

Learning Human-Robot Interaction: A Case Study on Human-Robot Handshaking

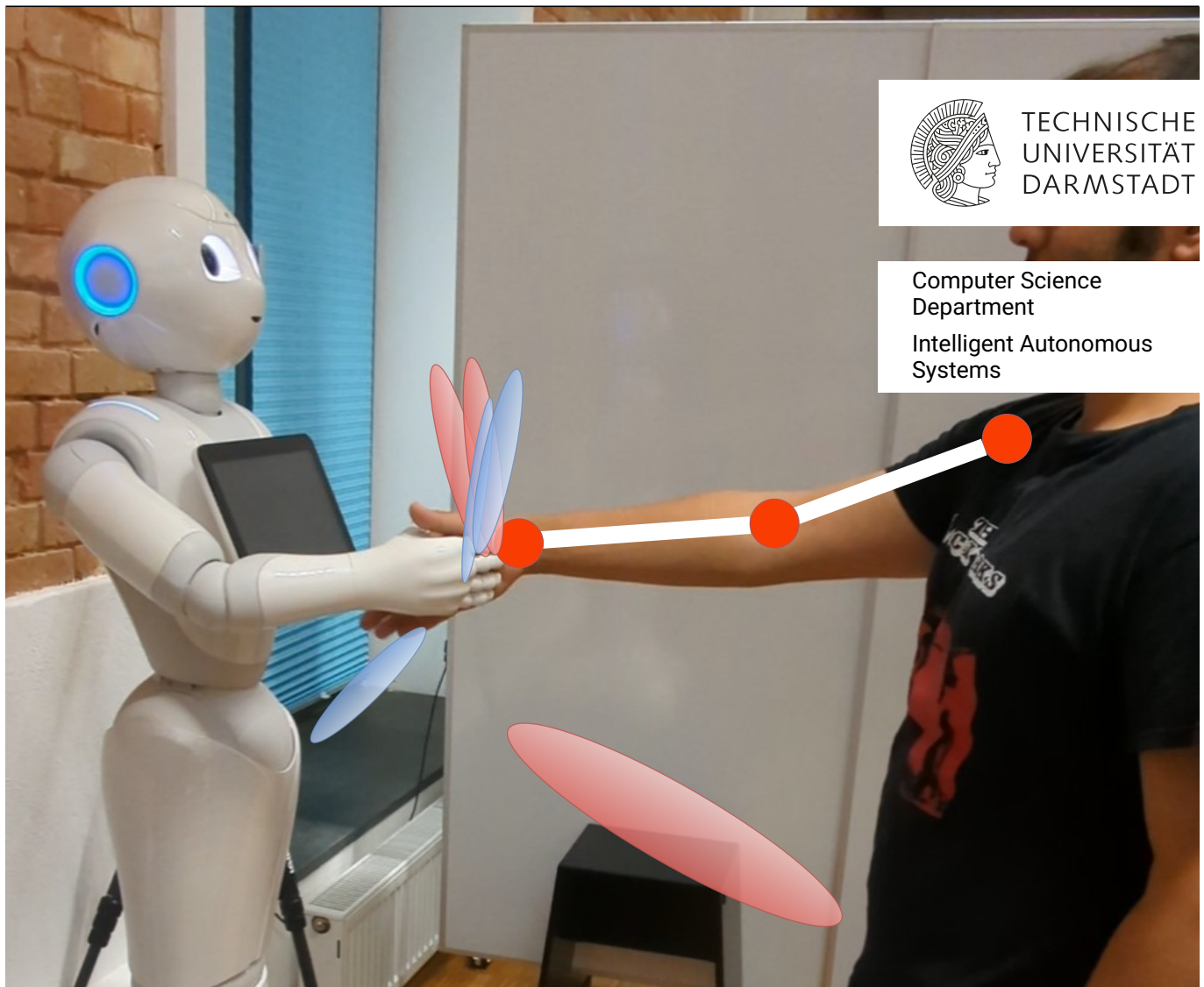
Lernen der Mensch-Roboter-Interaktion: Eine Fallstudie zum Mensch-Roboter-Händeschütteln

Zur Erlangung des akademischen Grades Doktor-Ingenieur (Dr.-Ing.)

Genehmigte Dissertation von Vignesh Prasad aus Hyderabad, India

Tag der Einreichung: 19. Oktober 2023, Tag der Prüfung: 01. Dezember 2023

1. Gutachten: Prof. Jan Peters (PhD)
2. Gutachten: Prof. Dr. Dr. Ruth Stock-Homburg
3. Gutachten: Prof. Dr. rer. nat. Yue Hu
Darmstadt, Technische Universität Darmstadt



Learning Human-Robot Interaction: A Case Study on Human-Robot Handshaking
Lernen der Mensch-Roboter-Interaktion: Eine Fallstudie zum
Mensch-Roboter-Händeschütteln

Accepted doctoral thesis by Vignesh Prasad

Date of submission: 19. Oktober 2023

Date of thesis defense: 01. Dezember 2023

Darmstadt, Technische Universität Darmstadt

Bitte zitieren Sie dieses Dokument als:

URN: urn:nbn:de:tuda-tuprints-190254

URL: <https://tuprints.ulb.tu-darmstadt.de/19025>

Jahr der Veröffentlichung auf TUprints: 2024

Dieses Dokument wird bereitgestellt von tuprints,

E-Publishing-Service der TU Darmstadt

<https://tuprints.ulb.tu-darmstadt.de>

tuprints@ulb.tu-darmstadt.de

Die Veröffentlichung steht unter folgender Creative Commons Lizenz:

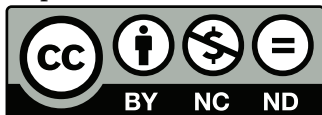
Namensnennung-Nicht kommerziell-Keine Bearbeitungen 4.0 International

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

This work is licensed under a Creative Commons License:

Attribution-NonCommercial-NoDerivatives 4.0 International

<https://creativecommons.org/licenses/by-nc-nd/4.0/>



For all the frontline workers who got us through the dark times during 2020-2021. I hope someday my work can help you in return.

Erklärungen laut Promotionsordnung

§ 8 Abs. 1 lit. d PromO

Ich versichere hiermit, dass zu einem vorherigen Zeitpunkt noch keine Promotion versucht wurde. In diesem Fall sind nähere Angaben über Zeitpunkt, Hochschule, Dissertationsthema und Ergebnis dieses Versuchs mitzuteilen.

§ 9 Abs. 1 PromO

Ich versichere hiermit, dass die vorliegende Dissertation – abgesehen von den in ihr ausdrücklich genannten Hilfen – selbstständig verfasst wurde und dass die „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Technischen Universität Darmstadt“ und die „Leitlinien zum Umgang mit digitalen Forschungsdaten an der TU Darmstadt“ in den jeweils aktuellen Versionen bei der Verfassung der Dissertation beachtet wurden.

§ 9 Abs. 2 PromO

Die Arbeit hat bisher noch nicht zu Prüfungszwecken gedient.

Darmstadt, 19. Oktober 2023

V. Prasad

Abstract

For some years now, the use of humanoid social robots in various situations has been on the rise. These are robots developed to interact with humans and are equipped with corresponding extremities. They already support human users in various industries, such as retail, gastronomy, hotels, education and healthcare. During such Human-Robot Interaction (HRI) scenarios, physical touch plays a central role in the various applications of social robots as interactive non-verbal behaviour is a key factor in making the interaction more natural. Shaking hands is a simple, natural interaction used commonly in many social contexts and is seen as a symbol of greeting, farewell and congratulations. Moreover, the act of handshaking, given its extended phase of physical contact allows one to convey complex emotions via the sense of touch. Giving an appropriate response, therefore, plays an important role in improving the naturalness of the interaction. Furthermore, having a timely response also yields a more natural interaction, where the robot is able to predict the human partner's movements and adapt its motion accordingly. Modelling the dynamics of such interactions is a key aspect of Human-Robot Interaction.

In this context, the main focus of this thesis is to understand how such a physically interactive behaviour affects an interaction with a social robot. The contributions of this thesis are as follows. We first perform a thorough analysis of existing works related to Human-Robot Handshaking exploring the modelling aspects for realising an effective handshakes and social aspects such as the acceptance of such behaviours, auxiliary elements, such as gaze or approach motions, human-likeness etc. We then incorporate these findings in a novel frameworks to realise a timely, adaptive and socially acceptable handshake on a humanoid social robot. We then explore how to extend this modularised form of learning towards a general framework for learning coordinated Human-Robot Interaction. We validate the effectiveness of the proposed frameworks through extensive experimental evaluations with human users who interact with a humanoid social robot equipped with our approaches.

As a first step, the existing state of Human-Robot Handshaking research is looked at and the works are categorised based on their focus areas. Following this, the major findings of these areas are drawn out and their pitfalls are analysed. It is mainly seen that synchronisation is key during the different phases of the interaction. Additional factors

like gaze, voice facial expressions etc. can affect the perception of a robotic handshake along with internal factors like personality and mood which can affect the way in which handshaking behaviours are executed by humans. Based on the findings and insights, possible ways forward for future research on such physically interactive behaviours are discussed.

In the case of handshaking and other similar physically interactive behaviours, having a timely response yields a more natural interaction, where the robot is able to predict the human partner's movements and adapt its motion accordingly. Modelling the dynamics of such interactions is a key aspect of Human-Robot Interaction. In this work, a framework is developed for robots to learn such interactions directly from human-human interactions, modular fashion by breaking down the interactions into their underlying segments and learning the sequencing between them. We do so using Hidden Markov Models to model the interaction dynamics via the latent embeddings learned by a Variational Autoencoder. We show how the interaction dynamics learned from Human-Human Interactions can help regularize the learning of robot trajectories and we explore the conditional generation of robot motions from human observations to enable learning suitable and accurate Human-Robot Interactions. We further explore how to adapt the generated motions for a spatially accurate and compliant handshaking behaviour, leading to a higher degree of acceptance by human users.

We further explore how the performance of the reactive motion generation can be improved by bridging the gap in the proposed framework by integrating the conditioning of the HMMs into the VAEs in a more principled manner. To this end, we demonstrate how Mixture Density Networks yield themselves as an extension of the underlying HMM conditioning. Such a structure inherently allows the model to capture the complex and multimodal nature of human behavior. We demonstrate how the proposed framework can enhance the prediction of the reactive motion generation by learning multiple latent policies which when combined enable the generation of more accurate interactions.

To summarise, the goals of this thesis are: (i) to further investigate the act of handshaking in the scope of physical Human-Robot Interactions, (ii) to develop a framework that can learn a library of such physically interactive behaviours to widen the social skills of a robot and (iii) to explore how the accuracy of generating realistic and natural interactive behaviors can be improved.

Zusammenfassung

Seit einigen Jahren nimmt der Einsatz humanoider sozialer Roboter in verschiedenen Situationen zu. Dabei handelt es sich um Roboter, die für die Interaktion mit Menschen entwickelt wurden und mit entsprechenden Extremitäten ausgestattet sind. Sie unterstützen bereits menschliche Nutzer in verschiedenen Branchen, wie z. B. im Einzelhandel, in der Gastronomie, in Hotels, im Bildungswesen und im Gesundheitswesen. Bei solchen Szenarien der Mensch-Roboter-Interaktion (HRI) spielt die körperliche Berührung eine zentrale Rolle in den verschiedenen Anwendungen sozialer Roboter, da interaktives non-verbales Verhalten ein Schlüsselfaktor ist, um die Interaktion natürlicher zu gestalten. Händeschütteln ist eine einfache, natürliche Interaktion, die in vielen sozialen Kontexten üblich ist und als Symbol für Begrüßung, Verabschiedung und Glückwünsche gilt. Darüber hinaus ermöglicht das Händeschütteln aufgrund der langen Phase des Körperkontakts die Vermittlung komplexer Emotionen über den Tastsinn. Eine angemessene Antwort spielt daher eine wichtige Rolle bei der Verbesserung der Natürlichkeit der Interaktion. Darüber hinaus führt eine rechtzeitige Reaktion auch zu einer natürlicheren Interaktion, bei der der Roboter die Bewegungen des menschlichen Partners vorhersagen und seine Bewegungen entsprechend anpassen kann. Die Modellierung der Dynamik solcher Interaktionen ist ein Schlüsselaspekt der Mensch-Roboter-Interaktion.

In diesem Zusammenhang liegt der Schwerpunkt dieser Arbeit darauf, zu verstehen, wie ein solches physisches Interaktionsverhalten eine Interaktion mit einem sozialen Roboter beeinflusst. Die Beiträge dieser Arbeit sind wie folgt. Zunächst führen wir eine gründliche Analyse bestehender Arbeiten zum Thema Mensch-Roboter-Handshake durch und untersuchen die Modellierungsaspekte für die Realisierung eines effektiven Handshakes sowie soziale Aspekte wie die Akzeptanz solcher Verhaltensweisen, Hilfselemente wie Blick- oder Annäherungsbewegungen, Menschenähnlichkeit usw. Diese Erkenntnisse fließen dann in ein neuartiges Framework ein, um einen zeitnahen, adaptiven und sozial akzeptablen Händedruck auf einem humanoiden sozialen Roboter zu realisieren. Anschließend untersuchen wir, wie diese modularisierte Form des Lernens zu einem allgemeinen Rahmen für das Lernen koordinierter Mensch-Roboter-Interaktion erweitert werden kann. Wir validieren die Wirksamkeit des vorgeschlagenen Rahmens durch umfangreiche experimentelle Auswertungen mit menschlichen Benutzern, die mit einem humanoiden sozialen Roboter

interagieren, der mit unseren Ansätzen ausgestattet ist.

In einem ersten Schritt wird der aktuelle Stand der Mensch-Roboter-Handshaking-Forschung untersucht und die Arbeiten werden nach ihren Schwerpunkten kategorisiert. Anschließend werden die wichtigsten Erkenntnisse aus diesen Bereichen herausgearbeitet und ihre Fallstricke analysiert. Es zeigt sich vor allem, dass die Synchronisation in den verschiedenen Phasen der Interaktion entscheidend ist. Zusätzliche Faktoren wie Blicke, Stimme, Mimik usw. können die Wahrnehmung eines Roboter-Handschlags beeinflussen, ebenso wie interne Faktoren wie Persönlichkeit und Stimmung, die sich auf die Art und Weise auswirken können, wie das Handshake-Verhalten von Menschen ausgeführt wird. Basierend auf den Erkenntnissen und Einsichten werden mögliche Wege für die zukünftige Forschung zu solchen physisch interaktiven Verhaltensweisen diskutiert.

Im Falle des Händeschüttelns und anderer ähnlicher physisch interaktiver Verhaltensweisen führt eine rechtzeitige Reaktion zu einer natürlicheren Interaktion, bei der der Roboter in der Lage ist, die Bewegungen des menschlichen Partners vorherzusagen und seine Bewegung entsprechend anzupassen. Die Modellierung der Dynamik solcher Interaktionen ist ein Schlüsselaspekt der Mensch-Roboter-Interaktion. In dieser Arbeit wird ein Rahmen für Roboter entwickelt, um solche Interaktionen direkt von Mensch-Mensch-Interaktionen zu lernen, und zwar auf modulare Weise, indem die Interaktionen in ihre zugrundeliegenden Segmente zerlegt werden und die Abfolge zwischen ihnen gelernt wird. Dazu verwenden wir Hidden Markov Modelle, um die Interaktionsdynamik über die latenten Einbettungen zu modellieren, die von einem Variational Autoencoder gelernt werden. Wir zeigen, wie die aus Mensch-Mensch-Interaktionen gelernte Interaktionsdynamik dazu beitragen kann, das Lernen von Robotertrajektorien zu regulieren, und wir untersuchen die bedingte Generierung von Roboterbewegungen aus menschlichen Beobachtungen, um das Lernen geeigneter und genauer Mensch-Roboter-Interaktionen zu ermöglichen. Darüber hinaus untersuchen wir, wie die generierten Bewegungen für ein räumlich genaues und nachgiebiges Handshaking-Verhalten angepasst werden können, was zu einem höheren Grad an Akzeptanz durch menschliche Benutzer führt.

Wir untersuchen weiter, wie die Leistung der reaktiven Bewegungsgenerierung verbessert werden kann, indem wir die Lücke im vorgeschlagenen Rahmenwerk schließen, indem wir die Konditionierung der HMMs auf prinzipiellere Weise in die VAEs integrieren. Zu diesem Zweck zeigen wir, wie sich Mixture Density Networks als eine Erweiterung der zugrunde liegenden HMM-Konditionierung ergeben. Eine solche Struktur ermöglicht es dem Modell, die komplexe und multimodale Natur des menschlichen Verhaltens zu erfassen. Wir zeigen, wie der vorgeschlagene Rahmen die Vorhersage der reaktiven Bewegungserzeugung verbessern kann, indem er mehrere latente Strategien erlernt, die in Kombination die Erzeugung genauerer Interaktionen ermöglichen.

Zusammenfassend kann man sagen, dass die Ziele dieser Arbeit folgende sind: (i) den

Akt des Händeschüttelns im Rahmen der physischen Mensch-Roboter-Interaktion weiter zu untersuchen, (ii) einen Rahmen zu entwickeln, der eine Bibliothek solcher physisch interaktiven Verhaltensweisen erlernen kann, um die sozialen Fähigkeiten eines Roboters zu erweitern und (iii) zu untersuchen, wie die Genauigkeit der Erzeugung realistischer und natürlicher interaktiver Verhaltensweisen verbessert werden kann.

Acknowledgements

From the bottom of my heart I would like to thank Georgia Chalvatzaki and Dorothea Koert without whom I could not have come this far with my work. Thank you for always being there to help me, advise me, to listen to my problems and to help solve them, for pushing me to go the extra mile and beyond and for being great role models to look up to. Without your help, I would not have come so far in my PhD journey and would not have been able to grow so much professionally and more importantly, stay sane while doing so. I especially thank Alap Kshirsagar, whose help came in at a vital time and played a crucial role towards the end of this thesis. I would like to thank my advisors Jan Peters and Ruth Stock-Homburg for giving me this opportunity to work on such interesting problems.

I am deeply grateful to the following organisations for funding the work done in this thesis: Interdisciplinary Research Forum (Forum Interdisziplinäre Forschung - FiF) at the Technical University of Darmstadt (FiF Project 2018#21 "Handshake Turing Test"), the German Research Foundation (DFG) (Grant No.: STO 477/14-1 "Social Robots at the Customer Interface"), the German Federal Ministry of Education and Research (BMBF) (Grant no.: 02L19C150 "KompAKI"), the Förderverein für Marktorientierte Unternehmensführung, Marketing und Personalmanagement e.V., the Leap in Time Stiftung.

I would not have been here at this point in my life without the guidance of my former manager at TCS R&I, Brojeshwar Bhowmick, and my former supervisor at IIIT-H, Prof. K. Madhava Krishna. Thank you for equipping me with the necessary skills to tackle my PhD and my research life. I am deeply grateful for all the love and support that my Parents and my wife, Nikita, have shown me. Thank you for always being there and being my support system. I am so grateful that I have you in my life.

A lot of the work presented in this thesis would not have been possible without the help of students and collaborators along the way. I thank you all for all your help and I hope that it was as much a pleasure for you to work with me as it was for me to work with you! I would especially like to thank Louis Sterker, Mark Baierl, Michel Kohl, Yasmin Göksu, Antonio De Almeida Correia and Fabain Hahne whose work contributed to the eventual development of some of the chapters in this thesis, and the papers that resulted from them.

I am indebted to Sabrina Glindmeyer and the team at TU Darmstadt Welcome Center,

Inés Scherer, and Nanette Schreiber for helping me navigate all the hurdles that entail the beast of German bureaucracy. Thank you for making all the processing so smooth and easy. I would like to especially thank Sven Schultze, Lea Heitlinger, Suman Pal, and Arjun Datta with whom I could always share a bad day and vent out to and for all the trauma bonding. I would also like to thank Snehal Jauhri, Felix Kaiser, Aiswarya Menon, and Lisa Scherf for their help through this journey. I additionally thank all my colleagues, from IAS, MuP, and LIT for all their help.

From the other bottom of my heart I want to thank all my friends who have been my support system throughout my time in Darmstadt. Hats off to the team at Tutor International. Because of your help and efforts, I was able to build a strong social life and meet many people and learn about life in Germany and Darmstadt. Same goes for Martin, Brian and the meetup group. Ahmed, Andreas, Annie, Anjali, Ashwin, Haris, Sophie, thanks lot for all those fun times, for all those laughs, for all those aimless nights spent in Herrngarten, for always lifting my mood and showing me that happiness, laughter, and joy are just around the corner, be it at Krone or Ratskeller. My heartfelt gratitude goes out to Bollymotion: Huong, Pratik, Nupur Francesca, Mohit, Alex, Nora, Varun and all others for saving the dancer in me and always bringing him out. To my Lifelong Friends, Vanya, Ghosh, Kannan, Phani, Kavya, Shirin, Jaggi, Anjali, thank you for always being there for me, for always lending your ears for my bullshit, for having my back. Be it in the past, present or future, I know I can always count on you guys to help me get through life. I also thank the entire Comedy Fraternity Kirthy, Meilor, Sup Comedy/Comedy Community, Soness Stevens, Garv Malik, Aaye Bade Comics and Aditya Jain who allowed me to tap into my nerdiness and bridge my work and comedy. Thank you for keeping me sane and making me laugh when I needed it the most and for helping me sharpen my writing skills. I am especially grateful to my partners in crime, Mars Münster, Akash Pandey, and Mathis Ali Gonzalez for starting *So Darm Funny* with me. It inevitably taught me a lot and helped keep me sane during the final stages of the PhD when I needed it the most.

I would also like to thank the open-source community for being so awesome! I am extremely grateful to Roberto Calandra, Joseph Campbell, Emmanuel Pignat and Sebastian Gomez-Gonzalez, and Pierre Manceron for open-sourcing their works on Humanoid Teleoperation, Interaction Primitives, Hidden Markov Models and Inverse Kinematics without which a lot of the work in this thesis would not have been possible (or would have been greatly delayed). I would also like to thank the Open Access Fund of TU Darmstadt and ArXiv for enabling me to openly publish the results from this thesis and supporting Open Science.

Contents

1. Introduction	1
1.1. Motivation	1
1.2. Research Questions	4
1.3. Contributions	5
1.4. Outline	7
2. Human-Robot Handshaking: A Review	9
2.1. Prologue	9
2.2. Insights and Evaluation of Handshakes	11
2.2.1. Insights from Human-Human Handshake Interactions	11
2.2.2. Handshake Evaluation Methods	14
2.3. Reaching phase of Handshaking	17
2.4. Controlling Hand Grasps in Handshaking	18
2.5. Shaking Motions and Synchronisation between Partners	20
2.5.1. Central Pattern Generators (CPGs) and Related Models	21
2.5.2. Harmonic Oscillator Systems	22
2.5.3. Miscellaneous Shaking Systems	23
2.6. Human Responses to Social Aspects of Robotic Handshaking	24
2.6.1. External Factors in Handshaking	24
2.6.2. Influence on the Perceived Image of the Robot	26
2.6.3. Distinguishing Ability of Handshakes	26
2.6.4. Human-likeness of Robotic Handshakes	27
2.7. Discussion	28
2.8. Conclusion	31
3. MILD: Multimodal Interactive Latent Dynamics	32
3.1. Prologue	33
3.1.1. Related Work	35
3.1.2. Objectives and Contributions	37

3.2. Foundations	39
3.2.1. Variational Autoencoders	39
3.2.2. Hidden Markov Models	40
3.2.3. Inverse Kinematics	41
3.3. Multimodal Interactive Latent Dynamics	42
3.3.1. Learning Interaction Dynamics using HMMs as VAE priors	44
3.3.2. Conditional Training of HRI Dynamics from HHI	46
3.3.3. Inverse Kinematics Adaptation and Stiffness Modulation	47
3.4. Experiments and Results	49
3.4.1. Experimental Setup	50
3.4.2. Datasets	51
3.4.3. Conditioned Prediction Results	53
3.4.4. HRI User Study	60
3.4.5. Bimanual Robot-to-Human Handovers	69
3.5. Conclusion and Future Work	70
3.5.1. Limitations	71
3.5.2. Future Work	71
4. MoVEInt: Mixture of Variational Experts for Learning Human-Robot Interactions from Demonstrations	73
4.1. Prologue	74
4.2. Mixture of Variational Experts for Learning Human-Robot Interactions from Demonstrations	76
4.2.1. Mixture Density Networks	76
4.2.2. Overview	77
4.2.3. GMR-based Interaction Dynamics with MDNs	78
4.2.4. Robot Motion Embeddings	80
4.2.5. Reactive Motion Generation	80
4.3. Experiments and Results	82
4.3.1. Datasets	82
4.3.2. Implementation Details	82
4.3.3. Reactive Motion Generation Results	84
4.3.4. User study	84
4.4. Conclusion and Future Work	86
5. Conclusion	88
5.1. Summary	88
5.2. Future work	90

Appendices

A. Publication List **112**

A.1. Under Review 112

A.2. Journal Papers 112

A.3. Conference Papers 112

A.4. Short Papers/Extended Abstracts 113

B. Curriculum Vitae **115**

List of Figures

1.1. Publications on Human-Robot Handshaking	3
1.2. An overview of the contributions and outline of this thesis. First, insights into Human-Robot Handshaking are drawn from an extensive review of existing works to better shape robotic handshaking behaviours (Chapter 2). Using these insights, developing an end-to-end handshaking behavior that captures all the segments of handshaking is explored as a general approach for additionally segmenting and learning other such physically interactive behaviours (Chapter 3). This modular learning approach is further extended to learn different interactions in a more principled manner (Chapter 4). . .	6
2.1. Conceptual Framework for categorising works on Human-Robot Handshaking	10
2.2. Modified Conceptual Framework for Human-Robot Handshaking. (The proposed suggestions are shaded in grey.)	30
3.1. In this chapter, we explore learning coordinated HRI behaviors using Hidden Markov Models (HMMs) to learn the interaction dynamics over a representation space spanned by a Variational Autoencoder (VAE). The VAEs are trained with the HMMs as a prior to better incorporate the understanding of the dynamics. Moreover, we learn directly from Human-Human interactions by leveraging the kinematic similarities of Humans and Humanoid robots. During testing, the predicted trajectories are adapted with Inverse Kinematics to mitigate errors in the spatial accuracy arising either from prediction errors or from shortcomings in transferring the human’s actions to the robot.	34



3.2. Overview of our approach. We train VAEs to reconstruct the observations of the interactions agents $(x_{1:t}^h, x_{1:t}^r)$ with an HMM prior to learn a joint distribution over the latent space trajectories $p(z_{1:t}^h, z_{1:t}^r)$ of the interacting agents. During test time, the observed agent’s latent trajectory conditions the HMM to infer the robot agent’s latent trajectory $p(z_t^r z_{1:t}^h)$ which is decoded to generate the robot agent’s joint trajectory \hat{x}_t^r . To ensure the proximity of a robot’s hand to the human’s hand during test time, we additionally adapt the predicted trajectory using Inverse Kinematics to fulfill the contact-based nature of the interaction.	39
3.3. Overview of our training approach. (a) We first learn the interaction dynamics from Human-Human Interactions by training both the VAEs and HMMs alternatively by using the HMM as the VAE prior and then by using the VAE embeddings to train the HMM. (b) We subsequently use the model learned from Human-Human Interactions for learning the latent dynamics for HRI. We do so by regularizing the robot VAEs with the learned HMMs and additionally, by training the robot decoder to reconstruct samples from the HMM’s conditional distribution after observing the human partner. . .	43
3.4. Geometric similarities between the degrees of freedom of a human’s upper body and the humanoid robot Pepper. The shoulder roll, pitch, yaw, and elbow angle of a human can be directly mapped to Pepper’s joint angles. .	48
3.5. Sample Handshake HRI on the Pepper robot. The top row shows the result of the generated reactive motion after observing the human partner’s skeleton. The middle row shows the latent trajectories and the HMM segments over the first 3 dimensions of the latent space (Red - Human, Blue - Robot). The opacity of each cluster is the corresponding cluster probability, given by the HMM forward variable. The progression of the HMM forward variable is shown in the bottom row, with different colors denoting the different segments.	59
3.6. Sample Rocket HRI on the Pepper robot. The top row shows the result of the generated reactive motion after observing the human partner’s skeleton. The middle row shows the latent trajectories and the HMM segments over the first 3 dimensions of the latent space (Red - Human, Blue - Robot). The opacity of each cluster is the corresponding cluster probability, given by the HMM forward variable. The progression of the HMM forward variable is shown in the bottom row, with different colors denoting the different segments.	60

3.7. Setup of the User Study for interacting with the Pepper robot. As Pepper is quite short (1.2m), it is placed on a pedestal to match the height of a human partner. The camera behind Pepper tracks the human partner’s motion, which is used to generate Pepper’s motions.	61
3.8. Participant’s ranking of the algorithms. (MILD-IK - yellow stars, MILD - green lines, Base-IK - purple dots). Most participants ranked MILD-IK first, better than both Base-IK and MILD.	64
3.9. Boxplot of the user study responses. (* - $p < 0.05$, ** - $p < 0.01$)	66
3.10. Barplot of the user study responses. (* - $p < 0.05$, ** - $p < 0.01$)	67
3.11. Peculiarities of the different algorithms. The “Base-IK” approach reaches the human’s hand but awkwardly with the robot hand rotated inwards and the elbow pointing out. MILD maintains a human-like posture but falls short of reaching the partner’s hand due to the mismatch in motion retargeting. However, MILD-IK accurately reaches the human’s hand while maintaining a human-like posture.	68
3.12. Example of a Bimanual Robot-to-Human Handover Interaction generated by MILD. The top row shows the observed trajectory of the human in red, and the generated trajectory of the robot in blue. The progression of the interaction can be seen in the bottom row.	70
4.1. Target poses generated in a reactive manner by the mixture of policies learned by MoVEInt for a Handshake interaction with the humanoid robot Pepper. MoVEInt generates multiple policies (shown in green, magenta, and orange) based on human observations which are combined to generate suitable robot motions.	75
4.2. Overview of our approach “MoVEInt”. We train a reactive policy using a Mixture Density Network (MDN) to predict latent space robot actions from human observations. The MDN policy is used not just for reactively generating the robot’s actions, but also to regularize a VAE that learns a latent representation of the robot’s actions. This regularization ensures that the learned robot representation matches the predicted MDN policy and also ensures that the robot VAE learns to decode samples from the MDN policy.	77
4.3. Sample Human-Robot Interactions generated with the reactive motions generated by MoVEInt for a Bimanual Handover scenario.	85

4.4. Sample trajectories generated by MoVEInt for the Bimanual and Unimanual Handovers in the HHI-Handovers dataset in [103]. The 3D plots show the reconstructed trajectories and the 2D plots show the corresponding progression of $\alpha_i(\mathbf{x}_t^h)$ for the different components of the MDN. In the 3D plots, the observed trajectory of the receiver is shown in red and the generated trajectory of the giver is shown in blue and the giver’s corresponding ground truth is shown in black. The reconstruction of the individual latent components of the MDN are shown in green, magenta, and orange. It can be seen that the learned components correspond to different parts of the task space. For example, green denotes the hand locations for a unimanual handover, magenta denotes the hand locations for a bimanual handover, and orange denotes the static hand locations for the starting and ending neutral poses. In the 2D plot, it can be seen how the coefficients for components corresponding to bimanual (magenta) and unimanual (green) get activated based on the interaction being performed, while the component corresponding to a neutral pose (orange) gets activated at the beginning of the interaction while both partners are static. 85

List of Tables

2.1. Summary of Results of [122, 124, 189] (M - Male, F - Female. Values are reported as mean \pm standard deviation and only mean is reported when standard deviation is not available.)	13
2.2. Methods and Parameters used by different works to evaluate robotic handshaking.	15
2.3. Different types of Robot End Effectors used for Human-Robot Handshaking. *-simulated robot	16
2.4. Division of works studying the shaking motion.	20
3.1. Prediction MSE (in cm) for the second human partner's trajectories after observing the first human partner averaged over all joints and timesteps. (* - $p < 0.05$, ** - $p < 0.01$, Lower is better)	53
3.2. Differences in conditioning and sampling strategies of the variants of MILD.	55
3.3. Prediction MSE (in radians) for robot trajectories after observing the human interaction partner averaged over all joints and timesteps. (Lower is better, Lowest MSE is highlighted, Significance Values shown in Tables 3.4, 3.5, and 3.6)	55
3.4. p -values for pairwise MSE comparisons of the results reported on the HRI-Yumi scenario from the dataset of [29] in Table 3.3. For $p < 0.001$, we report zeros.	56
3.5. p -values for pairwise MSE comparisons of the results reported on the HRI-Pepper scenario from the dataset of [29] in Table 3.3. For $p < 0.001$, we report zeros.	57
3.6. p -values for pairwise MSE comparisons of the results reported on the HRI-Pepper scenario from the NuiSI dataset in Table 3.3. For $p < 0.001$, we report zeros.	58



3.7. Results of a one-way Repeated Measures ANOVA for each of the survey items. Values less than 0.01 are reported as 0. The last three columns show the mean differences between the different algorithms (1 - Base-IK, 2 - MILD, 3 - MILD-IK) which were analyzed using paired sample t-tests (* - $p < 0.05$, ** - $p < 0.01$).	65
4.1. Statistics of the different datasets used.	82
4.2. Prediction MSE for robot trajectories after observing the human partner averaged over all joints and timesteps. Results for the HHI scenarios are in cm and for the HRI scenarios are in radians. (Lower is better)	83
4.3. Number of successful handovers of each object by the Kobo robot to each user (total of 5 per object per user i.e. total of 20 per object and 15 per user).	86

List of Acronyms

AE	Autoencoder
BIP	Bayesian Interaction Primitive
CASA	Computers as Social Actors
CPG	Central Pattern Generator
DMP	Dynamic Movement Primitive
DTW	Dynamic Time Warping
EKF	Extended Kalman Filter
EMG	Electromyography
FSR	Force Sensing Resistor
GRU	Gated Recurrent Unit
HRI	Human-Robot Interaction
HHI	Human-Human Interaction
HMM	Hidden Markov Model
IK	Inverse Kinematics
IMU	Inertial Measurement Unit
IP	Interaction Primitive
LSTM	Long Short Term Memory Network
MDN	Mixture Density Network
MHLG	Model Human Likeness Grade
MS	Mutual Synchronization
PAD	Pleasure-Arousal-Dominance
pHRI	Physical Human-Robot Interaction
ProMP	Probabilistic Movement Primitive
RL	Reinforcement Learning
RNN	Recurrent Neural Network
VAE	Variational Autoencoder

1. Introduction

“The hands of those I meet are dumbly eloquent to me. The touch of some hands is an impertinence. I have met people so empty of joy, that when I clasped their frosty finger-tips, it seemed as if I were shaking hands with a northeast storm. Others there are whose hands have sunbeams in them, so that their grasp warms my heart. It may be only the clinging touch of a child’s hand; but there is as much potential sunshine in it for me as there is in a loving glance for others. A hearty handshake or a friendly letter gives me genuine pleasure”

Helen Keller, The Story of My Life

1.1. Motivation

In the context of Social Robots and Human-Robot Interaction (HRI), there has been an increase in the use of robots in social settings, and they already support human users in various industries, such as retail, gastronomy, hotels, education, and healthcare services [56, 68, 78, 211]. In HRI, physical presence plays an important role as it can influence the image of a robot as compared to a virtual presence [108]. According to the Computer-As-Social-Actor (CASA) paradigm [135], physical presence is also an important predictor of the mindless social response to robots by which humans put a robot in the same category as humans. These responses are supposedly triggered by social human-like cues as well, such as voice [137], face [134], and language style [136]. Additionally, the appearance of the robot, as opposed to just the physical presence, can also have an impact on the perception of robots. As the Uncanny Valley [127] hypothesises, uncanny feelings are triggered by highly human-like robots (androids) but not as realistic as a human being as compared to robots that are less human-like in appearance (humanoids). These uncanny feelings are hypothesised to get exaggerated even more in the case of movements. The affinity, of both appearance and movement, rises again only when the human-likeness becomes very close to that of a human being.

In the context of HRI, physical contact plays a central role in the various applications of social robots. One of the key reasons for this is that non-verbal behaviour, especially

touch, can be used to convey information about the emotional state of a person [70, 141, 208]. This enables a special kind of emotional connection to human users during the interaction [68, 185]. If, for example, one considers a future scenario in which an accompanying robot shares the habitat with humans, a key requirement for the robot would be its ability to physically interact with humans [201]. In such a case, it would be advantageous that humans feel more welcoming and be willing to interact and help a social robot with its task like they would help other humans. Among such interactions, handshaking is a common natural physical interaction and an important social behaviour between two people [173], that is used in different social contexts [36, 184, 66]. The importance of touch in HRI and the prolonged nature of the contact additionally make handshaking a more important interaction as compared to other interactions, like high fives, fist bumps or other Asian greeting behaviours which do not involve physical contact. Handshaking can, therefore, represent an important social cue according to the CASA paradigm for several reasons:

- It is one of the first and foremost non-verbal interactions which takes place and should, therefore, be part of the repertoire of a social robot.
- It plays an important role in shaping the impressions of others [36, 184, 22], which is used to develop an initial personality judgement of a person [11, 10].
- The shaking of hands is seen as a symbol of greeting, farewell, agreement, or congratulation. Socially, it symbolises acceptance and respect for another person [150]. The most common of these settings is "greeting" where it is usually the first non-verbal interaction taking place in a social context.
- The shaking of hands may help set the tone of any interaction, especially since the sense of touch can convey distinct emotions [70].
- A good robot handshake may lead to future cooperation and coexistence [150].

Such interactions can also be further enriched, such as in the possible scenario wherein a robot can monitor the biological attributes of a person (such as stress levels from blood flow) and thereby infer social information about a person from just a single handshake [51]. Having human-like body movements plays an important role in the acceptance of HRI wherein humans tend to look at robots more as social interaction partners [104]. In the case of humanoid robots, having realistic motions enable similar responses as humans [35]. Thus, having a good handshake can not only widen the expressive abilities of a social robot but also provide a strong first impression for further interactions to take place. Robot handshaking can additionally help improve the perception of the robot and enable humans

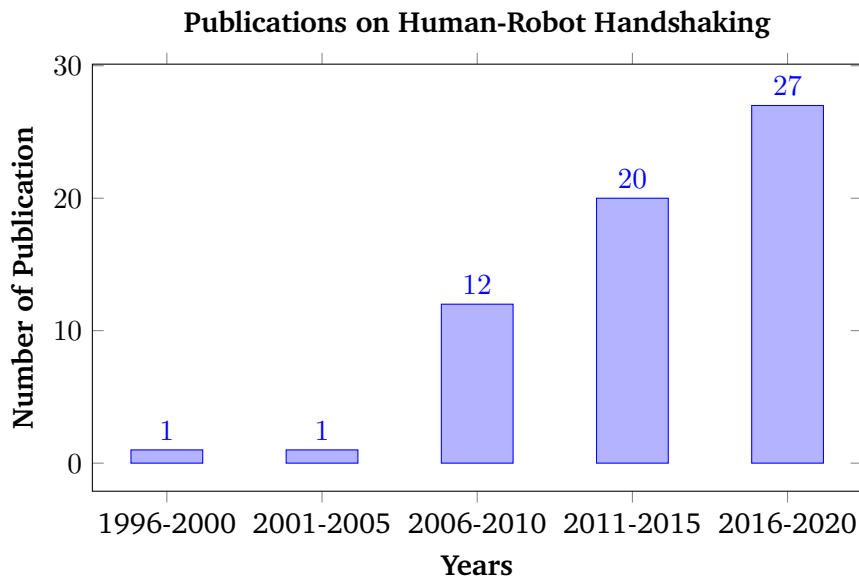


Figure 1.1.: Publications on Human-Robot Handshaking

to be more willing to help a robot [12] allowing for better integration of the robot into human spaces. To perform proper handshaking motions, a social robot should be able to detect and predict the movements of the human and react naturally. Therefore, for better acceptance and improved expressiveness, effective handshaking behaviours need to be present to make social robots feel more acceptable. This importance can be seen in Figure 1.1, in the rising trend of works on human-robot handshaking.

Finally, handshaking is just one kind of physically interactive behaviour that humans perform. It is therefore of additional interest to be able to generalise learnt behaviours to similar interactions, like fist-bumps, high fives among others, thereby building a library of such physical interactions. Such interaction libraries have been explored in the past [99, 98, 112] for learning robotic skills, which can additionally be combined with the ability to generate new physically interactive behaviours by sequencing learnt skills [109, 111, 110]. Such capabilities of being able to adapt to new interactions and potentially create new ones could improve the sociability of a robot.

1.2. Research Questions

Physically interactive behaviours can help yield a new dimension to social robots through the sense of touch. In this regard, the act of handshaking, due to the prolonged nature of the physical contact, can have a stronger impact compared to other interactions with a nearly instantaneous phase of contact. In addition to the emotions associated with such an interaction, the motions and the synchronisation also play an important role, not just in handshaking but for other similar physically interactive behaviours as well. It is therefore important for a social robot to understand both the motions and emotions of such interactions. Keeping these points in mind, this section presents the key research question aimed to be answered by this thesis.

What can we learn from existing Handshaking interactions?

As seen in Figure 1.1, there is a rising trend in works exploring the act of handshaking within the context of HRI. These works cover various approaches to craft different kinds of robotic behaviours and the personality traits portrayed by them, the acceptance of such an interaction with a robot, the specifics of the motions in different contexts and so on. Therefore it is important to first understand what these various works explore, obtain an overall classification of the works and draw insights from these works in order to develop socially acceptable robotic handshaking behaviours.

How can we learn seamless end-to-end Robot Handshaking behaviors?

Handshaking can be naturally divided into multiple segments. Once an intent to handshake has been established between the partners, they initially start moving in a timely manner towards a common point while adapting to the motions of one another. Once the hands are grasped, they begin an oscillatory movement after which the partners terminate the interaction and go back to a neutral pose. At each step of the interaction, the partners respond to one another and move in a timely, adaptive and mutually synchronoized manner. Understanding these segments and learning to move in a coordinated manner while progressing through the different segments is an important aspect for realizing a smooth and acceptable handshake behaviour.

How easily can a social robot generalise to other physical interactions beyond handshaking?

Many physically interactive behaviours have a similar progression and semantics as handshaking. Typically, there is a reaching phase where both partners bring their hands together, a contact phase which contains the core of the interaction and a receding phase which marks the end of the interaction. Therefore, an important skill for a social robot would be the ability to generalise to different physically interactive behaviours, like fist-bumps, hand waves, object handovers etc. For such social abilities, it is important that the robot recognises the interaction to perform, follows the motion of the human partner and reacts in a timely manner in order to ensure that the interaction is natural and well perceived.

1.3. Contributions

This section summarises the main contributions of this thesis aimed at tackling the research questions raised above in Section 1.2, an overview of which is shown in Figure 1.2. Specifically, this thesis contributes towards establishing a taxonomy of existing works on Human-Robot Handshaking, from which insights are drawn to (i) learn the underlying segments of such physically interactive behaviors and their subsequent adaptation for ensuring a natural interaction with a robot and (ii) exploring how to such a learning approach can be leveraged for blending multiple different underlying strategies for enabling more accurate Human-Robot Interactions.

In order to answer the first research question, an extensive review of existing Human-Robot Handshaking works is performed, in which, the following contributions are developed.

- A taxonomy over existing works on Human-Robot Handshaking.
- Understanding the different kinds of algorithms developed and how they affect the interaction.
- Extracting insights regarding the social and emotional aspects of handshaking.
- Proposals for future research based on open questions from existing works.

Coming to the second and third research questions, we explore the use of latent state space models for the learning and segmentation of physically interactive behaviors in an unsupervised manner from demonstrations. Specifically, in Chapter 3, we study the

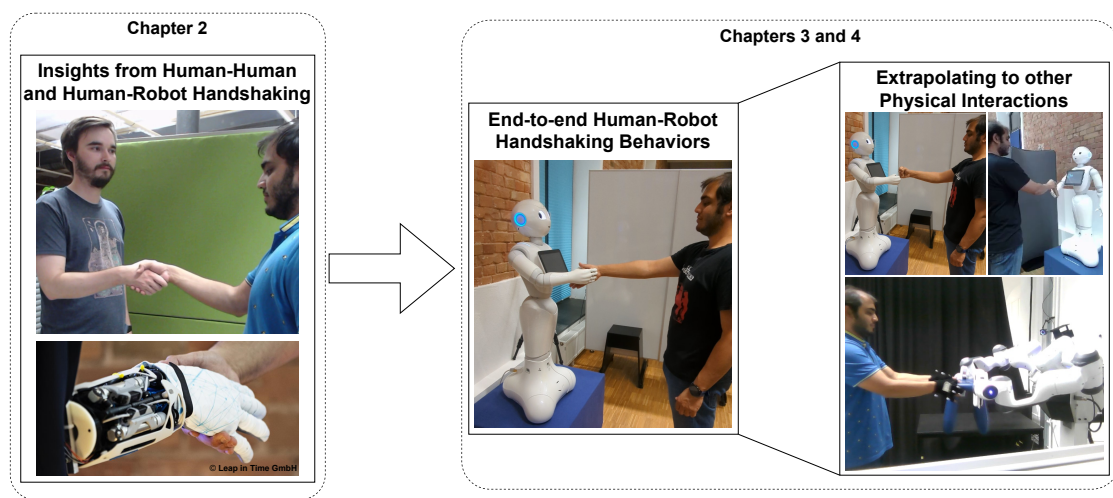


Figure 1.2.: An overview of the contributions and outline of this thesis. First, insights into Human-Robot Handshaking are drawn from an extensive review of existing works to better shape robotic handshaking behaviours (Chapter 2). Using these insights, developing an end-to-end handshaking behavior that captures all the segments of handshaking is explored as a general approach for additionally segmenting and learning other such physically interactive behaviours (Chapter 3). This modular learning approach is further extended to learn different interactions in a more principled manner (Chapter 4).

use of Hidden Markov Models (HMMs) which naturally yield themselves as a method for learning a joint distribution over a set of interaction trajectories [156, 31, 52]. We propose a method for learning Multimodal Interactive Latent Dynamics, “MILD” wherein we learn to segment the interaction demonstrations by training HMMs in the latent space learned by a Variational Autoencoder (VAE) trained to reconstruct the demonstrated trajectories. The VAE in turn uses the learned HMMs as an informative prior distribution, which we show helps improve the learning. We do so by first training on Human-Human Interaction data, which is then used to further regularize and improve the predictions of robot actions for enabling accurate reactive motion generation. We further explore the incorporation of Inverse Kinematics to ensure spatially accurate interactive behaviors. Specifically for handshaking, ensuring a compliant motion is key for realizing a natural interaction. Therefore, we perform stiffness modulation using the HMM segment predictions to enable a compliant and natural handshake motion during the segments involving physical contact. We validate our proposed approach via a user study where participants interact with a Humanoid Social Robot controlled by MILD. Our user study shows that MILD is able to generalize well to multiple interaction partners despite being trained on data from just two humans interacting with each other.

Taking a step further into the third research question, We aim to bridge the gap in the framework presented in Chapter 3 by incorporating the underlying mixture distribution learned by the HMMs into the learning process in a more principled manner. To this end, in Chapter 4, we propose “MoVEInt”, a method for learning a Mixture of Variational Experts for learning Interaction dynamics through a shared representation of a human and a robot. We show how the underlying reactive motion generation using HMMs can be approximated using Mixture Density Networks (MDNs) [24] by learning a set of latent policies which when blended, enable learning accurate Human-Robot Interactions.

1.4. Outline

After the introductory chapter (Chapter 1), the rest of the thesis is organised as follow.

Chapter 2 consists of a systematic review of Human-Robot Handshaking in order to extract insights that can be used for developing new robotic handshaking behaviours. A taxonomy of existing works is developed, and the works in each category are analysed and the major findings are discussed. The chapter concludes with suggestions for future works based on open problems in robotic handshaking. This chapter is based on the works in [160, 161].

Chapter 3 aims to examine how the underlying segments inherent to physical interactions like handshaking can be learned and how such a modular structure can be used

to generate acceptable Human-Robot Interactions. A framework is be developed which uses Hidden Markov Models (HMMs) for unsupervised segmentation of the interactions in the latent space learned by a Variational Autoencoder (VAE). We further demonstrate the ability of our approach to generate acceptable robot behaviors via a user study where participants interact with a Humanoid Social Robot controlled by our approach. This chapter is based on the works in [158, 157].

Chapter 4 extends the framework presented in Chapter 3 by exploring how to incorporate the reactive motion generation into the training in a more principled manner. We demonstrate how a formulation using Mixture Density Networks (MDNs) yields itself naturally for learning a latent mixture of policies that enable learning physical Human-Robot Interactions in a more accurate manner. This chapter is based on the work in [159].

Chapter 5 provides a conclusive summary of the work that has been done in this thesis and provides an outlook on future research directions.

2. Human-Robot Handshaking: A Review

Now that the importance of handshaking from a robotic standpoint has been established in Chapter 1.1, we first delve into works that tackle various aspects of Human-Robot Handshaking. This chapter aims to provide an extensive review of Human-Robot Handshaking, develop a taxonomy over existing works, analyse the findings of each of the categories and in the end provide some ways forward for future research in this area.

2.1. Prologue

We propose the following framework for categorising the different works, shown in Figure 2.1. We first discuss works that aim to model handshakes from human-human interactions. Along with that, we discuss the evaluation criteria used for analysing participants' feedback in experiments with humans interacting with the robot. This can be seen in Section 2.2. Following this, we go through each of the stages of handshaking, as depicted in [122, 198], namely reaching (Section 2.3), grasping (Section 2.4) and shaking (Section 2.5). We then explore more into the responses of participants to different handshaking methods along with various factors, such as their acceptability, their degree of preference, how external factors like voice and gaze affect the perception of robot handshakes and so on in Section 2.6. Finally, we discuss some of the shortcomings of existing works and propose some areas for future research in Section 2.7 and present some concluding remarks in Section 2.8. This was done as a broad categorisation of the works obtained after a search of digital libraries, Google Scholar, IEEE Xplore and ACM Digital Library using keywords that included "human-robot handshaking", "robotic handshaking" and "handshaking AND human-robot interaction / HRI" followed by a depth-first search-styled approach among the references and citations of the papers found. Of these, we included all the articles that were published in conferences or journals. The different works that model human handshaking interactions are looked into, in addition to the works in the different stages of the handshaking exchange and some social responses of participants. In this chapter, we dig deeper into these aspects and provide key findings and insights into the different areas of work in regards to Human-Robot Handshaking. In the end, we propose some

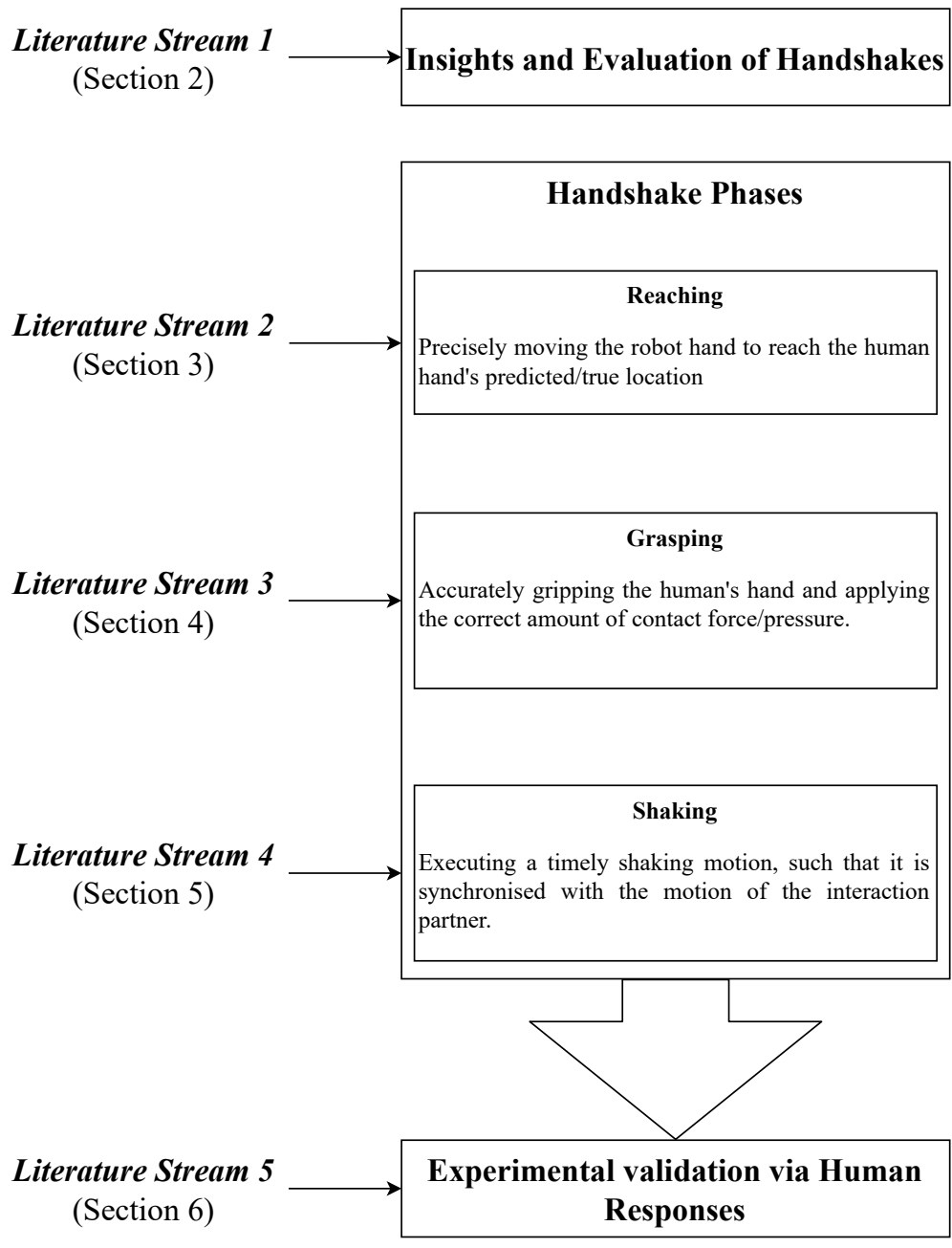


Figure 2.1.: Conceptual Framework for categorising works on Human-Robot Handshaking

ways forward based on aspects that still need to be worked on in order to realise a social robot that can truly capture the intricacies of such a physically interactive action.

2.2. Insights and Evaluation of Handshakes

Before going into robotic handshakes, we first discuss a few works that draw insights from human handshakes and some evaluation mechanisms used to measure the parameters related to human acceptance of robotic handshakes. Regarding the insights drawn, the main parameters looked at include trajectory profiles (mainly acceleration, velocity, contact forces) between participants when they shake hands and the mutual synchronisation of their movements while doing so. Regarding the evaluation criteria, they mainly relate to how human-like the handshake is and how the handshakes are perceived by humans who interact with the robot.

2.2.1. Insights from Human-Human Handshake Interactions

A group of researchers from Okayama Prefectural University, Japan conducted a series of studies [79, 80, 84, 81, 82, 83, 85, 207, 146, 147] to study handshaking interactions between humans and analyse how participants respond to different behaviours applied on a custom robotic arm. They use a VICON motion capture system to track markers placed on both shoulders, right elbow wrists and hands of both participants. Initially, Jindai et al. [85] observed that the motion of the responder was seen to be similar to that of the requester with a lag between their velocity profiles, which was found to be similar to a minimum jerk trajectory profile. Hence they applied a lag-based transfer function for generating the robot's reaching trajectory based on the human's reaching motion. It was previously shown that this type of "motion transfer" was emotionally acceptable to humans in an object-handover scenario [206]. Subsequently, the oscillatory motion profiles of the observed shaking behaviours were modelled as a spring and damper system [81, 207]. Following this, the interactiveness was studied by modelling the requester [83, 79] and responder [146, 147] behaviours with respect to delays in the motions and auxiliary behaviours like gaze and speech. The result of applying the key findings from the above works on their robot is explored in Section 2.6. Their main findings of the modelling can be summarised as follows.

- The reaching trajectories of the requestor and responder are similar in their trajectory profiles, which follow a minimum jerk trajectory model in [83], and the motion of one can be used to mimic the other, such as by using a transfer function [85].

-
- As expected, the shaking behaviour and the transition from reaching to shaking is modelled as a spring-damper system, given the oscillatory nature of the behaviour [207].
 - Leading handshake behaviours were preferred over a non-leading one [207] and a small delay (0.2s - 0.4s) between responding to a handshake request was better perceived.
 - In terms of auxiliary behaviours, using voice with a small or no delay was preferred and having the gaze shift steadily from the hand to the face was well perceived.

A group of researchers from the University of Lorraine, France explicitly study the mutual synchronisation (MS in short) between participants while shaking hands along with the forces exerted on the palms [122, 124, 125, 126, 189]. They firstly study the hand motions by having the participants wear a glove with an Inertial Measurement Unit (IMU) and 6 force sensors placed around the palm [125]. Tagne et al. [189] further investigate the joint motions as well (elbow and shoulder) with IMUs placed at each joint. The influences of a few different social settings, such as greeting, congratulating or sympathy, are then explored as well [124, 189]. The MS between participants is analysed using the Fourier analysis of the input signals (mainly accelerations). However, wavelet transforms are shown to qualitatively estimate the different stages of a handshake interaction as well [126].

Initially, the MS between participants was explored in a context-less setting along with the contact strength of the interaction [122]. The mean duration of a handshake was around 2.67 ± 0.87 seconds during which the average duration of grasping was similar across pairs (0.5 s) whereas the duration of shaking had a larger variation, from just below 1 s to almost 2.5 s. The frequency of shaking during MS peaked at around 4Hz. The average strength of the contact, which is the average of the forces measured by the sensors was 2.5N (no standard deviation was reported). This framework was extended to analyse differences in a few social contexts [124, 189] and gender-based differences[124]. Tagne et al. [189] observe the differences between 3 scenarios namely greeting, sympathy and congratulating. Melnyk and Hénaff [124] analyse similar social settings, namely greeting and consolation, and additionally analyse the trends across different gender-based pairings.

As seen in Table 2.1, a shorter duration was observed in greeting contexts. The duration in cases of sympathy and congratulations were similar. The grip strength shows contradictory results. Tagne et al. [189] saw the lowest grip strength in case of sympathy, followed by greeting and then, congratulations. Melnyk and Hénaff [124] found that it was slightly higher in consolation case although not significantly. In terms of gender-based pairings, it

Study	Setting	Duration (s)	Wrist Frequency during MS (Hz)	Grip Strength (N)	
Melnyk et al. [122]	None	2.67 ± 0.86	4.20 ± 1.00	2.50	
Tagne et al. [189]	Greeting	0.90 ± 0.26	2.43 ± 0.62	4.37 ± 2.4	
	Sympathy	1.30 ± 0.49	2.44 ± 0.69	3.12 ± 2.1	
	Congratulation	1.24 ± 0.40	2.66 ± 0.72	5.88 ± 3.2	
Melnyk and Hénaff [124]	Greeting	MM	0.73 ± 0.08	3.6 (median 3.5)	6.02 ± 0.99
		MF	1.48 ± 0.40		5.68 ± 0.8
		FF	1.95 ± 0.26		5.53 ± 1.17
	Consolation	MM	2.40 ± 0.23	3.7 (median 3.5)	6.08 ± 1.04
		MF	2.54 ± 0.42		6.18 ± 0.97
		FF	4.05 ± 0.53		6.29 ± 0.85

Table 2.1.: Summary of Results of [122, 124, 189] (M - Male, F - Female. Values are reported as mean \pm standard deviation and only mean is reported when standard deviation is not available.)

is seen that MM pairs shook for a lesser duration as compared to mixed pairings. Female pairs shook hands the longest. This is consistent with another study [144] as well. No conclusive correlations were found between gender and grip strength, contrary to previous studies[36, 144].

Unlike the above works that explicitly measure the stiffness and forces of the interactions, Dai et al. [47] indirectly model the stiffness of the elbow joint as a spring-damper system, like few other works described in Section 2.5. They measure the expansion/contractions of the muscles in the upper arm and forearm using EMG signals and thereby estimate the stiffness of the elbow using the biceps and triceps and use the forearm muscle measurements to observe indications of the grasping forces. Data was collected from 10 handshakes of which 5 were weak and 5 were strong. It was very evidently seen that muscle activation in the case of the strong handshake was higher than the weak condition.

The above works mainly study human-human handshaking to gain insights into the motion and forces involved in handshaking. However, works looking into how well robotic interfaces are suited for such haptic-heavy interactions are limited in number. In this regard, Knoop et al. [97] perform experiments to understand the contact area, contact pressure and grasping forces exerted by participants during handshaking and test out how a few robotic hands and custom finger designs comply with their observations from human-human handshaking interactions. Participants were asked to perform 3 handshakes of different strengths, namely, weak, normal and strong. A large variation was seen in the final contact positions of the fingers at the back of the hand. In the front, there is little variation across different handshakes as almost the whole palm is held during the contact.

This would imply that there is possibly no fixed grasping location for a handshake and that the palms should be sufficiently covered. during the interaction. A positive correlation was found between contact pressure and grasping force, which is a straightforward implication since the contact area doesn't vary during the interaction. They test out how a few robotic hands and custom finger designs compare with a human hand and argue that this study is useful for optimising robotic hand designs at a coarse level.

Major Findings. Handshaking is inherently a synchronous process, which is observed in the reaching motions by [85, 84] and the shaking by [122, 124, 189]. This would imply that both parties involved in a handshake try to achieve a common motion during the action. This kind of inherent similarity in the motions and the subsequent synchronisation can therefore be treated as an important aspect of making a handshaking behaviour more acceptable. The context of a handshake along with additional factors, like speech and gaze, play a role in the interaction as well. The factors studied and the measurements obtained (in terms of duration, frequency, relative grip strength etc.) can further be used to explicitly model robotic handshake behaviours to give them a "personality" of sorts or provide appropriate responses based on detected interaction contexts improving the social understanding of the robot.

2.2.2. Handshake Evaluation Methods

Given the differences in hardware and the evaluation criteria used by different works, it is difficult to converge on a single metric or scale for the task at hand. Moreover, different works evaluate different aspects of handshaking, using different robots. It is therefore difficult to come up with a common comparison baseline, although some studies evaluate their methods similarly. To this end, we collate some of the common evaluation metrics and methods used among different works in Table 2.2 and broadly categorise some of the different robotic interfaces in Table 2.3. Some common aspects are the aim to rate the acceptability of the handshaking interactions and the human-likeness or the naturalness of the handshaking. Moreover, most of the works that use a human-like end effector mainly have an inactive one, which could cause the interaction to seem more unnatural. For example, in the works of Wang et al. [203] and Giannopoulos et al.[60] who use a rod-like end-effector, the case when the robot arm is operated by a human, the handshake gets an average human-likeness rating of only 6.8/10 which is far from the maximum score.

The common metrics used, like in most psychological and human studies, are the seven-point or the five-point Likert scales, which is a bipolar scale that has a negative valued sentiment on one end and a positive valued one on the other, which allows for a nice representation, especially when averaging over the data. For example, an overall negative

Evaluation Method	Works	Evaluation Parameters
Bradley Terry Model	Jindai et al.[79, 80, 84, 81, 82, 83, 85], Ota et al. [146, 147], Yamato et al. [207]	Participant's preferences of handshakes
	Kasuga and Hashimoto [91]	Flexibility, naturalness, kindness, affinity
Seven point scale	Jindai et al.[79, 80, 84, 81, 82, 83, 85], Ota et al. [146, 147], Yamato et al. [207]	Handshake motion, Security, Velocity/comfort, politeness/Vitality
	Avelino et al. [12]	RoSaS (Warmth, competence, discomfort), Godspeed (anthropomorphism, animacy, likeability), closeness, willingness to help robot
	Mura et al. [129], Vigni et al. [197]	Quality, human-likeness, responsiveness, perceived leader, personality
Five point scale	Ammi et al. [5], Tsamalal et al. [194]	Valence, Arousal, Dominance of Visual, haptic and visuohaptic interactions
	Arns et al. [8]	Compliance, force feedback, overall haptics
	Christen et al. [41]	Naturalness of video of different simulated interactions
Score (out of 10)	Wang et al. [205, 204], Giannopoulos et al. [60]	Human likeness rating of Robot handshakes
	Dai et al. [47]	Naturalness
Model Human Likeness Grade	Avraham et al. [15], Karniel et al. [90], Nisky et al.[140]	Human likeness of proposed handshake models

Table 2.2.: Methods and Parameters used by different works to evaluate robotic handshaking.

		With force feedback	Without force feedback
Human-like hand (4/5 finger model)	Inactive	Dai et al.[48]	Campbell et al. [33], Jindai et al.[79, 80, 84, 81, 82, 83, 85], Kasuga and Hasimoto [91], Knoop et al. [97], Melnyk and Henaff [123], Nakanishi et al. [132], Orefice et al. [145], Ota et al. [146, 147], Stock-Homburg et al. [186], Vanello [196], Vinayavekhin et al. [198], Yamato et al. [207]
	Passively controlled	Arns et al. [8], Beaudoin et al. [20], Ouchi and Hashimoto [148], Pedemonte et al. [155, 154]	
	Actively controlled	Avelino et al. [12, 14], Ammi et al. [5], Christen et al.[41]*, Mura et al.[129], Tsamalal et al.[194], Vigni et al.[197]	
Gripper			Bevan and Fraser [23], Falahi et al. [55], Jouaiti et al.[88]*, Sato et al.[171]*
Rod-like end-effector			Avraham et al.[15], Giannopoulos et al.[60], Karniel et al.[90], Nisky et al.[140], Papageorgiou and Doulgeri[150], Wang et al.[204, 203]

Table 2.3.: Different types of Robot End Effectors used for Human-Robot Handshaking. *-simulated robot

average indicates an inclination towards the negative sentiment and a positive average indicates an inclination towards the positive sentiment. This, in contrast to comparing an absolute score (rating out of ten for example), can help indicate the sentiments of the participants better.

To use a more traditional test of computational intelligence, Karniel et al. [90] propose a Turing test for motor intelligence and come up with a metric called the Modern Human Likeness Grade (MHLG) which is used to indicate the human-likeness of different shaking behaviours in this kind of a mechanical Turing Test. This is based on the perceived probability by a participant of the model being a human shaking the stylus or the algorithm. Nisky et al. [140] propose different ways to perform this Turing test for motor intelligence. These are described in further detail in Section 2.5.3.

2.3. Reaching phase of Handshaking

We have already described the work of Jindai et al. [79, 83, 85] and Ota et al. [146, 147] above. To the best of our knowledge, these were the first works to model the hand reaching aspect and deploy it on a robot. As mentioned above they propose two models. One with a transfer function based on the human hand's trajectory with a lag element and the other is a minimum jerk trajectory model, which fits the velocity profiles and provides smooth trajectories by definition. These modelling choices imply that a smooth motion similar to the interaction partner is preferred with a small amount of delay between them. However, they do not have any study showing how these two models compare with each other.

More recent works model reaching using machine learning. Campbell et al. [33] use imitation learning to learn a joint distribution over the actions of the human and the robot. During testing time, the posterior distribution is inferred from the human's initial motion from which the robot's trajectory is sampled. Their framework estimates the speed of the interaction as well, to match the speed of the human. Christen et al. [41] use Deep Reinforcement Learning (RL) to learn physical interactions from human-human interactions. They use an imitation reward which helps in learning the intricacies of the interaction. Falahi et al. [55] use one-shot imitation learning to kinesthetically teach reaching and shaking behaviours based on gender and familiarity detected using facial recognition. However, it cannot be generalised due to the extremely low sample size. Vinayavekhin et al. [198] model hand reaching with an LSTM trained using skeleton data. They predict the human hand's final pose and devise a simple controller for the robot arm to reach the predicted location. In terms of smoothness, timeliness and efficiency, their method performs better than following the intermediate hand locations. However, it

performs worse than using the true final pose due to inaccuracies in the prediction.

Major Findings. Modelling of reaching behaviours draws heavily on learning from human interactions, unlike other robotic grasping/manipulation tasks, where a lot of it can be learnt from scratch. This provides a strong prior to help make the motions more human-like and can also be used to initialise [33] or guide [41] the learning.

2.4. Controlling Hand Grasps in Handshaking

One of the first remote handshaking systems, proposed by Ouchi and Hashimoto [148], was aimed at two people performing a handshake while on a telephone call with each other using a custom-made silicone-rubber based robotic soft hand. They measured the pressure exerted on the hand using a pneumatic force sensor which relays the force information over to the other user, who has a robotic hand as well. With this type of active haptic mechanism, they show that users better perceive the partner's existence during the call and that they were able to shake hands without feeling any transmission delay. This shows the effect that such haptic interactions have on the perception of the interaction partner.

Pedemonte et al. [155] design an anthropomorphic haptic interface for handshaking. It is an under-actuated robot hand with a passive thumb that is controlled based on the amount of force that is applied to it. It is a sensor-less model with a deformable palm that controls the closure of the fingers. A variable admittance controller is used to set the reference position for fingers based on the degree of deformation of the palm. Therefore the amount of force exerted by the robot hand on the human hand depends on the force exerted by the human, leading to a partial synchronisation in the grasping. It takes approximately 0.6s to close the fingers. Arns et al. [8] build upon this design using lower gear ratios and more powerful actuators to obtain a stronger grasping force and a faster interaction speed. They argue that the use of impedance control as opposed to admittance control helps improve responsiveness as well. A similar synchronisation is observed as in the previous work as the mechanisms are the same, in theory. The main difference is the speed of the interaction which is almost instantaneous in this case (less than 0.05s), making the interaction more realtime and natural.

Avelino et al. [13, 14] propose two models to develop a pleasant grasp for handshaking. This is extended to three different grasping models with different degrees of hand closure, corresponding to strong, medium and weak handshakes [13]. Force sensors present on the robots finger joints measure the interaction forces during grasping. It was found that female participants mainly preferred strong handshakes (85.7%). There was a larger variability among male participants. Since a simple position based control is employed, the force perceived depends on the hand sizes of the participants, which could be the

cause of the variability. This is addressed in [14], where an initial study is carried out where participants have to adjust their hand and the robot's grip until a preferable grasp is reached. This is done to find a suitable reference force distribution among the sensors on the robot hand. The finger joint positions are recorded as well. With this distribution, they compare a fixed handshake to a force control method. The force control is done with a PID controller whose set points are the average of the forces per sensor on each finger obtained from the previous data. Moreover, they combine this with a shaking motion presented in [205] that is described in Section 2.5. Participants had to rate the two handshakes based on various factors like scariness, arousal (boring/interesting), meaningfulness, excitement, strength/firmness, perceived enjoyment and safety, all on 7-point scales for each variable. Although both handshakes were evaluated positively overall, no significant differences were observed between them.

Vigni et al. [197] model the force exerted by the robot hand during handshaking based on the force exerted by the human, measured using force-sensitive resistors on the robot hand. The robot force is approximated from the degree of hand closure using a calibration experiment where participants are asked to mimic the force felt on their hand by a few open-loop handshakes of the robot. The human force is estimated by fitting a cubic polynomial to the sum of forces applied on the individual sensors. This is also done with a calibration experiment where participants were made to grasp a sensorised palm fitted with a load cell to measure the exerted force. They compare three different controllers based on the relationships between the exerted forces of the human and the robot namely linear, constant and combined (constant+linear). The latter two are used with two values of the constant force, weak and strong. Since humans have a small delay in reaction time, a controller delay of 120ms was observed to be more natural and was added to the behaviour.

The mean duration of handshakes was 2.2s with 24.8N as the mean sum of forces measured on the robot hand exerted by the humans. The participants (n=15) filled out a survey after interacting, rating the quality, human-likeness, responsiveness, perceived leader, and the perceived personality of the robot on a 7 point scale. The combined controllers were perceived better than the constant ones in terms of quality, human-likeness and responsiveness, with a significant difference between the weak variants. There was no significant effect in terms of who the perceived leader or follower was. However, it was observed that in the constant force cases, humans would adjust their force based on the robot's, showing that humans tend to follow the force exerted on their hands. These findings further emphasise the effect of mutual synchronisation in handshaking. In terms of personality, the stronger variants of the constant and combined controllers were perceived as more confident/extroverted, with a significant effect seen between the variants of the constant controller.

Major Findings. The main commonality among the above-mentioned works is that a force feedback mechanism is necessary to ensure good grasping since it enables a mutual synchronisation between the participants. To this end, although there is no force sensing mechanism as such in the hand designed by Pedemonte et al. [155] and Arns et al. [8], they still passively control the closure of the hand based on the deformation, thereby producing a similar synchronous behaviour. Vigni et al. [197] observed that the grip strength had an effect on the perception of the robot’s personality, which is consistent with the findings of Orefice et al. [144], from human-human handshakes. This can help in crafting behaviours to explicitly yield a personality to the robot, rather than observing such a personality passively. Additionally, encoding different types of such explicit behaviours can help the robot switch to adapt to the human interaction partner if necessary.

2.5. Shaking Motions and Synchronisation between Partners

In terms of shaking, one can easily say that there is a synchronisation that takes place between the participants while shaking. This is observed by the studies mentioned above in Section 2.2 as well and is one of the aims of most works that study the shaking aspect. They also look at reducing the interaction forces between the robot end-effector and the human hand, which is modelled using impedance/admittance control by some works.

We divide the works into 3 main categories: Central Pattern Generator (CPG) and Related Models, Harmonic Oscillator Systems and Miscellaneous Shaking Systems, as shown below in Table 2.4.

Central Pattern Generators and Related Models	Harmonic Oscillator Systems	Miscellaneous Shaking Systems
Jouaiti et al.[88] Kasuga and Hashimoto[91] Melnik et al.[9] Melnik and Henaff[123] Papageorgio and Doulgeri [150] Sato et al. [171]	Beaudoin et al.[20] Chua et al.[43] Dai et al. [47] Mura et al.[129] Wang et al.[205, 204] Yamato et al. [207] Zeng et al.[209]	Avraham et al. [15] Karniel et al.[90] Nisky et al.[140] Pedemonte et al.[154]

Table 2.4.: Division of works studying the shaking motion.

2.5.1. Central Pattern Generators (CPGs) and Related Models

Central Pattern Generators (CPGs) [76] are biologically inspired neuronal circuits that generate rhythmic output signals. One of the first works to develop an algorithm for the shaking phase proposed the idea of using a CPG-like neural oscillator to model the motion of the shoulder and elbow joints of a robot. They use the torque exerted on the joints as input and generate an oscillatory trajectory, that can be tuned by adjusting gains to amplify the input signal to go from active (high gain) to passive (low gain) [91].

One drawback of this method, as pointed out by Sato et al. [171] is that there are quite a few hand-tuned parameters. Therefore, they propose a polynomial approximation for the attractor model of the CPG and subsequently, a model for updating these parameters in an online fashion. A similar on-the-fly parameter update of the oscillator is done by Papageorgiou and Doulgeri [150], who use an impedance model to help tune the parameters of an internal motion generator modelled as a Hopf Oscillator [77]. Although this is not a CPG model, it shows similar synchronisation properties to produce rhythmic outputs like a CPG. The output of the impedance model and oscillators are used to update the oscillator parameters using Direct Least Squares in each iteration using n previous samples of the trajectory. This is unlike previous approaches where the adaptability of the CPG was an inherent trait.

In contrast, some works directly learn the CPG frequencies to enable a more online real-time approach. Melnyk et al. [9] build the CPG around an online learning mechanism that helps it sync with the human's motions directly. Their method dynamically adjusts to changes in the human's shaking frequency and synchronise with the human. Like the above-mentioned works, they too use the joint forces as an input to generate the motions. However, they only work on controlling a single degree of freedom. A similar model is proposed again by Melnyk and Henaff [123] using two different modes, joint positions and accelerations respectively as the inputs. Like their previous work, they control only one degree of freedom. Along similar lines, Jouaiti et al. [88] use a similar CPG model and incorporate dynamic plasticity [167] in it, making it easier to synchronise with the handshaking frequency. Moreover, they also propose learning the amplitude of the oscillations along with the frequency, thereby being more adaptive than previous approaches.

Major Findings. Overall, CPGs and oscillatory mechanisms synchronise well with the human's motion especially those that are dynamically learnt. They can also fare better than a conventional impedance control approach in terms of flexibility, naturalness, affinity and kindness of the perceived handshake [91]. Additionally, Jouaiti et al. [88] observed that incorporating plasticity in the CPG can help decrease the energy spent by the robot as well. One major drawback is the ability of such oscillatory mechanisms to converge

quickly to the required frequency, taking more than a few seconds even in the fastest cases. This would lead to unnatural handshaking behaviours that take too long to synchronise. Further research is required to increase the convergence speed of such mechanisms.

2.5.2. Harmonic Oscillator Systems

Harmonic oscillator models are those that employ harmonic systems, like spring-damper systems [48, 129, 207] or simpler sinusoidal motions [20, 204, 205, 209] to model the motion during shaking. Some works use both types of harmonic oscillator models in a two-step predictive and reactive system [43, 209]. Most works that employ harmonic oscillator models use them as reference motions for an impedance controller used to control the joint motions.

Beaudoin et al. [20] incorporate an impedance controller with different stiffness values using a sinusoidal reference trajectory with different frequencies and amplitudes along with the grasping model proposed by Arns et al. [8]. Dai et al. [48] develop a controller for a custom-made hand that controls the stiffness, viscosity and joint angles independently. Mura et al. [129] explore different shaking strategies w.r.t. robotic arm stiffness and their synchronisation with the human during handshaking. The parameters of the oscillations are estimated quickly in an online fashion using an Extended Kalman Filter from a fixed number of preceding frames. They compare three models of varying stiffness, namely high, low and variable based on the pressure exerted by the human, similar to Vigni et al. [197].

Wang et al. [205] propose an impedance control mechanism to model the handshaking mechanism and show how this can be learnt from human handshakes using least-squares minimisation. The reference trajectory for the model is generated using an amplitude of 10cm and mixed frequency components from 0 - 25Hz. Based on the human's response, the model parameters are fine-tuned when the human is being passive. When the human is being active, the model is used to estimate the interaction forces between the human and the robot and carry out the handshake while being passive. This low-level controller is expanded on in [204] where a high-level controller is used to generate reference trajectories for it. They first propose a new method using recursive least squares for a fast online estimation of the impedance parameters which are fed into an HMM that predicts the intention of the human i.e. active or passive from haptic data. The impedance parameters are used, rather than raw force inputs since they convey the state of the system.

Major Findings. The use of active impedance control was found to be more compliant to human motions as compared to simple position-based control [205]. Such active behaviours were rated better in terms of responsiveness than passive ones which also

had a significant effect on the perceived synchronisation [129]. Additionally, it was observed that the human partners would adapt their handshake to the robot's behaviour, even when a change was not explicitly mentioned [48]. This further shows the inherent synchronisation that takes place during handshaking, and that we as humans infer it from the interaction itself. There are still no studies that compare the perception of CPG-based shaking motions with harmonic oscillator motions. For a fair comparison, user studies with the same interface would be needed to analyse the perceptual differences between these methods.

2.5.3. Miscellaneous Shaking Systems

Karniel et al. [90] describe an experimental framework for a Turing test of motor intelligence for shaking behaviours. They do so on a 1D force-controlled haptic stylus, that is presented to a participant. In their test, the forces driving the participant's stylus is a linear combination of forces exerted by an experimenter and different proposed models. They develop a Model Human-Likeness Grade (MHLG) which measure how human-like the motions are from the participants' feedback. Nisky et al. [140] extend this to three different versions of the test. The first is a computer vs human test, where the participant is presented wither with a purely algorithmic handshake or a purely human handshake. However, this was not sensitive enough as the participants could almost always guess correctly when a human was shaking their hand. The second version is where the participants have to compare an algorithmic handshake with a noisy human one. The third is the weighted linear combination test proposed by Karniel et al. [90]. Unlike the "pure" test, the authors claim that the latter two variants are said to be better suited for this purpose.

Avraham et al. [15] make use of this "noise"-based Turing Test to compare 3 different shaking behaviours. The first is a tit-for-tat model that initially records the human's motion passively and then keeps replaying the same motion, assuming that the human's motion stays the same again. The second is a biologically inspired model that simulates a movement that could be generated by extensor and flexor muscles to ensure a low amount of overall interaction force. The final is a simple machine learning model that uses linear regression to learn the parameters of a linear combination of state variables with corresponding Gaussian kernels. It was found that the tit-for-tat model and the machine learning model fare similar to each other. They both fare much better than the biologically inspired model, which the authors argue can be improved by tuning the hyperparameters. While the proposed Handshaking Turing tests work for shaking a simple 1D stylus, it still needs to be seen how well these tests would fare on more complex robotic hardware.

Pedemonte et al. [154] introduce a mechanism for remote handshaking using the hand developed in [155]. They develop a vertical rail mechanism that the hand is mounted on to

support a vertical shaking motion that is passively controlled. The same mechanism is used by both the participants. This shaking motion along with the forces exerted on the hand is relayed to the opponent's hand and rail mechanism to allow a bilateral handshake to take place remotely. They show that their mechanism allows for realistic haptic interaction to take place remotely where the participants can adequately perceive each other's motions and forces.

2.6. Human Responses to Social Aspects of Robotic Handshaking

In previous sections, we have already talked about how some of the different works were perceived in HRI experiments. In this section, we expand further along similar lines and discuss works whose main aim was analysing the responses in such HRI experiments.

2.6.1. External Factors in Handshaking

While the sense of touch can convey emotional information, there are additional factors that enhance the perception of these feelings and the acceptance of the interaction. Below, we discuss some works that explore different external factors and present their findings to show the importance of fine-tuning these external factors, which although are subtle, have an impact on the way a handshake is perceived.

Amami et al. [5] and Tsalamal et al. [194] performed studies to explore how touch influences the perception of facial emotions. They used two haptic behaviours (strong and soft) combined with three visual behaviours namely happy (smiling), neutral and sad (frowning) which were displayed by the robot's lips. They test the interactions in three conditions, haptic-only, visual-only and visuo-haptic. The combination of visual expressions with a strong handshake showed higher arousal and dominance over all visual expressions, showing that a sense of touch can enhance robotic expressions. The majority of the comparisons between visuo-haptic and haptic-only cases were insignificant, which the authors argue could be due to the simplistic nature of the facial expression rendering. Vanello et al. [196] explore similar correlations between participant's perceptions while shaking hands with an artificial hand made of a plastic material while being presented with a visual stimulus of either a human or a robot face. While their experimental design to use fMRI data to understand such correlations is a useful one, their results cannot be deemed as conclusive since only three participants take part in their study.

Nakanishi et al. [132] explore social telepresence with a video screen equipped with a robot hand below it. They try out different visibility settings of the participant's hand in the frame and a robotic hand and compare one-way and two-way teleoperated handshake

settings. They build a hand that resembles a human hand using a soft sponge and gel-like covering and an artificial skin layer to make the appearance more human-like. It is also equipped with resistive wires that heat the fingers. The closure of the fingers is controlled by an external motor with wires connected to the fingertip that are pulled to extend or close the fingers. First, they look at different hand settings of the presenter where their hand would either be visible or out of the video frame (invisible) while shaking. They find that in the invisible case, the interaction was perceived better. Participants not only strongly felt that the presenter was in the same room but also had a strong feeling that they were shaking hands with the presenter in real life. The authors argue that though the visibility of synchronisation might be perceived as better, the visibility of the presenter's hand led to this effect getting cancelled out, which some subjects reported was due to the duplication of the hand i.e. seeing two hands at the same time, both the presenter's and the robot hand. After establishing the results of one-way handshaking, they tested out how participants felt when the presented had a robot hand that was controlled by the participants (two-way handshake). This was tested in two settings where the presenter's interaction was either visible or invisible to the participants. It was seen that the same feelings of physical closeness to the presenter and of shaking hands in real life were rated higher for the invisible two-way case, where participants knew that their handshake was being felt by the presenter. This could possibly be attributed to a perceived synchronisation of sorts which arises from the participant knowing that their actions are being perceived by their partner rather than the interaction being just in a one-way direction.

Jindai et al. [85] analyse various handshake motions generated by their model in two ways. They first fit a Bradley Terry Model [28] on paired comparisons of their interactions. Following that, they use a 7 point bipolar scale to test the participant's preferences w.r.t. the handshaking motion, velocity, relief, easiness, politeness and security. They additionally see how participants respond to voice [84] and gaze behaviours [83]. The study showed that a delay of 0.1 seconds between the voice and handshake motion of the robot was found acceptable. It was found that the most preferred behaviour was when the gaze shifts steadily from the hand while reaching out to the face after contact is established. The response models [146, 147] were tested based on the delay between the request and response motions and it was seen that starting the response a fraction of a second (0.2s to 0.4s) after the response was preferred. However, a larger delay of 0.6s was less preferable. The request model in [79] was additionally tried out with a human approaching the robot from a distance. They experimented with starting the request at different distances of the human from the robot. Apart from this, both the request and response models were combined with a similar transfer function as in [85] such that the robot requests a motion if the human doesn't. This type of behaviour was well perceived by humans as well (positive feedback on a 7-point bipolar scale).

2.6.2. Influence on the Perceived Image of the Robot

As mentioned in the introduction (Section 2.1), handshaking can impact first impressions. Therefore in the context of HRI, this can possibly help strengthen the perception of a robot for further interactions that take place. This is explored by the works described below, wherein the effect of robotic handshaking is studied on the specific tasks that a robot has to accomplish.

Avelino et al. [12] use their previously proposed handshaking model [13] to see how a handshake affects a subsequent interaction wherein the robot needs to perform a navigation task during which, it would need some assistance by the human. They found that participants who shook hands with the robot found it to be warmer and more likeable and were more willing to help the robot with its task. However, they argue that if the robot had an extremely human-like handshake, participants would not anticipate it to get stuck in a simple navigational task due to a mismatch between the behaviour during the experiment and the perceived behaviour according to the handshake.

Bevan and Fraser [23] perform an experiment to see the effect of handshaking on negotiations between participants, where one participant interacts with the other via telepresence on a Nao robot. It was seen that handshaking improved mutual cooperation, leading to a more favourable negotiation result for both parties. Haptic feedback for the telepresent negotiator didn't have a significant impact. They also found that handshaking did not affect the degree to which negotiators considered their opponent trustworthy, which they argue is possibly due to the childlike nature of the Nao robot.

2.6.3. Distinguishing Ability of Handshakes

In Section 2.2, it was shown that there were observable differences of gender and context on handshaking. In this section, we discuss some works that aim to model certain aspects of the participant, like their gender, personality and mood based on handshakes.

Orefice et al. [144] propose a model for making distinctions along the lines of gender and personality (introversion/extroversion), using a set of 20 parameters relating to acceleration, velocity, duration, pressure etc. They find that in male-male pairs, more pressure is applied than in male-female ones. Moreover, they found that female pairs have a longer duration and a lower frequency but the maximum speed of the oscillations is higher. They argue that some results could also be due to the hand sizes rather than gender since most of the females in their study had smaller hands. Coming to personality, they found that introverts reached a higher speed while shaking hands and extroverts would exert more pressure. Though they performed similar experiments in a human-robot handshaking scenario as well, the small number of participants ($n = 8$) makes their

results in this aspect inconclusive. Garg et al. [59] similarly aimed to classify people's personalities into weak and dominant. They use similar information extracted using a custom-made glove to measure accelerations, Euler angles and polar orientations. The features are ranked based on the Mutual Information followed by classification using K Nearest Neighbours, achieving a 75% accuracy. Orefice et al. [145] perform another longitudinal study with 11 participants over 16 non-consecutive days, that looks at how pressure variations while shaking hands reflect the mood of the participants. They use a custom-made glove with various pressure sensors and an accelerometer worn by both the participant and a Pepper robot. Before shaking hands, participants had to declare which mood (Calm, Relaxed, Cheerful, Excited, Tense, Irritated, Sad, and Bored) best described their current mood. Consistency in the mood was seen when participants shook hands with a human subject and with pepper, which was unexpected as one would expect an interaction with a robot to seem unrealistic or not as human-like. Overall, no significant results were found for most positive moods, except between "Calm" and "Cheerful" where the former had less pressure observed. In the case of negative moods, "Bored" handshakes had lower pressure than "Excited" and "Tense", which have more arousal than "Bored". In general, lower pressures were found with moods with lower arousal. This shows how handshaking can be used as an affective interaction to further increase the emotional understanding of robots.

2.6.4. Human-likeness of Robotic Handshakes

In the introduction (Section 2.1), we mentioned the importance of having human-like body movements, which plays an important role in HRI acceptance [35, 104, 127]. Therefore we analyse works that look at the human-likeness of robotic handshakes and draw insights that can help shape future experiments.

The social responses to the shaking models proposed by Wang et al. [205, 204] were analysed further in [203, 60]. Both studies perform their experiment on a robot with a rod as its end effector, in a bar like setting wherein participants have noise-cancelling headphones playing bar music and having ambient conversations. In both studies, participants had to perform around 6-7 handshakes in each of the three different handshake settings. First was the basic algorithm proposed in [205], the second was the interactive model proposed in [204] and third was a human operating the robot. After the handshakes, participants had to rate the human-likeness of the handshake from 1 (resembling a robotic handshake) and 10 (resembling a human handshake). In neither of the studies did the participants see the robot. In the study by Giannopoulos et al. [60], participants were first blindfolded and led to the robot whereas, in the study by Wang et al. [203], participants had a VR headset on which had a graphical rendering of a bar with a human model

rendered for the robot, who would walk up and request for a handshake with the virtual hand in the same position as the robot end effector in the real world. In both the studies, the human-operated handshake was rated the highest (6.8/10 in both), followed by the HMM-based handshake (5.9/10 in [60] and 5.3/10 in [205]). The least was the basic handshake proposed in [205] which was rated much lower than the other 2 alternatives (3.3/10 in [60] and 3.0/10 in [205]). Although the HMM-based handshake was rated closely as the human-operated one, they both were far from the maximum human-likeness score (10/10), which could be due to the rod-like end effector used. Having a similar experiment with a more sophisticated robot hand could yield the results to be more human-like. However, their studies conclude that having an adaptive handshake that matches the behaviour of the human ends up being perceived closer to human behaviour.

Stock-Homburg et al. [186] study whether a realistic android robot, that is modelled after a human, with soft silicone skin and pneumatically controlled joints can pass a hardware version of the Turing test. They have 15 participants blindfolded who have to interact with a human and the robot with their hand stretched out twice in a random order, leading to a total of 4 trials each. Although the robot is built to be as realistic as possible, the majority of the humans (11/15) correctly guessed the hand in the first attempt itself and by the last handshake, they were all able to guess the hand correctly. However, they only test a static interaction with the robot. For a better evaluation, comparing a handshake behaviour rather than just a static interaction could show better insights into the modelling of a human-like handshake.

2.7. Discussion

Overall, we have talked about the various works that look into human-robot handshaking, however, due to the differences in hardware and metrics used across different studies, it is difficult to come up with a common benchmark to evaluate these studies. Having said that, some qualitative conclusions can be drawn from analysing these studies. In general, over the different stages of handshaking, an element of synchronisation is present. In the reaching phase, this is seen in terms of the similarity of the requestor's and responder's motions, which is why most works modelling the reaching behaviour draw from human-human interactions. Although such behaviours can be learnt using Reinforcement Learning, human trajectories provide a strong prior to enable the learnt motions to be human-like. Following this, synchronisation in the grasping phase can be observed by matching the strength of the partner and can additionally affect the perceived personality of the robot, which highlights the affective nature of the interaction. Finally, the shaking phase is where the main element of synchronisation can be explicitly measured with the interaction forces

between the hands. This depends directly on how well the shaking motion adapts to that of the partner, leading to low levels of interaction forces as the synchronisation gets better. Though synchronisation is a major element of handshaking, in reality, it is difficult to be completely in sync, due to differences in hand shape and size, mental states etc. Therefore, a leader-follower situation can arise in the different stages as well, which could reflect on various personal attributes of the interaction partners.

Although there is a considerable amount of work on the topic, there are still some gaps in the current state of Human-Robot Handshaking research. Based on what we have already discussed above and their pitfalls, we propose the following suggestions/open areas for further research on Human-Robot handshaking.

Suggestion 1: One important, yet relatively difficult task is that of combining the different phases. For a more human-like perception, a proper transitioning would be required between each of the different phases. Only a few studies [41, 81, 129, 154, 207] look into combining pairs of different phases but still do not implement an end-to-end behaviour. There is still work that needs to be done to achieve a complete handshaking behaviour. Along with this, the termination of a handshake is an equally important criterion to make the interaction more socially acceptable. Current works neither take a smooth separation into account nor do they analyse the effects of it. A prolonged handshake or an untimely termination can possibly be perceived as unnatural and can affect the subsequent interaction [131]. Therefore an end-to-end handshaking behaviour should take this into account as well.

Suggestion 2: From the perspective of a robotic handshake, making use of contextual cues would be effective in having a successful impact on the handshake. As described in Section 2.2 and 2.6.3, different social contexts and moods have an effect on the handshake. Being able to detect such cues additionally requires further research in other fields like emotion recognition, intent recognition, etc. Some works try and estimate the mood/personality via the handshake [145, 204], but there is no explicit context detection in place, say for example by detecting it from facial expressions, or possibly from physiological data, and correspondingly using the insights from Section 2.2 for fine-tuning the handshake.

Suggestion 3: Developing better social robotic interfaces that have a force-sensing mechanism and performing closed-loop control can be more expressive, as seen in [5, 194, 197]. Currently, most works do not implement proper grasping control (as shown in Table 2.3), which is key for capturing the expressive ability of handshakes to its maximum. Additionally, the human-likeness of an interface is just as important for its perception. This can be seen in [60, 205] where even though a human was controlling the robot having a rod-like end effector, this mode of shaking only got a human-likeness score of 6.8/10 by the participants. Even sophisticated mechanisms, like the Android robot used

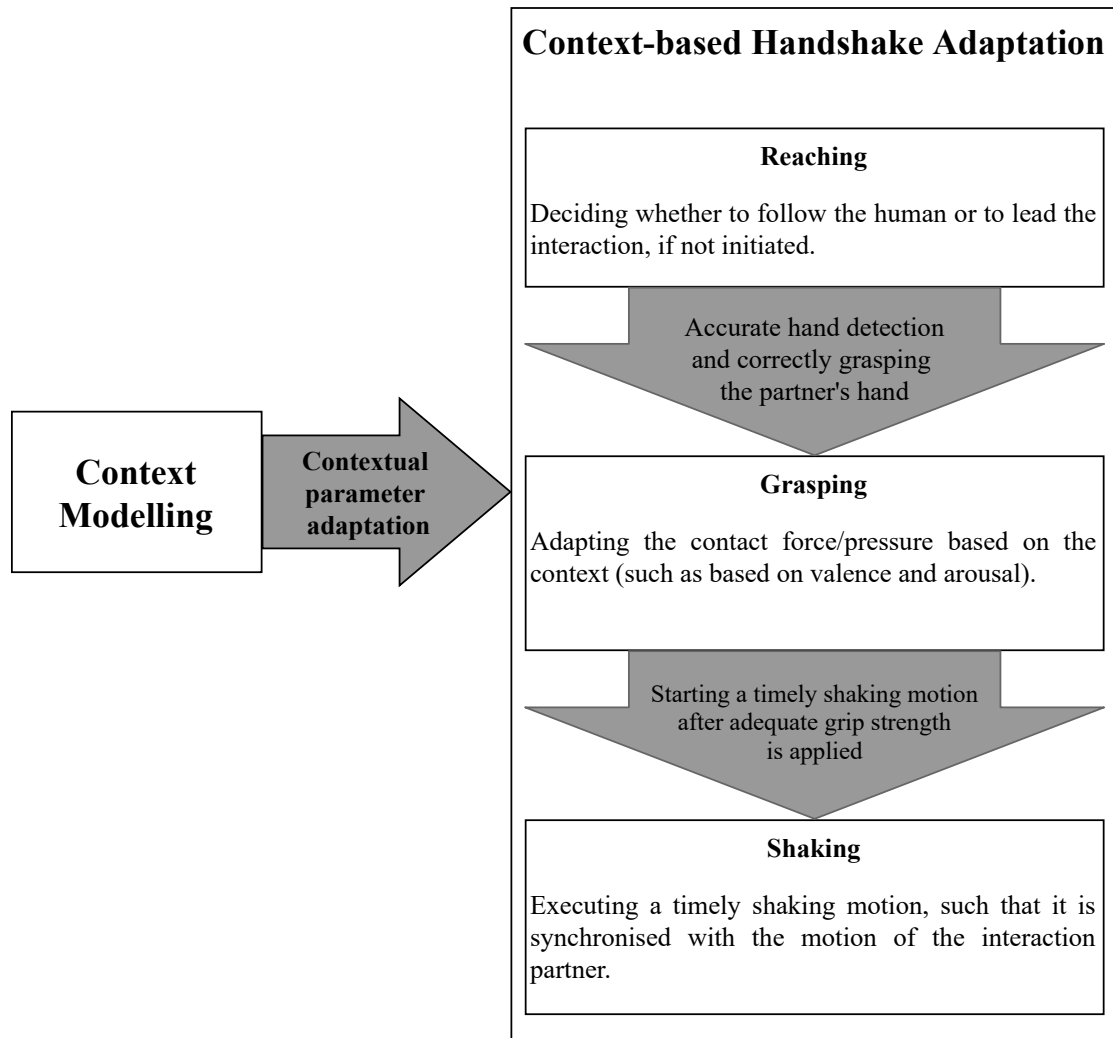


Figure 2.2.: Modified Conceptual Framework for Human-Robot Handshaking. (The proposed suggestions are shaded in grey.)

in [186] which has a soft skin-like layer and heated palms, are still easily distinguished from a human's hand. One workaround could be to use sensing gloves, like by Orefice et al. [145], which could help bridge the gap between a sophisticated interface and social robots.

Suggestion 4: Given that one of the main use-cases of handshakes in shaping first impressions is in business cases, the effect of robotic handshaking in such cases hasn't been properly explored. This is especially important given the use of social robots as front-line employees [78]. Bevan and Fraser [23] study a part in a negotiation context. They look at the impact of robotic handshaking on the impression of the negotiation partner who teleoperated the robot. Moreover, they conduct their study with a Nao robot, which can come across as very childlike and not be taken as seriously in such settings.

While suggestions 2 and 4 are still subjective, suggestions 1 and 3 are areas that can be objectively improved and incorporated to improve existing methods not just for Human-Robot handshaking but would be applicable towards other similar physically interactive behaviours as well. With these two suggestions, a modified framework of Human-Robot handshaking is shown in Figure 2.2 (the suggested aspects are shaded in grey). Here the main importance is given to contextual modelling, which influences parameters like strength, speed etc. by adapting them accordingly.

2.8. Conclusion

Handshaking is a versatile non-verbal interaction that plays an important role in social settings. In this chapter, we first draw insights from human-human handshaking regarding timing, trends in the grip strength and the synchronisation of shaking. We then explore the different phases of handshaking, namely reaching, grasping and shaking, while observing a common aspect of synchronisation between the phases. We finally discuss how handshaking affects the way the robot is perceived and propose some directions for future work. However, one thing to keep in mind is that handshaking is just one in so many physical interactions all of which vary over different cultures, age groups, geographic locations, contextual settings etc. To this end, learning different physically interactive behaviours, such as hand-claps/high-fives, fist bumps, or a combination of different touch-based interactions, would help improve the perception of the robot. Being able to distinguish and learn such new physically interactive behaviours on the go, building a skill library of sorts, rather than just a single one like handshaking, could improve the sociability of a robot.

3. MILD: Multimodal Interactive Latent Dynamics

Among the suggestions proposed in Chapter 2, the first suggestion for extracting, combining, and transitioning between the different phases of an interaction is key for enabling a synchronized behavior for not just Handshaking but other physical interactive behaviors in general. To enable seamless physical Human-Robot Interaction (HRI), a robot needs to react in a timely manner to a human’s actions and should accurately maintain physical proximity with the human in a coordinated manner. This chapter presents a method for learning interaction dynamics to enable well-coordinated HRI with a Humanoid Social Robot directly from Human-Human Interactions (HHI). A hybrid model is devised a hybrid model that uses Hidden Markov Models (HMMs) as latent space priors for a Variational Autoencoder (VAE) to learn a joint distribution over the dynamics of the interacting agents. We leverage the interaction dynamics learned from HHI demonstrations to regularize the learning of robot trajectories and to improve the predictions by incorporating the conditional generation of robot motions from human observations into the training. To improve the accuracy of the interaction, the generated robot motions are further adapted with Inverse Kinematics for achieving the desired physical proximity with the human partner. This helps mitigate issues arising from the mismatch in motion retargeting to the robot by combining the best of joint space learning and task space reachability. For a contact-rich interaction, we further use the HMM predictions to modulate the robot’s stiffness to enable a smooth, compliant, and realistic interaction. Our experimental evaluations show that our method captures the interaction dynamics well, with our human-conditioned reactive motion generation significantly improving the performance. We verify the effectiveness of our approach deployed on a Humanoid robot via a user study. Our method generalizes well to various humans despite being trained on data from just two humans. We find that users perceive our method as more human-like, timely, and accurate and rank our method with a higher degree of preference over other baselines. We additionally show the ability of our approach to generate successful interactions in a more complex scenario of Bimanual Robot-to-Human Handovers.

3.1. Prologue

Ensuring a synchronized and accurate reaction is an important aspect in Human-Robot Interaction (HRI) [4]. To do so, a human and a robot need to spatiotemporally coordinate their movements to reach a common target or perform a common task, that can be denoted as a *joint action* [175]. For realizing such coordinated jointly performed actions, spatial and temporal adaptation of motions in the shared physical space to accurately perform the action are key factors [176]. Such interpersonal coordination can enable a connection to one's interaction partner [119]. Therefore, well-executed and well-coordinated interactive behaviors can improve the perception of a social robot.

The paradigm of Learning from Demonstrations (LfD) is promising for HRI by learning joint distributions over human and robot trajectories in a modular, multimodal manner [52, 31, 156, 53, 99]. However, LfD approaches scale poorly with higher dimensions, which can be circumvented by incorporating Deep Learning for learning latent-space dynamics. Such Deep State-Space Models have shown good performance in capturing temporal dependencies of latent space trajectories, either using some kind of a forward propagation model [21, 46, 44, 29], learning the parameters for a dynamics model with LfD [89, 39, 37] or fitting parameterized LfD models to the underlying latent trajectories [50, 130]. In this chapter, we further explore this direction of latent space dynamics in the context of Human-Robot Interaction.

Furthermore, when learning trajectories in the joint configuration space of a robot, minor deviations in the joint space can cause a perceivable deviation in the robot's task space. While learning task space trajectories can mitigate problems from errors in the joint space, to do so, the demonstrated trajectories need to fall within the reachability of the robot. When this reachability assumption is violated, the task space proximity in HRI scenarios may not be effectively learned. Moreover, learning purely task space trajectories doesn't necessarily ensure human-like joint configurations, which is especially relevant for Humanoid robots. To this end, the combination of joint space learning and task space reachability can be achieved using Inverse Kinematics to supplement the learned behaviors [63, 162, 34].

In this chapter, as shown in Figure 3.1, we are interested in learning and adapting Deep LfD policies for spatiotemporally coherent interactive behaviors. We learn the interaction dynamics from HHI demonstrations using latent space Hidden Markov Models (HMMs). We show the efficacy of the learnt dynamics in real-world HRI scenarios. We do so in a manner that ensures spatially accurate physical proximity to the human partner and additionally enable compliant robot motions in contact-rich interactions like a handshake, thereby improving the perceived quality of the HRI behaviors.

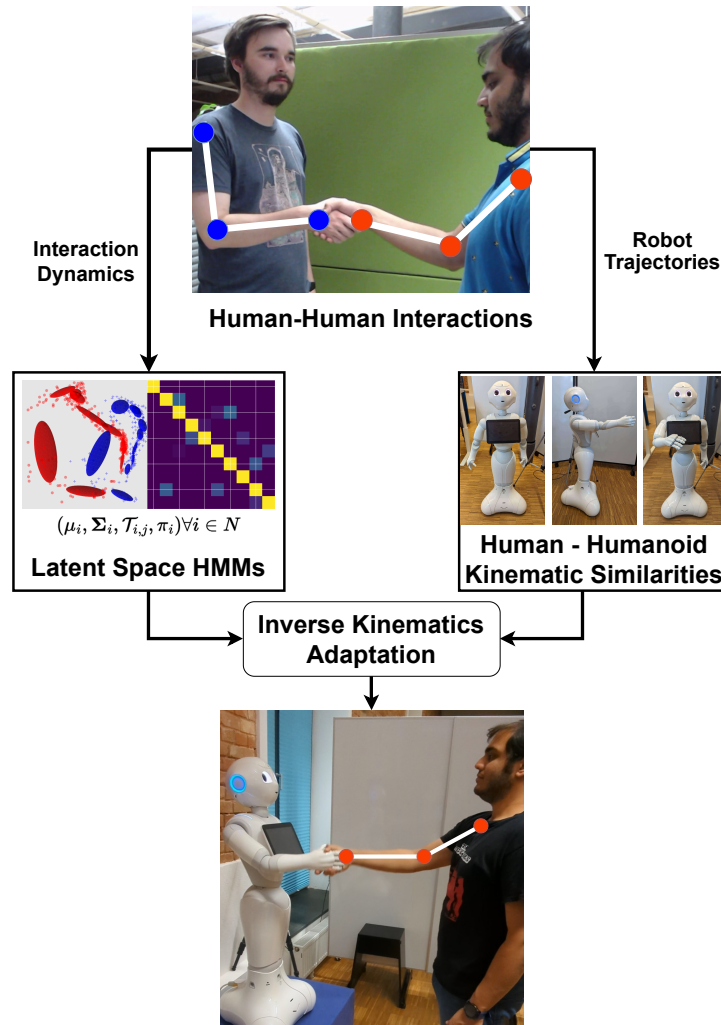


Figure 3.1.: In this chapter, we explore learning coordinated HRI behaviors using Hidden Markov Models (HMMs) to learn the interaction dynamics over a representation space spanned by a Variational Autoencoder (VAE). The VAEs are trained with the HMMs as a prior to better incorporate the understanding of the dynamics. Moreover, we learn directly from Human-Human interactions by leveraging the kinematic similarities of Humans and Humanoid robots. During testing, the predicted trajectories are adapted with Inverse Kinematics to mitigate errors in the spatial accuracy arising either from prediction errors or from shortcomings in transferring the human's actions to the robot.

3.1.1. Related Work

Learning HRI from Demonstrations

Early approaches for learning modular HRI policies modeled the interaction as a joint distribution with a Gaussian Mixture Model (GMM) learned over demonstrated trajectories of a human and a robot in a collaborative task [31]. The correlations between the human and the robot degrees of freedom (DoFs) can then be leveraged to generate the robot's trajectory given observations of the human for learning both proactive and reactive controllers [169, 30, 156]. Segmenting HRI demonstrations has also been shown using Graphical Models with Markov chain Monte Carlo [182, 181].

Along the lines of leveraging Gaussian approximations for LfD, Movement Primitives [152], which learn a distribution over underlying linear regression weight vectors, were extended for HRI by similarly learning a joint distribution over the weights of interacting agents [6, 116, 32]. Movement Primitive approaches can further be combined with GMMs for learning multiple task sequences seamlessly [53, 117, 99, 143, 112]. One drawback of the aforementioned primitive-based approaches is that they do not perform well on out-of-distribution data [213]. Additionally, for tasks that are loosely coupled in time, Movement Primitives can underperform [151]. These drawbacks therefore make them an unviable option for being able to generalize spatiotemporally in interactive behaviors. For a more extensive overview of Movement Primitive approaches for robot learning, we refer the reader to [191].

Vogt et al. [200] are similar in theory to our approach where they learn the interaction dynamics of Human-Human Interactions using an HMM over a low dimensional representation space of the human skeletons. In our approach, we show further performance improvements by incorporating human-conditioned reactive motion generation into the training pipeline. Additionally, Vogt et al. [200] train Interaction Meshes [74] for transferring the learned trajectories to a robot whose prediction is optimized during runtime to adapt to the human user's movements. As we work with Humanoid Robots, we follow a simpler approach by leveraging the structural similarity between a human and a humanoid. This allows us to map the joint motions directly [57] and adapt the motions to the interaction partner using Inverse Kinematics [162]. Through our user study, we find that our approach, while being simplistic, provides acceptable interactions with various human partners.

To make the learned distribution more robust, LfD methods are often learned via kinesthetic teaching which can be tedious for HRI tasks where one would need a variety of human partners. The need for extensive training data can potentially be circumvented by learning how humans adapt to a robot's trajectory [33]. A more general way, especially

in the context of Humanoid robots, is by learning from human-human demonstrations by leveraging the kinematic similarities between humans and robots [57]. While such approaches of motion retargeting leave some room for error due to minor differences between human and robot geometries, adapting the trajectories in the robot’s task space via Inverse Kinematics can improve the accuracy of such interactive behaviors [162, 198].

Integrating LfD with Deep Learning

Techniques at the intersection of LfD coupled with the use of Neural Network representations for higher dimensional data have grown in popularity for learning latent trajectory dynamics from demonstrations. Typically, an autoencoding approach, like VAEs, is used to encode latent trajectories over which a latent dynamics model is trained. In their simplest form, the latent dynamics can be modeled either with linear Gaussian models [89] or Kalman filters [21]. Other approaches learn stable dynamical systems, like Dynamic Movement Primitives [172] over VAE latent spaces [25, 38, 39, 37, 45]. Instead of learning a feedforward dynamics model, Dermy et al. [50] model the entire trajectory’s dynamics at once using Probabilistic Movement Primitives [152] achieving better results than [37].

Nagano et al. [130] demonstrated the use of Hidden Semi-Markov Models (HSMMs) as latent priors in a VAE for temporal action segmentation of motions involving a single human. They model each latent dimension independently, which is not favorable when learning interaction dynamics. To extend such an approach for learning HRI, one needs to consider the interdependence between dimensions. We extend this idea of using HSMMs as VAE priors for interactive tasks by exploiting the full rank of HSMM covariance matrices thereby capturing the dependencies between both agents by learning a joint distribution over the latent trajectories of interacting partners.

Interaction Modeling with Recurrent Neural Networks

When large datasets are available, Recurrent latent space models are powerful tools in approximating latent dynamics with some form of a forward propagating distribution [44, 67, 100, 113, 54]. Given their power of modeling temporal sequences, they yield themselves naturally for learning interactive/collaborative tasks in HRI [142, 212, 177]. To simplify the various scenarios of interaction tasks Oguz et al. [142] develop an ontology to categorize interaction scenarios and train an LSTM network for each case in a simple Imitation Learning paradigm. Similarly, Zhao et al. [212] explore using such LSTM-based policies for learning HRI in a simple collaboration scenario. Rather than simply regressing robot actions to human inputs, Bütepage et al. [29] explore the idea of learning shared latent dynamics in HRI in a systematic way. Bütepage et al. [29] initially learn latent em-

beddings of the human and robot motions individually using a VAE. Once the embeddings are learned, the latent interaction dynamics model of Human-Human Interactions (HHI) is learned using LSTMs. The learned HHI dynamics are used in conjunction with the learned robot embeddings to subsequently learn the robot dynamics from HRI demonstrations. However, we find in our experiments that the autoregressive nature of their approach, unfortunately, leads to a divergence in the performance since they only train it using ground truth demonstrations and not in an autoregressive manner. In contrast, rather than having an implicit shared representation, such as with an LSTM, the LfD approaches can be explicitly conditioned on the interaction partner’s observations [6, 52] which we find leads to improved predictive performance.

Unsupervised Skill Discovery

While skill discovery is not the main focus of our approach, there are some parallels between our approach and works on unsupervised skill discovery which we discuss below. In general, such approaches aim to partition a latent space into different skills which can then be sequenced for learning a variety of different tasks either from demonstrations or using Reinforcement Learning (an overview can be found in Sec. III-B in [128]). The works closest to our approach [190, 133, 165] explore this key idea of learning how to segment demonstrations into underlying recurring segments or “skills” by decomposing a latent representation of the trajectories and subsequently learning the temporal relations between these skills to compose the observed trajectories. Similar approaches are further explored in literature using different skill representations such as the Options framework [180] or Autoregressive HMMs [139, 102]. While the main focus of such approaches is the decomposition of demonstrations to learn re-usable atomic “skills”, our main aim is to explore how to do so in HRI settings via a joint distribution over the actions of interacting partners and subsequently how the learned skills/segments can be inferred during test time based on the dynamic observations of a human, rather than reproducing a given task or based on a static goal.

3.1.2. Objectives and Contributions

In this chapter, we first explore how Deep LfD can be used for learning spatiotemporally coherent interaction dynamics. Rather than using an uninformed, stationary prior for learning HRI [29], we show the use of Hidden Markov Models (HMMs) to model the latent space prior of a VAE as a joint distribution over the trajectories of both interacting agents. During test time, the robot trajectories can be conditionally generated based on the observed human trajectories in the latent space, showing that such an approach can capture

the latent interaction dynamics suitably well. We further explore how to systematically improve the predictive abilities of our model by incorporating reactive motion generation into the training process. We do so by exploring two different sampling approaches to incorporate the conditional distribution of the HMM into the training, rather than using the HMM just as a VAE prior. Furthermore, to deploy the proposed approach on the Humanoid Social Robot Pepper [149], we leverage the similarity between the kinematic structures of Pepper and a Human to effectively learn HRI behaviors from Human-Human Interactions. We further show how the predictions from the latent space HMMs can help supplement the robot controller by modulating the target motions via Inverse Kinematics for a more accurate interaction and modulating the joint stiffnesses for ensuring compliant motions during a contact-rich interaction like a handshake (Figure 3.2).

Specifically, the contributions of this chapter are as follows:

- (i) We demonstrate the use of Hidden Markov Models as informative latent space priors for a VAE thereby learning the interaction dynamics from demonstrations.
- (ii) We systematically explore how to incorporate human-conditioned reactive motion generation of the robot trajectories in the training process to improve the predictive abilities of our model.
- (iii) We integrate Inverse Kinematics in our pipeline to spatially adapt the predicted robot motions to the interaction partner, thereby ensuring suitable physical proximity and improved spatial accuracy of the learned behaviors.
- (iv) For contact-rich interactions like handshaking, we further perform stiffness modulation using the segment predictions from the HMM to enable realistic and compliant robot motions.
- (v) We validate our approach via a user study in a real-world HRI scenario where participants interact with a Pepper robot controlled by our approach. We find that our method ranks highly compared to other baselines and provides a perceivably more human-like, natural, timely, and accurate interaction, thereby showcasing its effectiveness. We further show to applicability of our approach on a more complex task of Bimanual Robot-to-Human Handovers.

The rest of the chapter is organized as follows. In Section 3.2, we explain the foundations needed to understand our work. We then explain our approach in Section 3.3. We highlight our experiments and results in Section 3.4 and present some concluding remarks and directions for future work in Section 3.5.

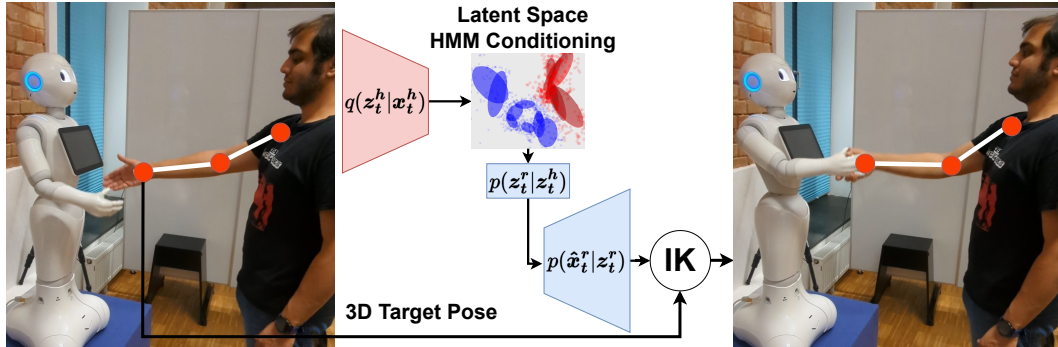


Figure 3.2.: Overview of our approach. We train VAEs to reconstruct the observations of the interactions agents $(\mathbf{x}_{1:t}^h, \mathbf{x}_{1:t}^r)$ with an HMM prior to learn a joint distribution over the latent space trajectories $p(z_{1:t}^h, z_{1:t}^r)$ of the interacting agents. During test time, the observed agent’s latent trajectory conditions the HMM to infer the robot agent’s latent trajectory $p(z_t^r | z_{1:t}^h)$ which is decoded to generate the robot agent’s joint trajectory $\hat{\mathbf{x}}_t^r$. To ensure the proximity of a robot’s hand to the human’s hand during test time, we additionally adapt the predicted trajectory using Inverse Kinematics to fulfill the contact-based nature of the interaction.

3.2. Foundations

We first explain Variational Autoencoders (VAEs) (Section 3.2.1) which is the main backbone of our approach, followed by Hidden Markov Models (HMMs) (Section 3.2.2), which are key for learning the interaction dynamics, and finally, we give an introduction to Inverse Kinematics (IK) (Sec 3.2.3) which is essential for improving the overall acceptance of our approach.

3.2.1. Variational Autoencoders

Variational Autoencoders (VAEs) [96, 166] are a type of neural network architecture that learns the identity function in an unsupervised, probabilistic way. The inputs “ \mathbf{x} ” are encoded into lower dimensional latent space embeddings “ \mathbf{z} ” that a decoder uses to reconstruct the original input. A prior distribution is enforced over the latent space during the learning, which is typically a standard normal distribution $p(\mathbf{z}) = \mathcal{N}(\mathbf{z}; \mathbf{0}, \mathbf{I})$. The goal is to estimate the true posterior $p(\mathbf{z}|\mathbf{x})$, using a neural network $q(\mathbf{z}|\mathbf{x})$ and is trained by minimizing the Kullback-Leibler (KL) divergence between them

$$KL(q(\mathbf{z}|\mathbf{x})||p(\mathbf{z}|\mathbf{x})) = \mathbb{E}_q[\log \frac{q(\mathbf{z}|\mathbf{x})}{p(\mathbf{x}, \mathbf{z})}] + \log p(\mathbf{x}) \quad (3.1)$$

which can be re-written as

$$\log p(\mathbf{x}) = KL(q(\mathbf{z}|\mathbf{x})||p(\mathbf{z}|\mathbf{x})) + \mathbb{E}_q[\log \frac{p(\mathbf{x}, \mathbf{z})}{q(\mathbf{z}|\mathbf{x})}]. \quad (3.2)$$

The KL divergence is always non-negative, therefore the second term in Eq. 3.2 acts as a lower bound. Maximizing it would effectively maximize the log-likelihood of the data distribution or evidence, and is hence called the Evidence Lower Bound (ELBO), which can be written as

$$\mathbb{E}_q[\log \frac{p(\mathbf{x}, \mathbf{z})}{q(\mathbf{z}|\mathbf{x})}] = \mathbb{E}_q[\log p(\mathbf{x}|\mathbf{z})] - \beta KL(q(\mathbf{z}|\mathbf{x})||p(\mathbf{z})). \quad (3.3)$$

The first term aims to reconstruct the input via samples decoded from the posterior. The second term is the KL divergence between the prior and the posterior, which regularizes the learning. To prevent over-regularization, the KL divergence term is also weighted down with a factor β . Further information can be found in [96, 166, 71].

3.2.2. Hidden Markov Models

A Hidden Markov Model (HMM) is used to model a sequence of observations $\mathbf{z}_{1:T}$ (robot joint angles, human skeletons, etc.) as a sequence of underlying hidden states such that they can “emit” the observations with a given probability. In mathematical terms, an HMM is characterized by a set of hidden states $i \in \{1, 2 \dots N\}$, each of which denotes a probability distribution, in our case a Gaussian with mean $\boldsymbol{\mu}_i$ and covariance $\boldsymbol{\Sigma}_i$, which characterize the emission probabilities of observations $\mathcal{N}(\mathbf{z}_t; \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$. An initial state distribution π_i denotes the initial probabilities of being in each state, and the state transition probabilities $\mathcal{T}_{i,j}$ describe the probability of the model going from the i^{th} state to the j^{th} state. The sequential progression via the probability of each hidden state given an observed sequence $\mathbf{z}_{1:t}$ is denoted by the forward variable of the HMM $\alpha_i(\mathbf{z}_t)$

$$\begin{aligned} \alpha_i(\mathbf{z}_t) &= \frac{\hat{\alpha}_i(\mathbf{z}_t)}{\sum_{j=1}^N \hat{\alpha}_j(\mathbf{z}_t)} \\ \hat{\alpha}_i(\mathbf{z}_t) &= \mathcal{N}(\mathbf{z}_t; \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \sum_{j=1}^N \alpha_j(\mathbf{z}_{t-1}) \mathcal{T}_{j,i} \\ \hat{\alpha}_i(\mathbf{z}_0) &= \pi_i \mathcal{N}(\mathbf{z}_0; \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \end{aligned} \quad (3.4)$$

where $\hat{\alpha}_i(\mathbf{z}_t)$ represents the non-normalized forward variable and π_i is the initial state distribution. The HMM is trained using Expectation-Maximization over the parameters

$(\pi_i, \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i, \mathcal{T}_{j,i})$ with the trajectory data. We refer the reader to [30, 156] for further details on HMMs in the context of robot learning.

In HRI scenarios, to encode the joint distribution between the human and the robot, we concatenate the Degrees of Freedom (DoFs) of both the human and the robot [31, 52] allowing the distribution to be decomposed as

$$\boldsymbol{\mu}_i = \begin{bmatrix} \boldsymbol{\mu}_i^h \\ \boldsymbol{\mu}_i^r \end{bmatrix}; \boldsymbol{\Sigma}_i = \begin{bmatrix} \boldsymbol{\Sigma}_i^{hh} & \boldsymbol{\Sigma}_i^{hr} \\ \boldsymbol{\Sigma}_i^{rh} & \boldsymbol{\Sigma}_i^{rr} \end{bmatrix} \quad (3.5)$$

where the superscript indicates the different agents (h -human, r -robot). Once the distributions are learned, given some observations of the human agent \mathbf{z}_t^h , the robot's trajectory can be conditionally generated using Gaussian Mixture Regression as

$$\mathbf{K}_i = \boldsymbol{\Sigma}_i^{rh} (\boldsymbol{\Sigma}_i^{hh})^{-1} \quad (3.6)$$

$$\hat{\boldsymbol{\mu}}_i^r = \boldsymbol{\mu}_i^r + \mathbf{K}_i (\mathbf{z}_t^h - \boldsymbol{\mu}_i^h) \quad (3.7)$$

$$\hat{\boldsymbol{\Sigma}}_i^r = \boldsymbol{\Sigma}_i^{rr} - \mathbf{K}_i \boldsymbol{\Sigma}_i^{hr} + \hat{\boldsymbol{\mu}}_i^r (\hat{\boldsymbol{\mu}}_i^r)^T \quad (3.8)$$

$$\hat{\boldsymbol{\mu}}_t^r = \sum_{i=1}^N \alpha_i (\mathbf{z}_t^h) \hat{\boldsymbol{\mu}}_i^r \quad (3.9)$$

$$\hat{\boldsymbol{\Sigma}}_t^r = \sum_{i=1}^N \alpha_i (\mathbf{z}_t^h) \hat{\boldsymbol{\Sigma}}_i^r - \hat{\boldsymbol{\mu}}_t^r (\hat{\boldsymbol{\mu}}_t^r)^T \quad (3.10)$$

$$p(\mathbf{z}_t^r | \mathbf{z}_t^h) = \mathcal{N}(\mathbf{z}_t^r | \hat{\boldsymbol{\mu}}_t^r, \hat{\boldsymbol{\Sigma}}_t^r) \quad (3.11)$$

where $\alpha_i(\mathbf{z}_t^h)$ is the forward variable calculated using the marginal distribution of the observed human agent.

3.2.3. Inverse Kinematics

Given a joint angle configuration \mathbf{y} , the end effector position can be calculated as $\mathbf{x}_{ee} = f(\mathbf{y})$ where $f(\mathbf{y})$ denotes the forward kinematics of the robot, which is based on the robot's geometry and calculates the end effector position through the hierarchy of intermediate transformations of each of the individual joints. For example, given the geometry of one's arm sizes, and the angles of the shoulder, elbow, and wrist joints, the position of the hand can be calculated through the position of the shoulder to the elbow to the hand.

Given a list of n joints $\mathbf{y} = \{y_1, y_2 \dots y_n\}$ and the corresponding relative transformations that the joints represent in the robot geometry $\mathbf{T}_1^0(y_1), \mathbf{T}_2^1(y_2) \dots \mathbf{T}_n^{n-1}(y_n)$ where $\mathbf{T}_j^i \in$

$SE(3)$ is the relative transformation between links i and j , the end effector pose T_{ee} can be calculated as

$$T_{ee} = \begin{bmatrix} \mathbf{R}_{ee} & \mathbf{x}_{ee} \\ \mathbf{0} & 1 \end{bmatrix} = T_n^0 = \prod_{i=1}^n T_i^{i-1}(y_i) \quad (3.12)$$

where $\mathbf{x}_{ee}, \mathbf{R}_{ee}$ are the end effector position and rotation.

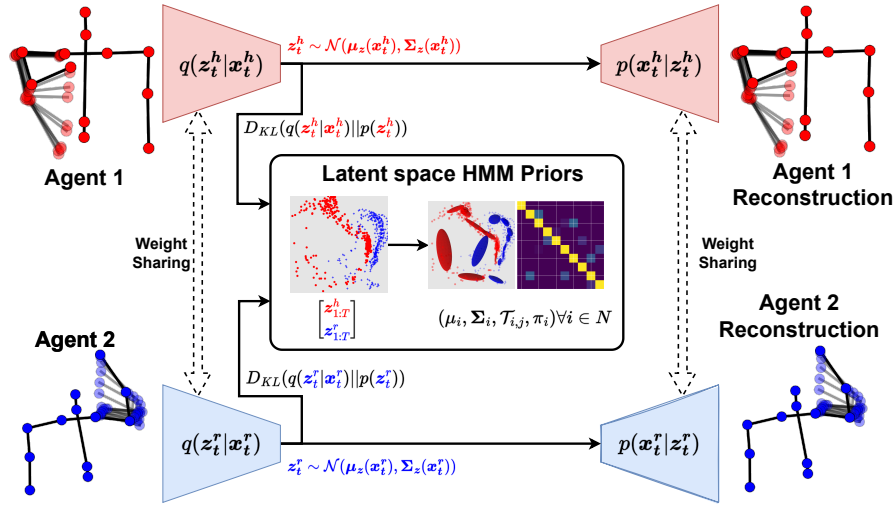
As the name suggests, Inverse Kinematics (IK) is the reverse process of estimating the robot’s joint angles from a given end effector pose, which can be solved via optimization. We aim to estimate $\mathbf{y} = f^{-1}(\mathbf{x}_{ee})$, which can be solved by finding the optimal joint configuration that minimizes the distance between the expected and predicted positions

$$\mathbf{y}^* = \arg \min_{\mathbf{y}} \|f(\mathbf{y}) - \mathbf{x}_{ee}\|^2. \quad (3.13)$$

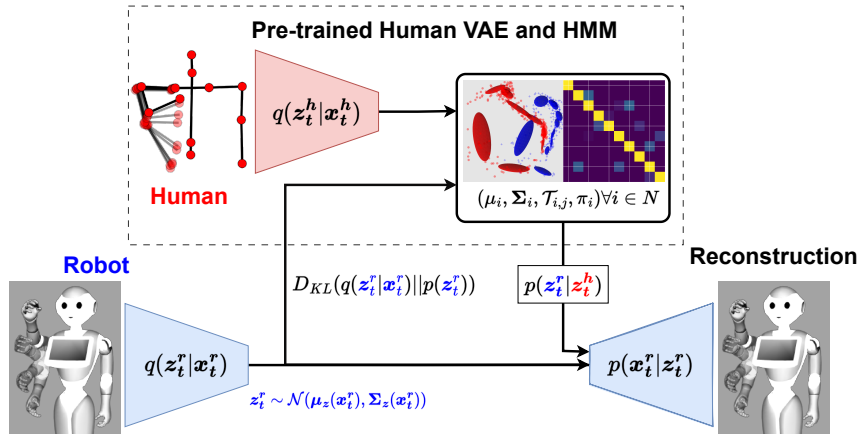
3.3. Multimodal Interactive Latent Dynamics

In this section, we introduce our overall approach, “MILD” for learning Multimodal Interactive Latent Dynamics which can be seen in Figure 3.3. We model the interaction dynamics in the latent space of a VAE using HMMs to model a joint distribution over the latent trajectories of both agents. The use of HMMs for the prior, as opposed to a unimodal Gaussian, enforces a multimodal modularized latent space with the HMM forward variable incorporating the learned transitions between the multiple modes. Once the underlying interaction dynamics is learned from Human-Human Interactions, the model is then used not just to regularize learning the robot motions but subsequently for reactively generating the response motions of the robot during training. This further incorporates the interdependence between the human and the robot during training thereby improving the predicted robot trajectories.

We start by explaining the key idea of MILD [158] which enhances the representation learning abilities of VAEs with HMMs as the prior distribution for modeling the interaction dynamics (Section 3.3.1). We then go through our key improvements on incorporating the human-conditioned HMM distribution into the training process in Section 3.3.2 and then explain the Inverse Kinematics adaptation for the generated robot motions in Section 3.3.3. For HRI scenarios, we use the superscripts h and r to denote the human and the robot variables respectively. For consistency, in the HHI scenarios, we use h and r to denote the variables of the first and second human partners respectively.



(a) Learning Interaction Models from Human-Human Interactions.



(b) Using Human-Human Interaction models for Human-Robot Interaction models.

Figure 3.3.: Overview of our training approach. (a) We first learn the interaction dynamics from Human-Human Interactions by training both the VAEs and HMMs alternatively by using the HMM as the VAE prior and then by using the VAE embeddings to train the HMM. (b) We subsequently use the model learned from Human-Human Interactions for learning the latent dynamics for HRI. We do so by regularizing the robot VAEs with the learned HMMs and additionally, by training the robot decoder to reconstruct samples from the HMM's conditional distribution after observing the human partner.

3.3.1. Learning Interaction Dynamics using HMMs as VAE priors

Typically in VAEs, the prior $p(z)$ is modeled as a stationary distribution. When it comes to learning trajectory dynamics, having meaningful priors can help learn temporally coherent latent spaces [39]. To this end, we explore using HMMs to learn latent-space dynamics in a modular manner by breaking down the trajectories into multiple phases and learning the sequencing between them. Since the HMMs learn a joint distribution over the latent trajectories of both interacting agents, we can conditionally generate the motion of the robot after observing the human agent. We do so by using the HMM distribution as the prior for the VAE. We then update the HMM at the end of each epoch by running expectation-maximization on the VAE embeddings, alternating between training the VAE and then subsequently updating the HMM.

The VAE prior at each time-step is calculated as the respective marginal distribution of the most likely HMM component at that timestep, given by the HMM forward variable (Eq. 3.4). However, the recurrent nature of estimating the forward variable, and consequently its gradients, leads to numerical instabilities from backpropagation through time. Hence, we use an approximation in the form of an unobserved forward variable by setting the likelihood term in Eq. 3.4 to unity. This unobserved forward variable provides a good approximation of the sequential progression of the hidden states based on the learned transitions which can be written as

$$\begin{aligned}\bar{\alpha}_i^t &= \frac{\hat{\alpha}_i^t}{\sum_{j=1}^N \hat{\alpha}_j^t} & \hat{\alpha}_i^t &= \sum_{j=1}^N \bar{\alpha}_j^{t-1} \mathcal{T}_{j,i} & \hat{\alpha}_i^0 &= \pi_i \\ i_t^* &= \arg \max_i \bar{\alpha}_i^t & & & & (3.14) \\ KL_t^h &= KL(q(z_t^h | \mathbf{x}_t^h) || \mathcal{N}(z_t^h; \boldsymbol{\mu}_{i_t^*}^h, \boldsymbol{\Sigma}_{i_t^*}^{hh})) \\ KL_t^r &= KL(q(z_t^r | \mathbf{x}_t^r) || \mathcal{N}(z_t^r; \boldsymbol{\mu}_{i_t^*}^r, \boldsymbol{\Sigma}_{i_t^*}^{rr}))\end{aligned}$$

where $\mathcal{N}(z_t^h; \boldsymbol{\mu}_{i_t^*}^h, \boldsymbol{\Sigma}_{i_t^*}^{hh})$ and $\mathcal{N}(z_t^r; \boldsymbol{\mu}_{i_t^*}^r, \boldsymbol{\Sigma}_{i_t^*}^{rr})$ are the marginal distributions of the most likely HMM component i_t^* at the given timestep t for each agent. The ELBO can then be reformulated as

$$\begin{aligned}\mathbb{E}_q[\log \frac{p(\mathbf{x}_t^h, \mathbf{x}_t^r, z_t^h, z_t^r)}{q(z_t^h, z_t^r | \mathbf{x}_t^h, \mathbf{x}_t^r)}] &= \mathbb{E}_q[\log p(\mathbf{x}_t^h | z_t^h)] \\ &+ \mathbb{E}_q[\log p(\mathbf{x}_t^r | z_t^r)] \\ &- \beta(KL_t^h + KL_t^r).\end{aligned}\tag{3.15}$$

The overall training procedure, shown in Algorithm 1 and Figure 3.3a, is as follows.

We train the VAE to reconstruct the input trajectory and use the marginal distributions of each agent from the HMM (Eq. 3.14) to regularize the latent posterior distribution. During a given epoch in the VAE training, the HMM parameters are fixed. After an epoch, the new HMM parameters are estimated using the learned encodings $z_{1:T}^h, z_{1:T}^r$ of the training trajectories. The HMMs are then fixed as the VAE prior for the next epoch. Currently, we learn a separate HMM per interaction. We defer learning the HMM selection via activity recognition to future work. For further details on training HMMs with expectation-maximization, or the use of HMMs in robot learning, we refer the reader to [49, 30].

Algorithm 1: Learning Latent Dynamics from Human-Human Interactions

Data: A set of trajectories with action labels $\mathbf{X} = \{\mathbf{X}_{1:T}^h, \mathbf{X}_{1:T}^r, c\}$ for $|\mathcal{C}|$ interactions

Result: VAE weights and $|\mathcal{C}|$ HMM parameters

Initialize VAE weights randomly

```

for  $c \in [1, |\mathcal{C}|]$  do
  for  $i \in [1, N]$  do
     $\mu_i^c \leftarrow \mathbf{0}$ 
     $\Sigma_i^c \leftarrow I$ 
  end
end
while not converged do
  for  $\mathbf{x}_{1:T}^h, \mathbf{x}_{1:T}^r, c \in \mathbf{X}$  do
    Compute VAE Posterior  $q(z_t^h | \mathbf{x}_t^h)$  and  $q(z_t^r | \mathbf{x}_t^r)$ 
    Reconstruct posterior samples
    Maximize ELBO (Eq. 3.15) to update VAE weights
  end
  for  $c \in [1, |\mathcal{C}|]$  do
     $\mathbf{X}^c \leftarrow$  set of demonstrations of Interaction  $c$ 
     $\mathbf{Z}^c \leftarrow \emptyset$ 
    for  $\mathbf{x}_{1:T}^h, \mathbf{x}_{1:T}^r, c \in \mathbf{X}^c$  do
       $z_{1:T}^h \sim q(\cdot | \mathbf{x}_{1:T}^h); z_{1:T}^r \sim q(\cdot | \mathbf{x}_{1:T}^r)$ 
       $\mathbf{Z}^c \leftarrow \mathbf{Z}^c \cup \begin{bmatrix} z_{1:T}^h \\ z_{1:T}^r \end{bmatrix}$ 
    end
    Train the  $c^{th}$  HMM with  $\mathbf{Z}^c$ 
  end
end

```

3.3.2. Conditional Training of HRI Dynamics from HHI

The HMMs learned from Human-Human demonstrations capture the overall latent interaction dynamics between two agents. These HMMs can, therefore, be used as an informative prior to learn robot motions to perform the given interaction. Hence, we use the marginal distribution of the second agent from the HMMs trained on the HHI demonstrations as the latent space prior for the robot VAE.

Although we regularize the VAEs with the HMM marginals (Eq. 3.14), during test time the decoder would see samples from the conditional distribution of the HMM after observing the human agent (Eq. 3.6-3.11) which would be finitely divergent from what the decoder would be trained to reconstruct in a normal autoencoding approach.

When the output spaces of both interaction partners are similar, such as in the HHI scenarios sharing the weights of the VAEs for can enable the decoder to learn to reconstruct the target distribution. However, given the difference in output spaces of the human and the robot in HRI scenarios, we instead train the robot decoder to additionally reconstruct samples from conditional distribution (Figure 3.3b).

Moreover, the VAE provides a confidence estimate of the posterior probability which can additionally be incorporated into the conditional distribution $p(\mathbf{z}_t^r | \mathbf{z}_t^h)$ as

$$\mathbf{K}_i = \Sigma_i^{rh} (\Sigma_i^{hh} + \Sigma_z(\mathbf{x}_t^h))^{-1} \quad (3.16)$$

$$\hat{\boldsymbol{\mu}}_i^r = \boldsymbol{\mu}_i^r + \mathbf{K}_i (\boldsymbol{\mu}_z(\mathbf{x}_t^h) - \boldsymbol{\mu}_i^h) \quad (3.17)$$

$$\hat{\Sigma}_i^r = \Sigma_i^{rr} - \mathbf{K}_i \Sigma_i^{hr} + \hat{\boldsymbol{\mu}}_i^r (\hat{\boldsymbol{\mu}}_i^r)^T \quad (3.18)$$

$$\hat{\boldsymbol{\mu}}_t^r = \sum_{i=1}^N \bar{\alpha}_i^t \hat{\boldsymbol{\mu}}_i^r \quad (3.19)$$

$$\hat{\Sigma}_t^r = \left[\sum_{i=1}^N \bar{\alpha}_i^t \hat{\Sigma}_i^r \right] - \hat{\boldsymbol{\mu}}_t^r (\hat{\boldsymbol{\mu}}_t^r)^T \quad (3.20)$$

$$p(\mathbf{z}_t^r | q_t^h) = \mathcal{N}(\mathbf{z}_t^r; \hat{\boldsymbol{\mu}}_t^r, \hat{\Sigma}_t^r) \quad (3.21)$$

where the terms in **magenta** are from the human VAE posterior $q_t^h = q(\mathbf{z}_t^h | \mathbf{x}_t^h)$ and the terms in **orange** are from the HMM.

Our modified ELBO at each timestep for training the robot VAE in the HRI scenario (Algorithm 2) can be written as

$$\begin{aligned} \mathbb{E}_q[\log \frac{p(\mathbf{x}_t^h, \mathbf{x}_t^r, \mathbf{z}_t^r)}{q(\mathbf{z}_t^r | \mathbf{x}_t^h, \mathbf{x}_t^r)}] &= \mathbb{E}_{q_r} \log p(\mathbf{x}_t^r | \mathbf{z}_t^r) \\ &- \beta K L_t^r \\ &+ \mathbb{E}_{\mathbf{z}_t^r \sim p(\mathbf{z}_t^r | q_t^h)} \log p(\mathbf{x}_t^r | \mathbf{z}_t^r) \end{aligned} \quad (3.22)$$

where the first term is the reconstruction term, $K L_t^r$ is the regularization term calculated according to Eq. 3.14, and the third term denotes the reconstruction of the conditional samples drawn from $p(\mathbf{z}_t^r | q_t^h)$ (Eq. 3.21). Both the reconstruction term and the conditional reconstruction are estimated in a Monte Carlo fashion by averaging the loss after reconstructing multiple samples from the corresponding latent distributions.

Algorithm 2: Learning HRI Dynamics

Data: A set of trajectories with action labels $\mathbf{X} = \{\mathbf{X}_{1:T}^h, \mathbf{X}_{1:T}^r, c\}$ for $|\mathcal{C}|$ actions, Human VAE and HMMs

Result: Robot VAE weights

Initialize VAE weights randomly

while *not converged* **do**

for $\mathbf{x}_{1:T}^h, \mathbf{x}_{1:T}^r, c \in \mathbf{X}$ **do**

 Compute VAE Posterior $q(\mathbf{z}_t^r | \mathbf{x}_t^r)$

 Compute Latent Conditional $p(\mathbf{z}_t^r | q_t^h)$ (Eq. 3.16-3.21)

 Reconstruct posterior and conditional samples

 Maximize ELBO (Eq. 3.22) to update VAE weights

end

end

During testing, we first encode the observations of the human agent \mathbf{x}_t^h and then condition the HMM $p(\mathbf{z}_t^r | q_t^h)$ to generate the latent trajectory of the second agent $\hat{\mathbf{z}}_t^r$ using Eq. 3.16 - 3.21, which is then decoded to obtain the actions of the second agent $\hat{\mathbf{x}}_t^r \sim p(\cdot | \hat{\mathbf{z}}_t^r)$.

3.3.3. Inverse Kinematics Adaptation and Stiffness Modulation

In typical LfD approaches, robot motions are learned via kinesthetic teaching [152, 172]. In HRI scenarios, kinesthetic teaching gives good results [52, 6, 32, 29] but can become tedious when trying to generalize to multiple human interaction partners. One way to circumvent this could be to execute randomized open-loop trajectories for the robot and learn how the human adapts to the robot motions [33].

Since we are dealing with a Humanoid robot, we learn the robot motions using the kinematic similarities between the human and the robot [57]. From the 3D positions of an arm, we extract the shoulder angles (yaw, pitch, and roll) and the bending of the elbow by using the geometry of the tracked skeleton, as seen in Figure 3.4. We defer the calculation of the wrist angle to our future work and instead use a fixed value for the wrist for each interaction.

Since we extract joint angles from the human demonstrations, some inaccuracies are present due to slight differences between the kinematic structure of the human skeleton and the robot. Given the difference in the arm dimensions of the Pepper robot and a human, the spatial accuracy of the learned behaviors is quite limited. Additionally, given the smaller size of the Pepper robot, its reachability is also limited, thereby limiting the ability to learn task space trajectories.

To bridge the mismatch in motion retargeting, we adapt the predicted motions during test time with Inverse Kinematics to reach the human partner’s hand while using the predicted motions as a prior. The use of Inverse Kinematics enables the physical proximity needed by the interaction while keeping the robot configuration close to the demonstrated behaviors [63]. Moreover, we do not need to use Inverse Kinematics all the time, but only in the segments involved in physical contact. We therefore inspect the underlying segments of the HMM manually to see which segments comprise of the contact-based portion of the trajectory and which ones do not, and subsequently perform the IK adaptation only in the contact-based segments.

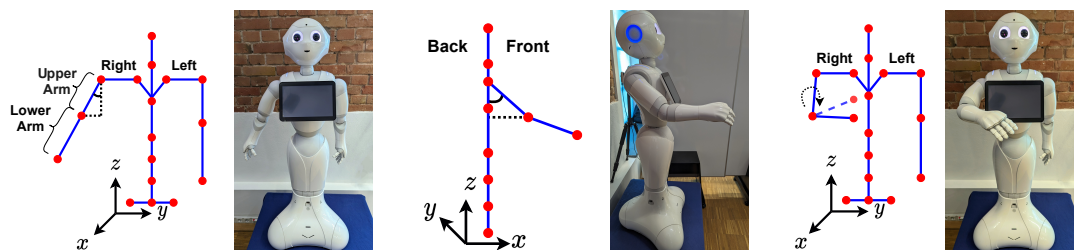


Figure 3.4.: Geometric similarities between the degrees of freedom of a human’s upper body and the humanoid robot Pepper. The shoulder roll, pitch, yaw, and elbow angle of a human can be directly mapped to Pepper’s joint angles.

Given an initial prediction of the robot’s joint angle distribution (μ_q, Σ_q) and a task space goal in 3D space (μ_x, Σ_x) (which in this case is the human partner’s hand position), we aim to find a joint angle configuration q^* that reaches μ_x , but does not stray too far

from μ_q [63] which boils down to the following optimization problem

$$\begin{aligned}
\mathbf{q}^* &= \arg \max_{\mathbf{q}} \mathcal{N}(f(\mathbf{q})|\mu_x, \Sigma_x) \mathcal{N}(\mathbf{q}|\mu_q, \Sigma_q) \\
&= \arg \min_{\mathbf{q}} \lambda_x (\mu_x - f(\mathbf{q}))^T \Sigma_x^{-1} (\mu_x - f(\mathbf{q})) + \\
&\quad \lambda_q (\mu_q - \mathbf{q})^T \Sigma_q^{-1} (\mu_q - \mathbf{q})
\end{aligned} \tag{3.23}$$

where $f(\mathbf{q})$ is the forward kinematics model of the robot to estimate the end-effector location given the joint angles \mathbf{q} , and λ_x, λ_q are relative weights balancing whether it is more important to stay close to the initial distribution or to reach the task space goal. In practice, the joint configuration prior μ_q comes from the VAE decoder $p(\mathbf{x}_t^r | \mathbf{z}_t^r)$ and μ_x comes from the human observation \mathbf{x}_t^h , and we set Σ_q and Σ_x to identity. We can then simplify Eq. 3.23 to

$$\mathbf{q}^* = \arg \min_{\mathbf{q}} \lambda_x \|\mu_x - f(\mathbf{q})\|^2 + \lambda_q \|\mu_q - \mathbf{q}\|^2 \tag{3.24}$$

where $\|\cdot\|$ denotes the Euclidean norm.

Moreover, we do not need to employ IK during every segment of the interaction. Once the training converges, we identify the segments from the underlying HMM that correspond to the contact-based phases of the interaction. Based on the HMM’s forward variable prediction from the human observation, we adapt the human-conditioned robot motions using IK if the current segment is contact-based. The incorporation of IK is further detailed in Algorithm 3.

In a contact-rich interaction like handshaking, along with moving in a coordinated manner, the robot must be compliant with the human’s motion when in contact. When executing handshaking trajectories on the robot, we use the forward variable not just for calculating the target pose of the robot, but also to detect when to lower the target joint stiffness once the robot enters the contact-based segments of the interaction. Without such as mechanism, the Pepper robot’s arm would not comply with the human’s hand, and the human would be unable to displace Pepper’s hand for executing a proper handshake.

3.4. Experiments and Results

In this section, we first provide our implementation details (Section 3.4.1) and the datasets used (Section 3.4.2). We then present the results of predicting the controlled agent’s trajectories after conditioning on the observed agent in Section 3.4.3 and finally, we discuss our user-study in Section 3.4.4 and show some additional results for a Bimanual Robot-to-Human Handover scenario in Section 3.4.5.

Algorithm 3: Conditioning on Human Observations and IK Adaptation

Data: An observation of the human agent $x_{1:t}^h$, trained VAE and HMM Models, a set of contact-based HMM states \mathcal{I} for the given interaction

Result: Conditioned Trajectory for the second agent $\hat{x}_{1:t}^r$

```
for  $t \in [1, T]$  do
  Encode the human observation  $q(z_t^h | x_t^h)$ 
  Compute Latent Conditional  $p(z_t^r | q_t^h)$  (Eq. 3.16-3.21)
  Decode the conditioned prediction  $\mu_q = p(\hat{x}_t^r | z_t^r)$ 
  if  $\arg \max_i \alpha_i(z_t^h)$  is a contact-based segment then
    |  $\hat{q}_t = \arg \min_q \lambda_x \|\mu_x - f(\hat{q}_t)\|^2 + \lambda_q \|\mu_q - q\|^2$ 
  else
    |  $\hat{q}_t = \mu_q$ 
  end
  Send  $\hat{q}_t$  to the robot controller.
end
```

3.4.1. Experimental Setup

The VAEs are implemented using PyTorch [153] with 2 hidden layers in the encoder and decoder each with sizes (40, 20) and (20, 40) respectively and a 5-dimensional latent space with Leaky ReLU activations [115] at all layers except the output layer. The weights are initialized using Xavier’s initialization [61]. The networks are trained with $\beta = 5 \times 10^{-3}$, 10 Monte Carlo samples per input data and using the Adam optimizer with weight decay [114] with a learning rate of 5×10^{-4} . In the HHI scenarios, we share the parameters of the VAEs for both human agents as the inputs are structurally similar. The networks were trained for 400 epochs. The best out of 4 seeds were used to report the results.

A separate HMM is trained for each interaction. The HMMs are implemented using a custom PyTorch version of PbDLib¹ [156]. Each HMM has 6 hidden states, which we found as the most numerically stable. The HMMs are initialized by splitting each of the latent trajectories in the training set into equally sized segments over time. To prevent numerical instabilities arising from vanishing values in the HMM covariance matrices, we add a small positive regularization constant of 10^{-4} to the diagonal elements. When sampling from the conditional distribution, due to numerical errors in computing the Cholesky decomposition of the covariance matrices, rather than adding a constant value to all elements, we add linearly increasing regularization based on the dimensions, going from 9.1×10^{-5} for the first dimension to 10^{-4} for the last dimension. Additionally, we

¹<https://gitlab.idiap.ch/rli/pbdlb-python>

adapt the covariance matrices of the conditional distribution by iteratively adding an even smaller constant that is proportional to the absolute value of the smallest eigenvalue of the covariance matrix to the diagonal elements [72].

We implement our approach on the Pepper robot [149], which is a 1.2m tall humanoid robot that has the same degrees of freedom in its arm as a human. For Inverse Kinematics, we use a modified version of the IKPy library [118]² which uses the least squares minimization in SciPy [199] for the IK optimization. We set $\lambda_x = 1$ and $\lambda_q = 0.01$ for the objective function in Eq. 3.24.

To prevent sudden changes in stiffness arising from misclassification of the segment, we disable back-transitions into the initial reaching segment. Additionally, during test time, since the forward variable is calculated using only the human partner’s latent state, there is a mismatch that arises in the segment prediction compared to using the full joint human-robot states for timesteps near the transition boundary between the initial reaching segment and the subsequent segments. Therefore, taking a leaf out of Transition State Clustering [101, 65], we learn an additional distribution over the states at the transition boundary that get misclassified. We then use this transition state distribution to detect when the interaction proceeds into the contact-based segments. Doing so gives a better indication of when to lower the robot’s stiffness and provides a more suitable interaction. Without such a scheme, due to the misclassification of the active segment, the joint stiffness would not always be lowered correctly, thereby resulting in a rigid and non-compliant handshake.

3.4.2. Datasets

Bütepage et al. [29]

The authors of [29] record HHI and HRI demonstrations of 4 interactions: Waving, Handshaking, and two kinds of fist bumps³. The first fistbump called “Rocket Fistbump”, involves bumping fists at a low level and then raising them upwards while maintaining contact with each other. The second is called “Parachute Fistbump” in which partners bump their fists at a high level and bring them down while simultaneously oscillating the hands sideways, while in contact with each other. Since our testing scenario involves the humanoid robot Pepper [149], we additionally extract the joint angles from one of the human partner’s skeletons from the above-mentioned HHI data for the Pepper robot using the similarities in DoFs between a human and Pepper [57, 162] (Figure 3.4), which we denote this as “HRI-Pepper”. Bütepage et al. [29] additionally record demonstrations

²<https://github.com/souljaboy764/ikpy>

³https://github.com/jbutepage/human_robot_interaction_data

of these actions with a human partner interacting with an ABB YuMi-IRB 14000 robot controlled via kinesthetic teaching. We call this scenario as “HRI-Yumi”.

In the HHI scenario (and the HRI-Pepper scenario), there are 181 trajectories (32 - Waving, 38 - Handshake, 70 - Rocket Fistbump, 49 - Parachute Fistbump) of which 80% of the trajectories (149 trajectories) are used for training and the rest (32 trajectories) for testing. In the HRI-Yumi scenario, there are 41 trajectories (10 each for Waving, Handshaking, and Rocket Fistbump and 11 for Parachute Fistbump) of which we use a similar split with 32 trajectories for training and 9 for testing. We use a time window of 5 observations as the input for a given timestep as done in [29]. We downsample the data to 20Hz to match our testing scenario.

We use the 3D positions and the velocities (represented as position deltas) of the right arm joints (shoulder, elbow, and wrist), with the origin at the shoulder, leading to an input size of 90 dimensions (5x3x6: 5 timesteps, 3 joints, 6 dimensions) for a human partner. For the HRI-Pepper and HRI-Yumi scenarios, we use a similar window of joint angles, leading to an input size of 20 dimensions (5x4) for the 4 joint angles of Pepper’s right arm and an input size of 35 dimensions (5x7) for the 7 joint angles of Yumi’s right arm.

Nuitrack Skeleton Interaction Dataset (NuiSI)

The Nuitrack Skeleton Interaction Dataset (NuiSI) ⁴ is a dataset which we collected ourselves of the same 4 interactions as in [29]. While the dataset in Bütepage et al. [29] has clean motions of the interaction partners, they require high-accuracy inertial sensors in whole-body suits. Therefore, for a more realistic setting, we capture the interaction partners using low-cost hardware.

While other datasets exist that perform skeleton tracking with a Kinect Camera [179, 195, 183], the participants in these datasets are recorded from the side. Therefore, due to partial occlusions, there is a high degree of noise in skeleton tracking, which made it difficult to train a suitable model for the interactions. Therefore, we record two human partners interacting with one another using one Intel Realsense D435 camera per partner (two in total), such that each camera captures a full frontal view of the interaction. We use Nuitrack [1] for tracking the upper body skeleton joints in each frame at 30Hz.

The data is first inspected manually to remove any trajectories where the tracking deteriorates. Finally, we have 44 trajectories (12 - Waving, 11 - Handshaking, 12 - Rocket Fistbump, 9 - Parachute Fistbump) of which we similarly use 80% of the trajectories for training (33 trajectories) and the rest (11 trajectories) for testing.

The skeletons are then rotated such that they follow the orientation shown in Figure 3.4

⁴<https://github.com/souljaboy764/nuisi-dataset>

with the x -axis in the forward direction, the y -axis going from right to left, and the z -axis in the upward direction. For training, the data is processed similarly as mentioned above with a window size of 5 time steps. Given the relatively low amount of samples, we finetune the network trained on the aforementioned dataset of [29]. As done with [29], we extract the joint angles for the Pepper robot from the skeleton trajectories for the HRI-Pepper scenario.

3.4.3. Conditioned Prediction Results

We test the conditioning ability of our approach compared to [29]⁵ to evaluate the accuracy of the generated motions of the controlled agent after observing the human interaction-partner. We evaluate the approaches over the four interactions of the dataset in [29] and our collected data (NuiSI). We calculate the Mean Squared Error (MSE) averaged over each joint and the 5-time step window.

Dataset	Action	MILD v1	[29]
[29]	Hand Wave	0.788 ± 1.226**	4.121 ± 2.252
	Handshake	1.654 ± 1.549*	1.181 ± 0.859
	Rocket Fistbump	0.370 ± 0.682	0.544 ± 1.249
	Parachute Fistbump	0.537 ± 0.579**	0.977 ± 1.141
NuiSI	Hand Wave	0.408 ± 0.538**	3.168 ± 3.392
	Handshake	0.311 ± 0.259**	1.489 ± 3.327
	Rocket Fistbump	1.142 ± 1.375**	3.576 ± 3.082
	Parachute Fistbump	0.453 ± 0.578**	2.008 ± 2.024

Table 3.1.: Prediction MSE (in cm) for the second human partner’s trajectories after observing the first human partner averaged over all joints and timesteps. (* – $p < 0.05$, ** – $p < 0.01$, Lower is better)

We do not train with the conditional loss in the Human-Human scenarios since we use shared weights, the decoder already learns to reconstruct the ground truth samples for the conditional distribution. Additionally, we found that the interplay between the shared weights and the conditional training discussed in Section 3.3.2 would cause the HMM posterior to collapse into a single unimodal distribution. Therefore, we only show comparisons of our vanilla approach presented in Section 3.3.1, which we denote as “MILD v1”.

⁵Results reported using our implementation of [29].

The results of MILD v1 in predicting the interaction partner’s trajectories in the HHI scenarios can be seen in Table 3.1. It can be seen that MILD v1 with its simplistic nature of using an HMM for the latent dynamics performs significantly better (as verified with a Mann-Whitney U Test) than [29] where the VAEs are trained with an uninformative standard normal distribution as a prior. Although additional LSTMs are employed in [29] to learn the latent dynamics, since the VAEs are trained with an uninformative prior, their approach fails to accurately reconstruct motions the learnt latent dynamics. In contrast, the latent dynamics of each segment of the interaction is captured well by the HMM which therefore acts as an informative prior for the VAEs which is reflected in the improved prediction accuracy of MILD v1.

Coming to HRI scenarios, starting from the initial incorporation of the HMM prior as shown in Section 3.3.1 (denoted as “MILD v1”), we explore two variants of the last conditional training term in Eq. 3.22. The first is with reconstructing the conditional latent predictions (i.e. the mean in Eq. 3.19) of samples drawn from the posterior distribution both with and without the use of the VAE posterior covariance (in Eq. 3.16). This can be summarized mathematically as $\mathbb{E}_{q(z_t^h | x_t^h)} \log p(x_t^r | \mu_t^r)$ where μ_t^r is calculated using Eq. 3.9 with samples drawn from $q(z_t^h | x_t^h)$. We explore this variant both without and with the posterior covariance in Eq. 3.16, denoted as “MILD v2.1” and “MILD v2.2” respectively

The second variant uses the posterior mean and covariance to calculate the conditional distribution and subsequently reconstruct samples drawn from the conditional distribution. This can be summarized mathematically as $\mathbb{E}_{p(z_t^r | q_t^h)} \log p(x_t^r | z_t^r)$ where $p(z_t^r | q_t^h)$ is calculated using Eq. 3.16 - 3.21 and the samples drawn from this distribution are reconstructed. We explore this variant both without and with the posterior covariance in Eq. 3.16, which we denote as “MILD v3.1” and “MILD v3.2” respectively.

The key differences between the different variants are highlighted in Table 3.2. To summarize, the prior means and covariances from the HMM are used by both variants in the conditioning, denoted by the orange terms in Eq. 3.16 - 3.21. The key difference between the variants is how the posterior distribution terms (denoted in magenta in Eq. 3.16 - 3.21) are used. In MILD v3.1 and MILD v3.2, we directly use the posterior mean and covariance as shown in Eq. 3.16 - 3.21. In the case of MILD v2.1 and MILD v2.2, we do not directly use the posterior mean and covariance, but first draw samples from the posterior distribution $z_t^h \sim \mathcal{N}(\mu_z(x_t^h), \Sigma_z(x_t^h))$ which are then used in the conditioning in Eq. 3.17 instead of the posterior mean $\mu_z(x_t^h)$ which can be written as $\hat{\mu}_i^r = \mu_i^r + K_i(z_t^h - \mu_i^h)$ to directly get the conditional samples that then get reconstructed.

As seen in Table 3.3, reconstructing samples from the conditional distribution (MILD v3.1 and v3.2) provides much better results as compared to conditioning samples drawn from the posterior (MILD v2.1 and v2.2), both of which perform better than MILD v1 that

Variant	Conditional Training Inputs	Samples given to Decoder
MILD v1	None	Only Posterior Samples
MILD v2.1	Posterior Samples	Posterior Samples and their corresponding conditioned outputs
MILD v2.2	Posterior Samples and Covariance	
MILD v3.1	Posterior Mean	Posterior Samples as well as Conditional Samples
MILD v3.2	Posterior Mean and Covariance	

Table 3.2.: Differences in conditioning and sampling strategies of the variants of MILD.

Dataset	Action	MILD v1	MILD v2.1	MILD v2.2	MILD v3.1	MILD v3.2	Bütepage et al. [29]
HRI-Yumi [29]	Hand Wave	1.705 ± 0.521	1.349 ± 1.972	1.641 ± 1.968	1.033 ± 1.204	1.143 ± 1.330	0.225 ± 0.302
	Handshake	0.290 ± 0.148	0.073 ± 0.040	0.068 ± 0.052	0.104 ± 0.056	0.123 ± 0.069	0.133 ± 0.214
	Rocket Fistbump	0.428 ± 0.175	0.236 ± 0.167	0.183 ± 0.122	0.130 ± 0.074	0.128 ± 0.071	0.147 ± 0.119
	Parachute Fistbump	0.425 ± 0.150	0.028 ± 0.042	0.033 ± 0.034	0.028 ± 0.034	0.028 ± 0.035	0.181 ± 0.155
HRI-Pepper [29]	Hand Wave	0.267 ± 0.152	0.161 ± 0.228	0.165 ± 0.189	0.103 ± 0.103	0.106 ± 0.105	0.664 ± 0.277
	Handshake	0.327 ± 0.253	0.111 ± 0.092	0.153 ± 0.154	0.061 ± 0.048	0.056 ± 0.041	0.184 ± 0.141
	Rocket Fistbump	0.161 ± 0.095	0.035 ± 0.068	0.035 ± 0.068	0.021 ± 0.037	0.018 ± 0.035	0.033 ± 0.045
	Parachute Fistbump	0.265 ± 0.178	0.116 ± 0.176	0.112 ± 0.181	0.095 ± 0.151	0.088 ± 0.148	0.189 ± 0.196
HRI-Pepper (NuisI)	Hand Wave	0.760 ± 0.325	0.050 ± 0.084	0.060 ± 0.087	0.046 ± 0.059	0.049 ± 0.059	0.057 ± 0.093
	Handshake	0.225 ± 0.114	0.025 ± 0.022	0.025 ± 0.020	0.021 ± 0.015	0.020 ± 0.014	0.083 ± 0.075
	Rocket Fistbump	0.354 ± 0.238	0.077 ± 0.095	0.080 ± 0.088	0.077 ± 0.067	0.079 ± 0.072	0.101 ± 0.086
	Parachute Fistbump	0.201 ± 0.072	0.032 ± 0.038	0.028 ± 0.040	0.025 ± 0.028	0.022 ± 0.027	0.049 ± 0.040

Table 3.3.: Prediction MSE (in radians) for robot trajectories after observing the human interaction partner averaged over all joints and timesteps. (Lower is better, Lowest MSE is highlighted, Significance Values shown in Tables 3.4, 3.5, and 3.6)

does not use conditional training. The samples drawn from the conditional distribution would be more representative of the type of samples that the decoder would see during run time. While in theory, enough samples from the posterior, when conditioned, can also estimate this distribution, empirically, this fails to match up to sampling from the conditional distribution. One argument for this is that reconstructing conditional samples enables learning a joint latent space more susceptible to the HMM conditioning.

Furthermore, the overall improvement in performance compared to MILD v1 and [29] highlights the advantage of incorporating conditional prediction into the training process for reactive motion generation. The importance of incorporating reactive motion generation into the training can also be seen in the improved performance of [29] over MILD v1 in the HRI scenarios as compared to the HHI scenarios. This improvement comes from the fact that MILD uses the HMMs just as a latent prior, whereas in [29], the authors explicitly train a separate HRI dynamics model for predicting the robot motions from the latent trajectories of both the human and the robot. During testing, however, the HRI dynamics network does not have access to the ground truth target of the robot which is used to train the network. Therefore, the HRI dynamics is predicted in an autoregressive manner, which deteriorates the performance due to out-of-distribution data. In this regard, it can be seen

		Version of MILD					
		v1	v2.1	v2.2	v3.1	v3.2	
HRI-Yumi [29]	Waving	[29]	0.	0.	0.	0.	0.
		MILD v1	–	0.	0.	0.	0.
		MILD v2.1	–	–	0.	0.284	0.337
		MILD v2.2	–	–	–	0.	0.
		MILD v3.1	–	–	–	–	0.675
	Handshake	[29]	0.	0.	0.397	0.	0.
		MILD v1	–	0.	0.	0.	0.
		MILD v2.1	–	–	0.018	0.	0.
		MILD v2.2	–	–	–	0.	0.
		MILD v3.1	–	–	–	–	0.008
	Rocket Fistbump	[29]	0.	0.	0.	0.976	0.999
		MILD v1	–	0.	0.	0.	0.
		MILD v2.1	–	–	0.	0.	0.
		MILD v2.2	–	–	–	0.	0.
		MILD v3.1	–	–	–	–	0.909
	Parachute Fistbump	[29]	0.	0.136	0.098	0.	0.
		MILD v1	–	0.	0.	0.055	0.
		MILD v2.1	–	–	0.785	0.	0.
		MILD v2.2	–	–	–	0.	0.
		MILD v3.1	–	–	–	–	0.003

Table 3.4.: p -values for pairwise MSE comparisons of the results reported on the HRI-Yumi scenario from the dataset of [29] in Table 3.3. For $p < 0.001$, we report zeros.

		Version of MILD					
		v1	v2.1	v2.2	v3.1	v3.2	
HRI-Pepper [29]	Waving	[29]	0.	0.	0.	0.	0.
		MILD v1	–	0.	0.	0.	0.
		MILD v2.1	–	–	0.	0.	0.
		MILD v2.2	–	–	–	0.	0.
		MILD v3.1	–	–	–	–	0.119
	Handshake	[29]	0.	0.	0.	0.	0.
		MILD v1	–	0.	0.	0.	0.
		MILD v2.1	–	–	0.	0.	0.
		MILD v2.2	–	–	–	0.	0.
		MILD v3.1	–	–	–	–	0.041
	Rocket Fistbump	[29]	0.	0.291	0.838	0.	0.
		MILD v1	–	0.	0.	0.	0.
		MILD v2.1	–	–	0.	0.	0.
		MILD v2.2	–	–	–	0.	0.
		MILD v3.1	–	–	–	–	0.
	Parachute Fistbump	[29]	0.	0.	0.	0.	0.
		MILD v1	–	0.	0.	0.	0.
		MILD v2.1	–	–	0.016	0.	0.
		MILD v2.2	–	–	–	0.	0.
		MILD v3.1	–	–	–	–	0.018

Table 3.5.: p -values for pairwise MSE comparisons of the results reported on the HRI-Pepper scenario from the dataset of [29] in Table 3.3. For $p < 0.001$, we report zeros.

		Version of MILD					
		v1	v2.1	v2.2	v3.1	v3.2	
HRI-Pepper (NuiSI)	Waving	[29]	0.	0.	0.139	0.158	0.309
		MILD v1	–	0.	0.	0.	0.
		MILD v2.1	–	–	0.	0.	0.
		MILD v2.2	–	–	–	0.004	0.530
		MILD v3.1	–	–	–	–	0.014
	Handshake	[29]	0.	0.	0.	0.	0.
		MILD v1	–	0.	0.	0.	0.
		MILD v2.1	–	–	0.862	0.095	0.003
		MILD v2.2	–	–	–	0.104	0.007
		MILD v3.1	–	–	–	–	0.065
	Rocket Fistbump	[29]	0.	0.	0.	0.	0.
		MILD v1	–	0.	0.	0.	0.
		MILD v2.1	–	–	0.036	0.	0.
		MILD v2.2	–	–	–	0.050	0.023
		MILD v3.1	–	–	–	–	0.685
	Parachute Fistbump	[29]	0.	0.	0.	0.	0.
		MILD v1	–	0.	0.	0.	0.
		MILD v2.1	–	–	0.026	0.017	0.
		MILD v2.2	–	–	–	0.808	0.018
		MILD v3.1	–	–	–	–	0.023

Table 3.6.: p -values for pairwise MSE comparisons of the results reported on the HRI-Pepper scenario from the NuiSI dataset in Table 3.3. For $p < 0.001$, we report zeros.

that explicitly incorporating such conditional out-of-distribution samples can lead to better results, as seen in the improved performance of the variants of MILD (v2.1 - v3.2).

We show a sample interaction of the learned behaviors, along with the progression of the HMM in the latent space for handshake (Figure 3.5) and rocket fistbump (Figure 3.6) on the Pepper robot. As it can be seen, the HMM captures the sequencing between the multiple modes of the latent space to generate suitable motions for real-world HRI scenarios. This is additionally validated via a user study (Section 3.4.4) which shows the ability of our approach to generalize well to various users, despite being trained on demonstrations of just two partners.

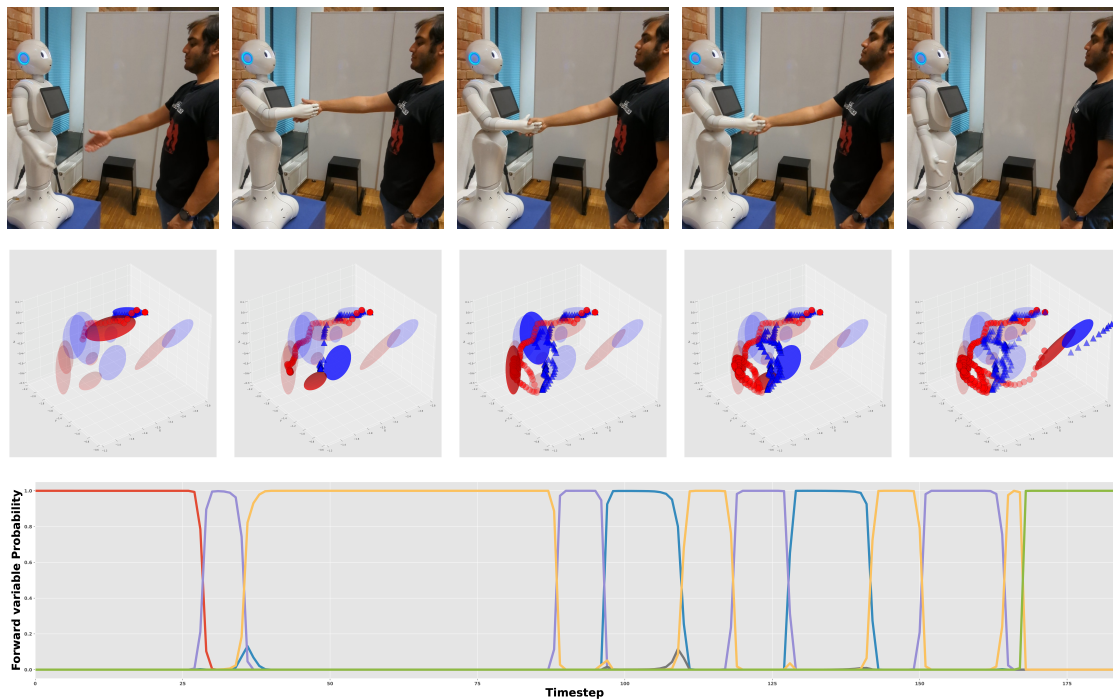


Figure 3.5.: Sample Handshake HRI on the Pepper robot. The top row shows the result of the generated reactive motion after observing the human partner’s skeleton. The middle row shows the latent trajectories and the HMM segments over the first 3 dimensions of the latent space (Red - Human, Blue - Robot). The opacity of each cluster is the corresponding cluster probability, given by the HMM forward variable. The progression of the HMM forward variable is shown in the bottom row, with different colors denoting the different segments.

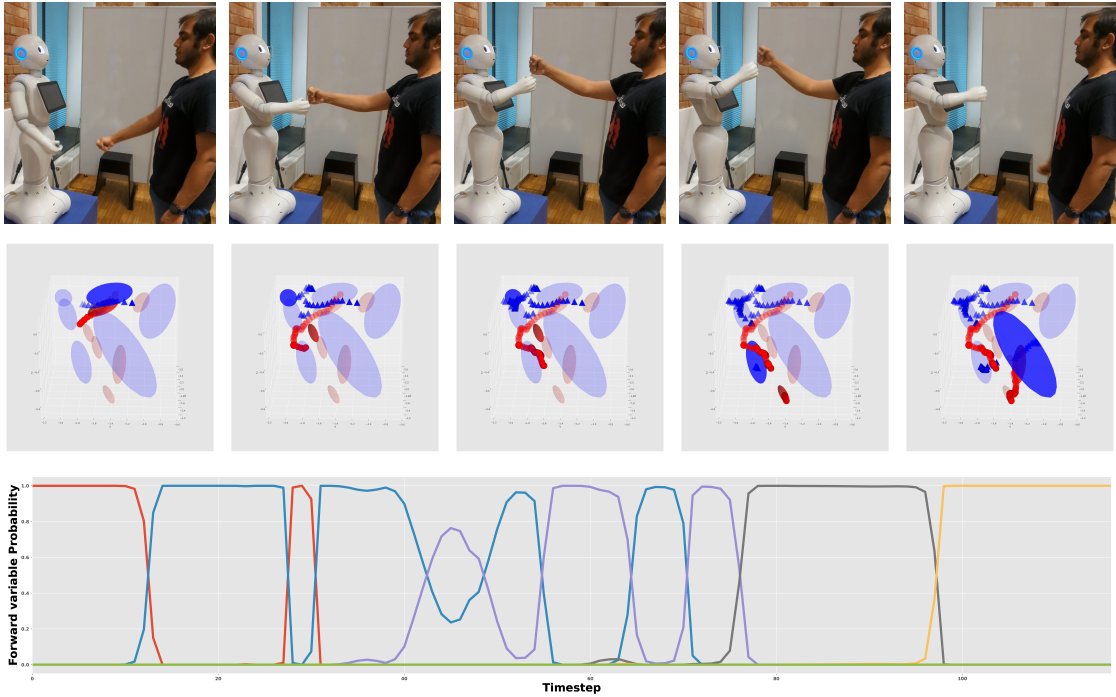


Figure 3.6.: Sample Rocket HRI on the Pepper robot. The top row shows the result of the generated reactive motion after observing the human partner’s skeleton. The middle row shows the latent trajectories and the HMM segments over the first 3 dimensions of the latent space (Red - Human, Blue - Robot). The opacity of each cluster is the corresponding cluster probability, given by the HMM forward variable. The progression of the HMM forward variable is shown in the bottom row, with different colors denoting the different segments.

3.4.4. HRI User Study

To see the effectiveness of our approach in producing acceptable physical behaviors, we conducted a user study where participants interacted with the robot. We evaluate our proposed approach both with and without IK adaptation, denoted as “MILD-IK” and “MILD” (while we use “MILD v3.2”, for ease of notation, we shorten it to “MILD”) against a baseline IK algorithm (Eq. 3.13) that uses the human’s hand pose as the target location (“Base-IK”). We run this study for two interactions, handshake and rocket fistbump.



Figure 3.7.: Setup of the User Study for interacting with the Pepper robot. As Pepper is quite short (1.2m), it is placed on a pedestal to match the height of a human partner. The camera behind Pepper tracks the human partner's motion, which is used to generate Pepper's motions.

Procedure

The study took place in a laboratory setting (Figure 3.7) where participants were initially guided to a desk to fill out a consent form followed by a pre-questionnaire which included demographic information, prior experiences with robots, their attitude towards robots [188], attitude towards physical interactions, and some personality questions to gauge extroversion [164]. We assessed these measures to gain additional information about our sample. Participants were then shown an instruction video of two humans performing an interaction (either a handshake or a rocket fistbump). Participants were then given the experiment protocol to read wherein they were instructed that they had to lead the interaction with the robot. Additionally, some general instructions were provided regarding the limits of the robot and how they should position themselves.

Participants were then guided behind a barrier where they would see the robot for the first time and would then stand at an adequate distance from the robot. When they were ready, an initial interaction was performed with a hard-coded motion. This is to get the

participant habituated to the way the robot moves and counter any novelty effects that may occur and to have the participant better understand how to perform the interaction with the robot. After the initial run, the participant was informed that the experimental trials would begin. The sequence of each trial was as follows:

1. Before each trial, the participant would see a video stream of the skeleton tracker to better position themselves.
2. Once the position was set and the tracking was stable, the experimenter would signal them to start the trial.
3. The participant would start the interaction and the robot would respond reactively.
4. Once the participant goes back to a neutral position with their hands by their side, the robot arm would go back to a neutral position, marking the end of the trial.
5. This process would then repeat once again after which the participant was asked to fill out a questionnaire.

This process constitutes a single session and was repeated 3 times (3 sessions in total) wherein the robot was controlled by one of the three aforementioned algorithms (Base-IK, MILD, MILD-IK). The algorithms were shown in randomized order to all participants to avoid sequence effects. The participant was neither informed of the randomization nor the algorithm they were interacting with. After all 3 sessions, the participant was asked to fill out a final questionnaire where they had to rank the sessions based on their preference and answer some open-ended questions about the sessions.

Participant Sample

A total of 20 users (8 female, 12 male) participated in our study and were recruited through the university environment. Of the 20 participants, 10 performed the Rocket fistbump (4 female, 6 male) and the rest performed a handshake. The mean age of the participants was 27.85 years (SD: 3.51). Participants had an average level of experience with robots overall on a scale from 1 (no experience at all) to 5 (a lot of experience) with a mean of 2.90 (SD: 1.45). They had quite a positive attitude towards robots on a scale from 1 (very negative) to 7 (very positive) with a mean of 5.90 (SD: 1.16). This is also consistent with the ratings regarding the attitudes towards robots for the following items on a scale of 1 (Strongly Agree) to 7 (Strongly Disagree). Participants largely had a high agreement towards feeling relaxed when interacting with robots (Mean: 5.60, SD: 1.05) and high levels of disagreement towards being paranoid when interacting with a

robot (Mean: 1.85, SD: 0.93) and towards feeling nervous standing in front of a robot (Mean: 1.80, SD:1 .06). Participants had a positive outlook towards physical interactions in general on a scale from 1 (distant) to 7 (open) with a mean of 6.10 (SD: 0.84). This was also confirmed with the Big 5 extroversion scale [164] with a mean extroversion of 4.75 (SD: 1.34) out of 7.

Methodology

The study followed a within-subject design where participants interacted with a Pepper robot controlled by each of the algorithms (Base-IK, MILD and MILD-IK) twice in a randomized order, leading to 6 HRIs per participant. Since humans perceive different types of physical touch differently [58], we aimed to remove any influence the type of interaction (handshake or fistbump) might have on how participants view the robot during different interactions. We wanted to keep the focus on understanding how our proposed algorithm is perceived by users. Therefore, each participant had to either perform a handshake or a fistbump, not both. Additionally, to prevent any sudden jumps in the motion of the robot, a weighted moving average filter was used to smoothen out the predicted motions of the robot.

We break the 6 trials into 3 sessions, where each session corresponds to two trials of a given algorithm. Each session was evaluated with 16 different items adapted from the Godspeed [17] and the SASSI [75] questionnaires, each rated on a 5-Point scale (1 - Strongly disagree, 5 - Strongly agree):

- The interaction with the robot was pleasant.
- The interaction with the robot was exciting.
- The interaction with the robot was human-like.
- The interaction with the robot was natural.
- The interaction with the robot was friendly.
- The interaction with the robot was comfortable.
- The interaction with the robot was well-timed.
- The interaction with the robot was accurate.
- The interaction with the robot was annoying.
- The interaction with the robot was awkward.

-
- The interaction with the robot was scary.
 - The robot interacted in an aggressive way.
 - I am satisfied with the way the robot interacted with me.
 - The second trial in the session was more effortless than the first.

At the end of the experiment, we asked the participants to rank the 3 sessions (algorithms) in their order of preference. We run a one-way Repeated Measures ANOVA to compare the responses of the different algorithms followed by paired sample t-tests for a post hoc analysis (Table 3.7).

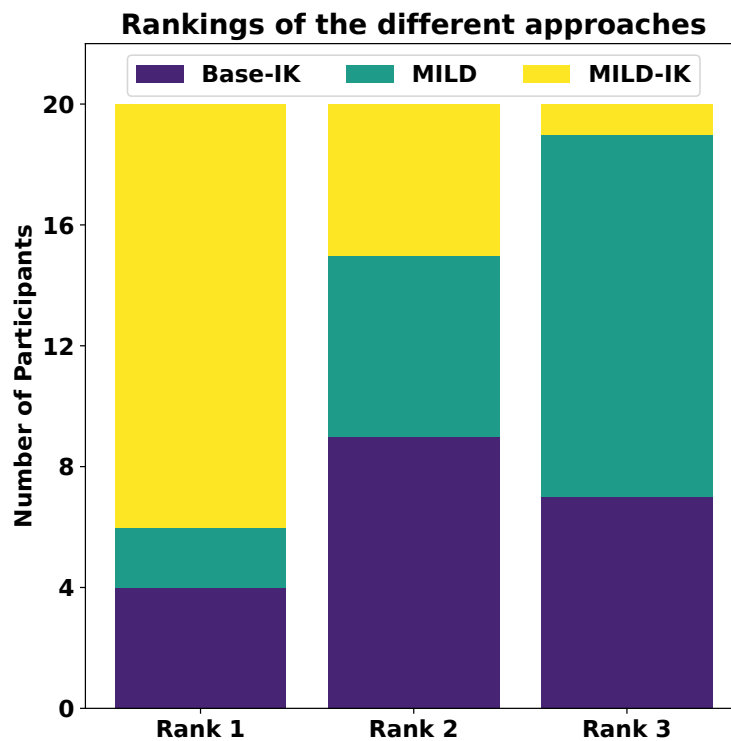


Figure 3.8.: Participant's ranking of the algorithms. (MILD-IK - yellow stars, MILD - green lines, Base-IK - purple dots). Most participants ranked MILD-IK first, better than both Base-IK and MILD.

Survey Item	Anova Results		Base-IK		MILD		MILD-IK		$\Delta\mu_{1,2}$	$\Delta\mu_{3,2}$	$\Delta\mu_{3,1}$
	$F_{2,38}$	p -value	Mean	SD	Mean	SD	Mean	SD			
Pleasantness	0.86	0.43	3.50	1.07	3.40	1.07	3.75	0.89	0.10	0.35	0.25
Excitement	1.32	0.28	3.85	0.85	3.45	1.02	3.75	0.89	0.40	0.30	-0.10
Human-likeness	4.89	0.01	2.60	1.20	2.45	1.16	3.25	1.22	0.15	0.80*	0.65*
Naturalness	3.46	0.04	2.95	1.24	2.55	1.20	3.25	1.18	0.40	0.70*	0.30
Friendliness	0.56	0.58	3.75	1.04	3.75	0.83	3.95	0.86	0.00	0.20	0.20
Comfort	2.79	0.07	3.35	1.06	3.35	0.96	3.85	0.96	0.00	0.50*	0.50
Timing	5.77	0.01	3.60	1.07	3.45	1.16	4.15	0.96	0.15	0.70**	0.55*
Accuracy	10.15	0.00	3.15	1.28	2.55	1.24	3.95	0.86	0.60	1.40	0.80
Annoyance	2.08	0.14	1.60	0.80	1.95	1.20	1.50	0.81	-0.35	-0.45*	-0.10
Aggression	0.77	0.47	1.15	0.48	1.30	0.64	1.25	0.62	-0.15	-0.05	0.10
Awkwardness	2.29	0.12	2.30	1.05	2.75	1.34	2.35	1.24	-0.45	-0.40	0.05
Scariness	0.59	0.56	1.15	0.48	1.25	0.54	1.20	0.51	-0.10	-0.05	0.05
Satisfaction	2.54	0.09	3.30	1.14	3.15	0.96	3.80	0.87	0.15	0.65*	0.50
Effortlessness between the two trials	4.15	0.02	3.35	1.19	3.40	1.16	4.05	1.24	-0.05	0.65*	0.70*

Table 3.7.: Results of a one-way Repeated Measures ANOVA for each of the survey items. Values less than 0.01 are reported as 0. The last three columns show the mean differences between the different algorithms (1 - Base-IK, 2 - MILD, 3 - MILD-IK) which were analyzed using paired sample t-tests (* - $p < 0.05$, ** - $p < 0.01$).

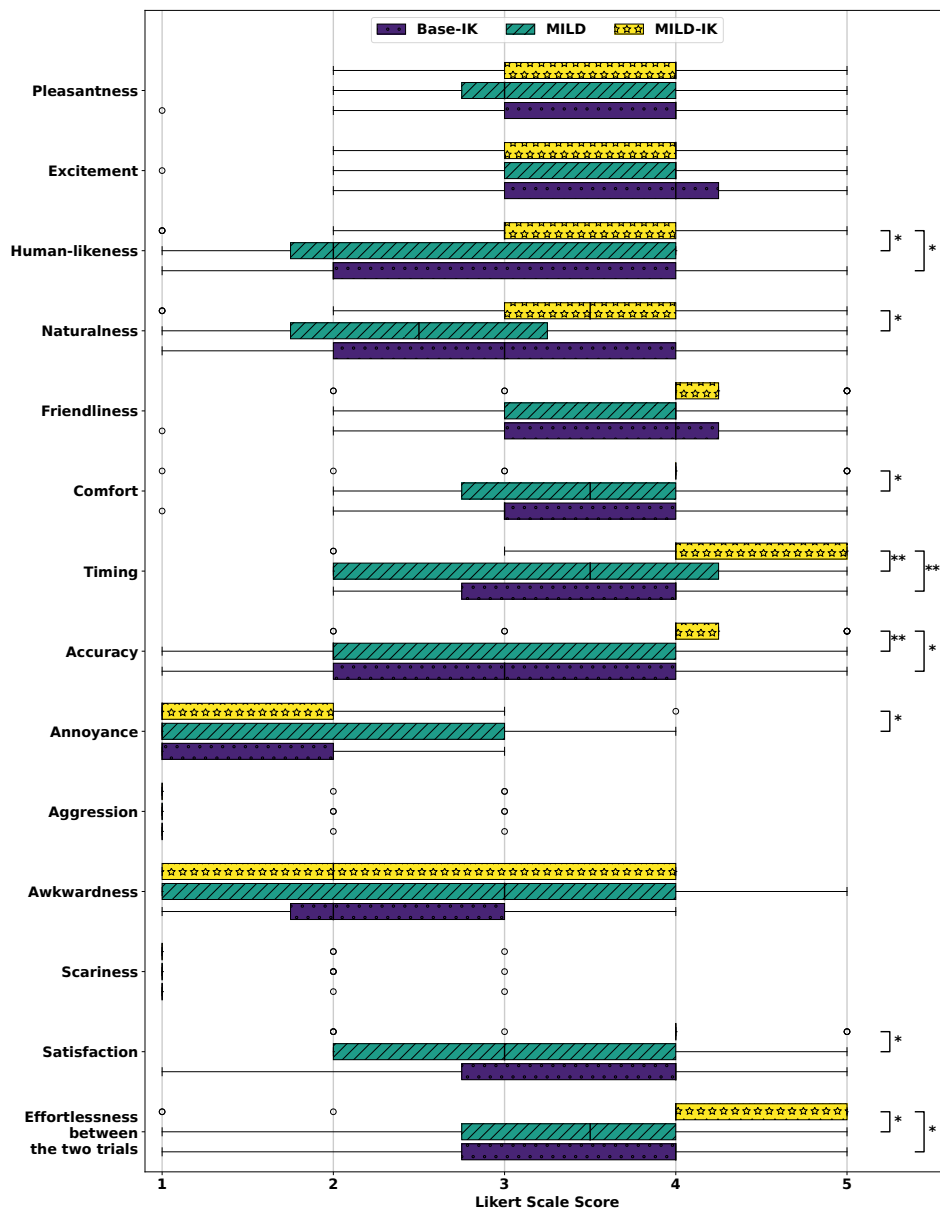


Figure 3.9.: Boxplot of the user study responses. (* - $p < 0.05$, ** - $p < 0.01$)

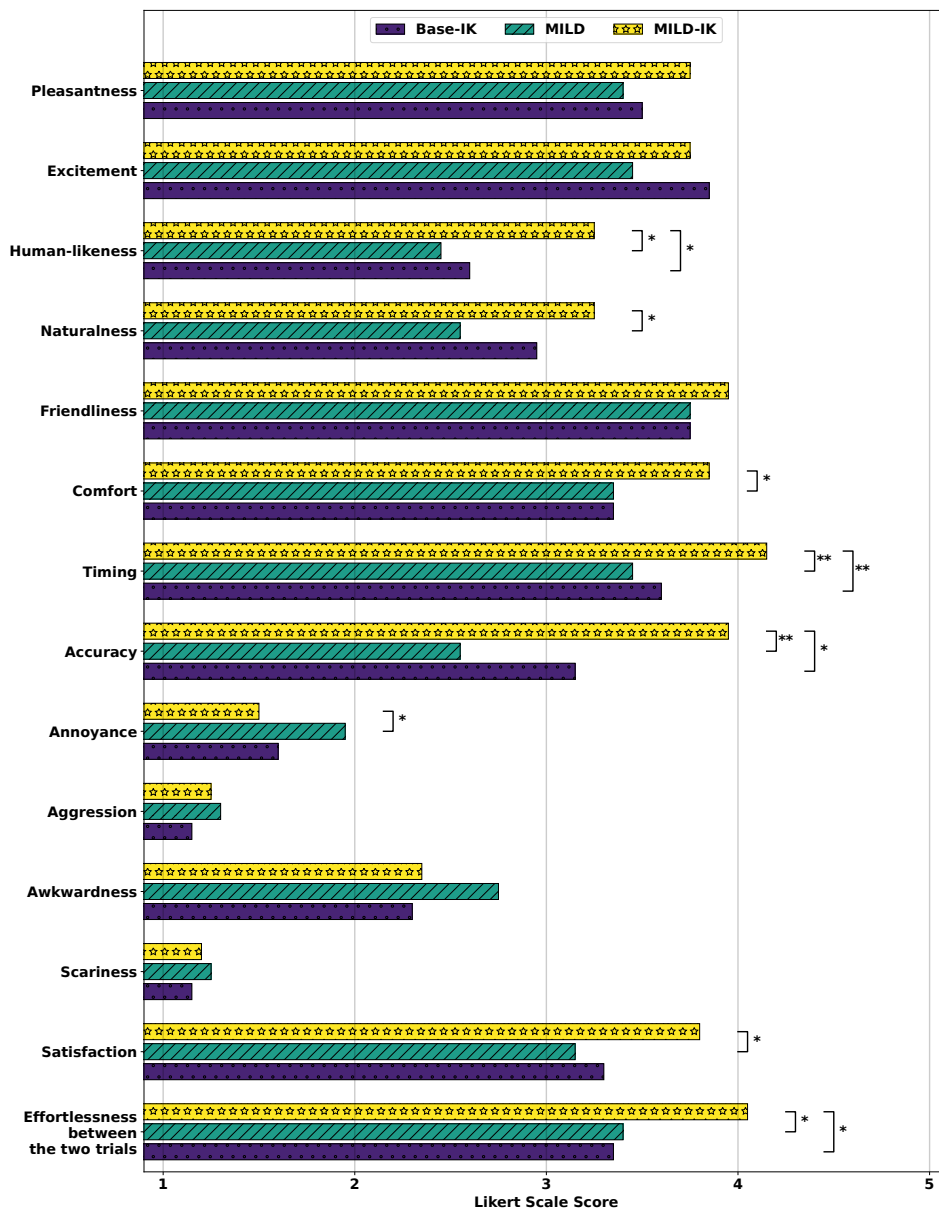


Figure 3.10.: Barplot of the user study responses. (* - $p < 0.05$, ** - $p < 0.01$)

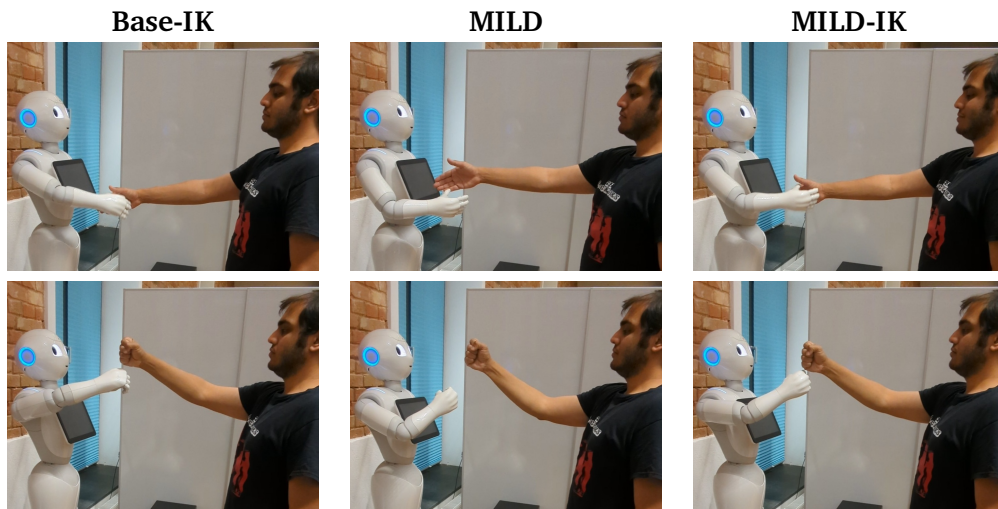


Figure 3.11.: Peculiarities of the different algorithms. The “Base-IK” approach reaches the human’s hand but awkwardly with the robot hand rotated inwards and the elbow pointing out. MILD maintains a human-like posture but falls short of reaching the partner’s hand due to the mismatch in motion retargeting. However, MILD-IK accurately reaches the human’s hand while maintaining a human-like posture.

Study Results

As seen in Figure 3.8, MILD-IK was by far ranked in the first place much more than MILD which was by far the least ranked, and Base-IK which was mostly ranked second. Just using MILD without IK was the least preferred among the algorithms, even though it was programmed keeping the interactivenss in mind, but due to the motion retargeting issues mentioned in Section 3.3.3, it is unable to reach the participant’s hand accurately, leading to a low acceptance by the participants which can further be seen in the results in Figures 3.9 and 3.10.

All approaches have similarly high levels of pleasantness, excitement, and friendliness and similarly low levels of annoyance, aggression, and scariness. This goes to show that the overall interaction scenario was well perceived by the participants and can additionally be attributed to an overall positive attitude toward robots.

Since the Base-IK approach does not have any prior over the IK solutions, it would lead to poses where the elbow is pointed outwards and the hand is turned sideways, which participants remarked was unnatural and awkward. Even though it was able to reach the human’s hand location, this unnatural pose of the robot hand prevented the Base-IK approach from being able to reach the human hand in a graspable manner, due to the

orientation of the end-effector. This is also reflected via a significantly lower accuracy rating for Base-IK and can further be seen in relatively lower (although non-significant) trends for naturalness, comfort, and satisfaction.

Participants found MILD awkward at times since the predicted motions would sometimes fall short of the partner’s hand and the robot would thereby not reach the participants’ hand correctly. This can especially be seen in the low accuracy that is given to MILD. This reachability issue gets mitigated with MILD-IK as it generates more natural poses. However, with MILD-IK, since the prediction of when to start the IK adaptation comes from the HMM, there would be a distinct moment when the robot hand would reach the human’s hand, which we attribute as a possible reason for a relatively higher awkwardness rating.

Overall, the high acceptability of MILD-IK is also reflected via significantly better ratings of the human-likeness, timing, and perceived accuracy. The effortlessness between two trials in an interaction was also rated significantly higher for MILD-IK showing that it was easier for participants to get habituated to the movements of the robot. MILD-IK also achieves a higher rating for timing, which is attributed to the ability of the HMM to generate the receding motion in a reactive manner, unlike Base-IK which would still try to reach the hand until the participant would go back to a neutral pose. Some of these peculiarities can be seen in Figure 3.11.

3.4.5. Bimanual Robot-to-Human Handovers

To further test the effectiveness of MILD on a more complex task, we showcase the ability of MILD to reactively predict motions for a Bimanual Robot in a Robot-to-Human handover scenario, wherein the robot has to hand over an object to the human partner. For this handover experiment, we use a Bimanual Franka Emika Panda robot setup, “Kobo” that runs a dual arm cartesian impedance controller to command each arm in the 3D task space. Therefore, we train MILD directly using task space trajectories from the Bimanual Handovers dataset in [103]. We re-scale the hand trajectories of the giver in the training data to fit the robot’s task space limits and train a model using the hand trajectories of both the giver and the receiver. Since we directly predict task space trajectories from Human-Human Interactions and do not need the inverse kinematics adaptation as done with Pepper, we use MILD v1 to learn the robot motions as the giver, with the human as the receiver. We find that MILD generates suitable response motions for various objects, an example of which can be seen in Figure 3.12. However, some failures still occur, mainly from the lack of incorporating object-specific information, such as the dimensions, object type, geometry, visual information, etc. While some failures can be mitigated with some post-processing optimization [62], this is currently out of the scope of our work.

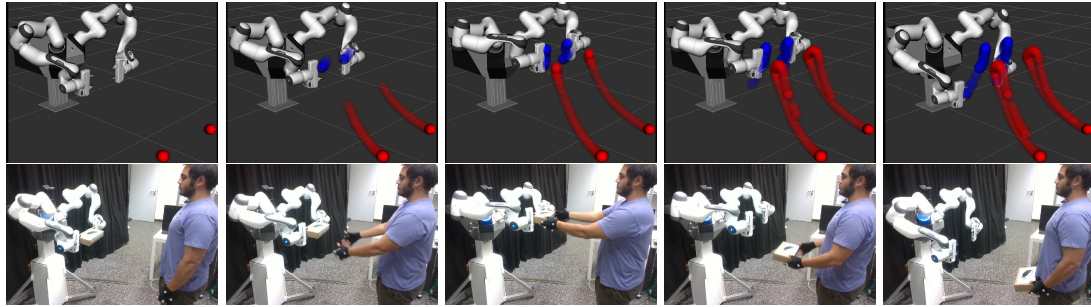


Figure 3.12.: Example of a Bimanual Robot-to-Human Handover Interaction generated by MILD. The top row shows the observed trajectory of the human in **red**, and the generated trajectory of the robot in **blue**. The progression of the interaction can be seen in the bottom row.

3.5. Conclusion and Future Work

In this chapter, we proposed a system for the adaptation and learning of real-world HRI from demonstrations of Human-Human Interactions (HHI). We first learned latent interaction dynamics from the HHI demonstrations in a modular manner using Hidden Markov Models (HMMs) which were then used to train models for HRI. We further improved the performance in HRI scenarios by exploring two strategies for incorporating the conditional distribution of the HMM into the training process for learning reactive motion generation which we found is an important aspect of achieving a competitive performance. We then showed the adaptation of the reactively generated robot trajectories during test time with Inverse Kinematics to enhance the task space performance of the predicted motions that are learned in joint space. For contact rich tasks like handshaking, we additionally showed how the HMM segment predictions can be used for stiffness modulation which improved the quality of the interaction. Through a user study, we found that our method is perceived better by human users in terms of human-likeness, timing, accuracy, and effortlessness. Our user study further validates the effectiveness of our approach in generalizing well to multiple users despite being trained on data from just two interaction partners. We additionally show the effectiveness of MILD in successfully generating Bimanual Robot-to-Human Handovers for different objects in a timely and reactive manner, which shows the usefulness of our approach even on more complex interaction scenarios.

3.5.1. Limitations

While the current approach generated acceptable and accurate response trajectories that, there are still some limitations to our approach which we highlight below. Currently, while training the VAE and HMM together as in the HHI scenario with the conditional loss as in the HRI scenario, we frequently encountered mode collapse of the HMM hidden states. This is why we resort to freezing the learned HMM and Human VAE for training on the HRI scenarios. This could stem from a bad approximation of the forward variable that does not accurately capture the progression of the hidden states but would rather favor just a few or a single state, thereby exacerbating the mode collapse. While this might be solved by propagating gradients through the forward variable, we found that this brings numerical issues arising from the recurrent nature of the forward variable computation, which leads to vanishing/exploding gradients (as in typical RNNs). Mitigating this issue would allow for a mathematically sound learning process.

Although we train the HMMs jointly over the latent spaces of both interacting agents, during testing, we use only the human observations. This can lead to inaccuracies in predicting the state of the interaction when solely using human observations as compared to the joint set of observations of both agents (as highlighted in the appendix). Incorporating the current state of the robot and the relative geometry between the human and the robot into the training process could improve the overall predictive performance of the network.

3.5.2. Future Work

On the practical side, coming to the interaction with the robot, we purposely left out auxiliary behaviors such as speech or gaze that make the robot more “alive” as our focus was to evaluate the different interaction algorithms. These auxiliary behaviors could help improve the overall perceived quality of the interaction. The influence of the inherent personality traits also affects the interaction [36]. Further research into quantifying such influences, for example, based on the mental states of the human partner [2], to adapt the robot actions and personalize the interactions leading can subsequently provide a more natural interaction and improve the perception of the robot. For handshaking, we had to incorporate our own mechanism to ensure compliance during the contact-based segments of the interaction. Alternatively, using an underlying Cartesian Impedance Controller can also help ensure compliance during the interaction. While some works have looked into such compliance in more static scenarios [27, 26], further research is required to adapt this to more contact-rich and dynamic tasks like handshaking for the Pepper robot.

From the learning side, the current bottleneck of our approach both in terms of training times and in achieving accurate predictions is the HMM. Currently, we train the

HMMs independently from the VAE in a separate step. In this regard, one could look at incorporating this into the variational framework in a more principled manner [7] or training the HMMs and VAEs in a more closely coupled manner by propagating the gradients of the expectation-maximization through the VAE could yield more suitable representations for the task at hand. Alternatively, we plan to look at going beyond HMMs by using neural variants for incorporating the multimodality, such as Mixture Density Networks [24] and additionally handle inputs from multiple sensors, which are important for functional HRI tasks in settings involving shared autonomy with a human. Additionally, incorporating a GAN-like discriminator [202] or diffusion-based architectures [138] could lead to performance improvements over a VAE, especially for more complex tasks.

4. MoVEInt: Mixture of Variational Experts for Learning Human-Robot Interactions from Demonstrations

In Chapter 3, we proposed an approach for learning latent interaction dynamics in a modular manner, however, as mentioned in Section 3.5.2, the main bottleneck of our approach is the HMM which is learned independently from the VAEs. One key takeaway though is that shared dynamics models are important for capturing the complexity and variability inherent in Human-Robot Interaction (HRI). Therefore, learning such shared dynamics models can enhance coordination and adaptability to enable successful reactive interactions with a human partner. Taking a step in the direction of better incorporating the multimodality into the training process, in this work, we propose “MoVEInt” an approach for learning a Mixture of Variational Experts for learning Human-Robot Interactions from demonstrations.

We propose a novel approach for learning a shared latent space representation for HRIs from demonstrations in a Mixture of Experts fashion for reactively generating robot actions from human observations. We train a Variational Autoencoder (VAE) to learn robot motions regularized using an informative latent space prior that captures the multimodality of the human observations via a Mixture Density Network (MDN). We show how our formulation derives from a Gaussian Mixture Regression formulation that is typically used approaches for learning HRI from demonstrations such as using an HMM/GMM for learning a joint distribution over the actions of the human and the robot. We further incorporate an additional regularization to prevent “mode collapse”, a common phenomenon when using latent space mixture models with VAEs. We find that our approach of using an informative MDN prior from human observations for a VAE generates more accurate robot motions compared to previous HMM-based or recurrent approaches of learning shared latent representations, which we validate on various HRI datasets involving interactions such as handshakes, fistbumps, waving, and handovers. Further experiments in a real-world human-to-robot handover scenario show the efficacy of our approach for generating successful interactions with four different human interaction partners.

4.1. Prologue

Ensuring a timely response while performing an interaction can enable a feeling of connectedness to one’s partner [119]. Therefore, learning to generate response motions for a robot in a timely manner for Human-Robot Interaction (HRI) is an important aspect of the interaction. One way to do so is by learning a shared representation space between the human and the robot [29, 86, 87, 158, 157, 200, 202]. An important aspect of such approaches, for learning HRI from demonstrations, is accurately capturing the multimodality of the underlying data to effectively capture the underlying skills and accurately generate response motions.

Vogt et al. [200] showed the use of shared latent spaces in an Imitation Learning (IL) approach by decomposing interactions into multiple segments in a lower dimensional space with an underlying Gaussian Mixture Model (GMM) and learning the sequence of key poses using a Hidden Markov Model (HMM) to define the progression of an interaction. In our previous work, MILD [158, 157], we explore learning a shared latent space model using a Variational Autoencoder (VAE) wherein we learn a joint distribution over the trajectories of both the human and the robot using an HMM with underlying Gaussian States to represent the multimodality of the demonstrations. Rather than extracting key poses as in [200], we generate the robot’s motion using Gaussian Mixture Regression (GMR) from the underlying HMM based on the human’s observations in a reactive manner. In doing so, we achieve better accuracy than using a recurrent representation of the shared latent dynamics [29].

Generative approaches have been used to jointly learn human and robot policies for collaborative tasks [202, 193, 178] and discover different underlying latent “strategies” of the human but only for a given task. In our work, we further explore how underlying latent strategies can be learned from different tasks in a dataset rather than just a single task by using a mixture distribution to predict different latent policies, which are then combined in a Mixture of Experts fashion. An example of this can be seen in Figure 4.1, where we show a handshake interaction with the Pepper robot. Trained on a dataset of different physical interactions like waving, handshakes, and fistbumps, we see the different policies that get predicted (the reconstructions of which are shown by the different colored arms of Pepper) which are then combined in the latent space yielding a suitable response motion.

To learn multiple latent policies and effectively combine them, we employ Mixture Density Networks (MDNs) [24] to capture the multimodality of the demonstrations. MDNs predict a mixture of Gaussians and the corresponding mixture coefficients yielding a multimodal prediction, rather than a unimodal distribution or a single output. MDN policy representations have been widely used in Imitation Learning and Reinforcement Learning (RL) on a variety of tasks such as autonomous driving [16, 105], entropy regularization

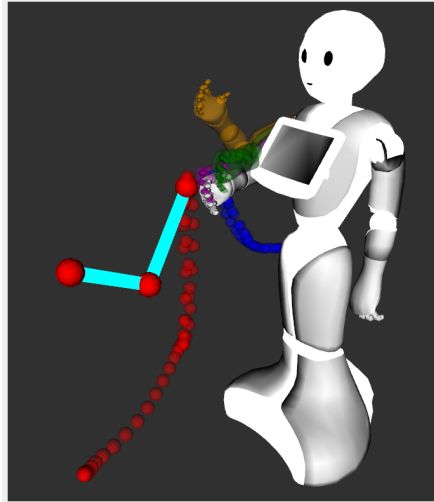


Figure 4.1.: Target poses generated in a reactive manner by the mixture of policies learned by MoVEInt for a Handshake interaction with the humanoid robot Pepper. MoVEInt generates multiple policies (shown in green, magenta, and orange) based on human observations which are combined to generate suitable robot motions.

for imitation learning [107], adapting to multiple goals [214], incorporating human intentions for robot tasks [192], predicting gaze behaviors for robot manipulation [95] or more generally as improved attention mechanisms [19, 3] or as forward models of the environment [64, 174, 210, 163]. With these key ideas for learning shared multimodal latent policy representations for HRI, the main contributions of this chapter are as follows.

We propose “MoVEInt”, a novel framework that employs a Mixture of Variational Experts for learning Human-Robot Interactions from demonstrations through a shared latent representation of a human and a robot. We learn latent space policies in a Mixture of Experts fashion via a Mixture Density Network (MDN) to encode the latent trajectory of a human partner, regularize the robot embeddings, and subsequently, predict the robot motions reactively. We show how our proposed formulation extends from Gaussian Mixture Regression (GMR) which is typically used in HMM-based approaches for learning HRI from demonstrations. We leverage the function approximation powers of Neural Networks to simplify the GMR formulation to a linear Mixture of Gaussian formulation.

Through our experiments, we see that our approach successfully captures the best of recurrent, multi-modal, and reactive representations for learning short-horizon Human-Robot Interactions from demonstrations.

We find that MoVEInt generates highly accurate robot behaviors without explicit action

labels, which is more natural as humans also internally infer what our interaction partner is doing and adapt to it without explicitly communicating the action being done. We validate our predictive performance on a variety of physical HRI scenarios such as handshakes, fistbumps, and robot-to-human handovers. We further demonstrate the efficacy of MoVEInt in a real-world interaction scenario for bimanual (dual-arm) robot-to-human handovers.

4.2. Mixture of Variational Experts for Learning Human-Robot Interactions from Demonstrations

In this section, we present MoVEInt, a novel framework that learns latent space policies in a Mixture of Experts fashion for modeling the shared dynamics of a human and a robot in HRI tasks. This process can be seen in Figure 4.2. We aim to model the dynamics of HRI tasks via shared latent representations of a human and a robot in a way that captures the multimodality of the demonstrations and subsequently predicts the robot’s motions in a reactive manner. To do so, we use an MDN that takes the human observations as input and predicts multiple latent policies, thereby enabling a multimodal output, and subsequently, the relative weights for each policy so that they can be effectively combined. For learning a shared latent representation between the human and the robot, we train a VAE over the robot motions and regularize the VAE with the predicted policy from the MDN, thereby learning the robot embeddings and the subsequent human-conditioned policy predictions in a cohesive manner.

We introduce MDNs in Section 4.2.1 and then motivate their use for HRI in Section 4.2.3 by showing the equivalence of MDNs with Gaussian Mixture Regression (GMR). We then explain learning the robot motion embeddings (Section 4.2.4), and then show how to train reactive policies for HRI (Section 4.2.5). We denote the human variables in **red** with the superscript h and the robot variables in **blue** with the superscript r .

4.2.1. Mixture Density Networks

Mixture Density Networks (MDNs) [24] is a probabilistic neural network architecture that encapsulates the representation powers of neural networks with the modular advantages that come with Gaussian Mixture Models (GMMs). MDNs parameterize a typical supervised regression problem of predicting an output distribution $p(\mathbf{y}|\mathbf{x}) = \mathcal{N}(\mathbf{y}|\boldsymbol{\mu}(\mathbf{x}), \boldsymbol{\sigma}^2(\mathbf{x}))$ as a multimodal distribution by predicting a set of means and variances $(\boldsymbol{\mu}_i(\mathbf{x}), \boldsymbol{\sigma}_i^2(\mathbf{x}))$ corresponding to the mixture distribution along with the mixture coefficients $\alpha_i(\mathbf{x})$ corresponding to the relative weights of the mixture distributions.

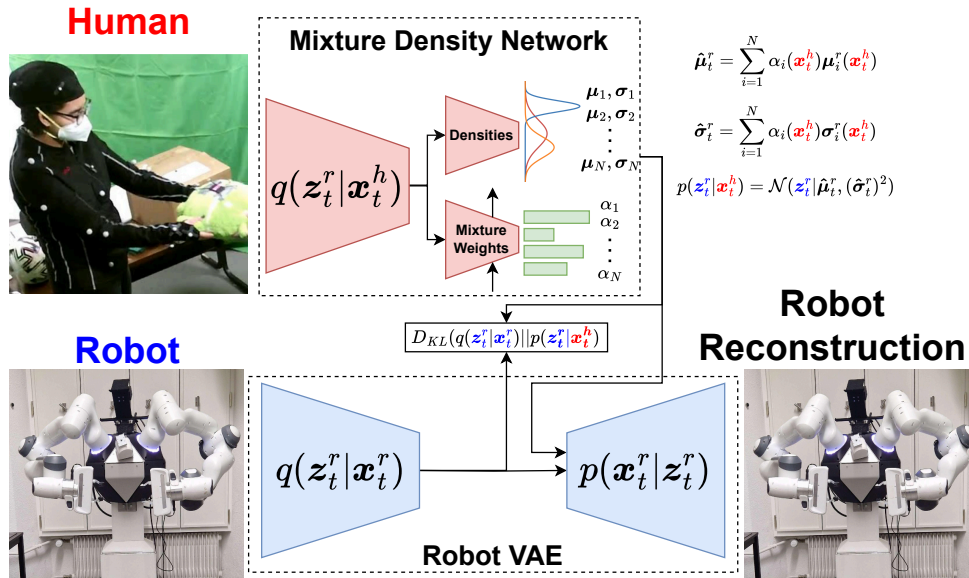


Figure 4.2.: Overview of our approach “MoVEInt”. We train a reactive policy using a Mixture Density Network (MDN) to predict latent space robot actions from human observations. The MDN policy is used not just for reactively generating the robot’s actions, but also to regularize a VAE that learns a latent representation of the robot’s actions. This regularization ensures that the learned robot representation matches the predicted MDN policy and also ensures that the robot VAE learns to decode samples from the MDN policy.

4.2.2. Overview

We aim to model the joint latent space dynamics in a Human-Robot Interaction scenario via a reactive policy. To do so, we train a Mixture Density Network that takes the human observations and the environment state (for eg. the object pose in a handover scenario) as input and predicts multiple policies corresponding to the different segments of the interaction and subsequently the relative weights for each segment. To enable a meaningful latent representation, we train a VAE over the Robot motions and regularize the VAE with the predicted policy from the MDN thereby learning the robot embeddings and the subsequent human-conditioned policy predictions in a cohesive manner. This process can be seen in Figure 4.2. For an easier understanding, we denote the human variables in **red** with the superscript h and the robot variables in **blue** with the superscript r .

4.2.3. GMR-based Interaction Dynamics with MDNs

Learning a joint distribution over the degrees of freedom of a human and a robot has been widely used in learning HRI from demonstrations [30, 52, 31, 116, 53, 158, 157]. With a joint distribution, GMR provides a mathematically sound formulation of predicting the conditional distribution of the robot actions. When using a Mixture of N Gaussian components $\{\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}$ that model a joint distribution of the Human and Robot trajectories, the distribution can be decomposed into the marginals for the human and the robot

$$\boldsymbol{\mu}_i = \begin{bmatrix} \boldsymbol{\mu}_i^h \\ \boldsymbol{\mu}_i^r \end{bmatrix}; \boldsymbol{\Sigma}_i = \begin{bmatrix} \boldsymbol{\Sigma}_i^{hh} & \boldsymbol{\Sigma}_i^{hr} \\ \boldsymbol{\Sigma}_i^{rh} & \boldsymbol{\Sigma}_i^{rr} \end{bmatrix} \quad (4.1)$$

In such a formulation, during test time, the robot motions can be generated reactively by conditioning the distribution using Gaussian Mixture Regression (GMR) [30, 187]

$$\begin{aligned} \mathbf{K}_i &= \boldsymbol{\Sigma}_i^{rh} (\boldsymbol{\Sigma}_i^{hh})^{-1} \\ \hat{\boldsymbol{\mu}}_i^r &= \boldsymbol{\mu}_i^r + \mathbf{K}_i (\mathbf{z}_t^h - \boldsymbol{\mu}_i^h) \\ \hat{\boldsymbol{\mu}}_t^r &= \sum_{i=1}^N \alpha_i(\mathbf{z}_t^h) \hat{\boldsymbol{\mu}}_i^r \\ \hat{\boldsymbol{\Sigma}}_i^r &= \boldsymbol{\Sigma}_i^{rr} - \mathbf{K}_i \boldsymbol{\Sigma}_i^{hr} + (\hat{\boldsymbol{\mu}}_i^r - \hat{\boldsymbol{\mu}}_t^r)(\hat{\boldsymbol{\mu}}_i^r - \hat{\boldsymbol{\mu}}_t^r)^T \\ \hat{\boldsymbol{\Sigma}}_t^r &= \sum_{i=1}^N \alpha_i(\mathbf{z}_t^h) \hat{\boldsymbol{\Sigma}}_i^r \\ p(\mathbf{z}_t^r | \mathbf{z}_t^h) &= \mathcal{N}(\mathbf{z}_t^r | \hat{\boldsymbol{\mu}}_t^r, \hat{\boldsymbol{\Sigma}}_t^r) \end{aligned} \quad (4.2)$$

where $\alpha_i(\mathbf{z}_t^h)$ is the relative weight of each component.

The covariance in the GMR Formulation in [30], which is derived from [187], can be obtained by simplifying Eq. 4.2 as

$$\begin{aligned}
\hat{\Sigma}_i^r &= \Sigma_i^{rr} - \mathbf{K}_i \Sigma_i^{hr} + (\hat{\boldsymbol{\mu}}_i^r - \hat{\boldsymbol{\mu}}_t^r)(\hat{\boldsymbol{\mu}}_i^r - \hat{\boldsymbol{\mu}}_t^r)^T \\
&= \Sigma_i^{rr} - \mathbf{K}_i \Sigma_i^{hr} + \hat{\boldsymbol{\mu}}_i^r (\hat{\boldsymbol{\mu}}_i^r)^T - \hat{\boldsymbol{\mu}}_i^r (\hat{\boldsymbol{\mu}}_t^r)^T \\
&\quad - \hat{\boldsymbol{\mu}}_t^r (\hat{\boldsymbol{\mu}}_i^r)^T + \hat{\boldsymbol{\mu}}_t^r (\hat{\boldsymbol{\mu}}_t^r)^T \\
\hat{\Sigma}_t^r &= \sum_{i=1}^N \alpha_i(\mathbf{z}_t^h) \hat{\Sigma}_i^r \\
&= \sum_{i=1}^N \alpha_i(\mathbf{z}_t^h) (\Sigma_i^{rr} - \mathbf{K}_i \Sigma_i^{hr}) \\
&\quad + \sum_{i=1}^N \alpha_i(\mathbf{z}_t^h) \hat{\boldsymbol{\mu}}_i^r (\hat{\boldsymbol{\mu}}_i^r)^T - \left(\sum_{i=1}^N \alpha_i(\mathbf{z}_t^h) \hat{\boldsymbol{\mu}}_i^r \right) (\hat{\boldsymbol{\mu}}_t^r)^T \\
&\quad - \hat{\boldsymbol{\mu}}_t^r \sum_{i=1}^N \alpha_i(\mathbf{z}_t^h) (\hat{\boldsymbol{\mu}}_i^r)^T + \hat{\boldsymbol{\mu}}_t^r (\hat{\boldsymbol{\mu}}_t^r)^T \sum_{i=1}^N \alpha_i(\mathbf{z}_t^h) \\
&= \sum_{i=1}^N \alpha_i(\mathbf{z}_t^h) (\Sigma_i^{rr} - \mathbf{K}_i \Sigma_i^{hr} + \hat{\boldsymbol{\mu}}_i^r (\hat{\boldsymbol{\mu}}_i^r)^T) - \hat{\boldsymbol{\mu}}_t^r (\hat{\boldsymbol{\mu}}_t^r)^T
\end{aligned} \tag{4.3}$$

since $\sum_{i=1}^N \alpha_i(\mathbf{z}_t^h) (\hat{\boldsymbol{\mu}}_i^r) = \hat{\boldsymbol{\mu}}_t^r$ and $\sum_{i=1}^N \alpha_i(\mathbf{z}_t^h) = 1$, therefore $-\hat{\boldsymbol{\mu}}_t^r \sum_{i=1}^N \alpha_i(\mathbf{z}_t^h) (\hat{\boldsymbol{\mu}}_i^r)^T + \hat{\boldsymbol{\mu}}_t^r (\hat{\boldsymbol{\mu}}_t^r)^T \sum_{i=1}^N \alpha_i(\mathbf{z}_t^h)$ becomes 0.

Revisiting Eq. 4.2, given the linear dependence of $\hat{\boldsymbol{\mu}}_i^r$ on \mathbf{z}_t^h , we can therefore consider $\hat{\boldsymbol{\mu}}_i^r$ as a direct output of a neural network. Although the covariance matrix has a quadratic relationship with \mathbf{z}_t^h , considering Neural Networks are powerful function approximations, we assume that our network can adequately approximate a diagonalized form of the covariance matrices $\hat{\Sigma}_i^r$ [120, 214]. To consider the temporal aspect of learning such trajectories from demonstrations, rather than using the Mixture Model coefficients, $\alpha_i(\mathbf{z}_t^h)$ can be approximated in a temporal manner using Hidden Markov Models (HMMs) [30] $\alpha_i(\mathbf{z}_t^h) = \mathcal{N}(\mathbf{z}_t^h; \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \sum_{j=1}^N \alpha_j(\mathbf{z}_{t-1}) \mathcal{T}_{j,i}$ where the parameters $(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i, \mathcal{T}_{j,i})$ are the means and covariances of the underlying Gaussian states and the transition probabilities between the states respectively.

Given that there exist parallels between HMMs and RNNs [18, 73, 40] and to capture the recurrence in predicting $\alpha_i(\mathbf{z}_t^h)$, we use a recurrent layer for predicting the mixing coefficients. Using the predictions of the mixture model parameters $(\boldsymbol{\mu}_i^r(\mathbf{x}_t^h), \boldsymbol{\sigma}_i^r(\mathbf{x}_t^h)^2)$ and coefficients $\alpha_i(\mathbf{x}_t^h)$ from the MDN ¹, we can re-write Eq. 4.2 as

¹Since the relation between \mathbf{x}_t^h and \mathbf{z}_t^h is deterministic, for ease of notation, we show the prediction of the MDN components with \mathbf{x}_t^h

$$\begin{aligned}
\hat{\boldsymbol{\mu}}_t^r &= \sum_{i=1}^N \alpha_i(\mathbf{x}_t^h) \boldsymbol{\mu}_i^r(\mathbf{x}_t^h) \\
(\hat{\boldsymbol{\sigma}}_t^r)^2 &= \sum_{i=1}^N \alpha_i(\mathbf{x}_t^h) \sigma_i^r(\mathbf{x}_t^h)^2 \\
p(\mathbf{z}_t^r | \mathbf{x}_t^h) &= \mathcal{N}(\mathbf{z}_t^r | \hat{\boldsymbol{\mu}}_t^r, (\hat{\boldsymbol{\sigma}}_t^r)^2)
\end{aligned} \tag{4.4}$$

which is then used as a prior for regularizing the VAE and training the decoder to reconstruct latent samples obtained after observing the human partner.

4.2.4. Robot Motion Embeddings

To learn a meaningful representation of the robot’s actions, we train a VAE to reconstruct the robot’s actions at each timestep. Typically, in VAEs, a standard normal distribution is used as the latent space prior $p(\mathbf{z}) = \mathcal{N}(\mathbf{0}, \mathbf{I})$. Rather than forcing an uninformative standard normal prior as in Eq. 3.3, we use the reactive policy predicted from the human observations by the MDN (Eq. 4.4) to regularize the VAE’s posterior $KL(q(\mathbf{z}_t^r | \mathbf{x}_t^r) || p(\mathbf{z}_t^r | \mathbf{x}_t^h))$, thereby learning a task-oriented latent space that is in line with the interaction dynamics. Our ELBO for training the robot VAE can be written as

$$\begin{aligned}
ELBO_t^r &= \mathbb{E}_{q(\mathbf{z}_t^r | \mathbf{x}_t^r)} [\log p(\mathbf{x}_t^r | \mathbf{z}_t^r)] \\
&\quad - \beta KL(q(\mathbf{z}_t^r | \mathbf{x}_t^r) || p(\mathbf{z}_t^r | \mathbf{x}_t^h))
\end{aligned} \tag{4.5}$$

Where β is a relative weight used to ensure numerical stability between the KL divergence term and the image reconstruction term [71].

4.2.5. Reactive Motion Generation

We aim to learn a policy for reactively generating the robot’s latent trajectory based on human observations $p(\mathbf{z}_t^r | \mathbf{x}_t^h)$. We do so in a Behavior Cloning Paradigm by maximizing the probability of the observed trajectories w.r.t. the current policy $\mathcal{L}_t^{BC} = -\mathbb{E}_{\mathbf{z}_t^r \sim p(\mathbf{z}_t^r | \mathbf{x}_t^h)} p(\mathbf{x}_t^r | \mathbf{z}_t^r)$ wherein we first draw samples from the current policy $\mathbf{z}_t^r \sim p(\mathbf{z}_t^r | \mathbf{x}_t^h)$ which we then reconstruct $p(\mathbf{x}_t^r | \mathbf{z}_t^r)$, thereby enabling the decoder to reconstruct latent samples obtained after observing the human, as done during test time.

However, as highlighted in [214], MDN policy representations are prone to mode collapses. Therefore, to ensure adequate separation between the modes so that we can learn a diverse range of actions, we employ a contrastive loss at each timestep.

The contrastive loss pushes the means of each mixture component further away, while maintaining temporal similarity by pushing embeddings that are closer in time nearer to each other. Further, as done in [214], we add entropy cost to ensure a balanced prediction of the mixture coefficients. Our separation loss can be written as

$$\begin{aligned}
\mathcal{L}_t^{sep} = & \underbrace{\sum_{i=1}^{N-1} \sum_{j=i+1}^N \exp(-\|\boldsymbol{\mu}_i^r(\mathbf{x}_t^h) - \boldsymbol{\mu}_j^r(\mathbf{x}_t^h)\|^2)}_{\text{separation of means}} \\
& + \underbrace{1 - \frac{1}{N} \sum_{i=1}^N \exp(-\|\boldsymbol{\mu}_i^r(\mathbf{x}_t^h) - \boldsymbol{\mu}_i^r(\mathbf{x}_{t-1}^h)\|^2)}_{\text{temporal closeness}} \\
& + \underbrace{\sum_{i=1}^N \alpha_i(\mathbf{x}_t^h) \ln \alpha_i(\mathbf{x}_t^h)}_{\text{entropy cost}}
\end{aligned} \tag{4.6}$$

Our final loss consists of the Behavior Cloning loss, the ELBO of the robot VAE, and the separation loss

$$\sum_{t=1}^T [\mathcal{L}_t^{BC} - ELBO_t^r + \beta \mathcal{L}_t^{sep}] \tag{4.7}$$

where β is the same KL weight factor used in Eq. 4.5. Our overall training procedure is shown in Algorithm 4.

Algorithm 4: Learning a Reactive Latent Policy for Human-Robot Interaction

Data: A set of human and robot trajectories $\mathbf{X} = \{\mathbf{X}^h, \mathbf{X}^r\}$

Result: MDN Policy Network, Robot VAE

while not converged **do**

for $\mathbf{x}_{1:T}^h, \mathbf{x}_{1:T}^r \in \mathbf{X}$ **do**

 Compute the MDN policy $p(\mathbf{z}_t^r | \mathbf{x}_t^h)$ (Eq. 4.4)

 Compute the robot VAE posterior $p(\mathbf{z}_t^r | \mathbf{x}_t^r)$ (Eq. 4.5)

 Reconstruct samples from $p(\mathbf{z}_t^r | \mathbf{x}_t^h)$ and $p(\mathbf{z}_t^r | \mathbf{x}_t^r)$

 Minimize the loss in Eq. 4.7 to update the network

end

end

During test time, given human observations \mathbf{x}_t^h , we compute the latent policy from the MDN $z_t^r = p(z_t^r | \mathbf{x}_t^h)$ which is then decoded to obtain the robot action $p(\mathbf{x}_t^r | z_t^r)$.

4.3. Experiments and Results

In this section, we first explain the datasets we use (Section 4.3.1), then we highlight the implementation details (Section 4.3.2), and finally present our results (Section 4.3.3).

4.3.1. Datasets

In addition to the datasets presented in Section 3.4.2, we additionally use the Human-Human Object Handovers Dataset [103], which is a dataset consisting of 12 pairs of participants performing object handovers with various objects. Each partner takes the role of the giver and the receiver, leading to 24 pairs of handover partners. The dataset consists of both Unimanual (single-arm) and Bimanual (dual-arm) handovers tracked using motion capture at 120Hz, which we downsample to 30Hz. We explore a Robot-to-Human handover scenario and use the giver’s observations corresponding to the robot and the receiver’s observations corresponding to the human.

Dataset	Downsampled FPS (Hz)	No. of Trajectories	
		Training	Testing
HHI, HRI-Pepper [29]	20	149	32
HRI-Yumi [29]	20	32	9
HHI, HRI-Pepper (NuiSI [157])	30	33	11
HHI-Handovers [103]	30	168	24

Table 4.1.: Statistics of the different datasets used.

4.3.2. Implementation Details

After downsampling the data to the frequencies mentioned in Table 4.1, we use a time window of observations as the input for a given timestep, similar to [29]. We use the 3D positions and the velocities (represented as position deltas) of the right arm joints (shoulder, elbow, and wrist), with the origin at the shoulder, leading to an input size of 90 dimensions (5x3x6: 5 timesteps, 3 joints, 6 dimensions) for a human partner. For the HHI-Handover scenario, we use both the left and right arm data (180 dimensions). For the HRI-Pepper and HRI-Yumi scenarios, we use a similar window of joint angles,

Dataset (units)	Action	MILD [157]	Bütepage et al. [29]	MoVEInt
HHI (Bütepage et al. [29]) (cm)	Hand Wave	0.788 ± 1.226	4.121 ± 2.252	0.448 ± 0.630
	Handshake	1.654 ± 1.549	1.181 ± 0.859	0.196 ± 0.153
	Rocket Fistbump	0.370 ± 0.682	0.544 ± 1.249	0.123 ± 0.175
	Parachute Fistbump	0.537 ± 0.579	0.977 ± 1.141	0.314 ± 0.348
HRI-Pepper (Bütepage et al. [29]) (rad)	Hand Wave	0.103 ± 0.103	0.664 ± 0.277	0.087 ± 0.089
	Handshake	0.056 ± 0.041	0.184 ± 0.141	0.015 ± 0.014
	Rocket Fistbump	0.018 ± 0.035	0.033 ± 0.045	0.007 ± 0.015
	Parachute Fistbump	0.088 ± 0.148	0.189 ± 0.196	0.048 ± 0.112
HRI-Yumi (Bütepage et al. [29]) (rad)	Hand Wave	1.033 ± 1.204	0.225 ± 0.302	0.147 ± 0.072
	Handshake	0.068 ± 0.052	0.133 ± 0.214	0.057 ± 0.044
	Rocket Fistbump	0.128 ± 0.071	0.147 ± 0.119	0.093 ± 0.045
	Parachute Fistbump	0.028 ± 0.034	0.181 ± 0.155	0.081 ± 0.082
HHI (NuiSI [157]) (cm)	Hand Wave	0.408 ± 0.538	3.168 ± 3.392	0.298 ± 0.274
	Handshake	0.311 ± 0.259	1.489 ± 3.327	0.149 ± 0.120
	Rocket Fistbump	1.142 ± 1.375	3.576 ± 3.082	0.673 ± 0.679
	Parachute Fistbump	0.453 ± 0.578	2.008 ± 2.024	0.291 ± 0.199
HRI-Pepper (NuiSI [157]) (rad)	Hand Wave	0.046 ± 0.059	0.057 ± 0.093	0.044 ± 0.048
	Handshake	0.020 ± 0.014	0.083 ± 0.075	0.011 ± 0.008
	Rocket Fistbump	0.077 ± 0.067	0.101 ± 0.086	0.045 ± 0.045
	Parachute Fistbump	0.022 ± 0.027	0.049 ± 0.040	0.017 ± 0.014
HHI-Handovers (Kshirsagar et al. [103]) (cm)	Unimanual	0.441 ± 0.280	1.133 ± 0.721	0.441 ± 0.221
	Bimanual	0.869 ± 0.964	0.990 ± 0.764	0.685 ± 0.643

Table 4.2.: Prediction MSE for robot trajectories after observing the human partner averaged over all joints and timesteps. Results for the HHI scenarios are in cm and for the HRI scenarios are in radians. (Lower is better)

leading to an input size of 20 dimensions (5x4) for the 4 joint angles of Pepper’s right arm and an input size of 35 dimensions (5x7) for the 7 joint angles of Yumi’s right arm. The reconstruction loss for the VAE and Behavior Cloning are calculated by decoding samples drawn from the respective distributions in a Monte Carlo fashion.

For the dataset from [29] and the NuiSI dataset [157], we use a VAE with 2 hidden layers each in the encoder and decoder with a dimensionality of (40, 20) and (20, 40) respectively, with Leaky ReLU activations and a 5D latent space. For the VAE posterior and the MDN outputs, we predict the mean and the log of the standard deviation (logstd). We add a regularization of 10^{-6} to the standard deviation. The MDN has a similar structure as the VAE encoder but predicts 3 different means and logstds. For the mixture coefficients α_i , to enable a recurrent nature of the predictions, the output of the MDN encoder is passed to a single-layer Gated Recurrent Unit (GRU) whose outputs are then passed through a linear layer followed by a softmax layer. For the NuiSI dataset, we initialize the model with the pre-trained weights trained on the dataset in [29]. For the Handover dataset [103], we use double the number of states in each layer, namely 80 and 40 hidden states and a 10D

latent space.

4.3.3. Reactive Motion Generation Results

To show the advantage of our approach, we compare MoVEInt with Bütepage et al. [29], who use an LSTM-based approach as a latent regularization for reactive motion generation, which is close to a unimodal version of our approach. Further, we compare MoVEInt to MILD [157] which uses HMMs to capture the multimodal latent dynamics of interactive tasks. The efficacy of MoVEInt can be seen via the low mean squared error of the predicted robot motions (Table 4.2).

We perform better than both MILD [157] and Bütepage et al. [29] on almost all interaction scenarios. We additionally want to highlight that on the HRI-Pepper scenarios, unlike MILD [157] and Bütepage et al. [29] where the pre-trained model from the HHI scenario is used, we train our model completely from scratch and still achieve better performance. Moreover, it is worth noting that both [29] and MILD are trained in a partially supervised manner using the interaction labels. In [29], a one-hot label denoting the interaction being performed is given as an input to the network for generalizing to different interactions, whereas in MILD, a separate HMM is trained for each interaction. In contrast, MoVEInt is trained on all the tasks in a given dataset without any labels in a purely unsupervised manner and still achieves competitive results on the different datasets.

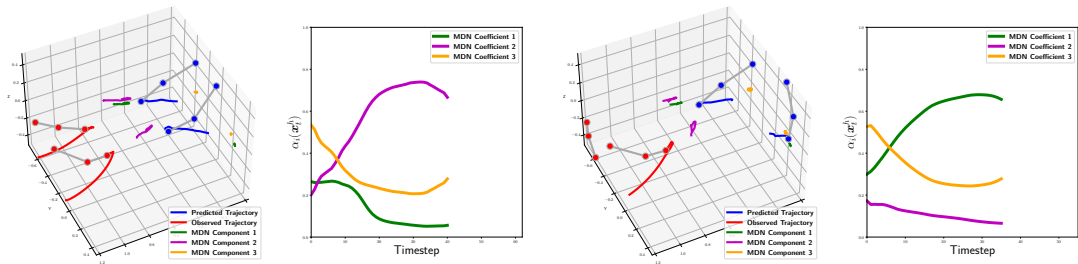
We additionally show some quantitative results of MoVEInt. We train a Handover model with just the hand trajectories whose predictions are used for reactive motion generation on a Bimanual Franka Emika Panda robot setup, “Kobo”, as shown in Figure 4.3. Some additional examples of the trajectories generated by MoVEInt for bimanual and unimanual handovers from the HHI-Handovers dataset [103] are shown in Figure 4.4a and 4.4b respectively. Since MoVEInt is trained on all the interactions in a corresponding dataset, which, when coupled with the separation loss, learns a diverse and widespread set of components that cover the various demonstrations, as can be seen by the reconstruction of the individual components. Combining the components in the latent space subsequently leads to an accurate and suitable motion generated reactively (shown in blue).

4.3.4. User study

To understand the effectiveness of our approach in the real world, we perform a feasibility study as a proof-of-concept with five users who perform bimanual robot-to-human handovers with the Kobo robot. We evaluate the ability of our approach to successfully generate a handover motion with three different objects where each participant interacts



Figure 4.3.: Sample Human-Robot Interactions generated with the reactive motions generated by MoVEInt for a Bimanual Handover scenario.



(a) Example of a generated Bimanual Handovers (b) Example of a generated Unimanual Handovers

Figure 4.4.: Sample trajectories generated by MoVEInt for the Bimanual and Unimanual Handovers in the HHI-Handovers dataset in [103]. The 3D plots show the reconstructed trajectories and the 2D plots show the corresponding progression of $\alpha_i(\mathbf{x}_t^h)$ for the different components of the MDN. In the 3D plots, the observed trajectory of the receiver is shown in red and the generated trajectory of the giver is shown in blue and the giver’s corresponding ground truth is shown in black. The reconstruction of the individual latent components of the MDN are shown in green, magenta, and orange. It can be seen that the learned components correspond to different parts of the task space. For example, green denotes the hand locations for a unimanual handover, magenta denotes the hand locations for a bimanual handover, and orange denotes the static hand locations for the starting and ending neutral poses. In the 2D plot, it can be seen how the coefficients for components corresponding to bimanual (magenta) and unimanual (green) get activated based on the interaction being performed, while the component corresponding to a neutral pose (orange) gets activated at the beginning of the interaction while both partners are static.

with the robot five times for each object (a total of 15 runs per participant). To maintain the object-centric nature of the interaction, we use a controller that tracks the mid-point of both end-effectors, thereby resembling tracking the object’s trajectory. This study was approved by the Ethics Commission at TU Darmstadt (Application EK 20/2023). For the study, we modify the Handovers dataset [103] such that the giver’s trajectories fall within the task space limits of the Kobo robot. We train a separate model to use only the receiver’s

hand trajectories and subsequently predict the end-effector trajectory for the robot. This is given to an Object-centric Inverse Kinematics controller that tracks the motion of the mid-point between the end-effectors.

As shown in Table 4.3, our approach can generate successful handover trajectories for different users and different objects. We observed some failure cases due to sudden jumps in the predicted robot motions resulting from inaccuracies in perceiving the human, which would overshoot the robot’s dynamic limits. However, this failure could be avoided by incorporating additional filters over the input and output data to MoVEInt. Some failures occurred because the object did not reach the exact vicinity of the human’s hand location. This failure could be avoided by incorporating object-related information such as the size or weight, allowing the robot to gauge better when the handover is executed. Sometimes, the robot would retreat before the human could grasp the object if sufficient time had passed. One reason for this hasty retreating behavior could be that the recurrent network’s hidden inputs overpower the observational input, causing the robot to follow the general motion of the handover seen during training and retreat. Such a failure could be mitigated by incorporating the robot state as part of the input.

ObjectUser ID	#1	#2	#3	#4	Total (per object)
Stool	5	4	5	5	19/20
Box	4	5	4	4	17/20
Bedsheet	4	5	3	3	15/20
Total (per user)	13/15	14/15	12/15	12/15	51/60

Table 4.3.: Number of successful handovers of each object by the Kobo robot to each user (total of 5 per object per user i.e. total of 20 per object and 15 per user).

4.4. Conclusion and Future Work

In this chapter, we presented “MoVEInt”, a novel deep generative Imitation Learning approach for learning Human-Robot Interaction from demonstrations in a Mixture of Experts fashion. We demonstrated the use of Mixture Density Networks (MDNs) as a multimodal policy representation in a shared latent space of the human and the robot. We showed how MoVEInt stems from the GMR-based formulation of predicting interaction dynamics used in HMM-based approaches to learning HRI. We showed how our MDN policy can predict multiple underlying policies and combine them to effectively generate response motions for the robot. We verified the efficacy of MoVEInt across a variety of

interactive tasks, where we found that MoVEInt mostly outperformed other baselines that either use explicitly modular representations like an HMM or simple recurrent policy representations. Our experimental evaluation showcases the versatility of MoVEInt, which effectively combines explicitly modular distributions with recurrent policy representations for learning interaction dynamics.

The main focus of this chapter is to explore the use of MDNs as a latent policy representation for simplistic short-horizon interactions like handshakes, handovers, etc., and show the feasibility of our approach in learning various interactive behaviors. One drawback is that we currently do not explicitly study how robust the approach is to unknown behaviors of the human, or how it performs on more complicated tasks.

Some next steps would be exploring how MoVEInt scales to longer horizon tasks such as bi-directional handovers, collaborative sequential manipulation, or proactive tasks where the robot leads the interaction. Some extensions for extrapolating to such tasks can be explored by using better Generative models [202, 178, 138] or adding explicit constraints for combining reactive motion generation and planning [69]. Alternatively, one could explore incorporating task-related constraints such as object information for handover grasps [42, 168], force information for enabling natural interactive behaviors [121, 93, 26], and ergonomic and safety constraints for a more user-friendly interaction [106, 170].

5. Conclusion

Physical non-verbal interactions play a key role in the perceived sociability of social robots. To this end, in this thesis, we proposed three main contributions towards studying and developing realistic and acceptable physical Human-Robot Interaction. First, we perform an extensive review of works on Human-Robot Handshaking to categorize the various existing works and understand what directions to pursue in the future. Based on this, we develop a framework for learning the underlying segments of such physical interactions such that they can be adapted to enable natural and acceptable interactions with a robot. We then improve upon this framework to incorporate reactive motion generation into the learning in a more principled manner, leading to the generation of more accurate interactive behaviors. In this chapter, we summarize the key contributions of this thesis and propose some directions for future work.

5.1. Summary

In Chapter 2, we performed a study of social robots and their interaction with humans, particularly focusing on handshaking, which is the key focus of this thesis. The exploration of handshaking as a fundamental interaction has led to significant insights and contributions in the field of Human-Robot Interaction (HRI). The research in this thesis began by conducting an extensive review of existing literature on Human-Robot Handshaking, which resulted in the development of a comprehensive taxonomy and the identification of various critical aspects influencing the design and implementation of robotic handshaking behaviors.

One of the key focus for realizing seamless robotic handshakes is the development of an end-to-end handshaking behavior, where the robot reacts in a coordinated and synchronized way to the human's motion, reaches the human hand accurately, and performs a compliant shaking motion followed by a timely termination of the interaction. Like handshaking, other physically interactive behaviors, like fistbumps or handovers, can also be broken down into similar segments which then need to be sequenced to realize an accurate interaction with a human. To this end, we develop a framework that

learns such underlying segments from demonstrations of interactions and subsequently generates corresponding robot motions in a reactive manner. In Chapter 3, we develop “MILD”, a framework for learning Multimodal Interactive Latent Dynamics aimed at learning real-world Human-Robot Interaction (HRI) from demonstrations. We focused on understanding the underlying dynamics of interactions from human demonstrations as a joint distribution using Hidden Markov Models (HMMs) in the latent space of a Variational Autoencoder (VAE). The dynamics learned from human demonstrations were subsequently utilized to train models for HRI, where we aimed to improve performance in various HRI scenarios. To achieve this, we explored two strategies for integrating the HMM’s conditional distribution into the training process, emphasizing the significance of reactive motion generation in achieving competitive performance. During testing, we demonstrated how reactively generated robot trajectories could be refined using Inverse Kinematics, which is particularly beneficial for tasks involving physical contact, such as handshaking. Our study also highlighted the effective use of HMM segment predictions for stiffness modulation, significantly enhancing the quality of the interaction. This was validated via a user study which confirmed the higher acceptance of our approach by human users in terms of human-likeness, timing, accuracy, and effortlessness. Moreover, the user study demonstrated the ability of our approach to generalize effectively across multiple users, despite being trained on data from only two interaction partners.

In Chapter 4, we further improve on the reactive motion generation capabilities by incorporating the HMM’s conditioning capabilities into the learning process in a more principled manner. We proposed “MoVEInt”, a novel deep generative Imitation Learning approach for learning a Mixture of Variational Experts for learning Human-Robot Interactions from demonstrations. We integrated Recurrent Mixture Density Networks (MDNs) as a latent policy representation that stems from a GMR-based formulation of the HMM conditioning for reactively generating robot motions based on the observed human partner. On a variety of different interactive tasks, we verified the effectiveness of MoVEInt where we found that MoVEInt largely outperformed other baselines that either use explicitly modular representations like an HMM or simple recurrent policy representations, showcasing the versatility of MoVEInt which combines the best of both worlds for learning interaction dynamics.

The thesis has significantly contributed to the understanding of the dynamics of physical interactions, emphasizing the importance of synchronized and adaptable motions in achieving natural and socially acceptable interactions between humans and social robots. The work done in this thesis contributed to the areas of Physical Human-Robot Interaction, Robot Learning for HRI, and Representation Learning by successfully developing and validating these frameworks through experimental evaluations with human participants. These findings and methodologies not only advance the current understanding of HRI

but also provide a solid foundation for future research in developing socially acceptable responsive robotic behaviors.

5.2. Future work

In this section, we outline some potential areas of future work based on the research done in this thesis. We highlight some key areas for improvement along with some avenues of extensions for each of the chapters in this thesis and then provide some overarching ideas in a broader context

Exploring the Role of Haptic Feedback

Haptic Feedback plays an important role in perceiving handshakes, as highlighted in Chapter 2. However, doing so requires specialized hardware that is capable of sensing the nuances of an intricate interaction like handshaking. Assuming the hardware barrier can be crossed, there is a larger need for developing datasets and frameworks to be able to learn how the emotional aspects of an interaction like a handshake can be learned by tactile feedback which can improve the social perception of a robot equipped with such a skill. One straightforward manner would be to apply the frameworks proposed in this thesis to learn the latent dynamics based on force interactions which can therefore be used to reactively control the robot grasp and subsequently be optimized with users' preferences[2]. While such datasets are not readily available for Handshaking, Handovers have been explored well in this regard [93, 94, 121], which can be a starting point for such a line of work. Moreover, in this thesis, the main focus of the developed frameworks was for reactive motion generation. However, enabling natural physical interactions requires a certain level of impedance control functionality to make the robot more adaptive and compliant with the human user [92]. The ability to learn such compliance from demonstrations would be truly beneficial for improving Human-Robot Interaction and Human-Robot Collaboration through the haptics of the interaction. This coupled with auxiliary behaviors like speech and gaze can help improve the overall perception of the interaction.

Extrapolating to Complex Collaborative Tasks

Currently, we found that using a Behavior Cloning approach for the frameworks generated an acceptable and accurate interaction given the relatively simplistic nature of the tasks. However, further research and experimentation are required to see how well the proposed

frameworks fare for more complex tasks such as sequential collaboration tasks, long-horizon planning, etc. In such cases, using an Adversarial approach like Co-GAIL [202] could enable learning diverse policies for more complex tasks, or for more proactive policies where the robot should lead the interaction.

Bibliography

- [1] 3DiVi. *Nuitrack*. <https://nuitrack.com/>. [Online; accessed 19-Oct-2023].
- [2] N. Abdulazeem and Y. Hu. “Human Factors Considerations for Quantifiable Human States in Physical Human-Robot Interaction: A Literature Review”. In: *Sensors* (2023).
- [3] P. Abolghasemi, A. Mazaheri, M. Shah, and L. Boloni. “Pay attention!-robustifying a deep visuomotor policy through task-focused visual attention”. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2019.
- [4] A. Ajoudani, A. M. Zanchettin, S. Ivaldi, A. Albu-Schäffer, K. Kosuge, and O. Khatib. “Progress and prospects of the human–robot collaboration”. In: *Autonomous Robots* (2018).
- [5] M. Ammi, V. Demulier, S. Caillou, Y. Gaffary, Y. Tsalamlal, J.-C. Martin, and A. Tapus. “Haptic human-robot affective interaction in a handshaking social protocol”. In: *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*. 2015.
- [6] H. B. Amor, G. Neumann, S. Kamthe, O. Kroemer, and J. Peters. “Interaction primitives for human-robot cooperation tasks”. In: *2014 IEEE international conference on robotics and automation (ICRA)*. IEEE. 2014.
- [7] O. Arenz, P. Dahlinger, Z. Ye, M. Volpp, and G. Neumann. “A Unified Perspective on Natural Gradient Variational Inference with Gaussian Mixture Models”. In: *Transactions on Machine Learning Research* (2023).
- [8] M. Arns, T. Laliberté, and C. Gosselin. “Design, control and experimental validation of a haptic robotic hand performing human-robot handshake with human-like agility”. In: *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2017, pp. 4626–4633.
- [9] M. A. Artem, K. M. Viacheslav, B. P. Volodymyr, and H. Patrick. “Physical human–robot interaction in the handshaking case: learning of rhythmicity using oscillators neurons”. In: *IFAC Proceedings Volumes* (2013).

-
-
- [10] J. Åström. “Introductory greeting behaviour: A laboratory investigation of approaching and closing salutation phases”. In: *Perceptual and Motor Skills* (1994).
- [11] J. Åström and L.-H. Thorell. “Greeting behaviour and psychogenic need: Interviews on experiences of therapists, clergymen, and car salesmen”. In: *Perceptual and motor skills* (1996).
- [12] J. Avelino, F. Correia, J. Catarino, P. Ribeiro, P. Moreno, A. Bernardino, and A. Paiva. “The power of a hand-shake in human-robot interactions”. In: *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018.
- [13] J. Avelino, T. Paulino, C. Cardoso, P. Moreno, and A. Bernardino. “Human-aware natural handshaking using tactile sensors for vizzy a social robot”. In: *Workshop on Behavior Adaptation, Interaction and Learning for Assistive Robotics at RO-MAN*. 2017.
- [14] J. Avelino, T. Paulino, C. Cardoso, R. Nunes, P. Moreno, and A. Bernardino. “Towards natural handshakes for social robots: human-aware hand grasps using tactile sensors”. In: *Paladyn, Journal of Behavioral Robotics* (2018).
- [15] G. Avraham, I. Nisky, H. L. Fernandes, D. E. Acuna, K. P. Kording, G. E. Loeb, and A. Karniel. “Toward perceiving robots as humans: Three handshake models face the turing-like handshake test”. In: *IEEE Transactions on Haptics* 5.3 (2012), pp. 196–207.
- [16] A. Baheri. “Safe reinforcement learning with mixture density network, with application to autonomous driving”. In: *Results in Control and Optimization* (2022).
- [17] C. Bartneck, D. Kulić, E. Croft, and S. Zoghbi. “Godspeed Questionnaire Series”. In: *International Journal of Social Robotics* (2008).
- [18] M. Baucum, A. Khojandi, and T. Papamarkou. “Hidden Markov models as recurrent neural networks: An application to Alzheimer’s disease”. In: *IEEE International Conference on Bioinformatics and Bioengineering (BIBE)*. 2021.
- [19] L. Bazzani, H. Larochelle, and L. Torresani. “Recurrent Mixture Density Network for Spatiotemporal Visual Attention”. In: *International Conference on Learning Representations (ICLR)*. 2016.
- [20] J. Beaudoin, T. Laliberté, and C. Gosselin. “Haptic Interface for Handshake Emulation”. In: *IEEE Robotics and Automation Letters* (2019).

-
-
- [21] P. Becker, H. Pandya, G. Gebhardt, C. Zhao, C. J. Taylor, and G. Neumann. “Recurrent kalman networks: Factorized inference in high-dimensional deep feature spaces”. In: *International Conference on Machine Learning*. PMLR. 2019, pp. 544–552.
- [22] F. J. Bernieri and K. N. Petty. “The influence of handshakes on first impression accuracy”. In: *Social Influence* (2011).
- [23] C. Bevan and D. S. Fraser. “Shaking hands and cooperation in tele-present human-robot negotiation”. In: *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE. 2015.
- [24] C. M. Bishop. “Mixture density networks”. In: (1994).
- [25] S. Bitzer and S. Vijayakumar. “Latent spaces for dynamic movement primitives”. In: *IEEE-RAS International Conference on Humanoid Robots*. 2009.
- [26] A. Bolotnikova, S. Courtois, and A. Kheddar. “Compliant robot motion regulated via proprioceptive sensor based contact observer”. In: *IEEE-RAS International Conference on Humanoid Robots (Humanoids)*. 2018.
- [27] A. Bolotnikova, S. Courtois, and A. Kheddar. “Contact observer for humanoid robot pepper based on tracking joint position discrepancies”. In: *IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. 2018.
- [28] R. A. Bradley and M. E. Terry. “Rank analysis of incomplete block designs: I. The method of paired comparisons”. In: *Biometrika* (1952).
- [29] J. Bütepage, A. Ghadirzadeh, Ö. Ö. Karadag, M. Björkman, and D. Kragic. “Imitating by Generating: Deep Generative Models for Imitation of Interactive Tasks”. In: *Frontiers in Robotics and AI* 7 (2020), p. 47.
- [30] S. Calinon. “A tutorial on task-parameterized movement learning and retrieval”. In: *Intelligent Service Robotics* (2016).
- [31] S. Calinon, P. Evrard, E. Gribovskaya, A. Billard, and A. Kheddar. “Learning collaborative manipulation tasks by demonstration using a haptic interface”. In: *International Conference on Advanced Robotics (ICAR)*. 2009.
- [32] J. Campbell and H. B. Amor. “Bayesian interaction primitives: A slam approach to human-robot interaction”. In: *Conference on Robot Learning*. 2017, pp. 379–387.
- [33] J. Campbell, A. Hitzmann, S. Stepputtis, S. Ikemoto, K. Hosoda, and H. B. Amor. “Learning Interactive Behaviors for Musculoskeletal Robots Using Bayesian Interaction Primitives”. In: *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2019.

-
-
- [34] P. Capdepuy, S. Bock, W. Benyaala, and J. Laplace. “Improving Human-Robot Physical Interaction with Inverse Kinematics Learning”. In: *International Conference on Social Robotics (ICSR)*. Springer. 2015.
- [35] T. Chaminade, D. W. Franklin, E. Oztop, and G. Cheng. “Motor interference between humans and humanoid robots: Effect of biological and artificial motion”. In: *Proceedings. The 4th International Conference on Development and Learning, 2005*. IEEE. 2005, pp. 96–101.
- [36] W. F. Chaplin, J. B. Phillips, J. D. Brown, N. R. Clanton, and J. L. Stein. “Handshaking, gender, personality, and first impressions.” In: *Journal of personality and social psychology* 79.1 (2000), p. 110.
- [37] M. Chaverroche, A. Malaisé, F. Colas, F. Charpillet, and S. Ivaldi. “A Variational Time Series Feature Extractor for Action Prediction”. In: *arXiv preprint arXiv:1807.02350* (2018).
- [38] N. Chen, J. Bayer, S. Urban, and P. Van Der Smagt. “Efficient movement representation by embedding dynamic movement primitives in deep autoencoders”. In: *IEEE-RAS International Conference on Humanoid Robots (Humanoids)*. 2015.
- [39] N. Chen, M. Karl, and P. Van Der Smagt. “Dynamic movement primitives in latent space of time-dependent variational autoencoders”. In: *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*. IEEE. 2016, pp. 629–636.
- [40] J. Chiu and A. M. Rush. “Scaling Hidden Markov Language Models”. In: *Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2020.
- [41] S. Christen, S. Stevsic, and O. Hilliges. “Guided Deep Reinforcement Learning of Control Policies for Dexterous Human-Robot Interaction”. In: *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2019.
- [42] S. Christen, W. Yang, C. Pérez-D’Arpino, O. Hilliges, D. Fox, and Y.-W. Chao. “Learning Human-to-Robot Handovers from Point Clouds”. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2023.
- [43] Y. Chua, K. P. Tee, and R. Yan. “Human-robot motion synchronization using reactive and predictive controllers”. In: *2010 IEEE International Conference on Robotics and Biomimetics*. IEEE. 2010.
- [44] J. Chung, K. Kastner, L. Dinh, K. Goel, A. C. Courville, and Y. Bengio. “A recurrent latent variable model for sequential data”. In: *Advances in neural information processing systems*. 2015, pp. 2980–2988.

-
-
- [45] A. Colomé, G. Neumann, J. Peters, and C. Torras. “Dimensionality reduction for probabilistic movement primitives”. In: *2014 IEEE-RAS International Conference on Humanoid Robots*. IEEE. 2014, pp. 794–800.
- [46] H. Dai, B. Dai, Y.-M. Zhang, S. Li, and L. Song. “Recurrent hidden semi-markov model”. In: *International Conference on Learning Representations (ICLR)*. 2016.
- [47] K. Dai, Y. Liu, M. Okui, R. Nishihama, and T. Nakamura. “Research of human-robot handshakes under variable stiffness conditions”. In: *2019 IEEE 4th International Conference on Advanced Robotics and Mechatronics (ICARM)*. IEEE. 2019.
- [48] K. Dai, Y. Liu, M. Okui, Y. Yamada, and T. Nakamura. “Variable viscoelasticity handshake manipulator for physical human–robot interaction using artificial muscle and MR brake”. In: *Smart Materials and Structures* 28.6 (2019), p. 064002.
- [49] A. P. Dempster, N. M. Laird, and D. B. Rubin. “Maximum likelihood from incomplete data via the EM algorithm”. In: *Journal of the royal statistical society: series B (methodological)* (1977).
- [50] O. Dermy, M. Chaveroche, F. Colas, F. Charpillet, and S. Ivaldi. “Prediction of Human Whole-Body Movements with AE-ProMPs”. In: *2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids)*. IEEE. 2018, pp. 572–579.
- [51] B. R. Duffy. “Anthropomorphism and the social robot”. In: *Robotics and autonomous systems* 42.3-4 (2003), pp. 177–190.
- [52] P. Evrard, E. Gribovskaya, S. Calinon, A. Billard, and A. Kheddar. “Teaching physical collaborative tasks: object-lifting case study with a humanoid”. In: *IEEE-RAS International Conference on Humanoid Robots (Humanoids)*. 2009.
- [53] M. Ewerton, G. Neumann, R. Lioutikov, H. B. Amor, J. Peters, and G. Maeda. “Learning multiple collaborative tasks with a mixture of interaction primitives”. In: *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2015, pp. 1535–1542.
- [54] O. Fabius, J. R. van Amersfoort, and D. P. Kingma. “Variational Recurrent Auto-Encoders”. In: *ICLR (Workshop)*. 2015.
- [55] M. Falahi, T. A. Shangari, A. Sheikhsafari, S. Gharghabi, A. Ahmadi, and S. S. Ghidary. “Adaptive handshaking between humans and robots, using imitation: Based on gender-detection and person recognition”. In: *2014 Second RSI/ISM International Conference on Robotics and Mechatronics (ICRoM)*. IEEE. 2014.

-
-
- [56] D. Feil-Seifer and M. J. Mataric. “Defining socially assistive robotics”. In: *9th International Conference on Rehabilitation Robotics, 2005. ICORR 2005*. IEEE, 2005.
- [57] L. Fritsche, F. Unverzag, J. Peters, and R. Calandra. “First-person tele-operation of a humanoid robot”. In: *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*. IEEE, 2015, pp. 997–1002.
- [58] A. Gallace and C. Spence. “The science of interpersonal touch: an overview”. In: *Neuroscience & Biobehavioral Reviews* 34.2 (2010), pp. 246–259.
- [59] V. Garg, A. Mukherjee, and B. Rajaram. “Classifying human-robot interaction using handshake data”. In: *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2017.
- [60] E. Giannopoulos, Z. Wang, A. Peer, M. Buss, and M. Slater. “Comparison of people’s responses to real and virtual handshakes within a virtual environment”. In: *Brain research bulletin* (2011).
- [61] X. Glorot and Y. Bengio. “Understanding the difficulty of training deep feedforward neural networks”. In: *Proceedings of the thirteenth international conference on artificial intelligence and statistics*. JMLR Workshop and Conference Proceedings, 2010, pp. 249–256.
- [62] Y. Göksu, A. De Almeida Correia, V. Prasad, A. Kshirsagar, D. Koert, J. Peters, and G. Chalvatzaki. “Kinematically Constrained Human-like Bimanual Robot-to-Human Handovers”. In: *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*. 2024, pp. 497–501.
- [63] S. Gomez-Gonzalez, G. Neumann, B. Schölkopf, and J. Peters. “Adaptation and robust learning of probabilistic movement primitives”. In: *IEEE Transactions on Robotics* 36.2 (2020), pp. 366–379.
- [64] D. Ha and J. Schmidhuber. “Recurrent world models facilitate policy evolution”. In: *Advances in Neural Information Processing Systems (NeurIPS)* (2018).
- [65] F. Hahne, V. Prasad, A. Kshirsagar, D. Koert, R. M. Stock-Homburg, J. Peters, and G. Chalvatzaki. “Transition State Clustering for Interaction Segmentation and Learning”. In: *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*. 2024, pp. 512–516.
- [66] P. M. Hall and D. A. S. Hall. “The handshake as interaction”. In: *Semiotica* (1983).
- [67] J. Han, M. R. Min, L. Han, L. E. Li, and X. Zhang. “Disentangled Recurrent Wasserstein Autoencoder”. In: *International Conference on Learning Representations (ICLR)*. 2021.

-
-
- [68] M.-J. Han, C.-H. Lin, and K.-T. Song. “Robotic emotional expression generation based on mood transition and personality model”. In: *IEEE transactions on cybernetics* 43.4 (2012), pp. 1290–1303.
- [69] K. Hansel, J. Urain, J. Peters, and G. Chalvatzaki. “Hierarchical policy blending as inference for reactive robot control”. In: *IEEE International Conference on Robotics and Automation (ICRA)*. 2023.
- [70] M. J. Hertenstein, J. M. Verkamp, A. M. Kerestes, and R. M. Holmes. “The communicative functions of touch in humans, nonhuman primates, and rats: a review and synthesis of the empirical research”. In: *Genetic, social, and general psychology monographs* (2006).
- [71] I. Higgins, L. Matthey, A. Pal, C. Burgess, X. Glorot, M. Botvinick, S. Mohamed, and A. Lerchner. “beta-vae: Learning basic visual concepts with a constrained variational framework”. In: *International Conference on Learning Representations (ICLR)*. 2016.
- [72] N. J. Higham. “Computing a nearest symmetric positive semidefinite matrix”. In: *Linear algebra and its applications* 103 (1988), pp. 103–118.
- [73] T. Hiraoka, S. Takase, K. Uchiumi, A. Keyaki, and N. Okazaki. “Recurrent neural hidden Markov model for high-order transition”. In: *Transactions on Asian and Low-Resource Language Information Processing* (2021).
- [74] E. S. Ho, T. Komura, and C.-L. Tai. “Spatial relationship preserving character motion adaptation”. In: *ACM Special Interest Group on Computer Graphics (SIGGRAPH)*. 2010.
- [75] K. S. Hone and R. Graham. “Towards a tool for the subjective assessment of speech system interfaces (SASSI)”. In: *Natural Language Engineering* (2000).
- [76] S. L. Hooper. “Central pattern generators”. In: *e LS* (2001).
- [77] E. Hopf. “Abzweigung einer periodischen Lösung von einer stationären Lösung eines Differentialsystems”. In: *Ber. Math.-Phys. Kl Sächs. Akad. Wiss. Leipzig* 94 (1942), pp. 1–22.
- [78] S Jeffrey and T DeSocio. “Meet Wally. The Room Service Robot of the Residence Inn Marriott at LAX”. In: *Fox11* (2016).
- [79] M. Jindai, S. Ota, Y. Ikemoto, and T. Sasaki. “Handshake request motion model with an approaching human for a handshake robot system”. In: *2015 IEEE 7th International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM)*. IEEE. 2015, pp. 265–270.

-
-
- [80] M. Jindai, S. Ota, H. Yamauchi, and T. Watanabe. “A small-size handshake robot system for a generation of handshake approaching motion”. In: *2012 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER)*. IEEE. 2012.
- [81] M. Jindai and T. Watanabe. “A handshake robot system based on a shake-motion leading model”. In: *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE. 2008.
- [82] M. Jindai and T. Watanabe. “A small-size handshake robot system based on a handshake approaching motion model with a voice greeting”. In: *2010 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*. IEEE. 2010.
- [83] M. Jindai and T. Watanabe. “Development of a handshake request motion model based on analysis of handshake motion between humans”. In: *2011 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*. IEEE. 2011.
- [84] M. Jindai and T. Watanabe. “Development of a handshake robot system based on a handshake approaching motion model”. In: *2007 IEEE/ASME international conference on advanced intelligent mechatronics*. IEEE. 2007, pp. 1–6.
- [85] M. Jindai, T. Watanabe, S. Shibata, and T. Yamamoto. “Development of a handshake robot system for embodied interaction with humans”. In: *ROMAN 2006-The 15th IEEE International Symposium on Robot and Human Interactive Communication*. IEEE. 2006.
- [86] A. Jonnavittula and D. P. Losey. “Learning to share autonomy across repeated interaction”. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2021.
- [87] R. P. Joshi, J. P. Tarapure, and T. Shibata. “Electric wheelchair-humanoid robot collaboration for clothing assistance of the elderly”. In: *IEEE International Conference on Human System Interaction (HSI)*. 2020.
- [88] M. Jouaiti, L. Caron, and P. Hénaff. “Hebbian plasticity in cpg controllers facilitates self-synchronization for human-robot handshaking”. In: *Frontiers in neurorobotics* (2018).
- [89] M. Karl, M. Soelch, J. Bayer, and P. Van der Smagt. “Deep variational bayes filters: Unsupervised learning of state space models from raw data”. In: *International Conference on Learning Representations (ICLR)*. 2017.

-
-
- [90] A. Karniel, I. Nisky, G. Avraham, B.-C. Peles, and S. Levy-Tzedek. “A Turing-like handshake test for motor intelligence”. In: *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*. Springer. 2010, pp. 197–204.
- [91] T. Kasuga and M. Hashimoto. “Human-robot handshaking using neural oscillators”. In: *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*. IEEE. 2005, pp. 3802–3807.
- [92] S. G. Khan, G. Herrmann, M. Al Grafi, T. Pipe, and C. Melhuish. “Compliance control and human–robot interaction: Part 1—Survey”. In: *International journal of humanoid robotics* 11.03 (2014), p. 1430001.
- [93] P. Khanna, M. Björkman, and C. Smith. “Human inspired grip-release technique for robot-human handovers”. In: *IEEE-RAS International Conference on Humanoid Robots (Humanoids)*. 2022.
- [94] P. Khanna, M. Björkman, and C. Smith. “A Multimodal Data Set of Human Handovers with Design Implications for Human-Robot Handovers”. In: *arXiv preprint arXiv:2304.02154* (2023).
- [95] H. Kim, Y. Ohmura, and Y. Kuniyoshi. “Memory-based gaze prediction in deep imitation learning for robot manipulation”. In: *IEEE International Conference on Robotics and Automation (ICRA)*. 2022.
- [96] D. P. Kingma and M. Welling. “Auto-encoding variational bayes”. In: *arXiv preprint arXiv:1312.6114* (2013).
- [97] E. Knoop, M. Bächer, V. Wall, R. Deimel, O. Brock, and P. Beardsley. “Handshakiness: Benchmarking for human-robot hand interactions”. In: *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2017, pp. 4982–4989.
- [98] D. Koert, S. Trick, M. Ewerton, M. Lutter, and J. Peters. “Incremental learning of an open-ended collaborative skill library”. In: *International Journal of Humanoid Robotics* 17.01 (2020), p. 2050001.
- [99] D. Koert, S. Trick, M. Ewerton, M. Lutter, and J. Peters. “Online learning of an open-ended skill library for collaborative tasks”. In: *2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids)*. IEEE. 2018, pp. 1–9.
- [100] R. Krishnan, U. Shalit, and D. Sontag. “Structured inference networks for nonlinear state space models”. In: *AAAI Conference on Artificial Intelligence (AAAI)*. 2017.

-
-
- [101] S. Krishnan, A. Garg, S. Patil, C. Lea, G. Hager, P. Abbeel, and K. Goldberg. “Transition state clustering: Unsupervised surgical trajectory segmentation for robot learning”. In: *The International Journal of Robotics Research (IJRR)* (2017).
- [102] O. Kroemer, C. Daniel, G. Neumann, H. Van Hoof, and J. Peters. “Towards learning hierarchical skills for multi-phase manipulation tasks”. In: *IEEE international conference on robotics and automation (ICRA)*. IEEE. 2015.
- [103] A. Kshirsagar, R. Fortuna, Z. Xie, and G. Hoffman. “Dataset of bimanual human-to-human object handovers”. In: *Data in Brief* (2023).
- [104] A. Kupferberg, S. Glasauer, M. Huber, M. Rickert, A. Knoll, and T. Brandt. “Biological movement increases acceptance of humanoid robots as human partners in motor interaction”. In: *AI & society* 26.4 (2011), pp. 339–345.
- [105] S. Kuutti, S. Fallah, and R. Bowden. “Adversarial mixture density networks: Learning to drive safely from collision data”. In: *IEEE International Intelligent Transportation Systems Conference (ITSC)*. 2021.
- [106] M. Lagomarsino, M. Lorenzini, M. D. Constable, E. De Momi, C. Becchio, and A. Ajoudani. “Maximising Coefficiency of Human-Robot Handovers through Reinforcement Learning”. In: *IEEE Robotics and Automation Letters (RA-L)* (2023).
- [107] K. Lee, S. Choi, and S. Oh. “Maximum causal tsallis entropy imitation learning”. In: *Advances in Neural Information Processing Systems (NeurIPS)* (2018).
- [108] J. Li. “The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents”. In: *International Journal of Human-Computer Studies* 77 (2015), pp. 23–37.
- [109] R. Lioutikov, O. Kroemer, G. Maeda, and J. Peters. “Learning manipulation by sequencing motor primitives with a two-armed robot”. In: *Intelligent Autonomous Systems 13*. Springer, 2016, pp. 1601–1611.
- [110] R. Lioutikov, G. Maeda, F. Veiga, K. Kersting, and J. Peters. “Inducing probabilistic context-free grammars for the sequencing of movement primitives”. In: *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2018, pp. 5651–5658.
- [111] R. Lioutikov, G. Maeda, F. Veiga, K. Kersting, and J. Peters. “Learning attribute grammars for movement primitive sequencing”. In: *The International Journal of Robotics Research* 39.1 (2020), pp. 21–38.
- [112] R. Lioutikov, G. Neumann, G. Maeda, and J. Peters. “Learning movement primitive libraries through probabilistic segmentation”. In: *The International Journal of Robotics Research* 36.8 (2017), pp. 879–894.

-
-
- [113] D. Liu, A. Honoré, S. Chatterjee, and L. K. Rasmussen. “Powering Hidden Markov Model by Neural Network Based Generative Models”. In: *European Conference on Artificial Intelligence (ECAI)*. IOS Press, 2020.
- [114] I. Loshchilov and F. Hutter. “Decoupled Weight Decay Regularization”. In: *International Conference on Learning Representations (ICLR)*. 2018.
- [115] A. L. Maas, A. Y. Hannun, and A. Y. Ng. “Rectifier nonlinearities improve neural network acoustic models”. In: *in ICML Workshop on Deep Learning for Audio, Speech and Language Processing*. 2013.
- [116] G. Maeda, M. Ewerton, R. Lioutikov, H. B. Amor, J. Peters, and G. Neumann. “Learning interaction for collaborative tasks with probabilistic movement primitives”. In: *2014 IEEE-RAS International Conference on Humanoid Robots*. IEEE. 2014, pp. 527–534.
- [117] G. Maeda, M. Ewerton, T. Osa, B. Busch, and J. Peters. “Active incremental learning of robot movement primitives”. In: *Conference on Robot Learning (CoRL)*. 2017.
- [118] P. Manceron. *IKPy*. 2022. URL: <https://doi.org/10.5281/zenodo.6551158>.
- [119] K. L. Marsh, M. J. Richardson, and R. C. Schmidt. “Social connection through joint action and interpersonal coordination”. In: *Topics in cognitive science* (2009).
- [120] G. J. McLachlan and K. E. Basford. *Mixture models: Inference and applications to clustering*. M. Dekker New York, 1988.
- [121] J. R. Medina, F. Duvallet, M. Karnam, and A. Billard. “A human-inspired controller for fluid human-robot handovers”. In: *IEEE-RAS International Conference on Humanoid Robots (Humanoids)*. 2016.
- [122] A. Melnyk, V. P. Borysenko, and P. Henaff. “Analysis of synchrony of a handshake between humans”. In: *2014 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*. IEEE. 2014.
- [123] A. Melnyk and P. Henaff. “Bio-inspired plastic controller for a robot arm to shake hand with human”. In: *2016 IEEE 36th International Conference on Electronics and Nanotechnology (ELNANO)*. IEEE. 2016.
- [124] A. Melnyk and P. Henaff. “Physical analysis of handshaking between humans: mutual synchronisation and social context”. In: *International Journal of Social Robotics* (2019).

-
-
- [125] A. Melnyk, P. Henaff, V. Khomenko, and V. Borysenko. “Sensor network architecture to measure characteristics of a handshake between humans”. In: *2014 IEEE 34th International Scientific Conference on Electronics and Nanotechnology (ELNANO)*. IEEE. 2014.
- [126] A. Melnyk, P. Hénaff, and A. Popov. “Analysis of a handshake between humans using wavelet transforms”. In: *2015 IEEE 35th International Conference on Electronics and Nanotechnology (ELNANO)*. IEEE. 2015.
- [127] M. Mori et al. “The uncanny valley”. In: *Energy* (1970).
- [128] T. Mu and H. Su. “Boosting Reinforcement Learning and Planning with Demonstrations: A Survey”. In: *arXiv preprint arXiv:2303.13489* (2023).
- [129] D. Mura, E. Knoop, M. G. Catalano, G. Grioli, M. Bächer, and A. Bicchi. “On the role of stiffness and synchronization in human–robot handshaking”. In: *The International Journal of Robotics Research* (2020).
- [130] M. Nagano, T. Nakamura, T. Nagai, D. Mochihashi, I. Kobayashi, and W. Takano. “HVGH: unsupervised segmentation for high-dimensional time series using deep neural compression and statistical generative model”. In: *Frontiers in Robotics and AI* (2019).
- [131] E. Nagy, T. Farkas, F. Guy, and A. Stafylarakis. “Effects of Handshake Duration on Other Nonverbal Behavior”. In: *Perceptual and Motor Skills* 127.1 (2020), pp. 52–74.
- [132] H. Nakanishi, K. Tanaka, and Y. Wada. “Remote handshaking: touch enhances video-mediated social telepresence”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2014, pp. 2143–2152.
- [133] S. Nasiriany, T. Gao, A. Mandlekar, and Y. Zhu. “Learning and Retrieval from Prior Data for Skill-based Imitation Learning”. In: *Conference on Robot Learning (CoRL)*. 2023.
- [134] C. Nass, K. Isbister, and E.-J. Lee. “Truth is beauty: researching embodied conversational agents”. In: *Embodied conversational agents*. MIT Press. 2001, pp. 374–402.
- [135] C. Nass and Y. Moon. “Machines and mindlessness: Social responses to computers”. In: *Journal of social issues* 56.1 (2000), pp. 81–103.
- [136] C. Nass, Y. Moon, B. J. Fogg, B. Reeves, and C. Dryer. “Can computer personalities be human personalities?” In: *Conference companion on Human factors in computing systems*. ACM. 1995.

-
-
- [137] C. Nass, Y. Moon, and N. Green. “Are machines gender neutral? Gender-stereotypic responses to computers with voices”. In: *Journal of applied social psychology* (1997).
- [138] E. Ng, Z. Liu, and M. Kennedy. “Diffusion co-policy for synergistic human-robot collaborative tasks”. In: *IEEE Robotics and Automation Letters* (2023).
- [139] S. Niekum, S. Osentoski, G. Konidaris, and A. G. Barto. “Learning and generalization of complex tasks from unstructured demonstrations”. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2012.
- [140] I. Nisky, G. Avraham, and A. Karniel. “Three alternatives to measure the human-likeness of a handshake model in a Turing-like test”. In: *Presence: Teleoperators and Virtual Environments* 21.2 (2012), pp. 156–182.
- [141] B. Ogden and K. Dautenhahn. “Robotic etiquette: Structured interaction in humans and robots”. In: *Procs SIRS 2000, 8th Symp on Intelligent Robotic Systems*. University of Reading. 2000.
- [142] O. S. Oguz, W. Rampeltshammer, S. Paillan, and D. Wollherr. “An ontology for human-human interactions and learning interaction behavior policies”. In: *ACM Transactions on Human-Robot Interaction (THRI)* (2019).
- [143] P. Oikonomou, A. Dometios, M. Khamassi, and C. S. Tzafestas. “Reproduction of human demonstrations with a soft-robotic arm based on a library of learned probabilistic movement primitives”. In: *IEEE International Conference on Robotics and Automation (ICRA)*. 2022.
- [144] P.-H. Orefice, M. Ammi, M. Hafez, and A. Tapus. “Let’s handshake and i’ll know who you are: Gender and personality discrimination in human-human and human-robot handshaking interaction”. In: *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*. IEEE. 2016.
- [145] P.-H. Orefice, M. Ammi, M. Hafez, and A. Tapus. “Pressure Variation Study in Human-Human and Human-Robot Handshakes: Impact of the Mood”. In: *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE. 2018.
- [146] S. Ota, M. Jindai, T. Fukuta, and T. Watanabe. “A handshake response motion model during active approach to a human”. In: *2014 IEEE/SICE International Symposium on System Integration*. IEEE. 2014.
- [147] S. Ota, M. Jindai, T. Sasaki, and Y. Ikemoto. “Handshake response motion model with approaching of human based on an analysis of human handshake motions”. In: *2015 7th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*. IEEE. 2015.

-
-
- [148] K. Ouchi and S. Hashimoto. “Handshake telephone system to communicate with voice and force”. In: *Proceedings 6th IEEE International Workshop on Robot and Human Communication. RO-MAN’97 SENDAI*. IEEE. 1997, pp. 466–471.
- [149] A. K. Pandey and R. Gelin. “A mass-produced sociable humanoid robot: Pepper: The first machine of its kind”. In: *IEEE Robotics & Automation Magazine* 25.3 (2018), pp. 40–48.
- [150] D. Papageorgiou and Z. Doulgeri. “A kinematic controller for human-robot handshaking using internal motion adaptation”. In: *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2015.
- [151] A. Paraschos, C. Daniel, J. Peters, and G. Neumann. “Using probabilistic movement primitives in robotics”. In: *Autonomous Robots* 42.3 (2018), pp. 529–551.
- [152] A. Paraschos, C. Daniel, J. R. Peters, and G. Neumann. “Probabilistic movement primitives”. In: *Advances in neural information processing systems*. 2013, pp. 2616–2624.
- [153] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimeshein, L. Antiga, et al. “Pytorch: An imperative style, high-performance deep learning library”. In: *Advances in neural information processing systems*. 2019, pp. 8026–8037.
- [154] N. Pedemonte, T. Laliberté, and C. Gosselin. “A haptic bilateral system for the remote human–human handshake”. In: *Journal of Dynamic Systems, Measurement, and Control* (2017).
- [155] N. Pedemonte, T. Laliberté, and C. Gosselin. “Design, control, and experimental validation of a handshaking reactive robotic interface”. In: *Journal of Mechanisms and Robotics* (2016).
- [156] E. Pignat and S. Calinon. “Learning adaptive dressing assistance from human demonstration”. In: *Robotics and Autonomous Systems (RAS)* (2017).
- [157] V. Prasad, L. Heitlinger, D. Koert, R. Stock-Homburg, J. Peters, and G. Chalvatzaki. “Learning multimodal latent dynamics for human-robot interaction”. In: *arXiv preprint arXiv:2311.16380* (2023).
- [158] V. Prasad, D. Koert, R. Stock-Homburg, J. Peters, and G. Chalvatzaki. “MILD: Multimodal Interactive Latent Dynamics for Learning Human-Robot Interaction”. In: *IEEE-RAS International Conference on Humanoid Robots (Humanoids)*. 2022.
- [159] V. Prasad, A. Kshirsagar, D. Koert, R. Stock-Homburg, J. Peters, and G. Chalvatzaki. “MoVEInt: Mixture of Variational Experts for Learning Human-Robot Interactions from Demonstrations”. In: *IEEE Robotics and Automation Letters (RA-L)* (2024).

-
-
- [160] V. Prasad, R. Stock-Homburg, and J. Peters. “Advances in Human-Robot Handshaking”. In: *International Conference on Social Robotics*. Springer. 2020.
- [161] V. Prasad, R. Stock-Homburg, and J. Peters. “Human-robot handshaking: A review”. In: *International Journal of Social Robotics* (2021).
- [162] V. Prasad, R. Stock-Homburg, and J. Peters. “Learning Human-like Hand Reaching for Human-Robot Handshaking”. In: *IEEE International Conference on Robotics and Automation (ICRA)*. 2021.
- [163] R. Rahmatizadeh, P. Abolghasemi, A. Behal, and L. Bölöni. “From virtual demonstration to real-world manipulation using LSTM and MDN”. In: *AAAI Conference on Artificial Intelligence (AAAI)*. 2018.
- [164] B. Rammstedt and O. P. John. “Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German”. In: *Journal of research in Personality* (2007).
- [165] D. Rao, F. Sadeghi, L. Hasenclever, M. Wulfmeier, M. Zambelli, G. Vezzani, D. Tirumala, Y. Aytar, J. Merel, N. Heess, et al. “Learning transferable motor skills with hierarchical latent mixture policies”. In: *International Conference on Learning Representations (ICLR)*. 2021.
- [166] D. J. Rezende, S. Mohamed, and D. Wierstra. “Stochastic Backpropagation and Approximate Inference in Deep Generative Models”. In: *31st International Conference on Machine Learning (ICML)*. Vol. 32. 2. PMLR, 2014, pp. 1278–1286.
- [167] L. Righetti, J. Buchli, and A. J. Ijspeert. “Dynamic hebbian learning in adaptive frequency oscillators”. In: *Physica D: Nonlinear Phenomena* 216.2 (2006), pp. 269–281.
- [168] P. Rosenberger, A. Cosgun, R. Newbury, J. Kwan, V. Ortenzi, P. Corke, and M. Grafinger. “Object-independent human-to-robot handovers using real time robotic vision”. In: *IEEE Robotics and Automation Letters (RA-L)* (2020).
- [169] L. Rozo, J. Silverio, S. Calinon, and D. G. Caldwell. “Learning controllers for reactive and proactive behaviors in human–robot collaboration”. In: *Frontiers in Robotics and AI* (2016).
- [170] M. Rubagotti, I. Tusseyeva, S. Baltabayeva, D. Summers, and A. Sandygulova. “Perceived safety in physical human–robot interaction—A survey”. In: *Robotics and Autonomous Systems* (2022).

-
-
- [171] T. Sato, M. Hashimoto, and M. Tsukahara. “Synchronization based control using online design of dynamics and its application to human-robot interaction”. In: *2007 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE, 2007, pp. 652–657.
- [172] S. Schaal. “Dynamic movement primitives-a framework for motor control in humans and humanoid robotics”. In: *Adaptive motion of animals and machines*. Springer, 2006, pp. 261–280.
- [173] D. Schiffrin. “Handwork as ceremony: The case of the handshake”. In: *Semiotica* (1974).
- [174] O. C. Schrempf and U. D. Hanebeck. “A generic model for estimating user intentions in human-robot cooperation”. In: *International Conference on Informatics in Control, Automation and Robotics*. SCITEPRESS, 2005.
- [175] N. Sebanz, H. Bekkering, and G. Knoblich. “Joint action: bodies and minds moving together”. In: *Trends in cognitive sciences* (2006).
- [176] N. Sebanz and G. Knoblich. “Prediction in joint action: What, when, and where”. In: *Topics in cognitive science* (2009).
- [177] F. Semeraro, A. Griffiths, and A. Cangelosi. “Human–robot collaboration and machine learning: A systematic review of recent research”. In: *Robotics and Computer-Integrated Manufacturing* (2023).
- [178] P. Sengadu Suresh, Y. Gui, and P. Doshi. “Dec-AIRL: Decentralized Adversarial IRL for Human-Robot Teaming”. In: *International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 2023.
- [179] A. Shahroudy, J. Liu, T.-T. Ng, and G. Wang. “Ntu rgb+ d: A large scale dataset for 3d human activity analysis”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 1010–1019.
- [180] T. Shankar and A. Gupta. “Learning robot skills with temporal variational inference”. In: *International Conference on Machine Learning (ICML)*. 2020.
- [181] T. Shu, X. Gao, M. S. Ryoo, and S.-C. Zhu. “Learning Social Affordance Grammar from Videos: Transferring Human Interactions to Human-Robot Interactions”. In: *IEEE International Conference on Robotics and Automation (ICRA)*. 2017.
- [182] T. Shu, M. S. Ryoo, and S.-C. Zhu. “Learning Social Affordance for Human-Robot Interaction”. In: *International Joint Conference on Artificial Intelligence (IJCAI)*. 2016.

-
-
- [183] T. Shu, M. S. Ryoo, and S.-C. Zhu. “Learning social affordance for human-robot interaction”. In: *arXiv preprint arXiv:1604.03692* (2016).
- [184] G. L. Stewart, S. L. Dustin, M. R. Barrick, and T. C. Darnold. “Exploring the handshake in employment interviews.” In: *Journal of Applied Psychology* 93.5 (2008), p. 1139.
- [185] R. Stock-Homburg. “Negative interaction spirals during service encounters : insights from human-human and human-robot interactions”. PhD thesis. University of Hagen, 2018.
- [186] R. Stock-Homburg, J. Peters, K. Schneider, V. Prasad, and L. Nukovic. “Evaluation of the handshake turing test for anthropomorphic robots”. In: *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 2020.
- [187] H. G. Sung. “Gaussian Mixture Regression and Classification”. PhD thesis. RICE UNIVERSITY, 2004.
- [188] D. S. Syrdal, K. Dautenhahn, K. L. Koay, and M. L. Walters. “The negative attitudes towards robots scale and reactions to robot behaviour in a live human-robot interaction study”. In: *Adaptive and emergent behaviour and complex systems* (2009).
- [189] G. Tagne, P. Hénaff, and N. Gregori. “Measurement and analysis of physical parameters of the handshake between two persons according to simple social contexts”. In: *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2016, pp. 674–679.
- [190] D. Tanneberg, K. Ploeger, E. Rueckert, and J. Peters. “SKID RAW: Skill Discovery from Raw Trajectories”. In: *IEEE Robotics and Automation Letters* 6.3 (2021), pp. 4696–4703.
- [191] M. Tavassoli, S. Katyara, M. Pozzi, N. Deshpande, D. G. Caldwell, and D. Praticchizzo. “Learning Skills from Demonstrations: A Trend from Motion Primitives to Experience Abstraction”. In: *IEEE Transactions on Cognitive and Developmental Systems* (2023).
- [192] M. Thabet, M. Patacchiola, and A. Cangelosi. “Sample-efficient deep reinforcement learning with imaginary rollouts for human-robot interaction”. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2019.
- [193] A. Tormos Llorente, V. Giménez Ábalos, M. Domènech Vila, D. Gnatyshak, S. Álvarez Napagao, and J. Vázquez Salceda. “Explainable agents adapt to human behaviour”. In: *International Workshop on Citizen-Centric Multi-Agent Systems (CMAS)*. 2023.

-
-
- [194] M. Y. Tsalamlal, J.-C. Martin, M. Ammi, A. Tapus, and M.-A. Amorim. “Affective handshake with a humanoid robot: How do participants perceive and combine its facial and haptic expressions?” In: *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE. 2015.
- [195] C. Van Gemeren, R. Poppe, and R. C. Veltkamp. “Spatio-temporal detection of fine-grained dyadic human interactions”. In: *Human Behavior Understanding: 7th International Workshop, HBU 2016, Amsterdam, The Netherlands, October 16, 2016, Proceedings 7*. Springer. 2016, pp. 116–133.
- [196] N. Vanello, D. Bonino, E. Ricciardi, M. Tesconi, E. P. Scilingo, V. Hartwig, A. Tognetti, G. Zupone, F. Cutolo, G. Giovannetti, P. Pietrini, D. De Rossi, and L. Landini. “Neural correlates of human-robot handshaking”. In: *19th International Symposium in Robot and Human Interactive Communication*. 2010, pp. 555–561.
- [197] F. Vigni, E. Knoop, D. Prattichizzo, and M. Malvezzi. “The Role of Closed-Loop Hand Control in Handshaking Interactions”. In: *IEEE Robotics and Automation Letters* 4.2 (2019), pp. 878–885.
- [198] P. Vinayavekhin, M. Tatsubori, D. Kimura, Y. Huang, G. De Magistris, A. Munawar, and R. Tachibana. “Human-like hand reaching by motion prediction using long short-term memory”. In: *International Conference on Social Robotics*. Springer. 2017, pp. 156–166.
- [199] P. Virtanen et al. “SciPy 1.0—Fundamental Algorithms for Scientific Computing in Python”. In: *arXiv e-prints*, arXiv:1907.10121 (2019), arXiv:1907.10121. arXiv: 1907.10121 [cs.MS].
- [200] D. Vogt, S. Stepputtis, S. Grehl, B. Jung, and H. B. Amor. “A system for learning continuous human-robot interactions from human-human demonstrations”. In: *IEEE International Conference on Robotics and Automation (ICRA)*. 2017.
- [201] B. Wang and M. Hoai. “Predicting body movement and recognizing actions: an integrated framework for mutual benefits”. In: *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*. IEEE. 2018.
- [202] C. Wang, C. Pérez-D’Arpino, D. Xu, L. Fei-Fei, K. Liu, and S. Savarese. “Co-gail: Learning diverse strategies for human-robot collaboration”. In: *Conference on Robot Learning*. PMLR. 2022.
- [203] Z. Wang, E. Giannopoulos, M. Slater, A. Peer, and M. Buss. “Handshake: Realistic human-robot interaction in haptic enhanced virtual reality”. In: *Presence: Teleoperators and Virtual Environments* (2011).

-
-
- [204] Z. Wang, A. Peer, and M. Buss. “An HMM approach to realistic haptic human-robot interaction”. In: *World Haptics 2009-Third Joint EuroHaptics conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*. IEEE. 2009.
- [205] Z. Wang, J. Yuan, and M. Buss. “Modelling of human haptic skill: A framework and preliminary results”. In: *IFAC Proceedings Volumes 41.2 (2008)*, pp. 14761–14766.
- [206] T Yamamoto, S Shibata, and M Jindai. “An Application of “KAN-SEI” Transfer Function to a Robot’s Handing Motion Over to a Human”. In: *Proceedings of The 9th World Multi-Conference on Systemics, Cvbernetics and Informatics*. 2005.
- [207] Y. Yamato, M. Jindai, and T. Watanabe. “Development of a shake-motion leading model for human-robot handshaking”. In: *2008 SICE Annual Conference*. IEEE. 2008.
- [208] S. Yohanan and K. E. MacLean. “The role of affective touch in human-robot interaction: Human intent and expectations in touching the haptic creature”. In: *International Journal of Social Robotics 4.2 (2012)*, pp. 163–180.
- [209] Y. Zeng, Y. Li, P. Xu, and S. S. Ge. “Human-robot handshaking: A hybrid deliberate/reactive model”. In: *International Conference on Social Robotics*. Springer. 2012, pp. 258–267.
- [210] H. Zhang, E. Heiden, S. Nikolaidis, J. J. Lim, and G. S. Sukhatme. “Auto-conditioned recurrent mixture density networks for learning generalizable robot skills”. In: *arXiv preprint arXiv:1810.00146 (2018)*.
- [211] T. Zhang, B. Zhu, L. Lee, and D. Kaber. “Service robot anthropomorphism and interface design for emotion in human-robot interaction”. In: *2008 IEEE International Conference on Automation Science and Engineering*. IEEE. 2008.
- [212] X. Zhao, S. Chumkamon, S. Duan, J. Rojas, and J. Pan. “Collaborative human-robot motion generation using LSTM-RNN”. In: *IEEE-RAS International Conference on Humanoid Robots (Humanoids)*. 2018.
- [213] Y. Zhou, J. Gao, and T. Asfour. “Learning via-point movement primitives with inter-and extrapolation capabilities”. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2019.
- [214] Y. Zhou, J. Gao, and T. Asfour. “Movement primitive learning and generalization: Using mixture density networks”. In: *IEEE Robotics & Automation Magazine (2020)*.



Appendices



A. Publication List

A.1. Under Review

V. Prasad, L. Heitlinger, D. Koert, R. Stock-Homburg, J. Peters, and G. Chalvatzaki. “Learning Multimodal Latent Dynamics for Human-Robot Interaction”. In: *IEEE Transactions on Robotics (T-RO)* (Under Review)

A.2. Journal Papers

V. Prasad, A. Kshirsagar, D. Koert, R. Stock-Homburg, J. Peters, and G. Chalvatzaki. “MoVEInt: Mixture of Variational Experts for Learning Human-Robot Interactions from Demonstrations”. In: *IEEE Robotics and Automation Letters (RA-L)* (2024)

V. Prasad, R. Stock-Homburg, and J. Peters. “Human-robot handshaking: A review”. In: *International Journal of Social Robotics* (2021)

A.3. Conference Papers

M. Gassen, F. Metzler, E. Prescher, V Prasad, L Scherf, F. Kaiser, et al. “I³: Interactive Iterative Improvement for Few-Shot Action Segmentation”. In: *IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. 2023

V. Prasad, D. Koert, R. Stock-Homburg, J. Peters, and G. Chalvatzaki. “MILD: Multimodal Interactive Latent Dynamics for Learning Human-Robot Interaction”. In: *IEEE-RAS International Conference on Humanoid Robots (Humanoids)*. 2022

V. Prasad, R. Stock-Homburg, and J. Peters. “Learning Human-like Hand Reaching for Human-Robot Handshaking”. In: *IEEE International Conference on Robotics and Automation (ICRA)*. 2021

V. Prasad, R. Stock-Homburg, and J. Peters. “Advances in Human-Robot Handshaking”. In: *International Conference on Social Robotics*. Springer. 2020

V. Prasad*, D. Das*, and B. Bhowmick. “Variational clustering: Leveraging variational autoencoders for image clustering”. In: *IEEE International Joint Conference on Neural Networks (IJCNN)*. 2020

V. Prasad and B. Bhowmick. “Sfmlearner++: Learning monocular depth & ego-motion using meaningful geometric constraints”. In: *IEEE Winter Conference on Applications of Computer Vision (WACV)*. 2019

V. Prasad, D. Das, and B. Bhowmick. “Epipolar geometry based learning of multi-view depth and ego-motion from monocular sequences”. In: *Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP)*. 2018

V. Prasad*, K. Yadav*, R. S. Saurabh, S. Daga, N. Pareekutty, K. M. Krishna, B. Ravindran, and B. Bhowmick. “Learning to Prevent Monocular SLAM Failure using Reinforcement Learning”. In: *Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP)*. 2018

M. Kaushik, V. Prasad, K. M. Krishna, and B. Ravindran. “Overtaking maneuvers in simulated highway driving using deep reinforcement learning”. In: *IEEE Intelligent Vehicles Symposium (IV)*. 2018

A.4. Short Papers/Extended Abstracts

Y. Göksu, A. De Almeida Correia, V. Prasad, A. Kshirsagar, D. Koert, J. Peters, and G. Chalvatzaki. “Kinematically Constrained Human-like Bimanual Robot-to-Human Handovers”. In: *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*. 2024, pp. 497–501

F. Hahne, V. Prasad, A. Kshirsagar, D. Koert, R. M. Stock-Homburg, J. Peters, and G. Chalvatzaki. “Transition State Clustering for Interaction Segmentation and Learning”. In: *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*. 2024, pp. 512–516

R. Stock-Homburg, J. Peters, K. Schneider, V. Prasad, and L. Nukovic. “Evaluation of the

handshake turing test for anthropomorphic robots”. In: *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 2020

V. Prasad, R. Jangir, R. Balaraman, and K. M. Krishna. “Data Driven Strategies for Active Monocular SLAM using Inverse Reinforcement Learning”. In: *International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*. 2017

B. Curriculum Vitae

[Google Scholar](#) [GitHub](#) [LinkedIn](#) [Personal Web](#)

Education

July 2019 – Dec. 2023	Technische Universität Darmstadt, Germany Ph.D. Computer Science
Aug. 2016 – Aug. 2017	International Institute of Information Technology, Hyderabad, India M. Sc. Computer Science and Engineering
Dec. 2015	Smart Robotics Lab, Hiroshima University, Higashihiroshima, Japan Exchange Student under the Sakura Science Program
Aug. 2012 – Aug. 2017	International Institute of Information Technology, Hyderabad, India B. Tech. (Honours) Computer Science and Engineering

Honours and Awards

Dec. 2018	Best Paper Award Indian Conference on Computer Vision, Graphics and Image Processing
Aug. 2015 – Dec. 2015	Dean's Merit List Award for Academic Excellence (top 15% in the batch)

Work Experience

July 2019 – Present	Technische Universität Darmstadt, Germany Research Assistant
Aug. 2017 – June 2019	Embedded System & Robotics Dept., TCS Research & Innovation, Kolkata, India Researcher in the Machine Vision Group
July 2016 – June 2017	International Institute of Information Technology, Hyderabad, India Research Assistant at the Robotics Research Center
July 2016 – Dec. 2016	Teaching Assistant for Multi-Agent Systems
July 2014 – Dec. 2014	Teaching Assistant for Structured Systems Analysis and Design
May 2015 – July 2015	Google Summer of Code Android Developer for the Organization "BuildmLearn"

Reviewing

Journals

IEEE Transactions on Robotics (T-RO)
IEEE Robotics and Automation Letters (RA-L)
International Journal of Social Robotics, Springer (IJSR)
Journal of Mathematical Imaging and Vision, Springer

Conferences

IEEE International Conference on Robotics and Automation (ICRA)
IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
IEEE-RAS International Conference on Humanoid Robots (Humanoids)
Conference on Robot Learning (CoRL)
ACM/IEEE International Conference on Human-Robot Interaction (HRI)

IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)

International Conference on Social Robotics (ICSR)

IEEE Intelligent Vehicles Symposium (IV)

IEEE Conference on Decision and Control (CDC)

Hawaii International Conference on System Sciences (HICSS)

Workshops

ACM ICMI Workshop on Interactive Multimodal Learning

Student Supervision

Master Theses

2023 Pi, R., Bluetooth Low Energy Localization for the Social Robot Zenbo

2023 Xu, R., Improving Markerless Hand Tracking using Software-Synchronized Visual-Inertial Odometry

2022 Comellas, O. H., Multi-modal Speaker Identification for Human-Robot Interaction

2022 Frisch, Y., Analysis of Self-supervised Keypoint Detection Methods for Robot Learning

2022 Yang, Z., Exploring Gripping Behaviours and Haptic Emotions for Human-Robot Handshaking

2021 Redkin, M., Personalizing Customer Interactions with Service Robots using Hand Gestures

2021 Kohl, M., Learning Latent Interaction Models using Interaction Primitives

Bachelor Theses

- 2023 Hahne, F., Hierarchical Hidden Markov Models for Interaction Segmentation and Learning
- 2023 Backstein, A., InteractionFlows: Learning Shared Latent Dynamics for Human-Robot Interaction
- 2023 Gao, H., Understanding Haptic Emotions for Human-Robot Handshaking
- 2022 Prescher, E., Visual Hierarchical Interaction Recognition and Segmentation
- 2022 Sterker, L., Social Affordance Segmentation and Learning using Hidden semi-Markov Models
- 2021 Gassen, M., Learning a library of Physical Interactions for Social Robots
- 2021 Scherbring, L., Analyzing the role of Physical Interactions on Service Robot Acceptance
- 2021 Ajmera, Y., Combining multiple ProMPs for end-to-end human-robot handshaking
- 2020 Baierl, M., Learning Action Representations For Primitives-Based Motion Generation

Integrated Projects

- 2023 Göksu, Y., Correia, A. D. A., Kinematically Constrained Humanlike Bimanual Robot Motion
- 2023 Gassen, M., Prescher, E., Metzler, F., I³: Interactive Iterative Improvement for Few-Shot Action Segmentation