C6):** At least one of the partners remains passive by not contributing much to the task. Force signals of agents in the object frame during interaction pattern C6 segment are presented in Figs. 4.3(c) and 4.3(f). Fig. 4.3(c) demonstrates the forces of agents in the direction along the motion. It is seen that one of the agents remains passive and does not apply significant amount of force to move the object.



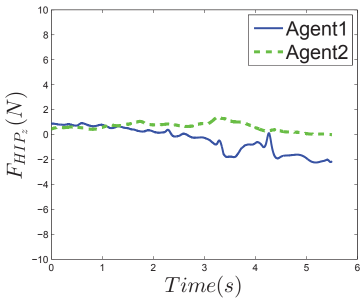
(a) F_{HIP_x} of C4



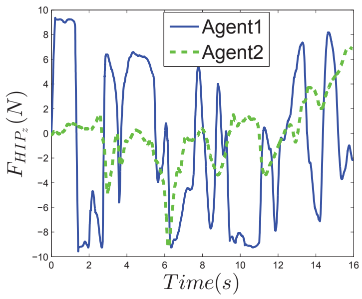
(b) F_{HIP_x} of C5



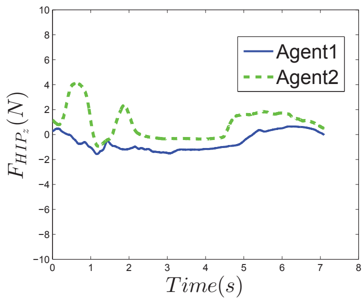
(c) F_{HIP_x} of C6



(d) F_{HIP_z} of C4



(e) F_{HIP_z} of C5



(f) F_{HIP_z} of C6

Figure 4.3: Agent force signals for pattern classes C4, C5 and C6 are presented in the object frame. The x direction presents the direction along the motion, whereas z direction presents the direction perpendicular to the motion. C4: Persistent conflict, C5: Jerky conflict, C6: Passive agreement.

Chapter 5

RECOGNITION OF HAPTIC INTERACTION PATTERNS

Our statistical pattern classification system possesses the structure given in Fig. 5.1. First, raw data is annotated by an expert to obtain a set of meaningful labeled interaction segments. Then, in order to avoid over-fitting, the data is split into 3 distinct parts, namely training, validation, and test sets. The training and validation sets are used to estimate parameters of the classifier, while the test set is used to assess the performance of the fully trained classifier. Then, features are extracted for each dataset, and model training is performed. Once the SVM is trained with the optimal parameters, it is used for the classification of patterns. The steps of our learning procedure is as follows:

1. Annotate raw data
2. Divide the data into training, validation, and test sets
3. Extract features from training, validation, and test sets
4. Select model parameters
5. Train the model using the training set
6. Evaluate the model using the validation set
7. Repeat steps 4 - 6 with different model parameters
8. Select the best parameters and train the model using the training and validation sets
9. Assess the final model using the test set

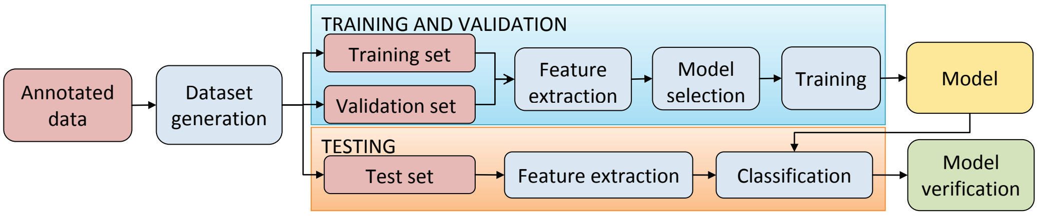


Figure 5.1: Stages of classifier learning.

5.1 *Annotation of Interaction Pattens*

After the experiment, we generated videos of the trials by simulating the recorded data in Matlab®environment. Regarding the videos, we manually annotated the data with the interaction behaviors using the ELAN annotation tool for annotating digital audio and video [1]. At the end of the annotation process, variable-length labeled interaction segments were obtained.

After annotation, we get a highly unbalanced dataset. The percentage of instances per interaction pattern class is shown in Fig. 5.2. The number of instances are particularly small in *Harmonious Rotation with Translation (C2)*, *Harmonious Braking (C3)*, and *Persistent Conflict (C4)* classes. The small number of instances for the C2 class can be explained due to the lack of need for rotation in the straight scene. For C3 and C4, the persistent nature of the patterns leads to longer continuous segments, which eventually stand as a single instance regardless of the length of the interaction.

5.2 *Identification of Meaningful Features*

The success of any pattern recognition system relies on the presence of informative features. At the end of the annotation process, we obtain a bulk set of labeled data, consisting of the agents' forces as well as variables related to the movement of the object, such as its position, orientation, linear and angular velocity, and acceleration. In order to infer which of the collected variables can be used for the recognition of interaction patterns, we conduct statistical analyses. We compute the means of several variables, and investigate their

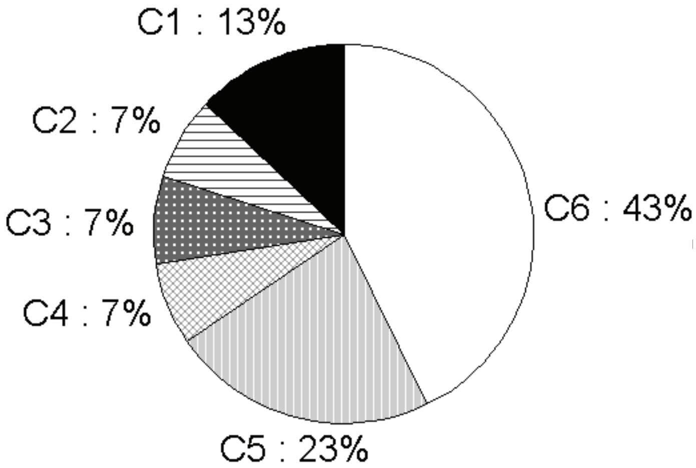


Figure 5.2: Percentage of instances per interaction pattern class in the dataset. C1: Harmonious translation, C2: Harmonious rotation with translation, C3: Harmonious braking, C4: Persistent conflict, C5: Jerky conflict, C6: Passive agreement.

descriptive power through one-way ANOVAs. We infer that statistically significant effects ($p < 0.001$) indicate descriptive features. Obviously, statistically significant differences in the feature values across classes does not necessarily imply high recognition accuracies during classifications. The predictive classification accuracies for each feature set are further discussed in Section 5.5. As a result of our analysis, six different descriptive variables are detected. Fig. 5.3 illustrates the means and the standard errors of means for each pattern class for each of the following variables:

5.2.1 Mean Magnitude of the Individual Forces Applied by the Agents

Individual forces exerted by the subjects ($\boldsymbol{F}_{HIP_1}$ and $\boldsymbol{F}_{HIP_2}$) are averaged over the duration of the interaction:

$$MF_{HIPs} = \frac{1}{2T} \sum_{u=1}^{2} \sum_{t=1}^{T} \|\boldsymbol{F}_{HIP_u}\| \quad , \tag{5.1}$$

where T is the length of the interaction sequence.

5.2.2 Mean Magnitude of the Net Force Applied by the Agents

The net force is the vector sum of the agent forces applied on the manipulated object. The mean magnitude of the net force exerted by the agents is calculated by:

$$MF_{net} = \frac{1}{T}\sum_{t=1}^{T} \|\boldsymbol{F}_{HIP_1} + \boldsymbol{F}_{HIP_2}\| \quad . \tag{5.2}$$

5.2.3 Mean Magnitude of the Interactive Force Acting on the Object

The interactive force, f_i acting on the object reflects the internal force that acts on the object. Interactive force is defined in the redundant task space [16] and occurs if the agents apply "compressive or tensile forces that do not contribute to the motion of the object" [10]. Interactive force is defined as:

$$f_i = \begin{cases} F_{HIP_{1x}} & \mathrm{sign}(F_{HIP_{1x}}) \neq \mathrm{sign}(F_{HIP_{2x}}) \\ & \wedge |F_{HIP_{1x}}| \leq |F_{HIP_{2x}}| \\ -F_{HIP_{2x}} & \mathrm{sign}(F_{HIP_{1x}}) \neq \mathrm{sign}(F_{HIP_{2x}}) \\ & \wedge |F_{HIP_{1x}}| > |F_{HIP_{2x}}| \\ 0 & \mathrm{sign}(F_{HIP_{1x}}) = \mathrm{sign}(F_{HIP_{2x}}) \quad , \end{cases} \tag{5.3}$$

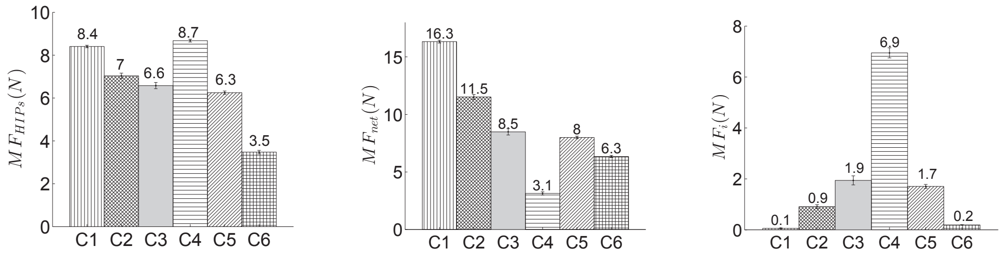
where $F_{HIP_{1x}}$ and $F_{HIP_{2x}}$ stand for the x components of the agent's applied forces in the object frame. The mean magnitude of the interactive force acting on the object (MF_i) is calculated as:

$$MF_i = \frac{1}{T}\sum_{t=1}^{T} |f_i| \quad . \tag{5.4}$$

5.2.4 Mean Magnitude of the Linear Velocity of the Object

The mean magnitude of the linear velocity of the object is calculated as follows:

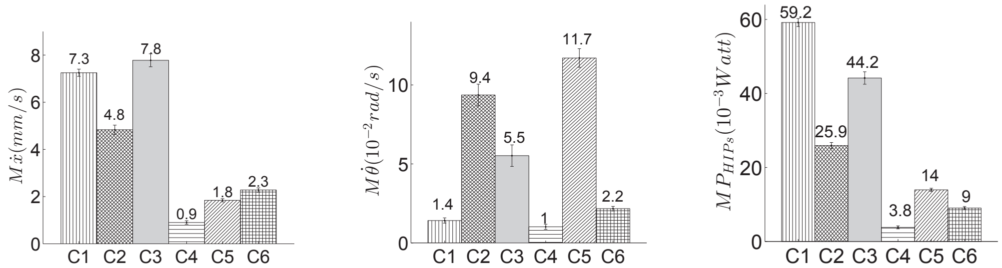
$$M\dot{x}_{obj} = \frac{1}{T}\sum_{t=1}^{T} \|\dot{\boldsymbol{x}}_{obj}\| \quad . \tag{5.5}$$



(a) Mean magnitude of the individual forces applied by the agents

(b) Mean magnitude of the net force applied by the agents

(c) Mean magnitude of the interactive force acting on the object



(d) Mean magnitude of the linear velocity of the object

(e) Mean magnitude of the angular velocity of the object about the y-axis

(f) Mean normalized power transferred by the agents to the object

Figure 5.3: Mean values of variables for each pattern class. The error bars indicate standard errors of the means. C1: Harmonious translation, C2: Harmonious rotation with translation, C3: Harmonious braking, C4: Persistent conflict, C5: Jerky conflict, C6: Passive agreement.

5.2.5 Mean Magnitude of the Angular Velocity of the Object about the y-axis

The mean magnitude of the angular velocity of the object ($\dot{\theta}_{obj}$) about the y-axis is calculated as follows:

$$M\dot{\theta}_{obj} = \frac{1}{T}\sum_{k=1}^{T}\left|\dot{\theta}_{obj}\right| \quad . \tag{5.6}$$

5.2.6 Mean Normalized Power Transferred by the Agents to the Object

The power transferred by agents to the object is calculated as follows:

$$P_{HIP_u} = \int_P \left(\boldsymbol{F}_{HIP_u} \cdot \mathrm{d}\boldsymbol{x}_{obj} + \left|M_{HIP_u}\mathrm{d}\theta_{obj}\right|\right), u = 1, 2, \tag{5.7}$$

where P is the path traversed by the object during the interaction segment. Keeping this in mind, the mean normalized power transferred by the agents to the object (MP_{HIPs}) is calculated as:

$$MP_{HIPs} = \frac{1}{2T}\sum_{u=1}^{2} P_{HIP_u} \quad , \tag{5.8}$$

5.3 Dataset Generation and Feature Extraction

The annotation process results in variable length interaction segments. However, in order to be used in classification, we need to represent the data using a fixed number of features for each annotated interaction segment. We compute features over 12 different regions of support (R1, R2, ..., R12). Time interval of regions (t_{int}), which indicates the beginning and end time of regions, are presented as:

$$t_{int} = \begin{cases} \frac{T}{2} \pm \frac{T}{2} & R1 \\ 0 \pm 500ms & R2 \\ T \pm 500ms & R3 \\ \frac{iT}{k} \pm 500ms & R4, 5, .., 12 \quad , \end{cases} \tag{5.9}$$

where T is the length of interaction segment, $2 < k < 5$ and $1 < i < k-1$. These regions of support are sampled over: 1) whole interaction segment, 2) the beginning and end of the segment, 3) different positions along the segment (see Fig. 5.4). R1 is sampled over whole interaction segment, while R2 and R3 are sampled over the beginning and end of the segment. And R4, R5,..., R12 are sampled over different positions along the segment. Same number of features are computed from each region of support.

We compute the mean, median, interquartile range, and standard deviation of the variables summarized in Table 5.1 over all regions of support. Then we use the result of this computation as our features. Each row of this table defines a separate feature set, which will be assessed for its discriminative power. Each feature set is extracted over all 12 regions of support.

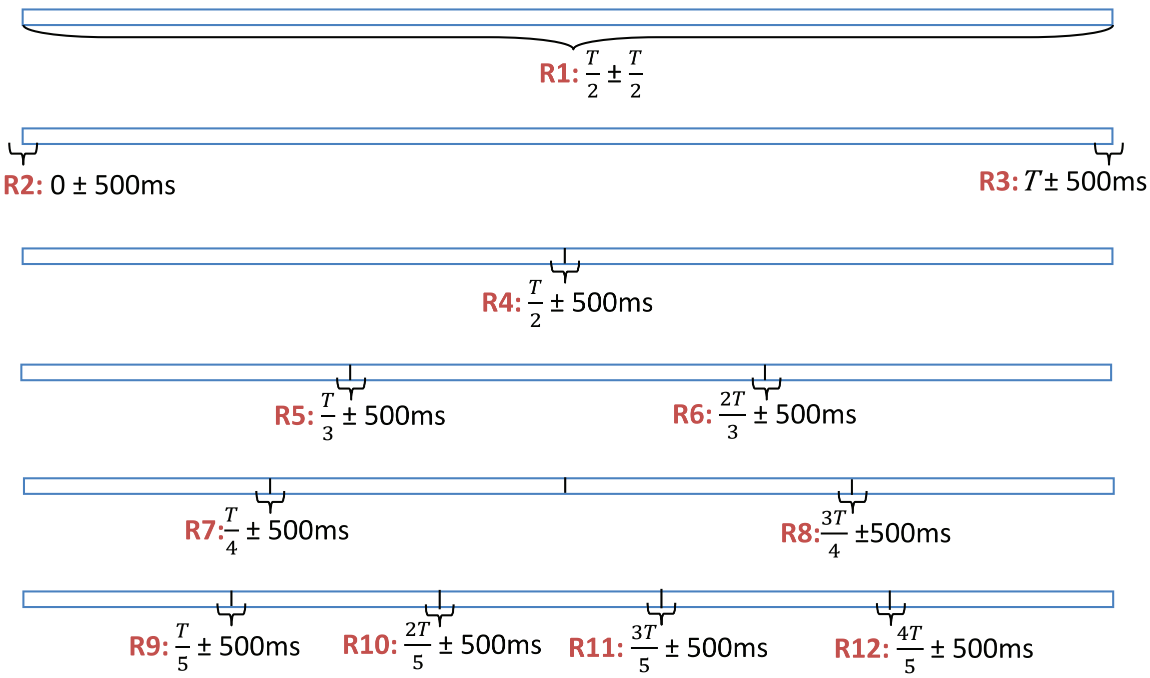


Figure 5.4: 12 different regions of support are demonstrated over the interaction segment. Equal number of features are computed over each region.

Table 5.1: Feature sets

Set Id	**Set Name**	**Features**	**Feature Count**
Set 1	Agent force-related	$\boldsymbol{F}_{HIP_1}, \boldsymbol{F}_{HIP_2}$	192
Set 2	Net force-related	$\boldsymbol{F}_{net}$	96
Set 3	Interactive force-related	f_i	48
Set 4	Velocity-Related	$\dot{\boldsymbol{x}}_{obj}, \dot{\theta}_{obj}$	144
Set 5	Power-Related	P_{HIP_1}, P_{HIP_2}	96
Total			576

5.4 Classifier Design

We utilize a multi-class Support Vector Machine (SVM) classifier with a Gaussian radial basis function kernel to recognize interaction patterns. In order to deal with the multi-class learning problem, we adopt the one-against-one strategy, which builds one SVM for each/every pair of classes[c)]. In order to obtain the optimal hyper-parameters, cost (C) and γ of the model, we perform model selection by 5-fold cross-validation using grid search.

5.5 Evaluation

For the evaluation of the classifier performance, we utilize the following metrics:

5.5.1 Normalized Confusion Matrix:

The normalized confusion matrix is a table which displays the correct and incorrect classification rates of each class. The values in the columns and rows respectively represent the number of instances in the predicted and the actual classes normalized by the class size. Hence, it clearly displays the classifier's confusion between two classes, if exists.

5.5.2 Correct Classification Rate (Accuracy):

The accuracy of classification is assessed by comparing the classification rate with ground truth labeling of the test set. The accuracy is defined as the number of correct classifications divided by the total number of examples in the test set.

5.5.3 Balanced Error Rate (BER):

BER is the average of the number of incorrect classifications for each class, normalized by the class size. The BER criteria is especially useful when the number of instances vary highly among different classes.

Chapter Notes

[c)]During the analysis, the SVM implementation provided within the LIBSVM toolbox for Matlab is used [4].

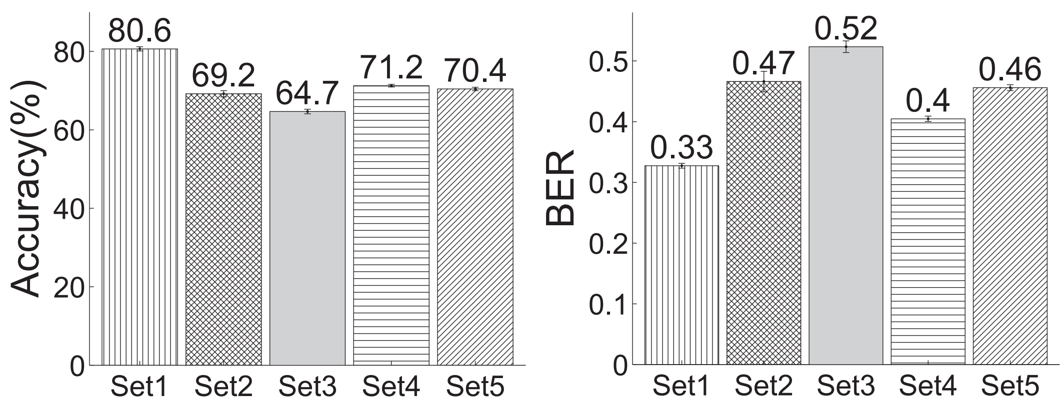
Chapter 6

RESULTS AND DISCUSSION

This section presents the classification results along with a discussion of them.

6.1 Classification Results Individual Feature Sets

Initially, we investigate the utility of using isolated feature sets for classifying the pattern classes. A separate model is trained with each feature set (Table 5.1) to discover how well these features capture the significant characteristics of the interaction pattern classes. The recognition performance of training with individual feature sets can be seen in Fig. 6.1, along with the confusion matrices in Fig. 6.2.



(a) Accuracy (b) Balanced Error Rate

Figure 6.1: Classification results of individual feature sets. Set 1: Agent force-related feature set, Set 2: Net force-related feature set, Set 3: Interactive force-related feature set, Set 4: Velocity-related feature set, Set 5: Power-related feature set.

The classifier trained with Set 1 (agent force-related features) achieves the highest classification performance with an accuracy of 80.6% and a BER of 0.33. On the other hand,

Set1	PREDICTED INSTANCES					
ACTUAL INSTANCES	C1	C2	C3	C4	C5	C6
C1	**0.94**	0.01	0.01	0.00	0.02	0.02
C2	0.03	**0.41**	0.02	0.01	0.43	0.10
C3	0.05	0.03	**0.14**	0.10	0.45	0.23
C4	0.00	0.00	0.05	**0.80**	0.15	0.00
C5	0.00	0.03	0.03	0.04	**0.78**	0.12
C6	0.00	0.00	0.00	0.00	0.03	**0.97**

(a) Agent force-related feature set

Set2	PREDICTED INSTANCES					
ACTUAL INSTANCES	C1	C2	C3	C4	C5	C6
C1	**0.95**	0.03	0.00	0.00	0.00	0.02
C2	0.05	**0.45**	0.05	0.00	0.32	0.13
C3	0.03	0.05	**0.16**	0.00	0.36	0.40
C4	0.00	0.00	0.00	**0.17**	0.11	0.72
C5	0.00	0.08	0.03	0.02	**0.60**	0.27
C6	0.00	0.00	0.00	0.01	0.10	**0.89**

(b) Net force-related feature set

Set3	PREDICTED INSTANCES					
ACTUAL INSTANCES	C1	C2	C3	C4	C5	C6
C1	**0.75**	0.00	0.00	0.00	0.03	0.22
C2	0.05	**0.10**	0.03	0.00	0.47	0.35
C3	0.08	0.01	**0.03**	0.06	0.49	0.33
C4	0.00	0.01	0.10	**0.55**	0.30	0.04
C5	0.02	0.06	0.04	0.04	**0.59**	0.25
C6	0.11	0.01	0.00	0.00	0.02	**0.86**

(c) Interactive force-related feature set

Set4	PREDICTED INSTANCES					
ACTUAL INSTANCES	C1	C2	C3	C4	C5	C6
C1	**0.88**	0.02	0.01	0.00	0.04	0.05
C2	0.06	**0.60**	0.02	0.00	0.20	0.12
C3	0.08	0.02	**0.62**	0.00	0.07	0.21
C4	0.00	0.00	0.00	**0.00**	0.12	0.88
C5	0.01	0.02	0.02	0.00	**0.61**	0.34
C6	0.00	0.01	0.01	0.00	0.11	**0.87**

(d) Velocity-related feature set

Set5	PREDICTED INSTANCES					
ACTUAL INSTANCES	C1	C2	C3	C4	C5	C6
C1	**0.86**	0.05	0.04	0.00	0.03	0.02
C2	0.11	**0.25**	0.10	0.00	0.47	0.07
C3	0.02	0.04	**0.64**	0.00	0.16	0.14
C4	0.00	0.00	0.02	**0.04**	0.10	0.84
C5	0.02	0.08	0.05	0.01	**0.58**	0.26
C6	0.00	0.01	0.01	0.00	0.05	**0.93**

(e) Power-related feature set

Combined Set	PREDICTED INSTANCES					
ACTUAL INSTANCES	C1	C2	C3	C4	C5	C6
C1	**0.95**	0.00	0.00	0.00	0.02	0.03
C2	0.03	**0.73**	0.00	0.02	0.14	0.08
C3	0.00	0.01	**0.70**	0.03	0.14	0.12
C4	0.00	0.00	0.02	**0.86**	0.08	0.04
C5	0.00	0.01	0.03	0.05	**0.74**	0.17
C6	0.00	0.00	0.01	0.00	0.07	**0.92**

(f) Combined feature set

Figure 6.2: Confusion matrices of classifiers trained with individual feature sets and the combined set. C1: Harmonious translation, C2: Harmonious rotation with translation, C3: Harmonious braking, C4: Persistent conflict, C5: Jerky conflict, C6: Passive agreement.

the classifier trained with Set 3 (interactive force-related features) yields the lowest performance with 64.7% accuracy and BER of 0.52.

Note that even though all classifiers achieve recognition accuracies higher than 60%, the BERs are comparatively high (≥ 0.3). Examining the confusion matrices in depth (see Fig. 6.2), we observe that each individual feature set is successful[d)] in recognizing at least 4 interaction patterns, but have confusions in one or two classes. Specifically, the classifiers trained individually with Sets 1 and 2 perform poorly in the classification of C3. In particular, agent-force related features in Set 1 suffer from confusion between C3 and C5, whereas net force-related features in Set 2 confuses C3 with both C5 and C6. As seen in Fig. 5.3, the mean magnitudes of individual forces are close to each other for C3 and C5, and so does the net force magnitudes of C3, C5, and C6. Hence the classifiers trained with these features are indeed expected to confuse the patterns as isolated features are not descriptive on their own for differentiating between these pattern classes. Similarly, it is no surprise for the classifier trained with the interactive-force related features in Set 3 to confuse C2 and C3 with C5 and C6. Finally, a similar case holds also for the Set 4's velocity- and Set 5's power-related features not being able to differentiate between C4 and C6.

6.2 Classification Results with the Combined Feature Set

The approach described above emphasizes the performance of isolated individual feature sets. However, some features can be used in combination to enhance the accuracy of the recognition of interaction patterns. We construct a *combined feature set*, comprising of all of the features in the aforementioned 5 feature sets from all regions of support. Using the combined feature set, we achieve an increased accuracy of 84.2% and a reduced BER of 0.19. The reduced BER value illustrates the increased discriminative power of the combined set in inhibiting the misclassifications. The confusion matrix of the classifier trained with the combined feature set is given in Fig. 6.2(f). Upon closer inspection, we observe that unlike the classifiers with individual feature sets, this model is able to recognize all of the interaction patterns without significant confusion. As seen in Fig. 6.2, combined feature set achieves the highest improvement for the classes C2 (Harmonious rotation with translation), C3 (Harmonious braking) and C4 (Persistent conflict) which had poor recognition performance with individual feature sets.

6.3 Selection of the Optimal Feature Set

The final step in our learning approach is to select the most informative features in the combined feature set and to find the most informative regions of support. This is motivated by the fact that the combined set gets quite large as a result of aggregating 5 individual feature sets from different regions of support. The combined set may contain some unnecessary and even irrelevant features, which may lead to inferior classification performance. Such features should be removed to enhance the recognition accuracy. Hence, we utilize the Minimum Redundancy Maximum Relevance (mRMR) feature selection algorithm to select most promising features [20].

The mRMR algorithm yields the k maximally relevant and minimally redundant features from a larger feature space of size K, consisting of 576 features in our case, where $k = 1, 2, ..., K$. In the end, the feature set that yields the highest accuracy is declared as the optimal feature set for the recognition of interaction patterns. Fig. 6.3 shows the classification accuracies against the number of features in the set. This diagram illustrates that the optimal feature set consists of 243 features. This optimal set achieves a performance even better than that of the combined feature set with an accuracy of 86% and a BER of 0.18. The confusion matrix of the classifier trained with the optimal feature set is given in Fig. 6.4. We observe that the classifier can successfully recognize all six of the interaction patterns.

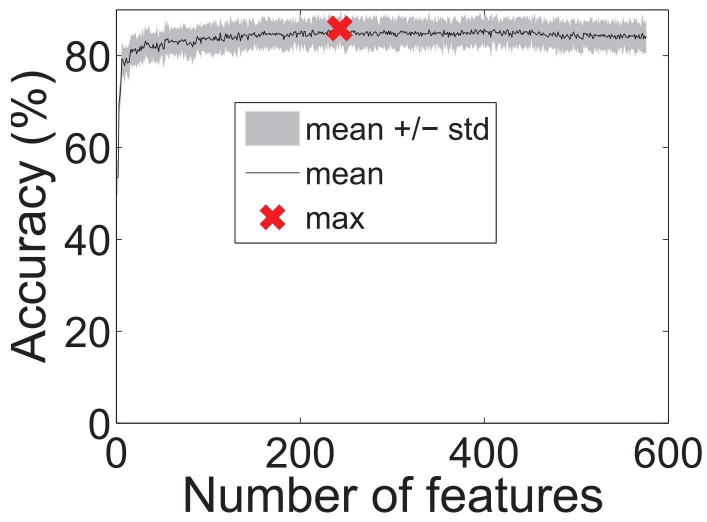


Figure 6.3: Classification accuracies for the feature sets, which are built incrementally using mRMR, plotted against the number of features in the features sets. The red circle denotes the optimum feature set, which yields the highest accuracy.

Fig. 6.5 presents the percentage of the features in the optimal feature set computed over the different regions of support explained in Section 5.3. 48 features are extracted from each region of support in the combined set. Although we do not conduct any statistical analysis, this figure gives an idea about the superiority of the regions of support. According to this figure, R1, which is sampled over whole interaction segment, is a superior region to recognize patterns, because of its large contribution to the optimal feature set. Also,

Optimal Set		PREDICTED INSTANCES C1	C2	C3	C4	C5	C6
ACTUAL INSTANCES	C1	**0.94**	0.00	0.01	0.00	0.03	0.02
	C2	0.03	**0.72**	0.02	0.03	0.15	0.05
	C3	0.00	0.01	**0.72**	0.04	0.12	0.11
	C4	0.00	0.00	0.00	**0.85**	0.08	0.07
	C5	0.00	0.02	0.01	0.04	**0.77**	0.16
	C6	0.00	0.00	0.00	0.00	0.04	**0.96**

Figure 6.4: Confusion matrix of classifier trained with the optimal feature set. C1: Harmonious translation, C2: Harmonious rotation with translation, C3: Harmonious braking, C4: Persistent conflict, C5: Jerky conflict, C6: Passive agreement.

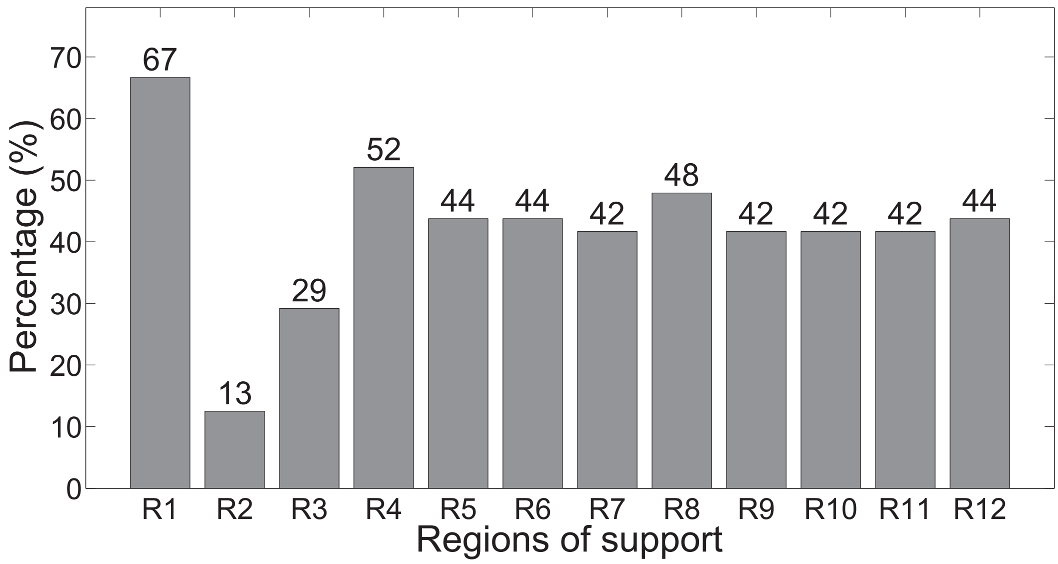
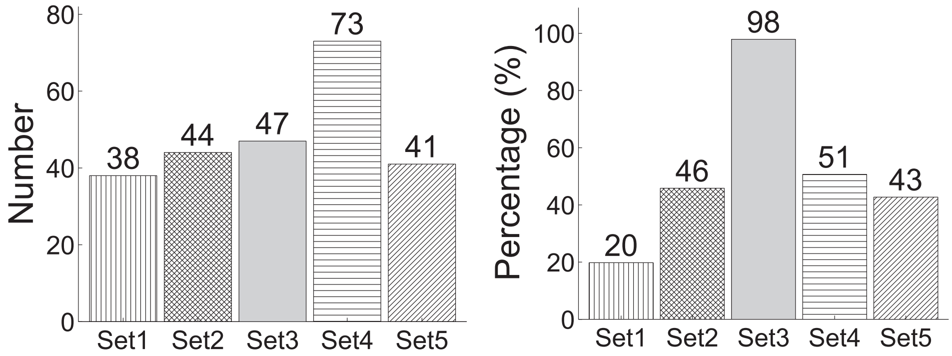


Figure 6.5: Percentage of features from different regions of support, contributing to the optimal feature set.

R4, R5,..., R12, which are sampled over different positions along the segment, are also informative regions, since they contribute to the optimal feature sets moderately. On the other hand, R2 and R3, which are sampled over the beginning and end of the segment, contribute slightly to the optimal feature set. Thus R2 and R3 are poor regions for extraction of features.



(a) Number of features. (b) Percentages of features.

Figure 6.6: Number and percentage of features from individual feature sets, contributing to the optimal feature set. Set 1: Agent force-related feature set, Set 2: Net force-related feature set, Set 3: Interactive force-related feature set, Set 4: Velocity-related feature set, Set 5: Power-related feature set.

Fig. 6.6 presents the number and the percentage of the features in the optimal feature set taken from the individual sets. At first glance, Fig. 6.6(a) gives the impression that Set 4 (velocity-related features) is a superior feature representation because of its large contribution to the optimal feature set; however this is misleading and is partly due to the high number of features in the initial set. The percentages of features contributed by each individual feature set provides more meaningful information about the superiority of feature representations. As demonstrated in Fig. 6.6(b), almost all of the features in Set 3 (interactive force-related features) eventually contribute to the optimal feature set. On the other hand, almost half of Set 4 is discarded during feature selection.

In this study, we demonstrate that feature sets presented in Section 5.3 are complementary. Moreover, we illustrate the significance of feature selection in accomplishing higher recognition accuracies. As happened in our case, the inclusion of many features may actually diminish recognition performance unless all are collectively relevant. Also, in addition to the improvement in the recognition accuracy, balanced error rate decreases. However, it is worth noting that there is a trade-off among the processing required for optimal feature selection and the resulting gain in the accuracy and decrease in BER.

Chapter Notes

[d)]We consider a classification to be unsuccessful in case that the correct classification rate is lower than random recognition rate, which is $1/6$ in our case.

Chapter 7

CONCLUSIONS

The thesis presents the results of the experimental study with 20 dyads who collaboratively manipulate an object, to identify haptic interaction patterns. This work is a first step discovering patterns in dyadic haptic interaction between humans. Specifically, this study presents a taxonomy of conflict-originated interaction patterns and a method for the classification of these patterns in physical collaboration scenarios, where two humans communicate through the haptic channel. Six different interaction patterns were identified based on the interaction of 20 human dyads who transport a virtual object to certain goal positions in a haptics-enabled simulation environment. Time-series data of the human-human interaction was divided into segments, each of which was labeled by an expert, who monitored the interaction from outside. We proposed five distinct feature sets, four of which consist of haptic features, to recognize the interaction patterns. We demonstrate that haptic features exhibit significant information about the interaction between partners, and the classifier trained with a combination of haptic and velocity-related features achieves a correct classification rate of 86% when recognizing human-human interaction patterns.

In this work, we propose a machine learning algorithm, which enables classification of human interaction patterns in dyadic tasks involving haptic interactions. We believe that the ideas we present here are generic for any kind of physical task, and given training data, can be generalized to a plethora of different tasks and systems, for both human-human and human-robot interaction. One shortcoming of this technique is its being offline. In the future, we intend to apply different learning methods to enable online intention prediction during an ongoing collaboration between two humans. Additionally, we intend to investigate the use of more sophisticated features in all force, velocity and power domains to enhance classification accuracy. Our final goal is to develop a robot, which can infer human interaction patterns in real-time and collaborate with its human partner(s) accordingly in complex object manipulation tasks.

Chapter 8

CONTRIBUTIONS AND FUTURE DIRECTIONS

This thesis has explored intrinsic properties of human-human interaction, which is defined as interaction patterns. A method trained with five different feature sets to classify the interaction patterns and a taxonomy of these patterns are proposed. Our contributions can be summarized as follows

1. An expert-labeled human interaction dataset is generated to train the machine learning method for classification of the patterns. Interactions of 20 dyads, who collaboratively carry an object in an haptics-enabled virtual environment, are observed from outside and time-series data from these interactions are labeled with interaction patterns manually.

2. A taxonomy of human interaction patterns is proposed in the light of information gained by studying the expert-labeled dataset.

3. Five different feature sets which are force-, velocity-, power-related information are proposed for the classification of these patterns.

4. Our evaluation suggests that each individual set is successful in recognizing at least 4 of 6 interaction patterns.

5. Utilizing multi-class SVM classifiers, a correct classification rate up to 86% is accomplished for the identification of the interaction patterns, which is obtained by fusing all feature sets by mRMR algorithm, even though individual feature sets fail to recognize some interaction patterns.

Our contributions enable to comprehend how humans interact to perform dyadic joint object manipulation tasks. This study demonstrates the capability of the proposed machine learning technique in classifying human interaction patterns automatically, almost as good as a human expert does manually. According to our study, possible future works can be summarized as follows

• The same experiment without providing any haptic feedback to the users could be conducted in order to compare the recognition results with the results of the current study.

• Clustering algorithms could be utilized in the future for automated segmentation of interaction between two humans, after reducing noise on the interaction data.

• It would be a good idea to investigate different learning methods to enable online intention prediction during ongoing collaboration between two humans.

• More sophisticated features in all the force, velocity and power domains could be investigated to improve the recognition accuracy.

• New experiments could be designed in order to test the utility of the proposed approach in different settings.

• An intelligent robot, which comprehends how humans interact and assist accordingly, could be developed. This robot could either mimic the interaction patterns of one of the partners or interfere in the interactions of humans as an assistant.

BIBLIOGRAPHY

[1] A. L. Berezm. Review of eudico linguistic annotator (elan). In *Language Documentation and Conservation*, volume 1, pages 283–289, December 2007.

[2] A. Billard, S. Calinon, R. Dillmann, and S. Schaal. Robot programming by demonstration. In B. Siciliano and O. Khatib, editors, *Springer Handbook of Robotics*, pages 1371–1394. Springer, 2008.

[3] A. Bussy, P. Gergondet, A. Kheddar, F. Keith, and A. Crosnier. Proactive behavior of a humanoid robot in a haptic transportation task with a human partner. In *RO-MAN, 2012 IEEE*, pages 962–967. IEEE, 2012.

[4] C.-C. Chang and C.-J. Lin. Libsvm: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.*, 2(3):27:1–27:27, May 2011.

[5] B. Corteville, E. Aertbeliën, H. Bruyninckx, J. D. Schutter, and H. V. Brussel. Human-inspired robot assistant for fast point-to-point movements. In *IEEE International Conference on Robotics and Automation, ICRA*, pages 3639–3644, 2007.

[6] V. Duchaine and C. M. Gosselin. General model of human-robot cooperation using a novel velocity based variable impedance control. In *Second Joint EuroHaptics Conference, 2007 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pages 446–451. IEEE, 2007.

[7] P. Evrard, E. Gribovskaya, S. Calinon, A. Billard, and A. Kheddar. Teaching physical collaborative tasks: object-lifting case study with a humanoid. In *Humanoid Robots, 2009. Humanoids 2009. 9th IEEE-RAS International Conference on*, pages 399–404. IEEE, 2009.

[8] P. Evrard and A. Kheddar. Homotopy switching model for dyad haptic interaction in physical collaborative tasks. In *WHC'09: IEEE World Haptics Conference*, pages 45–50, Washington, DC, USA, 2009.

[9] T. Flash and N. Hogan. The coordination of arm movements: an experimentally confirmed mathematical model. *The Journal of Neuroscience*, 5(7):1688–1703, 1985.

[10] R. Groten, D. Feth, H. Goshy, A. Peer, D. Kenny, and M. Buss. Experimental analysis of dominance in haptic collaboration. In *Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on*, pages 723–729, 2009.

[11] R. Groten, D. Feth, R. Klatzky, and A. Peer. The role of haptic feedback for the integration of intentions in shared task execution. *IEEE Transactions on Haptics*, 6(1):94–105, 2013.

[12] N. Jarrassé, T. Charalambous, and E. Burdet. A framework to describe, analyze and generate interactive motor behaviors. *PLOS One*, 7(11):e49945, 2012.

[13] A. Kheddar. Human-robot haptic joint actions is an equal control-sharing approach possible? In *Human System Interactions (HSI), 2011 4th International Conference on*, pages 268–273, 2011.

[14] A. Kucukyilmaz, T. Sezgin, and C. Basdogan. Conveying intentions through haptics in human-computer collaboration. In *WHC'11: IEEE World Haptics Conference*, pages 421–426, June 2011.

[15] A. Kucukyilmaz, T. Sezgin, and C. Basdogan. Intention recognition for dynamic role exchange in haptic collaboration. *Haptics, IEEE Transactions on*, 6(1):58–68, 2013.

[16] M. Lawitzky, A. Mörtl, and S. Hirche. Load sharing in human-robot cooperative manipulation. *Proc. of IEEE Int. Symposium in Robot and Human Interactive Communication*, pages 185 –191, 2010.

[17] A. Melendez-Calderon, V. Komisar, G. Ganesh, and E. Burdet. Classification of strategies for disturbance attenuation in human-human collaborative tasks. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 2364–2367. IEEE, 2011.

[18] A. Moertl, M. Lawitzky, A. Kucukyilmaz, T. M. Sezgin, C. Basdogan, and S. Hirche. The role of roles: Physical cooperation between humans and robots. *Int. J. Robotic Res.*, 31(13):1656–1674, 2012.

[19] S. Oguz, A. Kucukyilmaz, T. Sezgin, and C. Basdogan. Haptic negotiation and role exchange for collaboration in virtual environments. In *Haptics Symposium, 2010 IEEE*, pages 371–378, 2010.

[20] H. Peng, F. Long, and C. Ding. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(8):1226–1238, 2005.

[21] M. Rahman, R. Ikeura, and K. Mizutani. Control characteristics of two humans in cooperative task and its application to robot control. In *Industrial Electronics Society, 2000. IECON 2000. 26th Annual Confjerence of the IEEE*, volume 3, pages 1773–1778. IEEE, 2000.

[22] K. B. Reed and M. A. Peshkin. Physical collaboration of human-human and human-robot teams. *IEEE Trans. Haptics*, 1(2):108–120, 2008.

[23] O. Schrempf, U. Hanebeck, A. Schmid, and H. Woern. A novel approach to proactive human-robot cooperation. In *Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop on*, pages 555–560, 2005.

[24] N. Stefanov, A. Peer, and M. Buss. Role determination in human-human interaction. In *WHC'09: IEEE World Haptics Conference*, pages 51–56, Washington, DC, USA, 2009.

[25] T. Takeda, Y. Hirata, and K. Kosuge. Dance step estimation method based on HMM for dance partner robot. *IEEE Transactions on Industrial Electronics*, 54(2):699 – 706, 2007.

[26] T. Tsumugiwa, R. Yokogawa, and K. Hara. Variable impedance control with virtual stiffness for human-robot cooperative peg-in-hole task. In *Intelligent Robots and Systems, 2002. IEEE/RSJ International Conference on*, volume 2, pages 1075–1081. IEEE, 2002.

[27] Z. Wang, A. Peer, and M. Buss. An HMM approach to realistic haptic human-robot intearction. In *Third Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, 2009.

Appendix A

PARAMETERS USED IN THE EXPERIMENT

Table A.1: Object and board information for both scenes(Section 3.1)

Scene	Object Mass(kg)	Object Dimensions(mm)	Board Dimensions(mm)
Straight Scene	4	24 x 9	200 x 30
Bifurcated Scene	4	24 x 9	200 x 82

Table A.2: Spring and damper coefficients of the physical model (Section3.2) used in the experiment to model the forces of users

Variable	Value
Kp	0.25 N/mm
Kd	0.001 Ns/mm

Table A.3: Static and kinetic friction coefficient values for both force and moment (Section3.2)

	Force	Moment
Static	0.19	0.20
Kinetic	0.15	0.19

Appendix B

PHYSICAL MODEL OF THE EXPERIMENT

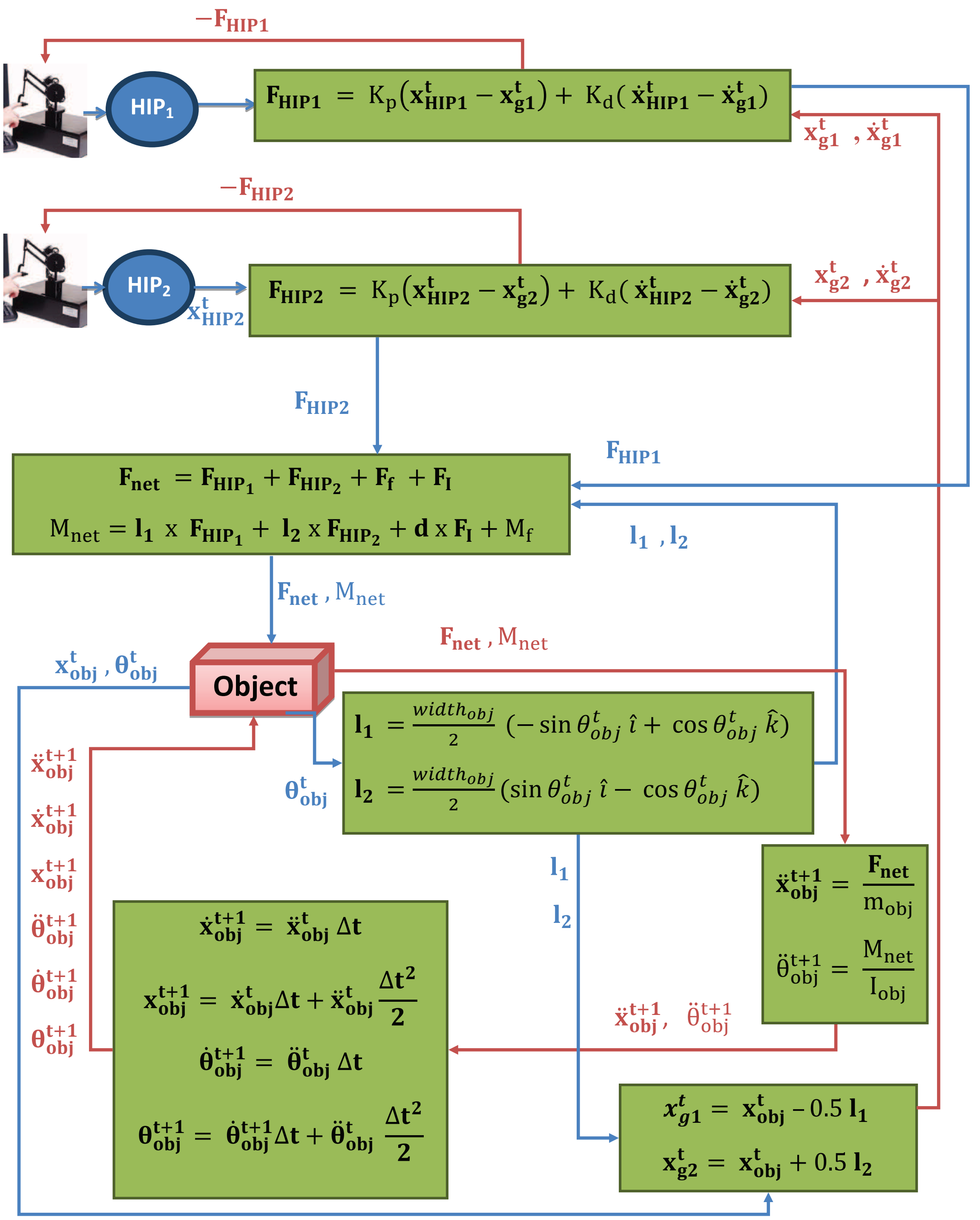


Figure B.1: Detailed physical model of the interaction between humans during dyadic joint object manipulation (Section 3.2)

Appendix C

INSTRUCTIONS USED IN THE EXPERIMENT

• Thank you for agreeing to participate in this study.

• Please read through these instructions and ask any question you may have before the experiment begins.

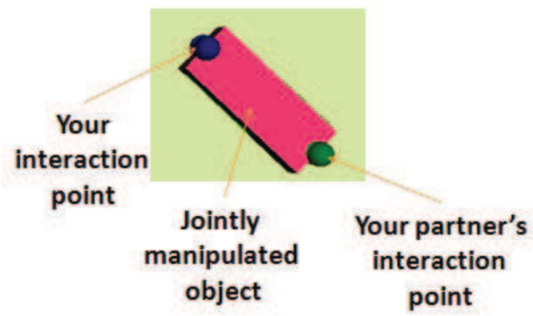
• Please turn off any electronic devices before the experiment begins.

• This experiment requires you to manipulate a virtual object with a partner.

• Your partner will sit in the opposite room, as a result he/she wont be able to hear you and vice versa. However you will be able to monitor his/her movements through the motion of the jointly manipulated object.

• The object you will manipulate in this experiment is a rectangular block as shown in the figure.

• The interaction point you will be able to move will be colored BLUE\GREEN, whereas that of your partner will be colored GREEN\BLUE.



• You will need to coordinate with your partner in order to generate smooth movements.

• There will be a force feedback device on the table that enables your interaction with the scene.

• You need to hold the stylus and move your hand right / left / forward / backward to move your interaction point and the jointly manipulated object.

• There might be occasional sudden small jumps when working with the device.

• Please do not panic, and hold the stylus between your fingers firmly(as shown in the picture).

• The device will stabilize momentarily.

INITALIZATION OF DEVICE

• Before starting the experiment, you will be asked to calibrate the device using a "Phantom Test" application.

• In order to calibrate, please hold the device in neutral position and iterate as the application instructs you.

THE GOAL

• Throughout the experiment, you will be asked to translate the block to a certain position

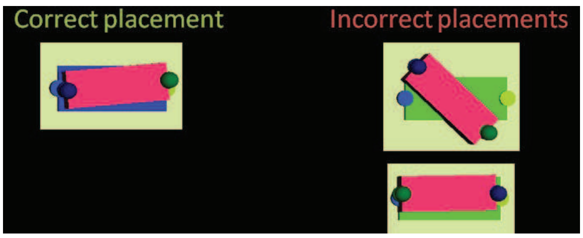
• A typical translation task is shown in FigureC

• Here, your goal position is marked with a green rectangle.

• When you reach the target rectangle it will turn blue.

• Your goal is to stay on this target to the count of 5.

• Beware: The target configuration contains orientation information!



TRIALS

• A trial ends successfully when you and your partner reach the target rectangle and wait for it to the count of 5.

• In case you cannot reach the target after a long time the trial will end abruptly, indicating unsuccessful task completion.

• Please note that unsuccessful task completion is not desirable; hence try to complete the task by achieving the goals as instructed to the best of your ability.

HOW TO MOVE THE OBJECT

• The object moves as a rigid body.

• You can apply forces on the object by moving the haptic device.

Translating the object	Rotating the object
The parallel component of a force produces translation	The perpendicular component of a force produces rotation

THE GOAL

• Note that the goal of your partner may be different than that of yours.

• If there is severe conflict during your operation, a warning will be displayed: CONFLICT!

• Your aim is to finish the task quickly and without conflicts.

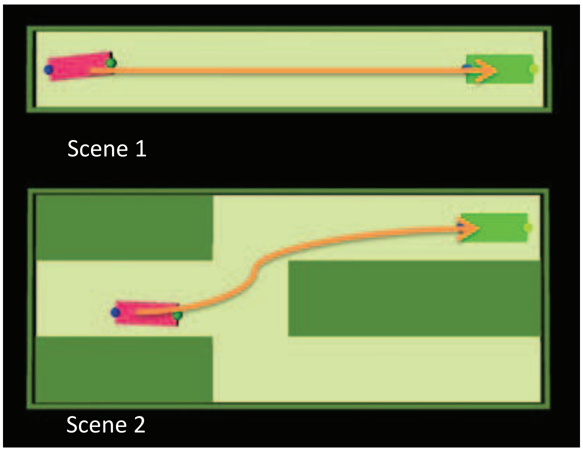
• Hence, if you face a conflict with your partner try to observe his/her movements and resolve the conflict.

• You should avoid hitting the boundaries and obstacles while moving on your path.

• In case of collisions, the collided boundaries and/or obstacles will be highlighted.

INSTRUCTIONS

• You will be asked to perform the task in two different scenes.



• In the first scene, you will be asked to translate the block on a straight path.

• The second scene will require you to turn a corner (without hitting boundaries) in order to reach the correct parking location.

• During the experiment, you will perform 2 sets per scene, for a total of 4 sets.

• Sets 1 and 3 will be practice sets, where you will have the chance to familiarize yourself with the scenes.

Scene1	Scene2
Set1(2-6 trials)	Set3(2-6 trials)
Set2(14 trials)	Set4(19 trials)

• During the experiment, you will perform 2 sets per scene, for a total of 4 sets. Sets 1 and 3 will be practice sets, where you will have the chance to familiarize yourself with the scenes.

• The current trial number will be indicated on top of the game screen

Thank you for your participation