Ricardo Filipe Alves Martins

Development of techniques for haptic exploration and recognition of objects

- a contribution to autonomous robotic hands -

Tese de doutoramento em Engenharia Electrotécnica e de Computadores, ramo de especialização em automação e robótica, orientada pelo Professor Doutor Jorge Manuel Miranda Dias e co-orientada pelo Professor Doutor Miguel de Sá e Sousa Castelo-Branco, apresentada ao Departamento de Engenharia Electrotécnica e de Computadores da Universidade de Coimbra.

Agosto de 2016

UNIVERSIDADE DE COIMBRA

Faculty of Science and Technology
University of Coimbra

Development of techniques for haptic exploration and recognition of objects

- a contribution to autonomous robotic hands -

Thesis submitted to the Department of Electrical and Computer Engineering of the Faculty of Science and Technology of the University of Coimbra in partial fulfilment of the requirements for the Degree of Doctor of Philosophy.

Ricardo Filipe Alves Martins

Coimbra, August 2016

Research work developed under the coordination of:

supervision

professor **Jorge Manuel Miranda Dias**

Associate Professor at the Department of Electrical and Computer Engineering

University of Coimbra, Portugal.

Professor at the Department of Electronical and Computer Engineering

Khalifa University, United Arab Emirates.

co-supervision

professor **Miguel de Sá e Sousa Castelo-Branco**

Associate Professor at the Faculty of Medicine

University of Coimbra, Portugal.

This research work was financially supported by Ricardo Filipe Alves Martins and by an individual Ph D. scholarship (SFRH/BD/65990/2009) funded by the Portuguese science agency FCT - Fundação para a Ciência e a Tecnologia (Foundation for Science and Technology).

Acknowledgment

I would like to express my deepest gratitude to my supervisor, Prof. Jorge Dias, for the opportunity of developing my PhD studies at the *Institute of Systems and Robotics* (ISR), *University of Coimbra* (UC). His encouragement, guidance, and proposals and assistance in exploring new ideas were fundamental to making me a better researcher and improving the quality of my PhD work. The opportunity to participate in European projects, and the integration of meetings and teams presenting proposals of new national and international projects also broadened my perspectives regarding international research and academia. They also prepared me for the challenges I may encounter after the PhD studies.

I would like to thank my co-supervisor, Prof. Miguel Castelo-Branco, for providing theoretical support concerning the human perception and cognition mechanisms. His optimistic words of encouragement were of great help and inspiration during the final stages of my PhD studies.

My sincere appreciation is extended to the *Artificial Perception for Intelligent Systems and Robotics laboratory* (AP4ISR) of the *Institute of Systems and Robotics* (ISR) at the *University of Coimbra* (UC) by providing me administrative support and experimental conditions to develop the work presented in this thesis. This research work was financially supported by an individual PhD scholarship (SFRH/BD/65990/2009) funded by the Portuguese science agency *FCT - Fundação para a Ciência e a Tecnologia* (Foundation for Science and Technology).

My acknowledgements also go to other senior members of the AP4ISR laboratory. I would like to thank Prof. João Filipe Ferreira for introducing me to the theoretical background and formalism of probabilistic modelling applied to artificial perception, and for his persistence and continuous support during the implementation and reviewing of several research papers submitted during my PhD studies. My sincere appreciation goes out to Prof. Jorge Lobo for his kindness during all these years, and for the diversity of scientific challenges he presented during the activities of the FP7-HANDLE project, which greatly improved my professional skills (technical documentation, integration of data acquisition devices, signal processing, and programming).

I would like to extend my sincerest thanks and appreciation to several of my laboratory colleagues at AP4ISR for their scientific and personal support, friendship, and

discussions: Nuno Semedo, Rita Catarino, Amilcar Ferreira, Luís Santos, Diego Faria, Pedro Trindade, Pedro Machado, David Bina, Luís Almeida, José Marinho, João Monteiro, Kamrad Khoshhal, Hadi Aliakbarpour, Jafar Hosseini, Alexandre Malhão, João Quintas, Pablo Lanillos, João Santos, João Martins, Paula Lopes, Paulo Drews Jr, Carlos Simplício, Gonçalo Augusto, José Sousa, André Araújo, Rosana Herrera, Hussam Al-Hussein, José Prado, and Diana Guardado. I would like to make a much deserved special mention to four colleagues. Diego Faria gave me the opportunity to collaborate in several scientific papers when I arrived at AP4ISR. Diego introduced me to the scientific field of robotic learning of dexterous manipulation skills from human demonstrations. Thank you to Luís Santos for the fruitful technical and scientific discussions, and sharing his expertise in probabilistic modelling and technical writing, which contributed to an increased quality of the work presented in this thesis. I am also grateful to Rita Catarino, Luís Santos, and Pedro Trindade for the encouragement and support during difficult moments of my family life.

Since January 2014, I also worked with the *IBILI-UC Visual Neuroscience Laboratory* team headed by Prof. Miguel Castelo-Branco. I am thankful for the opportunity of working in a new and challenging research field (clinical investigation and neuroscience), which established the bridge between medical and engineering sciences. This job also helped me to fund the last two years of my PhD studies by providing an after-work activity. I would like to show my appreciation to my colleagues at *IBILI/ICNAS-UC* for helping me to become more confident and for improving my teamwork and communication skills. Thanks for the extensive discussions about science as well as other funny and non-serious subjects.

I am greatly thankful to João Duarte, Joana Crisóstomo, Tiago Marçal, Lúcia Batata, Paulo Barbeiro, Sofia Pinto, Emília Neto, Joana Caldas, and Catarina Pereira for being amazing friends through all these years. Whenever I needed your support and advice, you were there, ready to help me.

I would like to thank all my family members for supporting me throughout the past years. A special mention to my father, mother, and brother, who have continued to encourage me to pursue my PhD studies and finish this thesis, despite all the unforeseen circumstances in our lives.

That's Life

That's life (that's life), that's what all the people say
You're ridin' high in April, shot down in May
But I know I'm gonna change that tune
When I'm back on top, back on top in June

I said that's life (that's life), and as funny as it may seem
Some people get their kicks stompin' on a dream
But I don't let it, let it get me down
'cause this fine old world, it keeps spinnin' around

I've been a puppet, a pauper, a pirate, a poet, a pawn and a king
I've been up and down and over and out and I know one thing
Each time I find myself flat on my face
I pick myself up and get back in the race

That's life (that's life), I tell you I can't deny it
I thought of quitting, baby, but my heart just ain't gonna buy it
And if I didn't think it was worth one single try
I'd jump right on a big bird and then I'd fly

I've been a puppet, a pauper, a pirate, a poet, a pawn and a king
I've been up and down and over and out and I know one thing
Each time I find myself layin' flat on my face
I just pick myself up and get back in the race

That's life (that's life), that's life and I can't deny it
Many times I thought of cuttin' out but my heart won't buy it
But if there's nothin' shakin' come this here July
I'm gonna roll myself up in a big ball a-and die

From: *That's Life*, Frank Sinatra, 1966. Writers: Dean Kay and Kelly Gordon.

Abstract

During the past few years, a new generation of robotic platforms has begun being integrated in distinct environments (e.g. domestic, healthcare, and entertainment). The robotic platforms must execute autonomously a great variety of tasks (to and in cooperation with humans) in uncertain and dynamic environments. To overcome these challenges, the robotic platforms are equipped with a high diversity of sensory (e.g. monocular and stereo cameras; microphones; and force, torque, and tactile sensing arrays) and actuation apparatus (e.g. dexterous robotic arms and hands, touch screens, humanoid heads, and audio speakers). The research works presented in this thesis are related to the subject of robotic dexterous manipulation and haptic exploration of objects (rigid and soft).

This thesis contributes to the development of robotic platforms with autonomous dexterous manipulation capabilities by studying the human manipulation and haptic exploration skills, presenting several approaches to translate and transfer them to a robotic platform. The study of the human visual and somatosensory systems, the neuronal and functional units supporting the sensory processing pipeline, as well as the behavioural patterns participating in the action-perception loop were used as guidelines and benchmarks throughout the thesis during the formulation and evaluation of three artificial perception applications.

Toward the first application, this thesis presents an approach to model the human strategies executed during a dexterous manipulation task. The human hand is instrumented with a tactile sensing array. The thesis proposes a symbolic description of the tasks using grasping primitives. Each grasping primitive is described by the hand-object contact interaction signature. During the human demonstration of two different dexterous manipulation tasks, the sequence of grasping primitives is recognized by a Bayesian model. The statistical relations emerging from the analysis of the sequence of grasping primitives are used to define the model of the task.

The research works presented in this manuscript contribute to a second application consisting of an artificial perception system to discriminate in-hand explored objects with different hardness properties. The human hand is instrumented with a tactile sensing array and a motion tracking sensor. A Bayesian model integrates features (contact intensity, contact area, and contact indentation) extracted from the sensory data acquired

during the press-and-release exploration of the objects. The cutaneous and kinesthetic cues are integrated by a Bayesian model so the system can learn to discriminate between three distinct materials (haptic memory). The learned parameters are used to infer the perceived hardness properties of unknown objects based on the haptic memory of the system.

The final contribution of this thesis is concerning the implementation of a probabilistic approach to perform active haptic exploration of surfaces using dexterous robotic hands (simulation environment). The proposed approach represents the structure of an unknown surface as a probabilistic grid. As long as the haptic exploration of the surface progresses, haptic cues regarding texture and compliance are integrated by a Bayesian model and used to infer the category of material of that region of the workspace. The approach showed an excellent capability to discriminate between ten different types of materials (haptic stimulus). Based on this perceptual representation of the workspace, the robotic system infers the next region of the unknown workspace that should be explored. This decision is made by integrating bottom-up and top-down cues related to the haptic saliency of the stimulus, uncertainty of the current perceptual representation of the workspace, inhibition-of-return mechanisms, objectives of the task, and the current structure of the exploration path. The Bayesian models involved in this approach were tested on a planar surface, during the detection and following haptic discontinuities between three different materials. The following of haptic discontinuity was performed with good structural accuracy. The tactile attention mechanisms of the system demonstrated a high specificity, following the discontinuities of interest and ignoring the others. The role and impact of the different cues (haptic saliency, inhibition-of-return, uncertainty, and structure of exploration path) was also studied by removing each of these components from the Bayesian models.

Resumo

Nos últimos anos, uma nova geração de plataformas robóticas tem sido integrada em novos tipos de cenários (ex: ambiente doméstico, unidades de saúde, locais de entretenimento). Neste tipo de cenários as plataformas robóticas necessitam de executar autonomamente uma grande variedade de tarefas (para e em cooperação com Humanos), sendo confrontadas com ambientes dinâmicos e imprevisíveis. De forma a ultrapassar estes desafios, as plataformas robóticas são equipadas com uma grande variedade de sensores (ex: câmaras monocular e estéreo, microfones, sensores de força, torque e de tacto) e interfaces/actuadores (mãos e braços robóticos, ecrãs tácteis, cabeças humanóides, sintetizadores de áudio). Os trabalhos de investigação apresentados nesta tese estão relacionados com a manipulação e exploração háptica de objectos (rígidos e moles).

Com esta tese pretende-se contribuir para o desenvolvimento de plataforms robóticas com capacidade autónoma de manipulação, através do estudo da perícia Humana em tarefas de manipulação e exploração háptica. So propostas diferentes abordagens para replicar estas habilidades Humanas em platformas robóticas. O estudo do sistema visual e somatosensorial Humano, das unidades neuronais e funcionais envolvidas no processamento sensorial, assim como dos padrões comportamentais intervenientes nos mecanismos do ciclo acção-percepção, foram considerados como referências durante a formulação, desenvolvimento e avaliação do desempenho de três aplicações de percepção artificial.

Numa primeira aplicação, propõe-se uma metodologia para modelizar as estratégias utilizadas por Humanos durante uma tarefa de manipulação. A mão Humana é instrumentada com uma série de sensores de tacto. Nesta tese propõe-se uma descrição simbólica das tarefas, utilizando diferentes primitivas para efectuar essa modelização. Cada primitiva é descrita pelo perfil de contacto (região e intensidade do contacto) entre a mão e o objecto. Durante a demonstração de dois tipos de tarefas de manipulação, a sequência temporal de primitivas é reconhecida por um modelo Bayesiano. O modelo de cada um dos tipos de tarefa é extraído a partir das relações de estatísticas de causalidade que se estabelecem entre tipos de primitivas consecutivas que tenham sido inferidas pelo modelo Bayesiano proposto.

Esta tese apresenta ainda um segundo sistema de percepção artificial que discrimina objectos com diferentes características de rigidez, durante tarefas de exploração háptica. A

mão Humana é instrumentada com uma série de sensores de tacto e um sensor de rastreamento de movimento. Desenvolveu-se um modelo Bayesiano que integra simultaneamente descritores (intensidade de contacto, área de contacto, nível de indentação do contacto) extraídas dos dados sensoriais adquiridos durante a exploração dos objectos através de movimentos de palpação. As componentes cutaneas e quinestésica da interacção dedo-objecto são integradas pelo modelo Bayesiano de forma a que este estime os parâmetros do modelo e aprenda a discriminar três materials diferentes (memória háptica do sistema) As características de rigidez de objectos desconhecidos ao sistema são inferidas pelo modelo Bayesiano, baseando-se na memória háptica do sistema desenvolvida anteriormente.

Por fim, esta tese apresenta um sistema de percepção artificial e acção relacionado com a implementação de uma abordagem probabilística para executar a exploração háptica activa de superfícies usando mãos robóticas (ambiente de simulação). Na abordagem proposta a estrutura espacial das superfícies desconhecidas pelo sistema é representada por uma grelha probabilística. À medida que a exploração háptica progride, descritores relacionados com características de textura e complacência são integrados por um modelo Bayesiano e usados para inferir a categoria de material existente naquela região da superfície explorada. Esta abordagem foi testada experimentalmente, sendo capaz de discriminar com um elevado desempenho 10 materials diferentes. A partir da representação perceptual do espaço de trabalho inferida pelo modelo Bayesiano referido anteriormente, o sistema robótico estima qual será a próxima região da superfície a ser explorada. Esta decisão é tomada integrando simultaneamente informação relacionada com a saliência háptica e incerteza associada às diferentes regiões da representação perceptual actual da superfície, mecanismos de inibição-de-retorno, objectivos da tarefa de exploração e estrutura do percurso de exploração efectuado pelo sistema robótico até ao momento. O sistema proposto foi testado numa superfície plana, durante a execução de tarefas de exploração háptica relacionadas com o seguimento de discontinuidades entre três materiais diferentes. As discontinuidades entre os três materiais, correspondentes aos objectivos das diferentes tarefas, foram seguidas com uma boa precisão estrutural. Os mecanismos de atenção táctil demonstraram uma grande especificidade, fazendo com que o sistema seguisse as discontinuidades de interesse para a tarefa, ignorando as restantes descontinuidades existentes no espaço de tabalho. O impacto da contribuição de cada uma das componentes (saliência háptica, inibição-de-retorno, incerteza, estrutura do percurso explorado) foi estudado através da análise do desempenho do sistema quando o efeito de cada uma dessas componente era neutralizado na formulação dos modelos Bayesianos.

Contents

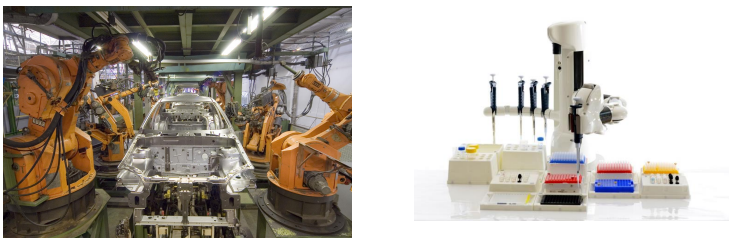
Chapter 1

Introduction

1.1 Motivation

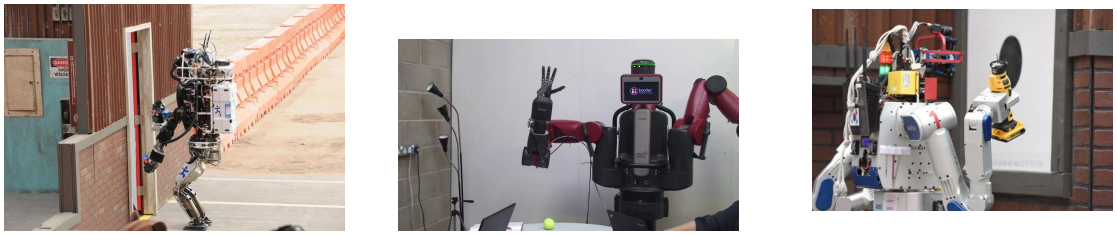
During the past few decades, robotic platforms (e.g. Figure 1.1) have been integrated into industrial and laboratory environments, with the primary task of performing repetitive and monotonous tasks quickly and accurately. Typically, the movements and tasks required to successfully fulfil the job for which the robots were preprogrammed required little or no sensory feedback or interaction with humans during execution time [Nof, 1999]. This type of robotic platform was primarily designed to be integrated in clean environments, rely on complete task information, and, if required, to be reprogrammed for new task execution. This type of platform was not designed to interact directly with humans and to handle the execution of tasks in uncertain environments.

Currently, various types of robotic platforms (e.g. Figure 1.2) have begun being heavily introduced in different environments (including domestic, healthcare, entertainment, and education) in which the robots must manage new challenges, such as the ability to interact with persons and objects in the environment [Zollo et al., 2013]. These new classes of environments are dynamic, unpredictable, and cannot be completely known in



(a) (b)

Figure 1.1: Integration of robotic systems in controlled environments. a) *KUKA IR 160/60* (KUKA Robot Group, Augsburg, Germany) operating at an industrial car factory. b) *Andrew Alliance* (Andrew Alliance S.A., Vernier, Switzerland) anthropomorphic robot handling liquids.



(a) (b) (c)

Figure 1.2: Integration of robotic systems on daily-life environments. a) *ATLAS* robot (Boston Dynamics, Waltham, USA) opening a door. b) *Baxter* robot (Rethink Robotics, Boston, USA) grasping objects on table. c) *DRC-HUBO* robot (Rainbow Co., Daejeon, South Korea) operating a power drill.

advance. These factors have led to the development of mobile robotic platforms with multi-modal sensing systems, and complex actuation systems and interaction interfaces such as active vision systems, audition, multi-articulated arms, dexterous robotic hands, touch displays, and microphones. These multi-modal modules, actuators, and interfaces provide a framework to develop artificial perception systems to deal autonomously with the dynamics of the environments and wide variety of objects, and to interact safely with humans.

One of the key elements contributing to the performance of that type of robotic platform is the ability to perform autonomous grasping, manipulation, exploration, and characterization of partially known objects in the environment [Saudabayev and Varol, 2015]. To achieve these objectives, the tendency in the field of robotic research is to move the development of robotic hands from simple grippers toward human-inspired articulated hands (with a mechanical structure, integration on robotic arms, and various degrees of freedom) and introduction on the robotic hand of sensing devices such as tactile, temperature, and force/torque sensors (e.g. Figure 1.3).

The greater relevance of the study and development of robotic hands and issues relating to autonomous manipulation is revealed by the increasing number of projects and funding resources applied to this scientific field. In the past years, several collaborative large-scale research projects (primarily between European institutions) were funded: HANDLE [HANDLE, 2009] (Developmental pathway towards autonomy and dexterity in robot in-hand manipulation), GRASP [GRASP, 2008] (Emergence of Cognitive Grasping through Introspection, Emulation and Surprise), DEXMART [DEXMART, 2008](DEXterous and autonomous dual-arm/hand robotic manipulation with sMART sensory-motor skills: A bridge from natural to artificial cognition), GeRT [GeRT, 2011] (Generalizing Robot Manipulation Tasks), and THE [THE, 2010](The Hand Embodiment).

As presented in Figure 1.4, most of these large-scale research projects in cognitive

(a) (b)

Figure 1.3: Dexterous robotic hands. a) *Shadow* (Shadow Robot Company Ltd, London, UK) robotic hand equipped with *Syntouch Biotac* (SynTouch LLC, Los Angeles, USA) sensory fingertips. b) *DLR* (German Aerospace Center, Cologne, Germany) robotic hand.

robotics follow an approach that develops its engineering solutions to the topic of dexterous robotic manipulation and exploration by integrating principles from the human studies sciences. Neurobiology and neuroscience provide functional models describing the processes involved in the sensing, transduction, and sensory processing pipeline, perception, cognition, and motor capabilities. Neuropsychology and cognitive science systematically describe and categorise the standard human behaviours, which are used as benchmarks to define protocols to evaluate robotic skills [Cheng, 2014].

The objectives of these projects can be grouped in two main classes. One class is related to the development of hardware and sensors, as well as its integration within the structure of the robotic hand or the structure of a main robotic platform. A second class concerns the development of algorithms and techniques dedicated to the modelling and implementation of artificial perception-to-action and action-to-perception systems based on hardware platforms provided by the first class of approaches. This work intends to contribute to the second class of approaches.

The integration of this new generation of robotic hands onto robotic platforms places new challenges concerning the motion (fingers, palm, coordination fingers-fingers and fingers-palm) of the robotic hand, with a certain number of degrees of freedom. However, the introduction of sensing devices opens new possibilities, allowing the replication of human motor strategies and perceptual capabilities concerning the ability to perform dexterous, finer manipulation, and exploration skills. These capabilities are boosted due to the integration of haptic data (flexure level, tactile, force/torque, and temperature) in the control loop of the robotic manipulation [Yousef et al., 2011] .

This new class of skills is related to in-hand manipulation tasks that consist of internal consecutive (re-)grasping and release of the object to perform its reorientation, fine positioning, or more complex interaction such as sequential rotation of the object. Hu-

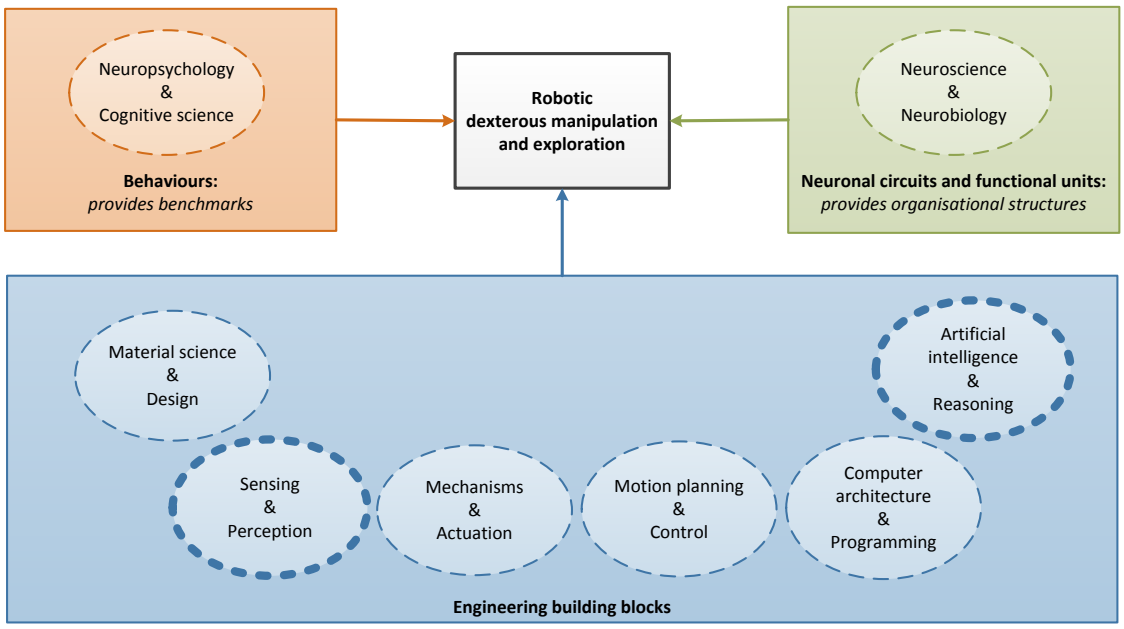


Figure 1.4: Overview of the multidisciplinary approach followed to develop robotic dexterous manipulation and exploration capabilities. The fields of intervention of this thesis are highlighted in bold. Adapted from [Cheng, 2014].

mans also use the in-hand manipulation skills to perform in-hand exploration of objects to complete the construction and definition of the model of the object (e.g. shape, size, superficial texture, superficial friction coefficient, weight, and softness) [Lederman, 1994], when the properties are not completely known in advance or to complete partial information about the object model provided by other sensing modalities (e.g. vision) [Lacey and Sathian, 2014], [Stone and Gonzalez, 2015].

Haptic perception plays a relevant role in this type of skills [Dahiya et al., 2010] by providing sensory cues about the regions of the hand contacting the object, the temporal sequence, and object-hand contact dynamics [Johansson and Flanagan, 2009]. The contributions of this thesis relate to the development of methodologies used to implement artificial haptic perception skills and to establish the interdependence of those skills with actions.

1.2 Thesis outline

This chapter, *Introduction*, describes the context, motivation, and applications of the research works presented throughout this thesis.

Chapter 2, *Fundamentals*, proposes and describes the probability theory, probabilistic

grids, and information theory formalisms used by this thesis. The graphical and Bayesian programming notations used to describe the Bayesian models are described as well.

The following chapter, *Recording human manipulation and exploration movements*, reviews the human somatosensory apparatus and processing pipeline involved in the dexterous manipulation and haptic exploration of objects. Special focus is dedicated to the extraction of haptic features, tactile attention mechanisms, and hand motion patterns. Chapter 3 also presents an overview of the current benchmarking guidelines used to evaluate robotic manipulation and exploration tasks.

The experimental area of the AP4ISR laboratory (Artificial Perception for Intelligent Systems and Robotics) used to record multi-modal datasets of human demonstrations of dexterous manipulation and exploration tasks is presented in chapter 4, *Recording human manipulation and exploration movements*. The characteristics of the different data acquisition devices are detailed. The implementation of the global data acquisition architecture is explained. The applications of the datasets recorded during this PhD study are listed. Several software tools developed to promote the acquisition, integration, and demonstration of the datasets are also documented.

Chapter 5, *Recognition of grasping primitives using tactile sensory data*, chapter 6, *Categorization of soft objects during haptic exploration tasks*, and chapter 7, *Active haptic exploration of surfaces using dexterous robotic hands* present different but complementary approaches to endow robotic platforms with dexterous manipulation and exploration capabilities. Artificial haptic perception plays a central role in these three chapters. All three chapters are arranged similarly: introduction; related works discussed using a table to compare this thesis with other works; proposed approach illustrated by two schemes: simplified overview and detailed overview of the flow of data, variables, and algorithms; sections covering methods and algorithms; experimental results; and main conclusions.

Chapter 5, *Recognition of grasping primitives using tactile sensory data*, proposes an approach to recognize different grasp shape primitives during the human demonstration of dexterous manipulation tasks. Chapter 6, *Categorization of soft objects during haptic exploration tasks*, implements an approach to recognize the category of material during the human haptic exploration of objects. Finally, chapter 7, *Active haptic exploration of surfaces using dexterous robotic hands*, demonstrates a system (simulation environment) which performs active haptic exploration of surfaces using dexterous robotic hands. The contribution of tactile attention mechanisms to improve the exploration performance is demonstrated.

1.3 List of deliverables

During the PhD studies conducted in the development of this thesis, several deliverables were produced. The contribution of each deliverable to this thesis is described in the list presented next. All the deliverables are available on line at `http://www.rmartins.net/phd-docs/`.

1.3.1 Peer-reviewed international journals

- `Contributes to chapter 7` ⇒ [Martins et al., 2017] **R. Martins**, J. F. Ferreira, M. Castelo-Branco , J. Dias - "Integration of touch attention mechanisms to improve the robotic haptic exploration of surfaces" - Neurocomputing, Volume 222, 26 January 2017, Pages 204-216.
  DOI: 10.1016/j.neucom.2016.10.027

- `Contributes to chapter 5` ⇒ [Faria et al., 2012] D. R. Faria, **R. Martins**, J. Lobo, J. Dias - "Extracting Data from Human Manipulation of Objects Towards Improving Autonomous Robotic Grasping" - Robotics and Autonomous Systems, Special Issue on Autonomous Grasping, Volume 60, Issue 3, Pages 396-410, March 2012.
  DOI: 10.1016/j.robot.2011.07.020

1.3.2 Peer-reviewed proceedings of international conferences

- `Contributes to chapter 7` ⇒ [Martins et al., 2014] **R. Martins**, J. F. Ferreira, J. Dias - "Touch attention Bayesian models for robotic active haptic exploration of heterogeneous surfaces". Proceedings of 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2014), pages 1208-1215, Chicago, USA, Sept. 14-18, 2014.
  DOI: 10.1109/IROS.2014.6942711

- `Contributes to chapter 7` ⇒ [Martins et al., 2012b] **R. Martins**, J. F. Ferreira, J. Dias - "Touch attention Bayesian models for object feature extraction in robotic blind manipulation". Proceedings of 32nd International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering (MaxEnt 2012), Munich, July, 2012.

- `Contributes to chapter 7` ⇒ [Martins et al., 2013] **R. Martins**, J. F. Ferreira, J. Dias - "Touch attention Bayesian models for robotic active haptic exploration".

Proceedings of 2nd Workshop on Recognition and Action for Scene Understanding (REACTS 2013), pages 45-58, York, UK, 30 August, 2013.
ISBN: 978-84-616-7092-5

- `Contributes to chapter 6` ⇒ [Martins et al., 2012a] **R. Martins**, D. R. Faria, J. Dias - "Representation framework of perceived object softness characteristics for active robotic hand exploration". Proceedings of 7th ACM/IEEE International Conference on Human Robot Interaction (HRI2012) - Workshop on Advances in Tactile Sensing and Touch based Human-Robot Interaction, Boston, USA, March 5-8, 2012.
DOI: 10.13140/RG.2.1.2739.3445

- `Contributes to chapter 5` ⇒ [Martins et al., 2010] **R. Martins**, D. R. Faria, J. Dias, "Symbolic Level Generalization of In-hand Manipulation Tasks from Human Demonstrations using Tactile Data Information". Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2010): Workshop on Grasping Planning and Task Learning by Imitation, Taipei, Taiwan - October 2010.
DOI: 10.13140/2.1.3782.2401

1.3.3 Peer-reviewed poster in international conferences

- `Contributes to chapter 6` ⇒ **R. Martins**, J. Dias - "Representation framework of perceived object softness characteristics for active robotic hand exploration". In HANDLE Workshop, Benicassim, Spain, February, 2012.

1.3.4 Research collaborations as co-author

- [Diego Faria, 2012] ⇒ D. R. Faria, **R. Martins**, J. Lobo, J. Dias. "A Probabilistic Framework to Detect Suitable Grasping Regions on Objects". In 10th IFAC Symposium on Robot Control (SYROCO 2012), Dubrovnik, Croatia, September, 2012.
DOI: 10.3182/20120905-3-HR-2030.00090

- [Faria et al., 2011] ⇒ D. R. Faria, **R. Martins**, J. Lobo, J. Dias. "Manipulative Tasks Identification by Learning and Generalizing Hand Motions". In DoCEIS'11 - 2nd Doctoral Conference on Computing, Electrical and Industrial Systems. Costa da Caparica - Portugal, February, 2011.
DOI: 10.1007/978-3-642-19170-1-19

- [Faria et al., 2010b] ⇒ D. R. Faria, **R. Martins**, J. Dias - Grasp Exploration for 3D Object Shape Representation using Probailistic Map - In Proceedings of the

DoCEIS'10 - Doctoral Conference on Computing, Electrical and Industrial Systems. Lisbon, February, 2010. Springer
DOI: 10.1007/978-3-642-11628-5-23

- [Faria et al., 2010a] ⇒ D. R. Faria, **R. Martins**, J. Lobo, J. Dias - Probabilistic Representation of 3D Object Shape by In-Hand Exploration - in Proceedings of The 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'10, pp. 1560-1565 - Taipei, Taiwan - October 2010.
DOI: 10.1109/IROS.2010.5649286

- [Faria et al., 2009] ⇒ D. R. Faria, **R. Martins**, J. Dias - Human reach-to-grasp generalization strategies: a Bayesian approach - Workshop at Robotics: Science and Systems 2009, Workshop: "Understanding the Human Hand for Advancing Robotic Manipulation" - July 28, 2009 - Dillon Eng Seattle, WA, USA.

- [Faria et al., 2010c] ⇒ D. R. Faria, **R. Martins**, J. Dias - Learning Motion Patterns from Multiple Observations along the Actions Phases of Manipulative Tasks - to appear in IEEE/RSJ IROS'2010: Workshop on Grasping Planning and Task Learning by Imitation, Taipei, Taiwan - October 2010.

1.3.5 Technical report

- `Contributes to chapters 4, 5, 6` ⇒ [Martins, 2010] Distributed synchronization of multi-modal data acquisition devices using NTP (network time protocol).

- `Contributes to chapter 4` ⇒ [Martins, 2013] Installing controller area network (CAN-Bus) drivers and compiling code on Ubuntu.

- `Contributes to chapter 4` ⇒ [Martins, 2012b] Experimental evaluation and calibration protocol of *Tekscan Grip* system.

- `Contributes to chapters 4, 5, 6` ⇒ [HANDLE-UC, 2009] Protocol for the corpus of sensed grasp and handling data: storage of multimodal datasets.

- `Contributes to chapter 3` ⇒ [Martins, 2008] Modelling the Human body and hand: kinematic structure, degrees-of-freedom.

1.3.6 Software tools and documentation

- `Contributes to chapter 3` ⇒ [Martins, 2009a] 3D interactive demonstrator of human body and hand: kinematic structure, degrees-of-freedom.

- `Contributes to chapter 4` ⇒ [Martins, 2012a] 3D visualization tool of instrumented Rubik cube touch data.
- `Contributes to chapters 4, 5, 6` ⇒ [Martins, 2009g] *importDatasetTB*: toolbox for integrating data in MATLAB.
- `Contributes to chapter 4` ⇒ [Martins, 2012c] Annotation tool for multi-modal Human grasping datasets.
- `Contributes to chapters 4, 6` ⇒ [Martins, 2009b] Distributed data acquisition architecture: software client for *CyberGlove* (data glove).
- `Contributes to chapter 4` ⇒ [Martins, 2009c] Distributed data acquisition architecture: software client for instrumented Rubik cube (instrumented object).
- `Contributes to chapter 4` ⇒ [Martins, 2009d] Distributed data acquisition architecture: software client for instrumented sensing can (instrumented object).
- `Contributes to chapter 4` ⇒ [Martins, 2009f] Distributed data acquisition architecture: software client for *Tekscan Grip* system (tactile sensing array).
- `Contributes to chapters 4, 6` ⇒ [Martins, 2009e] Distributed data acquisition architecture: software client for *Polhemus Liberty* system (6D motion tracking).

1.3.7 Datasets

- `Contributes to chapter 4` ⇒ `Human demonstration of dexterous manipulation tasks:`
  - Dexterous manipulation of a laboratory pipette.
  - Thumb movement during manipulation tasks.
  - Screwdriver in-hand rotation.
  - In-hand manipulation of toys.
  - Grasp the Wii remote and press a button.
  - Fill a toy sorting box with objects.
  - Pick up a pen and write.
  - Pick an object and slide.

Chapter 2

Fundamentals

The computational models and engineering solutions proposed in this thesis for performing the dexterous manipulation and haptic exploration of surfaces follow the principles involved in human perception, cognition, and action.

As presented in Figure 2.1, the perceptual process starts a sequence of sub-mechanisms that work together to estimate a representation of the environment. The perceptual representation is then used to infer a reaction strategy to those stimuli coming from the environment.

Several models have been proposed to explain and describe the mechanisms involved in human perception and how they are integrated into global human behaviour. Humans perceive in order to act on the environment and, the actions performed with environment elements affect the perception of the environment: the so-called action-perception loop (Figure 2.1).

Although most of the time, the sensory signals are ambiguous and corrupted with noise, humans have a remarkable capability to create successful perceptual representations which they use to guide their actions [Ernst and Bulthoff, 2004]. To explain this capability, Hermann von Helmholtz proposed an approach to model the perception mechanisms (Figure 2.1), introducing a principle designed by unconscious inference [Westheimer, 2008]. The principle states that humans perceive a specific state of the environment, choosing the state which is most likely to have caused the pattern of stimulus that the human subject has received through the sensory apparatus. Additionally, although sensory data is diverse, it is not sufficient to uniquely determine what is perceived. Prior knowledge must be used, which introduces constraints to the process of inference from ambiguous sensory signals.

The next sections present the fundamentals of the formalism of probability theory used to model the state of robotic systems, human agents, and the environment (section 2.1). This chapter also presents the formalism of probabilistic grids (section 2.2), used in this thesis to represent the workspace surrounding the robotic and human agents. The information exchanged between the different modules of the methods proposed in this

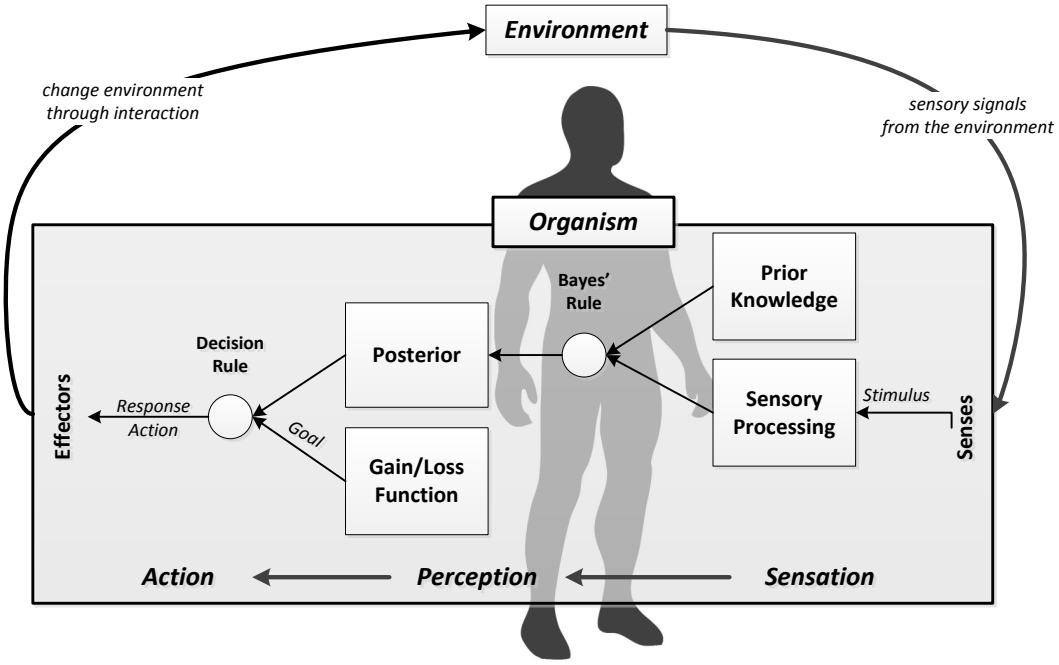


Figure 2.1: Representation of the fundamental mechanisms underlying the action-perception loop. Adapted from [Ernst and Bulthoff, 2004].

work is described using the formalism of information theory (section 2.3).

2.1 Probabilistic modelling

In robotics, different formalisms (such as first-order predicate logic and probability theory) have been followed to represent knowledge and describe reasoning applications. An extensive review of these formalisms can be found in [Hertzberg and Chatila, 2008].

This work uses the probability theory to represent the knowledge of the state of the robotic system (and other agents) and its surroundings, following an approach analogous to the work [Knill and Richards, 1996] to model human perception mechanisms and reasoning. This formalism has been used extensively in robotics. The increasing interest in this formalism is related to its ability to deal with the incompleteness of the description of the system (inaccurate modelling, relevant effect of hidden variables) and uncertainty of the available data (multimodal noisy data) [Thrun et al., 2005], [Ferreira and Dias, 2014c]. A new generation of computer architectures [Faix et al., 2015] and programming languages [Lebeltel et al., 2004] is also being developed to optimize and generalize the implementation of the methods described using this formalism.

The probabilistic methods proposed in this thesis follow the principles of the Bayesian probability theory.

2.1.1 Bayes rule

In this thesis, several descriptors such as robot state, multi-modal sensor measurements, and surrounding environment state, are represented by continuous or discrete random variables. Each variable is defined for a specific domain (possible values). A random variable or logical operation of random variables is characterized by a probability density function (continuous variables) or probability mass function (discrete variables), which assigns a probability ($[0, 1]$) to each value of the domain of the random variable. The work [Chung and AitSahlia, 2012] presents an extensive introduction to the basic concepts of probability theory.

Let C denote a random variable and c denote a specific value of the domain of C. In this abstract formulation of a problem, C represents a potential cause of an event of interest E (with e being a specific value of this variable).

During the probabilist modelling of a problem, the random variables establish statistical independence relations between them. If E and C are considered independent, then E does not influence C. This type of influence is modelled by a conditional probability, as shown in equation 2.1 .

$$P(C|E) = P(C) \tag{2.1}$$

However, in robotics it is common that a random variable carries information about other random variables. Considering that assumption and returning to the example presented previously, equation 2.2 can be formulated.

$$P(C|E) = \frac{P(C, E)}{P(E)} = \frac{P(E|C)P(C)}{P(E)} \tag{2.2}$$

Equation 2.2 describes the Bayes rule. It expresses the relation between $P(C|E)$ and its inverse $P(E|C)$. C expresses the quantity to be inferred, using the knowledge of evidence E. The factor $P(C)$ represents the *prior probability distribution*, expressing the information available about C before the incorporation of the evidence E.

The probability distribution $P(C|E)$ is denoted as the *posterior probability distribution*. It describes the knowledge of C after integrating the data E and the *a-priori* information about C. The element $P(E|C)$ represents the *likelihood probability distribution*, which expresses the knowledge about how the variable C influences E. In robotics, this factor is also termed the *generative model*. The *likelihood probability distribution* is determined analytically, or it can result from a training period. The data acquired during

this training period is used to learn the parameters of the probability distribution function $P(E|C)$. The literature [Ferreira and Dias, 2014b] presents different methods to perform the probabilistic learning of $P(E|C)$, such as *Maximum Likelihood* (ML) and *Expectation Maximization* (EM).

The factor $P(E)$ is a normalization constant. It guarantees that $P(C|E)$ sums up to 1, for all the domain of C. In some contexts, this parameter is not represented for simplicity purposes.

Additional details about the determination of $P(E|C)$ and $P(C)$ are given throughout this thesis as they are used to model each problem presented in this manuscript.

2.1.2 Bayesian inference

The *posterior probability distribution* $P(C|E)$ is used as a source of information to perform a decision about which state of C should be chosen. The work of [Ferreira and Dias, 2014a] presents different approaches to define the decision rule of the inference process.

In this thesis, the decision rule is formulated directly in the *posterior probability distribution* $P(C|E)$ by applying the *Maximum a-Posteriori* (MAP) principle.

The inferred value $\hat{c}$ of C is determined by selecting the argument of $P(C|E)$ which provides the highest value of probability, as presented in equation 2.3 .

$$\begin{aligned} \hat{c} &= \arg\max_{c} P(C|E) \\ \hat{c} &= \arg\max_{c} \frac{P(E|C).P(C)}{P(E)} \\ \hat{c} &= \arg\max_{c} P(E|C).P(C) \end{aligned} \tag{2.3}$$

2.1.3 Representing the Bayesian models

Throughout this thesis, several Bayesian models are formulated and described to provide a solution to the different challenges that are identified. In this manuscript, the Bayesian models are represented and characterized by using two different, but complementary, formalisms: Bayesian network and Bayesian program.

Bayesian network

A Bayesian network is a directed acyclic graph. The random variables of the Bayesian model are represented as nodes. The probabilistic (causal) relationships between pairs of

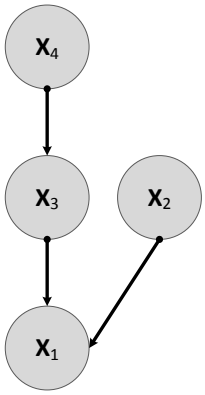


Figure 2.2: Graphical representation of a Bayesian model described by the random variables X_1, X_2, X_3, X_4 represented in the nodes. The causal dependencies are represented by the arrows (directed arcs).

random variables are represented as directed arcs. This representation approach provides an appealing visual description of the dependence relationship between random variables.

The dependence relationships expressed by the structure of a Bayesian network are used to simplify the formulation of the joint probability distribution function. These simplifications allow the design of efficient learning and inference algorithms based on simpler conditional probability distributions.

If the set of nodes which have arcs terminating at X_i is described by $parents(X_i)$, then equation 2.4 can be formulated. Let us consider a Bayesian model with N random variables $X_i, \ldots, X_N$ represented in a graph.

$$P(X_1, X_2, \ldots, X_N) = \prod_{i=1}^{N} P(X_i | parents(X_i)) \tag{2.4}$$

Equation 2.5 describes the joint probability distribution function for $N = 4$ and implementing the statistical causal dependence relations illustrated in the Bayesian network of Figure 2.2.

$$P(X_1, X_2, X_3, X_4) = P(X_1|X_2, X_3).P(X_3|X_4).P(X_2).P(X_4) \tag{2.5}$$

Bayesian program

The Bayesian program is a mathematical formalism and methodology used to organize and systemize the description of a Bayesian model. This approach facilitates the analysis and comparison of the properties of different Bayesian models and the respective

Bayesian Program

Description

Specification

- **Relevant variables:**
  Selection of the random variables to model the problem being studied.
- **Decomposition:**
  The global joint probability distribution function is factored as the product of simpler probability distribution functions, expressing the conditional independence relations of the random variables.
- **Parametric forms:**
  A custom mathematical function or a family of probability distribution functions (e.g. uniform, normal, Bernoulli, Poisson, Binomial) is associated to each of the factors identified during the decomposition stage.

- **Identification:**
  The free parameters of the custom mathematical functions or of the families of probability distribution functions are estimated/learned from experimental data acquired during training sessions (or other sources of information).

- **Question:**
  This stage of the formalism describes the inference problem to be answered by the Bayesian model. The Bayesian model is queried concerning information about a given random variable. The answer provided by the Bayesian model (posteriori probability distribution function) is submitted to a Bayesian decision theory criterion (e.g. MAP)

(a)

Figure 2.3: Schematic description of the organization of the formalism used by a Bayesian program to describe a Bayesian model.

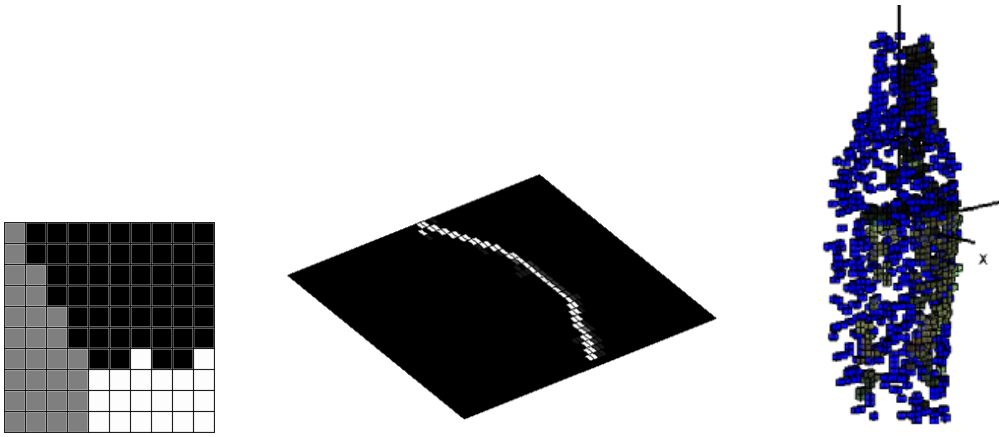
computational implementations.

This methodology represents the Bayesian model as it follows an approach consisting of different stages, which are described in Figure 2.3.

2.2 Probabilistic grids

In chapters 6 and 7 of this thesis, the environment surrounding the agent (robotic system or human) performing the haptic exploration is represented by a bi-dimensional probabilistic grid. The workspace is divided into a uniform grid of square cells.

A property of interest of the environment is associated to each cell with coordinates (i, j) and described by a random variable $X_{(i,j)}$. This approach considers that the repre-



(a) (b) (c)

Figure 2.4: Example of probabilistic grids used in robotics research. Probabilistic grid representing: a) the occupancy of a 2D environment [Rocha et al., 2005]. b) the haptic discontinuity between two regions of a surface [Martins et al., 2014]. c) the 3D shape of a hand explored object. Occupancy state is fused with color information from an artificial vision system [Faria et al., 2010a].

sentations of the cells of the grid are independent from each other.

The representation framework allows the integration/fusion of multi-modal data using probabilistic modelling techniques; this approach can deal with uncertainty of the data sources. The grid structure also provides the ability to represent heterogeneous environments (e.g. spatial discontinuities of the property of the environment being represented).

In several previous works (e.g. [Rocha et al., 2005], [Faria et al., 2010a], [Elfes, 1989]), the bi-dimensional probabilistic grids were typically used to represent the state of the workspace regions as *empty* or *occupied*. In this work, the state $X_{(i,j)}$ of each cell represents multivalued properties of the workspace regions, such as category of material.

The state $X^k_{(i,j)}$ of each cell is updated at each time iteration k by integrating new sensory measurements $Z^k_{(i,j)}$ acquired at that region of the workspace. At the time instant $k = n$, the cell (i, j) has integrated n sensory measurements $\mathbf{Z}^n_{(i,j)} = (Z^1_{(i,j)}, Z^2_{(i,j)}, \ldots, Z^n_{(i,j)})$. At that time instant, the state of each cell of the probabilistic grid is described by equation 2.6.

$$P(X^n_{(i,j)}|\mathbf{Z}^n_{(i,j)}) = \frac{P(Z^n_{(i,j)}|X_{(i,j)}).P(X^{n-1}_{(i,j)}|\mathbf{Z}^{n-1}_{(i,j)})}{P(\mathbf{Z}^n_{(i,j)}|\mathbf{Z}^{n-1}_{(i,j)})} = \Theta.P(Z^n_{(i,j)}|X_{(i,j)}).P(X^{n-1}_{(i,j)}|\mathbf{Z}^{n-1}_{(i,j)}) \quad (2.6)$$

The parameter Θ is a normalization constant. Consecutive sensory measurements $Z^k_{(i,j)}$ and $Z^{k-1}_{(i,j)}$ are considered independent.

According to equation 2.6, the updated representation of the state of each cell of

the grid $P(X^n_{(i,j)}|\mathbf{Z}^n_{(i,j)})$, after a new sensory measurement, is given by $P(Z^n_{(i,j)}|X_{(i,j)})$ and $P(X^{n-1}_{(i,j)}|\mathbf{Z}^{n-1}_{(i,j)})$.

The factor $P(Z^n_{(i,j)}|X_{(i,j)})$ represents the likelihood probability distribution function, which expresses the sensor measurements model. It models the knowledge available of how the sensor measurements are affected by the possible state of the cell/workspace (i, j).

Alternatively, the factor $P(X^{n-1}_{(i,j)}|\mathbf{Z}^{n-1}_{(i,j)})$ describes the state of the cell (i, j) at the previous time iteration $n - 1$. It encodes a complete summary of all past integration of sensory data by that cell of the grid.

2.3 Information theory and entropy

Several chapters of this thesis use random variables to model the proposed approaches and transfer information between the different modules. This PhD thesis uses Shannon entropy [Shannon, 2001a] to quantify the information encoded by a probability distribution function $P(X)$ of a random variable X. The formulation for discrete random variables is presented in equation 2.7.

$$H(X) = E[-log(X)] = \sum_x -P(x)\log_2(P(x)) \tag{2.7}$$

Higher values of entropy express lower levels of information (e.g. uniform probability distribution). Lower values of entropy encode higher levels of information (e.g. certain event).

Different approaches for calculating the entropy of continuous probability distribution functions are presented in [Gelfand and Yaglom, 1993].

Chapter 3

Dexterous manipulation and exploration: from Humans to robots

The human hand is a powerful tool. The dexterous manipulation and haptic exploration movements are some of the most useful strategies to interact with the surrounding environment.

Human manipulation and exploration capabilities result from a combination of a multimodal sensorimotor system [Flanagan et al., 2006] and a powerful effector apparatus. The conjugation and coordination of these two elements is controlled by the human nervous system following an action-perception loop architecture. Successful manipulation strategies require the capability to predict the appropriate motor commands to grasp, in-hand manipulate, transport, and release the object, and to predict and evaluate the sensory events caused by the motor commands.

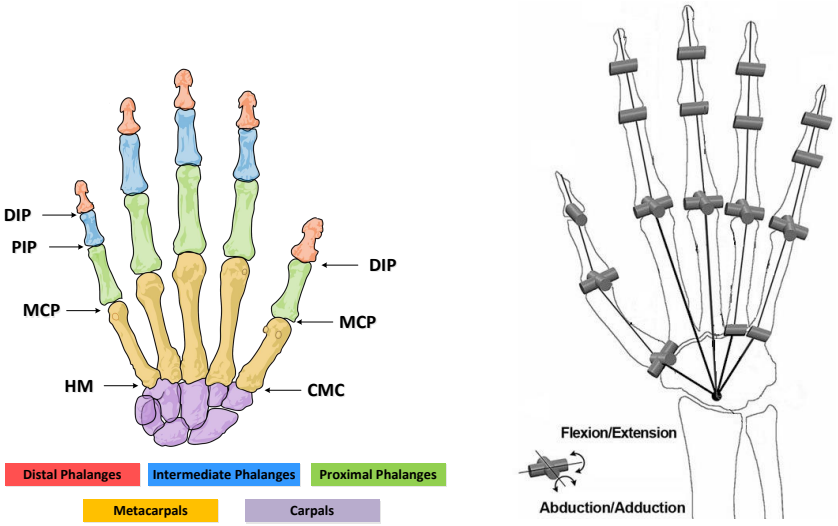
The following sections elaborate on some of the elements involved in the action-perception loop: specifically those related to dexterous manipulation and haptic exploration.

3.1 The human hand

3.1.1 Anatomical structure

The human ability to make and use tools to interact with the environment and other persons is one of the main evolutionary factors that distinguishes humans from other animals. The human hand plays a fundamental role in these capabilities [Jones and Lederman, 2006].

The musculoskeletal system of the human hand consists of bones (skeleton), muscles, tendons, ligaments, and joints. This system is responsible for maintaining the posture and shape of the hands, and providing the ability to move and to produce dexterous movements (interaction with the environment). A schematic representation of the bones



(a) (b)

Figure 3.1: Human hand. a) bones (represented by coloured regions) and joints (specified by arrows) [Wikipedia, 2014]. b) kinematic model of the human hand describing the joints, links, and degrees of freedom [Li et al., 2010].

and joints of the human hand is presented in Figure 3.1a.

Previous work [Lee and Kunii, 1995] proposed that the dexterity of the human hand is supported by a total of 23 internal degrees of freedom of movement. The degrees of freedom are provided by several joints. The elementary movements of the hand can be described by flexion/extension and adduction/ abduction movements. The flexion/extension movements correspond to the rotations toward and away from the palm. The adduction/abduction motions are used to describe the movement of joining and separation of the fingers.

A kinematic model of the human hand was proposed by [Lee and Kunii, 1995], and it is illustrated in Figure 3.1b. The index, middle, ring, and little fingers have four degrees of freedom. The distal interphalangeal (DIP) joint and proximal interphalangeal (PIP) joint have 1 degree of freedom, allowing flexion/extension movements. The metacarpophalangeal (MCP) joint has the remaining two degrees of freedom, allowing flexion/extension movements, as well as adduction/abduction movements. The mechanical structure of the thumb is different from the other four fingers. It has five degrees of freedom that are distributed by two joints: two at the carpometacarpal (CMC) joint (flexion/extension and adduction/abduction), two at the metacarpophalangeal (MCP) joint (flexion/extension and adduction/abduction), and one at the interphalangeal joint (flexion/extension). The curve and fold movement of the palm are accomplished by two internal degrees of freedom located in the transition region between the proximal region of the fingers and the palm.

Figure 3.2: Screenshot of the 3D virtual visualization tool implemented to demonstrate the kinematic model of the human hand. The software tool is available online [Martins, 2009a] (url: `http://www.rmartins.net/phd-docs/st01/`).

Several simplifications of the model may be implemented to reduce the number of free degrees of freedom, considering the objectives of the application for which the model is going to be used. One common simplification considers that the flexion/extension level of the distal interphalangeal joint of the fingers is dependent (by a 2/3 factor) on the proximal interphalangeal joint flexure/extension level [Lee and Kunii, 1995].

The hand is actuated by two distinct groups of muscle [Jones and Lederman, 2006]. The intrinsic muscles are thenar, hypothenar, interossei, and lumbrical muscles. The extrinsic muscles (long flexors and extensors) are located on the forearm.

Kinematic model of the human hand: 3D virtual demonstration tool

Using the the kinematic model of the human hand presented previously, a software tool was developed to demonstrate the kinematic model (various links, types of joints, and degrees of freedom) of a human hand. The software tool was implemented using *Python* programming language and the computer graphics library *Vizard VR Software Toolkit*.

The software tool allows the use of the graphical interface to select a joint and a degree of freedom, which is explored interactively by changing the flexure level by moving a slider. Users also can change the viewing perspective of the 3D avatar. The links are represented by blue lines and the joints by red dots.

The *Python* source code and a compiled version (Windows operating system) of the software tool are available online [Martins, 2009a] (url: `http://www.rmartins.net/phd-docs/st01/`). A screen shot of the graphical interface of the software tool is pre-

sented in Figure 3.2. The technical report *Modelling the Human body and hand: kinematic structure, degrees-of-freedom* [Martins, 2008] provides additional details about the kinematic models of the human hand and body. The proposed models were used to build the 3D virtual tool. The technical report also is available online (url: `http://www.rmartins.net/phd-docs/tr01/`).

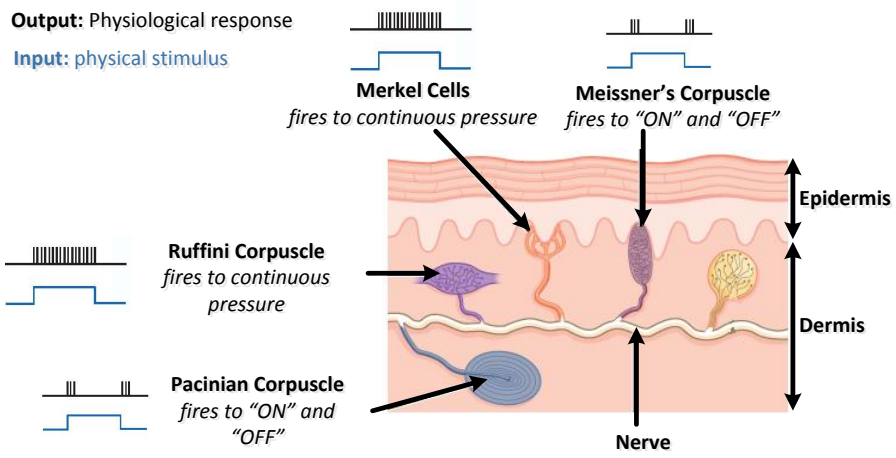


Figure 3.3: Schematic representation of the four classes of mechanoreceptors which can be found on the skin of the human hand. Summary of the different functions [Goldstein, 2002]

.

3.1.2 Sensing apparatus

The somatosensory system is a diverse sensory system that contains the cutaneous senses. It is responsible for perceptions such as touch, which are typically caused by the mechanical stimulation of the skin; proprioception, which is the ability to sense the position of the body and limbs; and kinesthesis, which is the ability to sense the movement of the body and limbs.

The human sense of touch is distinct from the other four sensory modalities (vision, olfaction, taste, and hearing) in several ways. Receptors for touch are varied and have different distributions throughout the skin, unlike other sensory modalities (e.g. hearing) which are confined to localized and specialized structures on the human body. The perceptual apparatus (receptors, neural pathways, etc.) mediating touch responds to different types of stimulation.

Touch is involved in the perception of external stimuli through different types of sensory receptors [Goldstein, 2002]. Mechanoreceptors are involved in the discrimination of stimuli related to pressure and vibration. Nociceptors are related to the sensing of pain.

However, nociceptors also respond to intense sensory stimulation of heat and pressure, which are associated with extreme conditions of temperature and mechanical stimulation. Thermoreceptors participate in the sensing of heat transfers.

The research works presented in this thesis are dedicated to the development of artificial perception mechanisms that model the tactile perception of mechanical stimulus during the manipulation and exploration of objects.

The different types of mechanical stimulation (pressure, stretching, and vibration) are detected by different categories of mechanoreceptors, which are represented in Figure 3.3: Ruffini cylinder, Merkel receptor, Meissner corpuscle, and Pacinian corpuscle. The distinction between mechanoreceptors is made by the location on the anatomical structure of the skin. They differ in terms of spatial acuity, dimension of receptive field, and the manner in which the nervous fibers associated with each mechanoreceptor respond to stimuli with different mechanical characteristics. The properties of the response can be associated with a specific function during the execution of dexterous manipulation and haptic exploration tasks. Table 3.1 summarizes the characteristics and functions of the different mechanoreceptors. The work [Romano et al., 2011] proposed a methodology to model and implement the different tactile sensing capabilities of the mechanoreceptors using the PR2 robotic platform equipped with a tactile sensing array on the gripper of the robot.

Table 3.1: Mechanoreceptors: physiological integration, stimulation, and function [Dahiya et al., 2010]

Classification Basis	**Pacinian Corpuscle**	**Ruffini Corpuscle**	**Merkel Cells**	**Meissner's Corpuscle**
Type	FA II	SA II	SA I	FA I
Adaptation Rate	Fast	Slow	Slow	Fast
Spatial Acuity (mm)	10+	7+	0.5	3-4
Vibration (μm)	0.01	40	8	2
Indentation Threshold (μm)	0.08	300	30	6
Stimuli Frequency (Hz)	40-500+	100-500+	0.4-3	3-40
Effective Stimuli	Temporal changes in the skin deformation	Sustained downward pressure. Lateral skin stretch. Skin slip.	Spatial deformation. Sustained pressure. Curvature, edge, corner.	Temporal changes in skin deformation.
Sensory Function	High-frequency vibration detection. Tool use.	Finger position. Stable grasp. Tangential force. Motion direction .	Pattern/form detection; texture perception. Tactile flow perception.	Low-frequency vibration and motion detection. Grip control. Tactile flow detection.

The fibers associated with the sensory receptors, which are integrated on the skin, follow a path to the somatosensory area of the cortex, as illustrated in Figure 3.4. The fibers conducting the electrical signals from the sensory receptors enter the spinal cord through the dorsal root. The signals travel through the spinal cord along two major pathways: the medial lemniscal pathway and the spinothalamic pathway.

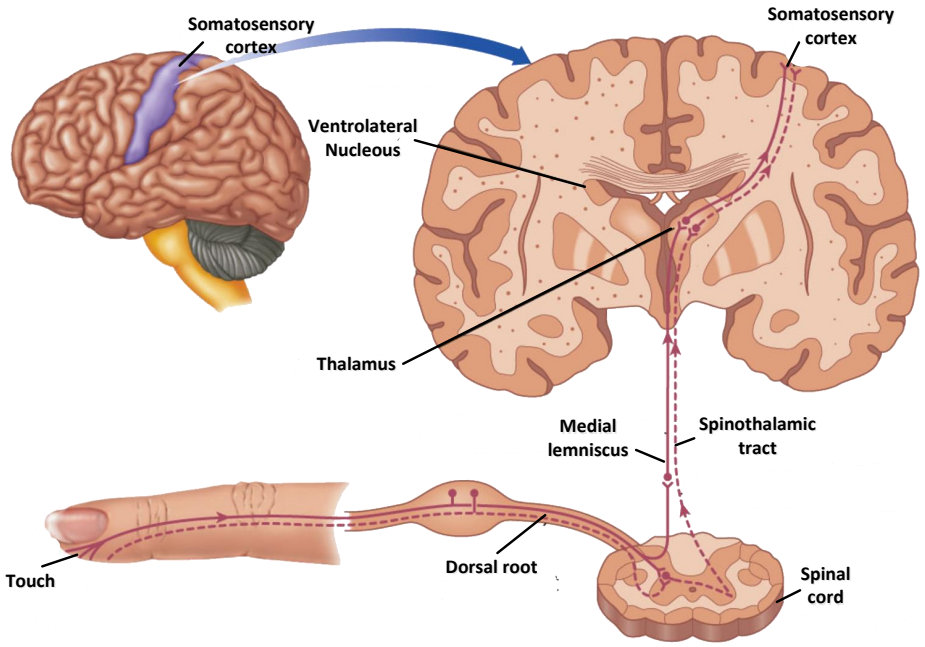


Figure 3.4: Somatosensory pathways: from fingertip to cortex [Goldstein, 2002]

.

The lemniscal pathway transports signals related to sensing proprioception and kinesthesis (the ability to sense the position of the body and limbs and the ability to sense the movement of the limbs, respectively) and perceiving touch (mechanical stimulus). The spinothalamic pathway transmits signals related to temperature and pain. Both pathways synapse in the ventrolateral nucleus of the thalamus and then send fibers to the somatosensory cortex in the parietal lobe [Goldstein, 2002].

3.2 Planning and control of dexterous manipulation tasks

When humans intend to perform a manipulation task with an object placed in the surrounding environment, the head and eye gaze are oriented toward the target object. The category of object and the spatial pose are analysed to plan the configuration of hand and arm during the reach-to-grasp or transport movement. During the contact of the hand with the object, the proper amount of force must be applied by the fingers to avoid dropping the object and breaking it [Johansson and Flanagan, 2009].

The mechanisms involved in the visual and somatosensory processes responsible for perception and action during manipulation and haptic exploration tasks are detailed in sections 3.2.1 and 3.2.2.

3.2.1 Reach-to-grasp and transport movements

A model was proposed in [Oztop and Arbib, 2002] that described the neuronal mechanisms and functional relations involved in the planning and control of execution of reach-to-grasp and transport movements.

The online visuo-motor planning and control of the movement integrate visual information provided by two main pathways [Goodale and Milner, 1992]. The ventral pathway (known as *what* stream) is related to recognition, categorization, and assessment functions. The dorsal pathway (known as *where/how* stream) is dedicated to the estimation of the position, orientation, and shape of the target object of the manipulation task.

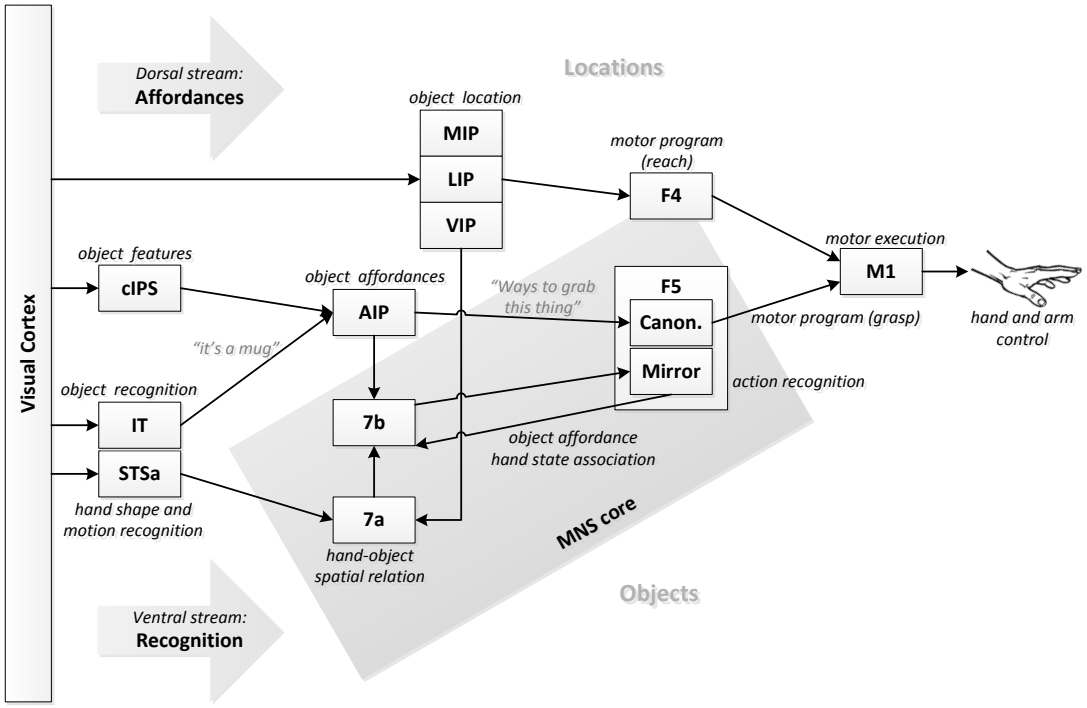


Figure 3.5: Oztop and Arbib model [Oztop and Arbib, 2002] describing the role of the mirror-neuron system in reach-to-grasp manipulation movements guided by vision [Arbib et al., 2008]

.

The ventral and dorsal visual information streams work closely together in the planning and control of the dexterous manipulation movement. They also are closely associated with the somatosensory system in both programming and executing this control [Oztop and Arbib, 2002] (e.g. object modelling and recognition, initial hand configuration and grasping forces, and online movement evaluation). These relations are detailed in section 3.2.2.

The model proposed by the work [Oztop and Arbib, 2002] also integrates the represen-

tation of a mirror neuron system (MNS). Their work extends the previous FARS model (Fagg - Arbib - Rizzolatti - Sakata) [Fagg and Arbib, 1998]. The mirror neuron system, initially discovered in macaque monkeys, is a fundamental element to understanding the neuronal processes involved in the reaching and grasping movements of human manipulation of objects. The MNS links the visual processes of the superior temporal sulcus (STS) to the parietal regions and premotor regions (F5). Recently, it was demonstrated that the premotor regions (F5) integrate neurons involved with grasping capabilities. The two main classes of neurons on F5 are the mirror neurons and the canonical neurons. The mirror neurons discharge when the monkeys observe other monkeys hand movements which are similar to those whose execution is associated with the firing of the neuron. The canonical neurons fire when the monkey performs a specific action and also when the monkey sees a possible target object of such action. The canonical neurons do not fire when the monkey sees other monkeys or humans performing that action.

Object features extracted by cIPS (caudal intraparietal sulcus), together with the object recognition output provided by IT (inferotemporal cortex), task analysis and work memory, are processed by AIP (anterior interparietal area) to extract grasp affordances. The object affordances are sent to the canonical neurons of F5 that choose a particular prehensile pattern to grasp the object. The location of the object is estimated by MIP (medial intraparietal sulcus), LIP (lateral intraparietal sulcus) and VIP (ventral intraparietal sulcus). These regions provide parameters to the motor programming area of F4 which estimates the characteristics of the reach movement. The motor cortex M1 (F1) integrates the grasp and reach information provided by the canonical neurons of F5 and F4 regions, respectively, and coordinates the execution of the proposed motor program.

The quality of the execution of the defined motor program is controlled. The mirror neurons of F5 recognize the grasping actions being performed to the object, while MIP, LIP, and VIP provide outputs related to the object location, which are combined with hand shape and motion recognition provided by STSa (superior temporal sulcus) in 7a, providing a description of the hand/object spatial relation. The evaluation of the spacial relation between hand/object, action being performed with the object, and selected object affordance is made in 7b, which receives inputs from F5 mirror neurons, 7a, AIP, and STSa. The output of 7b is transmitted to the F5 in order to, if required, readjust the current motor program being executed.

3.2.2 In-hand manipulation and haptic exploration movements

The work [Dijkerman and Haan, 2007] proposes a model, presented in Figure 3.6, to describe the somatosensory processing mechanisms. The model describes a parallel processing approach, which is somewhat analogous to the parallel processing described previously

that is happening in the visual system.

Two streams are proposed. A "what" pathway involved in the processes related to somatosensory perception and a "how" pathway participating in action mechanisms. As illustrated in Figure 3.6, both pathways start in the thalamus, where the main somatosensory inputs concerning touch and proprioception terminate. The initial somotosensory processing stages occur in SI (primary somatosensory cortex). Simple and complex features are extracted (location and duration). Different theories about the extraction of haptic features have been proposed and are the object of several ongoing studies. The two main approaches describing different mechanisms integrating sensory data are illustrated in Figure 3.7.

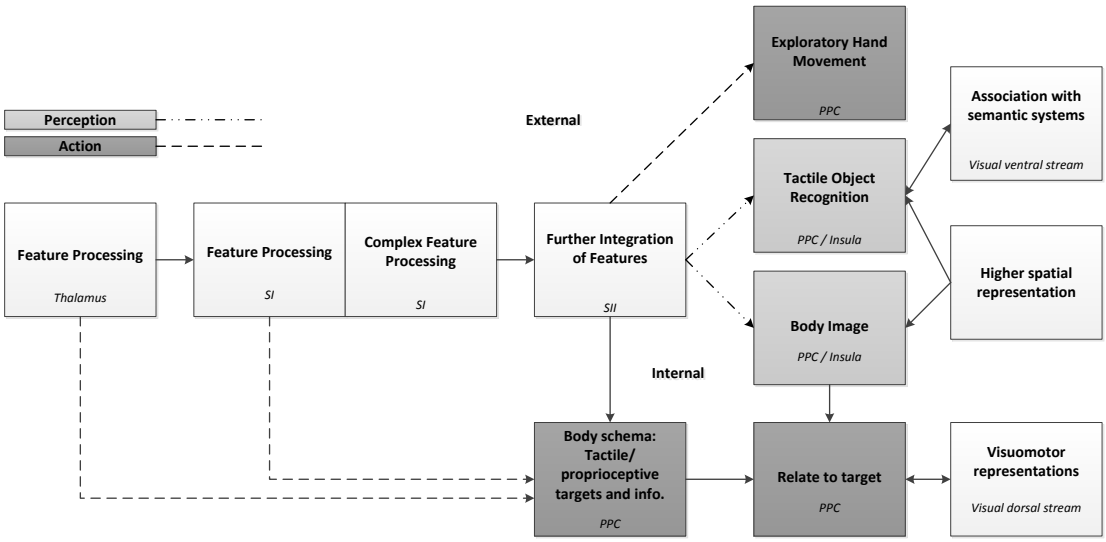


Figure 3.6: Schematic representation of the Dijkerman model [Dijkerman and Haan, 2007], describing the somatosensory processing pipeline occurring in the cortex, supporting action and perception mechanisms.

After this common initial stage, the two different somatosensory processing pathways diverge, according to the stimulus characteristics and the purpose of processing.

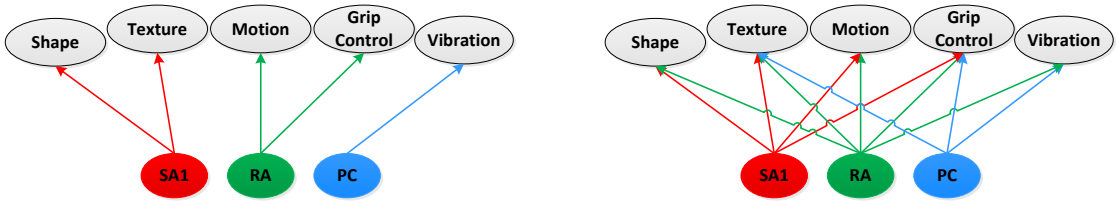
The "what" pathway projects from SI via SII (secondary somatosensory cortex) to the insula. PPC (posterior parietal cortex) is involved in the spatial and temporal integration of information (velocity of the stimulus). This pathway provides mechanisms to recognize objects using somatosensory data.

The "how" pathway projects from SI to SII, terminating in PPC. This pathway is involved in the planning and control of action mechanisms. PPC is subsequently involved in crossmodal integration (interaction with the vision system) and the preparation of movements.

The "what" (perception) and "how" (action) pathways are not as independent as they

are in the vision system. Haptic perception requires a close cooperation between action and perception, because the sensing apparatus needs to contact directly with the stimulus / object.

The features extracted during the somatosensory processing pipeline can be integrated in high-order association areas to infer complementary properties of objects. All the mechanisms described previously are regarding the perception of external haptic stimulus (environment surrounding the subject).



(a) (b)

Figure 3.7: Integration of somatosensory afferent inputs during the haptic features extraction pipeline. a) submodality segregation. b) submodality convergence [Saal and Bensmaia, 2014]

.

At a behavioural level, the tactile recognition of external objects requires close coordination between action-related and perception-related somatosensory processes. This model also demonstrates that the ventral and dorsal streams of the visual system work closely together not only in the programming and control of skilled manipulation movements, but they also have intimate associations with the somatosensory system in both programming and executing this control [Dijkerman and Haan, 2007].

As suggested in [Okamura et al., 2001], the in-hand exploration strategies and grasping movements used by humans are inherently coupled, as shown in Figure 3.8. The execution of the manipulation movements requires an object model to plan the grasping configuration of the hand. By having a more complete and accurate model of the object, the planning of the movements performed with the object can be controlled precisely. If the model of the object is incomplete, typically humans improve the object model by performing in-hand exploration of the object. However, manipulation is required for exploration. To be able to use one or more fingers to explore the surface of the object by moving the fingers over the uncharacterised regions of the object (such as regions not accessible to vision), two or more fingers are required to stabilize the object by grasping it.

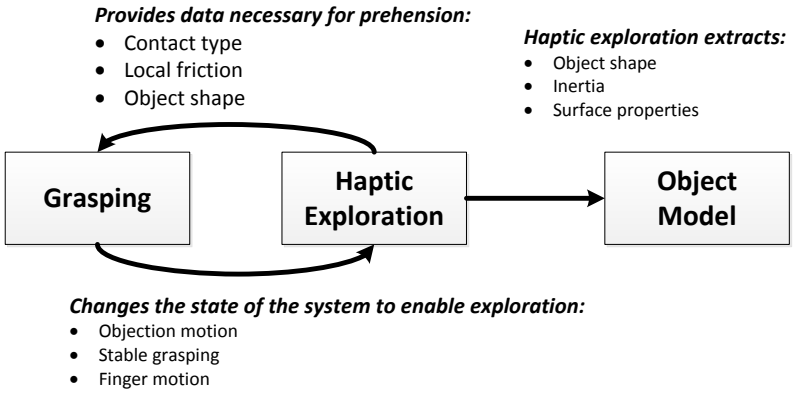


Figure 3.8: Interdependence and complementarity between manipulation and haptic exploration during the representation and progressive update of object model [Okamura et al., 2001]

.

3.3 Attention mechanisms in somatosensory system

Humans are integrated and interact with the surrounding environment, which provides multiple sources of noisy sensory data. The sensory overload is not all sensed, perceived, and processed by the human sensory and neural systems [Corbetta and Shulman, 2002] [Petersen and Posner, 2012].

By definition, the attention mechanisms have been identified as the set of processes which allow humans to focus the concious awareness and limited processing resources of the brain on the relevant part of the enormous multi-modal sensory inputs. The attention mechanisms are implemented at different levels of the sensory processing and cortical pipelines. At the physiological level, attention mechanisms are manifested by different processes such as enhanced synchronization, scale of magnitude of response, and variation of thresholds.

The processes involved in the attention mechanisms can be categorized differently, depending on the effect being analysed. Attention mechanisms can be considered top-down (endogenous) or bottom-up (exogenous) [Shipp, 2004]. Top-down attention mechanisms suppress, enhance, or prioritize the representation of sensory information, taking into consideration only the objectives of the task (e.g. haptic exploration of a surface, searching for a cold object, and ignoring the textures of the surface). Bottom-up attention mechanisms prioritize and mobilize resources to analyse unexpected external events and salient sensory signals (e.g. vibration of a cell-phone in the pocket while the subject is resting).

A complementary approach to categorizing the attention mechanisms considers the domain of the sensory inputs being integrated: space-based vs. feature-based attention

mechanisms. Space-based attention mechanisms concern processes that enhance or suppress in different ways the sensory information coming from sensing apparatus located in distinct regions. A typical example demonstrating the activation of these mechanisms is the exploration of a surface using the fingertips of the human hand. Stimuli presented to the fingertips are processed faster, with higher priority and detail than stimuli presented in the back of the hand (clothes, watch, bracelet) during the exploration movements. Feature-based attention mechanisms enhance or suppress processes involved in the analysis of specific features of the incoming sensory data (e.g. search and recognition of faces in vision; search for a smooth surface during a haptic exploration task).

The attention mechanisms in vision [Amso and Scerif, 2015] have been studied widely. In the touch / somatosensory system, research is not as advanced. Many research topics are still undergoing active and intense research [Chapman, 2009], [Mller and Giabbiconi, 2008]. This section summarizes the relevant processes involved in the touch attention process and how these mechanisms are integrated in the somatosensory processing pipeline, influencing the perception and action behaviours.

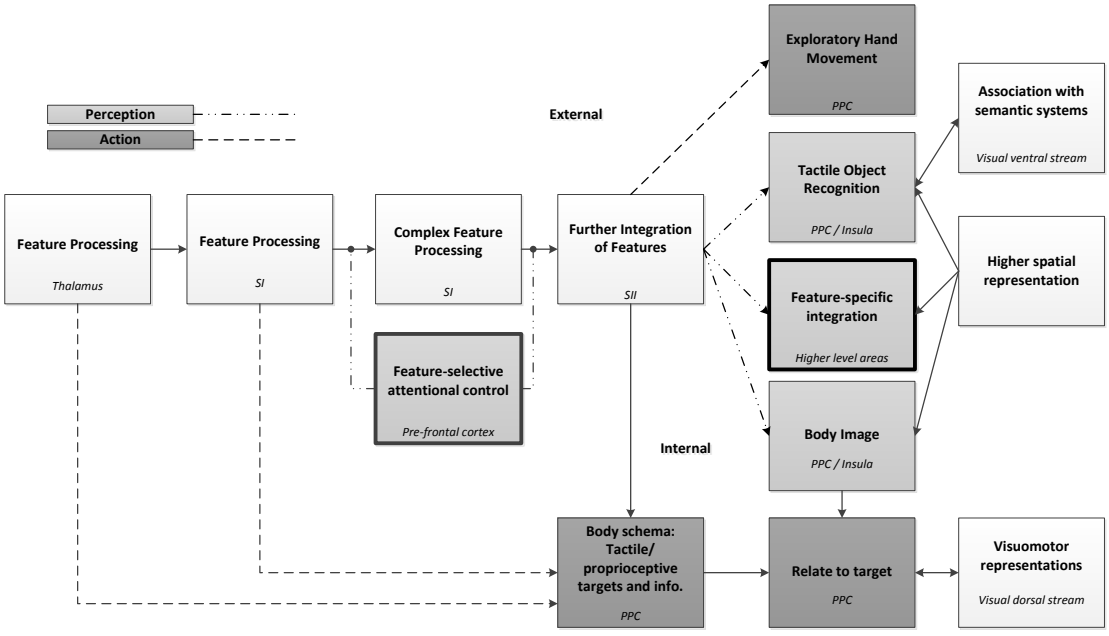


Figure 3.9: Integration of the tactile attention mechanisms (feature-based) proposed by [Wacker, 2011], in the Dijkerman model [Dijkerman and Haan, 2007].

In robotics, as with humans, the artificial sensing elements of touch can be integrated in distinct locations of the robotic system (palm, fingers, torso, arms, feet, legs) and with different extensions and densities [Dahiya et al., 2010]. The selection and hierarchy of the sensory inputs coming from all different sources would be modelled by artificial space-based attention mechanisms. However, in this work, the touch sensing elements are only

integrated in the fingertips of the robotic hands. Thus, space-based attention mechanisms are not studied in this thesis. The focus of this section is on the feature-based attention mechanisms.

The feature-based touch-attention mechanisms have been studied using fMRI [Wacker, 2011] . The authors focused the research on the analysis of the role of feature-specific high-order areas for tactile perception and the functional impact of top-down modulation in the processing of haptic features. [Wacker, 2011] extended the model of somatosensory processing for perception and action proposed by [Dijkerman and Haan, 2007] that was presented in section 3.2.2. The extensions are highlighted in bold in Figure 3.9.

In the box highlighted on the left side of Figure 3.9, the top-down modulation of touch attention is represented by signals coming from the pre-frontal cortex. These signals modulate the processing streams between somatosensory areas. Each processing stage of somatosensory features is modulated by feature selective attention, which promotes the extraction and processing of haptic features relevant for the objectives of the haptic exploration or in-hand manipulation tasks. The features considered relevant are extracted and transmitted to feature-specific and/or posterior parietal areas (PPC) to improve the perceptual representation of the object being explored.

The box highlighted in bold on the right side of Figure 3.9 integrates in the model [Dijkerman and Haan, 2007] high-order areas responsible for feature-specific processing. These areas fuse multi-modal haptic attributes in a modality independent representation of the object being perceived (e.g. representation of texture of object, by integrating haptic and visual information).

The inhibition-of-return (IOR) is a behaviour effect that has been described in several works studying perception and action, with seminal work by [Posner and Cohen, 1984]. This effect was identified across different sensory modalities. The work [Klein, 2000] presents an extensive review. The IOR is described as a simple, but powerful, effect. When a salient stimulus is presented at a specific location, the attention mechanisms promote the perception and processing of that stimulus. However, if consecutive stimulation is presented in that region, the enhanced perception of the first stimulus inhibits the processing of the next ones.

The subject of IOR mechanisms in touch is an active research topic. Different theories are proposed to explain the mechanisms underlying IOR effect. The most accepted view describes the IOR as an effect caused by the attention mechanisms which inhibit attention to return to a region attended previously. Alternative or complementary formulations were proposed, presenting mechanisms at the sensory perception or motor control that caused the IOR effect. Studies [Jones and Forster, 2014] [Jones, 2011] present an extensive review of this active research topic: IOR on touch.

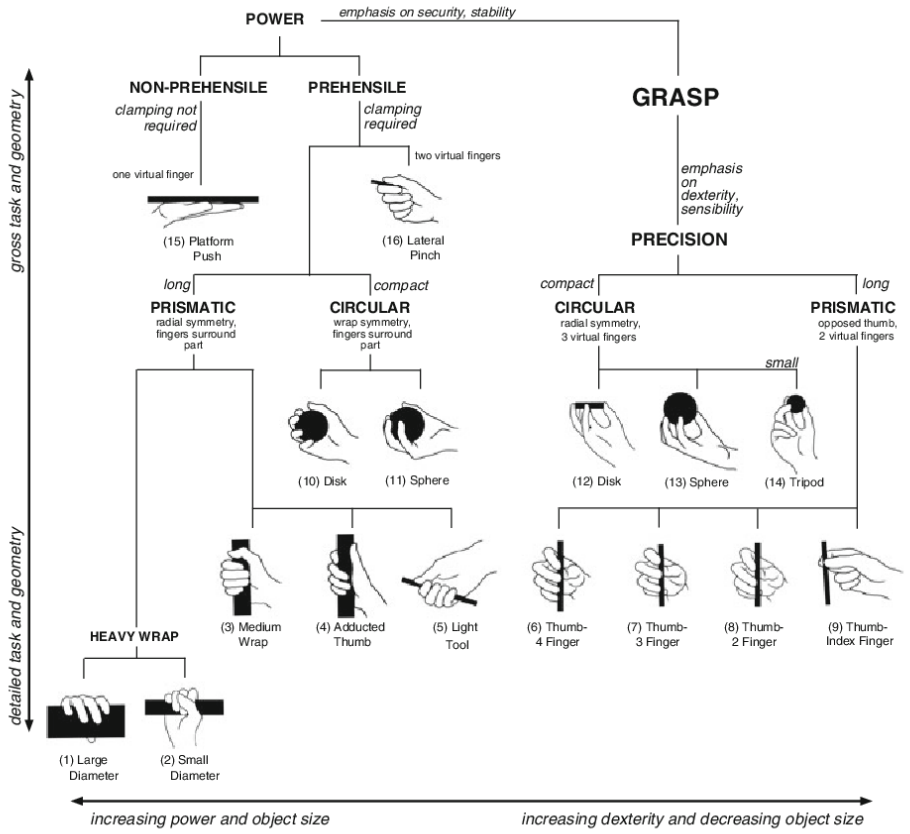


Figure 3.10: Grasp taxonomy proposed by Cutkosky [Cutkosky, 1989].

3.4 Categorization of manipulation movements

3.4.1 Grasping patterns

The dexterity of the human hand provides the capability to perform many different grasp types. This set of possible grasp types has been analysed to model the human strategies used to grasp different objects. These studies have been been applied in multiple contexts: human physical therapy and rehabilitation; and development and evaluation of artificial robotic and prosthetic hands. Several grasping taxonomies have been proposed in the literature. In general, the taxonomies are derived from a statistical study of human hand movements while performing the tasks of a typical day.

[Napier, 1956] characterized the human grasping movements in two general categories: power grasps and precision grasps. In the power grasp, the object is held inside a palm by opposition of the fingers, which exert a pressure on the object. In the precision grasp, the object is grasped by the fingertips. The power grasp is used to stabilize the grasped object. The precision grasp is used in tasks where precision and dexterity are fundamental

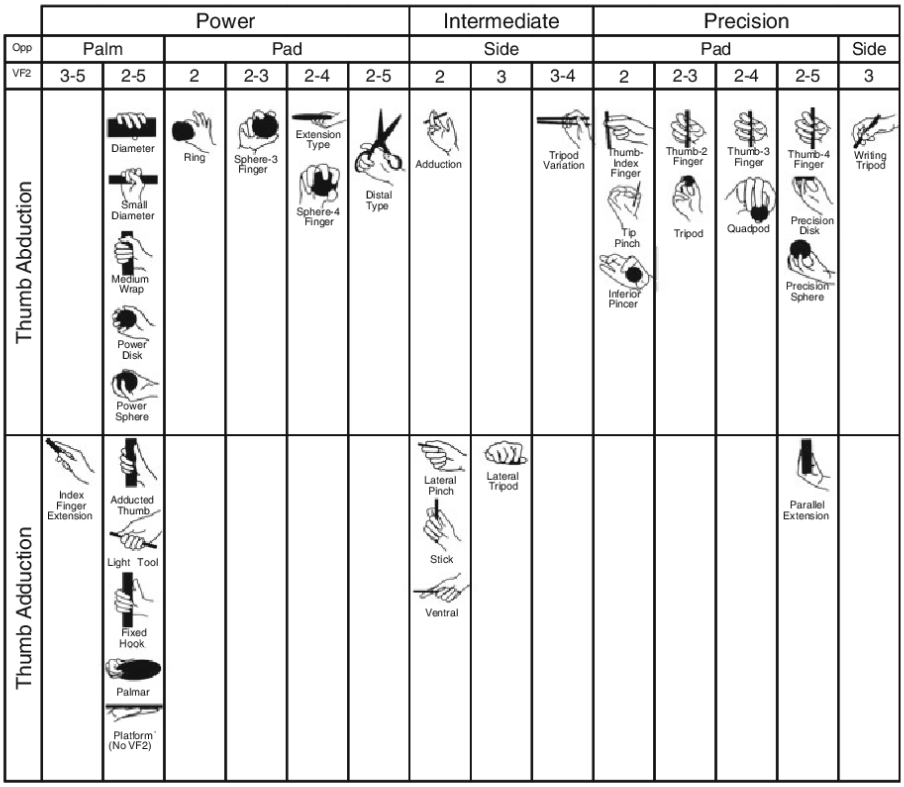


Figure 3.11: Grasp taxonomy proposed by the GRASP project [GRASP, 2008]

.

to the success of the task.

[Kamakura et al., 1980] proposed fourteen grasping patterns under four categories. The categorization of the prehensile postures of the hand is made considering the different regions of the hand contacting the object. Each tactile signature is the consequence of the mechanical configuration of the hand and the implicit characteristics of the object being held.

[Cutkosky, 1989] proposed a hierarchical taxonomy suitable to be integrated in grasp planning algorithms for robotic hands (Figure 3.10). The search for the most appropriate grasp type requires information about task requirements and object shape.

The GRASP project [GRASP, 2008] reviewed 14 different works, identifying a total of 33 different grasping patterns. Based on the properties of the grasping pattern, the grasps were rearranged in a new grasping taxonomy, represented in Figure 3.11. Each grasp in that taxonomy is characterized by parameters such as the type of grasp, opposition type, and thumb position.

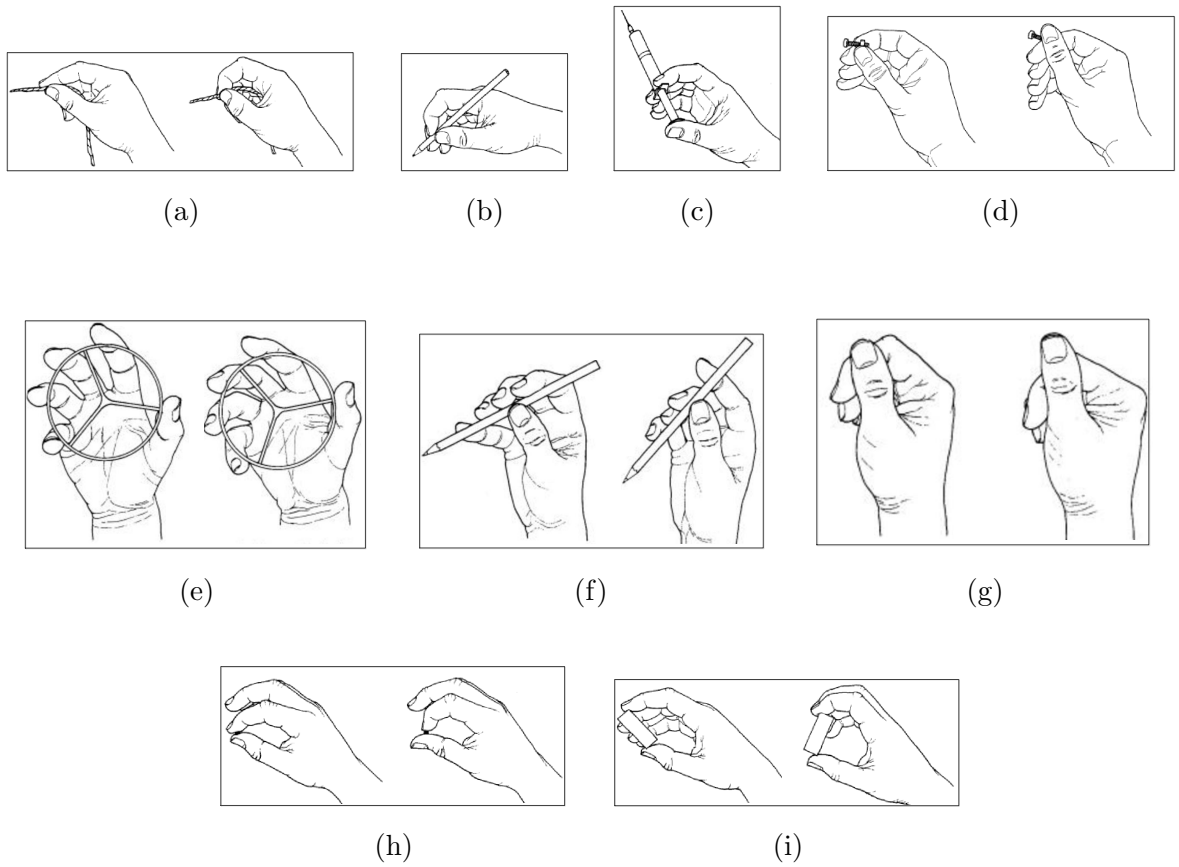


Figure 3.12: Taxonomy proposed by Elliot and Connolly [Elliott and Connolly, 1984] describing the different types of in-hand manipulation patterns. a) Pinch. b) Dynamic tripod. c) Squeeze. d) Twiddle. e) Rock. f) Radial roll. g) Index roll. h) Full roll. i) Rotary step.

3.4.2 In-hand manipulation patterns

The in-hand manipulation movements are skills performed by humans to manipulate the objects within the hand after an initial grasp. This skill has been studied by different research areas (e.g. human development researchers, occupational therapists, and robotic researchers). Several reference studies have tried to describe, categorize, and test the high diversity of possible in-hand movements performed by the human hand.

[Exner, 1992] developed a taxonomy to classify the in-hand manipulation movements. Exner described five classes of in-hand manipulation movements: finger to palm translation, palm to finger translation, shift, simple rotation, and complex rotation. Pont [Pont et al., 2009] extended Exners classification system with additional categories of movements: finger to palm translation to achieve stabilization, palm to finger translation, simple shift, complex shift, simple rotation, and complex rotation.

Elliot and Connolly [Elliott and Connolly, 1984] extended the previous works and pro-

posed a system to classify the intrinsic hand movements used to manipulate an object within the hand. The study proposed four categories of intrinsic in-hand manipulation movements: simple synergies, reciprocal synergies, sequential patterns, and palmar combinations. For each category of movements, several individual movement patterns are suggested, involving different fingers.

Table 3.2 and Figure 3.12 describe the different classes of in-hand manipulation movements proposed in [Elliott and Connolly, 1984] by Elliot and Connolly .

Table 3.2: Description and demonstration of in-hand manipulation patterns [Elliott and Connolly, 1984]

In-hand manipulation patterns	**Description**	**Demonstration**
Pinch	The object, typically small, is held between the pulp surfaces of the opposed thumb and index finger.	Figure 3.12a
Dynamic tripod	The object is grasped between the radial distal surface of digit and the pulp surfaces of the thumb and index.	Figure 3.12b
Squeeze	With the thumb opposed and all the digits relatively extended, the object is squeezed by synergetic flexion of the digits.	Figure 3.12c
Twiddle	Abduction of the thumb combined with metacarpophalangeal extension and some ulnar deviation of the index.	Figure 3.12d,
Rock	The pattern described in twiddle is executed with the thumb fully opposed and with increasing recruitment of the ulnar digits, moving in synchrony with the index but having a greater excursion.	Figure 3.12e
Radial roll	With the thumb in relatively slight opposition, the object is rolled between the ball of the thumb and the radial surface of the distal phalanx of the index finger.	Figure 3.12f
Index roll	With the thumb in full opposition, the object is rolled between the pulp surfaces of the thumb and index finger.	Figure 3.12g
Full roll	The pattern is similar to Index Roll, but using additional digits, often all five.	Figure 3.12h
Rotary step	The object is stepped around, with brief pauses while the position of the digits is readjusted, using sequential or phased set of movements.	Figure 3.12i

3.4.3 Haptic exploration patterns

The characterization of objects using the hand is not a passive process [Gibson, 1962]. The object is explored actively, promoting the interaction between fingers and palm with the object, building a perceptual representation of the object. The results of psychophysical studies suggest that the movements of the fingers and palm are not performed randomly, but depend on the characteristics of the object being extracted. [Lederman and Klatzky, 1987] observed that when subjects were asked to discriminate a specific characteristic of the object (e.g. texture, hardness, or weight), different types of movements, named haptic exploration patterns, were identified. The stereotypical exploration patterns and the corresponding characteristics are represented in Figure 3.13 and described in Table

3.3. Recently, the invariance of exploration patterns was verified and quantified using motion-tracking, force, and tactile sensing technologies [Jansen et al., 2013].

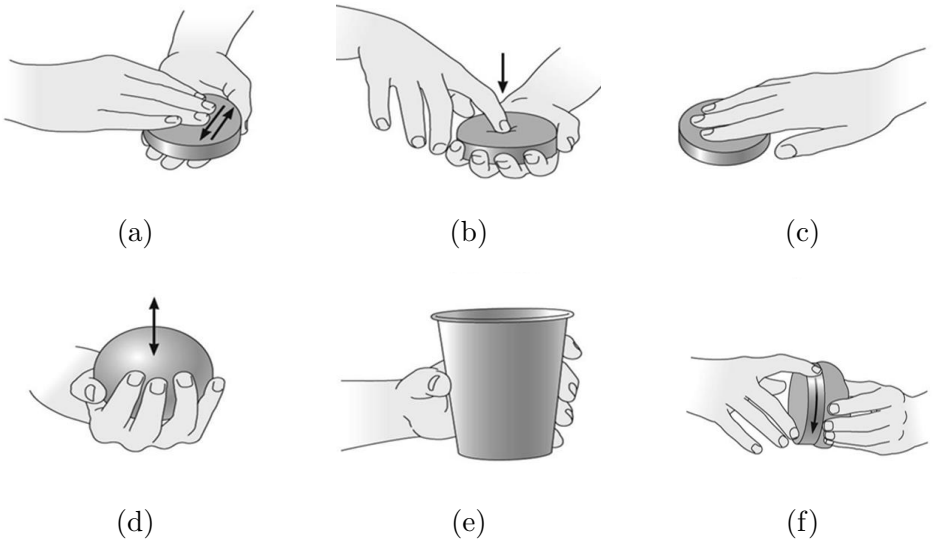


Figure 3.13: Taxonomy describing the haptic exploration patterns used to extract object features [Lederman and Klatzky, 1987]. a) lateral motion. b) pressure. c) static contact. d) unsupported holding. e) enclosure. f) contour following.

3.5 Benchmarking robotic manipulation and exploration skills

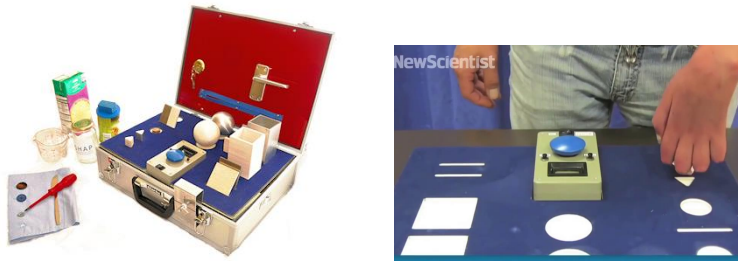
Different research and industrial sectors (e.g. informatics, manufacturing, and computer vision) have been interested in implementing benchmark methodologies, and in some of those fields, the benchmarks are established, widely recognized, and accepted. Due to an increase in research projects, diversity of robotic platforms, and industrial development, the robotics community is actively discussing and drafting benchmark protocols dedicated to robotic dexterous manipulation [Bonsignorio and del Pobil, 2015].

The definition of the benchmark protocols in robotics is a challenging task. The field of robotic dexterous manipulation integrates multiple research areas (advanced sensors, mechanical design, planning, control, and artificial perception) and can be applied in very different scenarios, executing multiple tasks with specific objects [del Pobil et al., 2014]. Currently, two main approaches are being followed and used in complementary ways to develop the benchmark protocols for robotic dexterous manipulation.

One of the proposed approaches employs robots for the existing clinical protocols to evaluate and score the human dexterous manipulation skills and capabilities.

Table 3.3: Description and demonstration of the haptic exploration patterns [Lederman and Klatzky, 1987]

Exploratory strategy	**Object property**	**Movement**	**Demonstration**
Lateral motion	Surface texture	The skin is passed laterally across a surface, producing shear force.	Figure 3.13a
Pressure	Compliance or hardness	Force is exerted on the object against a resisting force; for example, by pressing into the surface, bending the object, or twisting.	Figure 3.13b
Static contact	Apparent temperature	The skin surface is held in contact with the object surface, without motion; typically a large surface (like the whole hand) is applied. This EP gives rise to heat flow between the skin and the object.	Figure 3.13c
Unsupported holding	Weight	The object is held while the hand is not externally supported; typically this EP involves lifting, hefting, or wielding the object.	Figure 3.13d
Enclosure	Volume; Global shape	The fingers (or other exploring effector) are molded closely to the object surface.	Figure 3.13e
Contour following	Exact shape	Skin contact follows the gradient of the object's surface or is maintained along edges when they are present.	Figure 3.13f



(a) (b)

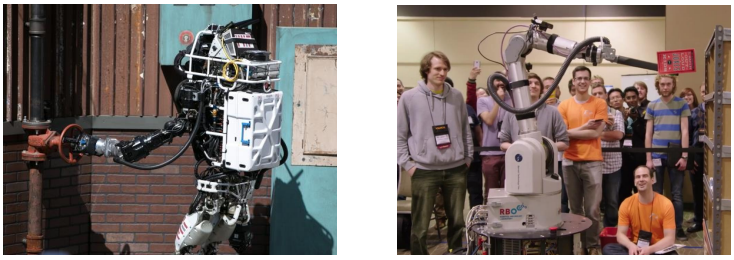
Figure 3.14: Southampton Hand Assessment Protocol (SHAP) [Adams et al., 2009]: a) complete kit. b) demonstration of assessment of hand function using the kit.

In clinical practice [Pont et al., 2008], different methods have been proposed to evaluate in-hand manipulation skills: the in-hand manipulation test quality [Miles Breslin and Exner, 1999], the test of in-hand manipulation (TIHM) [Case-Smith, 2000], and observation protocol on in-hand manipulation and functional skill development [Humphry et al., 1995]. An extensive review of the dexterity assessment of human hands in clinical practice is presented by [Yancosek and Howell, 2009] .

The Southampton Hand Assessment Procedure [Adams et al., 2009] SHAP is a clinical test originally developed to assess the physical functionality of hand and arm prostheses. However, it has been extended to evaluate the hand skills in humans. The SHAP consists of a kit of general objects and proposes a protocol to evaluate the hand functions during the execution of fourteen daily living activities. The participants are

asked to perform the tasks using one of six different grasps. The kit includes software which generates a score to evaluate the impairment level of the hands.

Additionally, the protocols ARAT [Yozbatiran et al., 2008] (a standardized approach to performing the action research arm test) and GRASSP [Kalsi-Ryan et al., 2012] (development of the graded redefined assessment of strength, sensibility, and prehension) are used to access and score the function of upper limbs and the human hand.



(a) (b)

Figure 3.15: a) ATLAS robot [ATLAS, 2013] performing a manipulation task during the DARPA robotic challenge. b) Robotic platform during the Amazon Picking Challenge [Wurman and Romano, 2015].

In addition to the clinical tests, researchers of the robotics community started implementing and discussing benchmark protocols specifically designed for robotic platforms. During scientific forums such as EURON (EUropean RObotics research Network), the robotics community proposed three different scenarios and benchmarking measures involving pick and place tasks, reorientation, and object removal tasks [EURON, 2008].

The DEXMART project (DEXterous and autonomous dual-arm/hand robotic manipulation with sMART sensory-motor skills: A bridge from natural to artificial cognition) presented [DEXMART, 2008], an extensive and detailed guideline to the implementation of benchmarking methodologies [DEXMART, 2009]. The guidelines were designed to evaluate the project outcomes from the data acquisition level to the complete system level. For each of the tasks presented in the document, the objectives and control points of the tasks are defined, as well as the reference outcomes of the task that should be evaluated.

[Matheus and Dollar, 2010] describes a set of tasks compiled from different works in the literature which have been used to evaluate the performance of manipulation robots using objects of daily living. [Calli et al., 2015] also reviews ongoing projects related to the assessment of the dexterous manipulation capability of robotic platforms and proposes a set of objects and tasks to benchmarks the robotic skills. The set of objects is described in detail (3D models with RGB data) to facilitate the replication of the experiments worldwide. [Feix et al., 2013] presents a metric to score the anthropomorphic motion capability of artificial hands and compare them with the human hand.

Recently, several international and large-scale challenges also tried to gather together the robotic community working in (dexterous) manipulation and haptic exploration. Typically, the events elaborate a list of tasks and scoring rules, and promote the competition between different teams (universities, research centres, companies, abd consortia). This allows a good benchmark between the approaches and methods proposed by the different teams. Among others, the main events proposing challenges requiring manipulation tasks are The Amazon Picking Challenge 2015 [Wurman and Romano, 2015], DARPA's robotics challenge [Guizzo and Ackerman, 2015], and RoCKIn 2015: Robot Competitions Kick Innovation in Cognitive Systems and Robotics [Amigoni et al., 2015].

Chapter 4

Recording Human manipulation and exploration movements

One of the most popular approaches for endowing robotic systems with human-like capabilities (e.g. dexterous manipulation, locomotion, and social interaction) is to learn those strategies from human demonstrations. Extensive overviews are presented in [Argall et al., 2009] and [Billard et al., 2008]. The research works provided in this PhD thesis replicate human-like capabilities for dexterous in-hand manipulation and haptic exploration of objects/surfaces.

Due to the mechanical and functional complexity of the human hand, the recording of human dexterous manipulation and exploration movements is a challenging task. Typically, a multi-modal approach (e.g. external RGB cameras, depth sensors, data gloves, motion trackers, and tactile arrays) is followed to capture all the diverse movements and contact interactions between the human hand and the objects [Faria et al., 2012]. However, care should be taken to integrate such a high number of devices simultaneously in order not to constrain the natural movements and sensing capabilities (e.g. touch feedback) of the human hand. A complementary approach consists of using instrumented objects instead of over-instrumenting the human hand (e.g. [Matsuo et al., 2009], [Lobo

(a) (b) (c)

Figure 4.1: Overview of the experimental area at the *Artificial Perception for Intelligent Systems and Robots* AP4ISR laboratory: a) table. b) data acquisition devices and objects. c) data acquisition operator controlling the central computer.

et al., 2011]).

This chapter presents the experimental area of *Artificial Perception for Intelligent Systems and Robots* AP4ISR laboratory, describing the data acquisition devices (section 4.2) used to record manipulation tasks. The data acquisition software tools and data acquisition architecture developed for this purpose, during the PhD studies reported by this thesis, are also described (see section 4.1).

Several datasets of manipulation tasks recorded during the PhD studies are presented in section 4.3. The contribution of these datasets to robotics research is also reported. Multiple software tools were developed to promote and ease the integration and annotation of datasets (section 4.4).

4.1 Experimental area and data acquisition architecture

Figure 4.1 shows an overview of the experimental area of of *Artificial Perception for Intelligent Systems and Robots* AP4ISR laboratory and the data acquisition devices available to record data regarding human demonstrations of manipulation tasks.

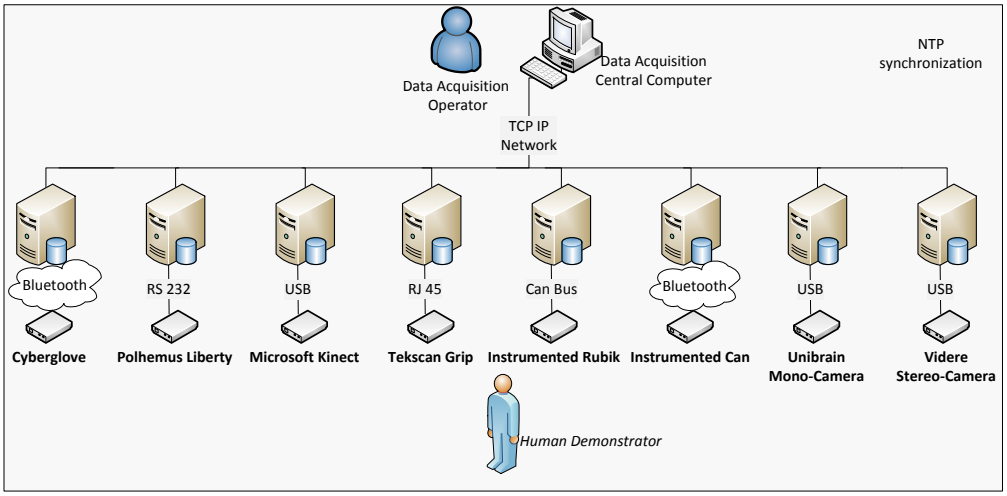


Figure 4.2: Representation of the data acquisition architecture implemented in the experimental area of *Artificial Perception for Intelligent Systems and Robots* AP4ISR laboratory.

A distributed data acquisition architecture was implemented on the experimental area. This approach was followed due to the high number of data acquisition devices that can be used simultaneously in a single session. The data acquisition architecture is described in Figure 4.2.

In this data acquisition architecture, each computer can be connected to one or more

data acquisition devices. For simplicity in Figure 4.2, only one data acquisition device is connected to each computer. All the computers involved in a data acquisition session are connected by a TCP/IP computer network. One of the computers of this architecture works as server and the remaining ones (computers having data acquisition devices connected to them) work as clients. Software clients for each of the data acquisition devices (RGB cameras, motion trackers, data gloves, tactile sensors, instrumented objects) were developed, as well as software for the server computer.

The main concept behind this approach is to have the server software coordinating all the data acquisition sessions. At the beginning of the data acquisition session, the server software is responsible for acknowledging the connection requests made by the clients software. The server builds a list of all the clients connected to it. Each client software manages both the data acquisition from the respective device, as well as the storage of the data.

The structure of the dataset is based on XML files. Each data acquisition device stores the data using a specific and predetermined structure of XML files. This allows easy integration of the datasets by other software applications and the transfer of data between partners. The structure of the XML files for each data acquisition device is detailed in the technical report *Protocol for the corpus of sensed grasp and handling data: storage of multi-modal datasets* [HANDLE-UC, 2009] expanded upon during the PhD studies, under the scope of the *HANDLE* project. The document is available online (url: `http://www.rmartins.net/phd-docs/tr02/`).

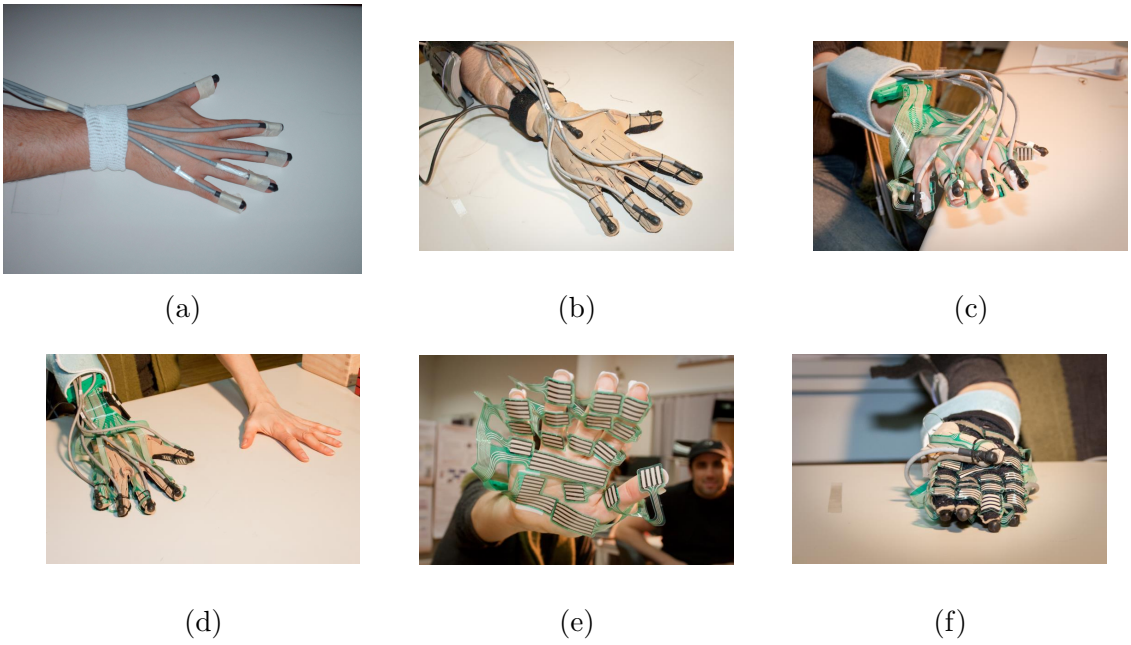
All the data acquired from the different devices is time-stamped. To enable a common temporal reference between the different computers involved in the data acquisition session, the Network Time Protocol (NTP) is used for clock synchronization of the different computers. NTP is a free, widely available protocol designed to synchronize the clocks of computers over a network. The steps involved on setting-up the NTP synchronization are detailed in the technical report *Distributed synchronization of multi-modal data acquisition devices using NTP (network time protocol)* [Martins, 2010] examined during the PhD studies under the scope of the *HANDLE* project. This tutorial is available online (url: `http://www.rmartins.net/phd-docs/tr03/`).

4.2 Data acquisition devices

4.2.1 *Cybersystems Cyberglove II*

Data glove systems are devices designed to acquire data about movements of the hand: specifically, the level of flexure of the joints of fingers, palm, and wrist. A survey by

[Dipietro et al., 2008] provides a detailed description of the historical evolution of the research (technologies, main features, and materials) in the field of data glove development. The type of glove, number of sensors and their position on the glove, as well as the glove material differ among the various commercial devices available on the market. One of the most popular high-end models is the *Cybersystems CyberGlove II*.



(a) (b) (c)

(d) (e) (f)

Figure 4.3: Human hand instrumented using different data acquisition devices (*Cybersystems CyberGlove II*, *Polhemus Liberty*, *Tekscan Grip* system).

The *Cybersystems CyberGlove II* [CyberGloveII, 2008] (CyberGlove Systems LLC, San Jose, CA, U.S.) is a wireless data glove. *Cybersystems CyberGlove II* is equipped with 22 piezo resistive bend sensors. Each glove's finger (index, middle, ring, and little) has two bend sensors (located in the metacarpophalangeal (MCP) and proximal interphalangeal joint regions), an abduction/adduction sensor (located in the MCP region), and an additional sensor in each finger (index, middle, ring, and little) to measure the distal interphalangeal joint flexure. The *Cybersystems CyberGlove II* also has sensors to measure the thumb crossover, palm arch, wrist flexure, and abduction/adduction.

The *Cybersystems CyberGlove II* is powered by a battery attached to the forearm. A wireless module (Bluetooth) transmits the data acquired by the glove to a host computer. The glove has mounting provisions in the wrist region for motion tracking sensors. However, it can be integrated easily with other types of sensing devices such as tactile sensing arrays, or motion trackers on wrist and fingertips, as described in Figure 4.3.

4.2.2 *Polhemus Liberty*

Most of the data gloves available on the market only provide information about the level of flexion/bending of the fingers and wrist joints. However, some applications also require the determination of the global position and orientation of the hand (wrist) in space (reach-to-grasp trajectories, manipulation trajectories), as well as the relative location of the joints of the hand or the fingertips, palm global orientation, and position. This type of data can be provided by motion tracking devices. These devices can be used in conjunction with the data gloves (Figure 4.3).

The *Polhemus Liberty* system [Polhemus-Liberty, 2008] (Polhemus, Colchester, VT, U.S.) is an electromagnetic motion tracking device used to record the pose (6DOF) (position and orientation) of magnetic sensors within a restricted area. The main components of this device are the *system central unit*, the sensors, and the *source*.

The *system control unit* contains the hardware and software responsible for the algorithms supporting the emission and sensing of magnetic fields and interpretation of the sensed fields. This unit is responsible for determining the position and orientation of each sensor. The unit communicates with a computer through a *RS-232* or *USB* connection. The *source* of the *Polhemus Liberty* device is responsible for the emission of the magnetic field. The source defines the inertial reference frame for sensor measurements.

Each sensor is connected to the *system control unit* by a cable. The position and orientation of the sensors is determined by considering the magnetic field sensed by each sensor. The *Polhemus Liberty* can have up to eight sensors connected to the *system control unit*, of which the position and orientation can be updated up to 240 times per second.

This type of motion tracker has a limited operational range around the magnetic field source, typically between 0.5 and 2 m. However, there are no requirements for direct line of sight between the sensors and the magnetic source. This device requires that the environment, where the data will be acquired, does not have metallic materials near or between the magnetic field source and the magnetic sensors. This requirement ensures that the metallic materials do not distort the emitted magnetic field, thus avoiding the introduction of large errors in the measurements.

4.2.3 *Tekscan Grip*

The devices presented previously were used to determine the position, orientation, and level of flexure of the hand or specific segments of the hand. However, analysis of the force/pressure applied by the hand during the execution of grasping and manipulation tasks is also critical. This analysis will be important to determine the temporal profile of the pressure/forces applied by the human hand (fingertips and palm) while performing

different types of grasps. This analysis can also provide information about the fingers that are more active in different types of grasps and about the sequence of segments of the hand that come into contact with the object.

The tactile sensing device *TekScan Grip* [Tekscan-Grip, 2010] (Tekscan Inc, Boston, MA, U.S.) is a system specifically designed to acquire the pressure applied by the different regions of the human hand (fingers, thumb, and palm) during the execution of tasks that require grasping movements. The device consists of a flexible thin film (0.1 mm) with embedded pressure sensors and electronics connecting the sensors to a data acquisition module (*TekScan VersaTek*) attached to the wrist region. The *TekScan VersaTek* module acquires data at up to 850 Hz and provides a *USB* connection to transfer the acquired data to a computer.

The *TekScan Grip* system is composed of five segments which can be attached to the ventral regions of the fingers, thumb, and palm of the human hand (total of 361 sensing elements). The sensing regions have a spatial resolution of about 6.2 sensing elements per square centimeter, and each element can sense pressures up to 50 psi (344.7 KPa). This device can be used in either the right or left hand.

The *TekScan Grip* system can be attached directly to the human hand or to a data glove (e.g. *Cybersystems CyberGlove II*), as demonstrated in Figure 4.3.

The performance, characteristics, and calibration methods of the *TekScan Grip* system were tested extensively and documented in the technical report *Experimental evaluation and calibration protocol of Tekscan Grip system* [Martins, 2012b] available online (url: `http://www.rmartins.net/phd-docs/tr05/`).

4.2.4 Instrumented objects

The previous sections described the sensors/devices that can be attached to the human hand. These devices give a human hand-centred perspective.

However, it can be useful to acquire data related to the movement/pose of the object during the task being performed; this is the object-centred approach. This type of data can be obtained by integrating external sensors on the object (attaching motion trackers or tactile arrays), or making custom-designed objects with those types of sensors embedded in the design (e.g. instrumented Rubik cube and instrumented sensing can). The object-centred approach does not constrain the natural movements of the human hand.

Instrumented Rubik cube

The instrumented Rubik cube has the same dimensions and colors as a standard Rubik cube, as presented in Figure 4.4a. The instrumented Rubik cube [HANDLE-SHADOW,

(a) (b) (c)

Figure 4.4: Instrumented objects. a) Rubik cube. b) sensing soda can. c) *Nintendo Wiimote*

.

2009] (Shadow Robot Company, London, UK) has 54 tactile sensing regions (six in each side of the cube) that provide non-calibrated 12-bit-resolution tactile data. Each side of the instrumented Rubik cube is equipped with a three-axis accelerometer, providing acceleration data with eight-bit resolution. The instrumented Rubik cube is powered by a cabled external source and has a wired CAN-BUS communication interface. A technical report *Installing controller area network (CAN-Bus) drivers and compiling code on Ubuntu* was published and is available online [Martins, 2013] (url: `http://www.rmartins.net/phd-docs/tr04/`).

Instrumented sensing can

The instrumented sensing can [HANDLE-SHADOW, 2010] (Shadow Robot Company, London, UK) is designed to mimic a standard 330 ml soda can (Figure 4.4b). It has a total of 40 (4×10) independent tactile sensing regions distributed using ten sensing vertical panels. Each panel contains four tactile sensing elements (eight-bit resolution per sensing element) and has a three-axis accelerometer attached to it (ten-bit resolution per channel).

The instrumented sensing can is portable. It is powered by an internal set of four AAA batteries, and it has a Bluetooth communication interface.

Nintendo Wiimote

The *Nintendo Wiimote* [Nintendo, 2006] (Nintendo Co. Ltd, Kyoto, Japan) is a useful instrumented object that can be employed during the data acquisition sessions (Figure 4.4c). The *Nintendo Wiimote* is a portable device (Bluetooth interface) of an appropriate size for handling tasks; it contains embedded sensors (three-axis accelerometer and infrared camera sensor) and several buttons.

The accelerometer provides information about the movements of the device. The infrared camera sensor has an integrated multi-object tracking engine that can determine the position (x,y) and size of up to four simultaneous infrared light sources [Lee, 2008]. The buttons can be used to acquire user inputs during the execution of the task.

4.2.5 *Microsoft Kinect*

The *Microsoft Kinect* [Microft-Kinect, 2009] (Microsoft Corporation, Redmond, WA, U.S.) is a device initially designed to be used by the gaming and entertainment industry as a human-computer interface [Smisek et al., 2013]. However, due to its versatility and low price, it became a revolutionary device in other areas such as robotics, physical rehabilitation and therapy, art installations, and do-it-yourself (DIY) homemade projects [Zhang, 2012].

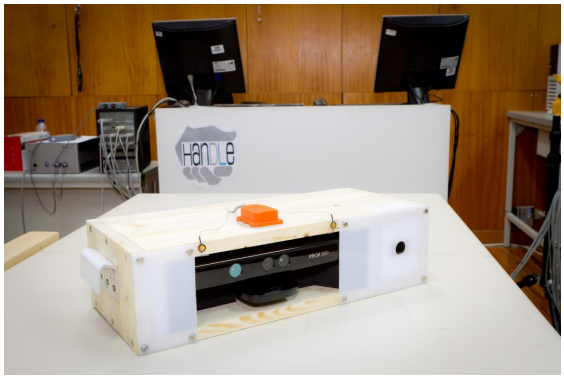


Figure 4.5: *Microsoft Kinect* integrated in a wood box.

The *Microsoft Kinect* device integrates different types of sensors (Figure 4.5). An infrared (IR) projector of light patterns and IR camera (structured light principles) provides a depth map (1280×1024) of the environment surrounding the device. The depth data is fused with the output of a monocular *RGB* camera (1280×1024), generating a *RGB-D* map (cloud points with color). The *Microsoft Kinect* also incorporates an array of four microphones. They can be used to implement voice recognition systems and speaker spatial tracking applications.

Several software development kits (SDK) that have been released take advantage of the features of the *Microsoft Kinect*: face recognition, human body tracking, and 3D reconstruction.

4.2.6 *Videre* camera

The *Videre STH-MD CS3* [Videre, 2006] is a digital stereo camera consisting of two 1.3 MP CMOS sensors mounted on a metallic (aluminium) package, as shown in Figure 4.6b. This camera has a fixed 9 cm baseline.

The stereo camera can acquire colour and monochromatic data at different frame rates (30 Hz for 640x480, 15 Hz for 1024x768 colour only, and 7.5 Hz for 1280x960). The camera is powered through a firewire cable, which is also used for data transmission and control (synchronization, exposure, gain, and colour balance).

(a) (b)

Figure 4.6: RGB cameras: a) *Unibrain* camera. b) *Videre* camera.

4.2.7 *Unibrain* camera

The *Unibrain Fire-I Digital* [Unibrain, 2007] (Unibrain, Athens, Greece) is a firewire colour digital camera (CCD sensor), as presented in Figure 4.6a. This device can acquire video at different frame rates (30, 15, 7.5, 3.75 frames per second), different resolutions (640x480, 320x240, 160x120) and video coding modes (YUV, RGB 24-bit, monochromatic eight-bit).

The *Unibrain Fire-I Digital* can be powered by the firewire cable (when the camera is connected to a desktop computer) or through a DC adapter (when the camera is connected to a laptop computer). The camera is housed in a plastic polymer package.

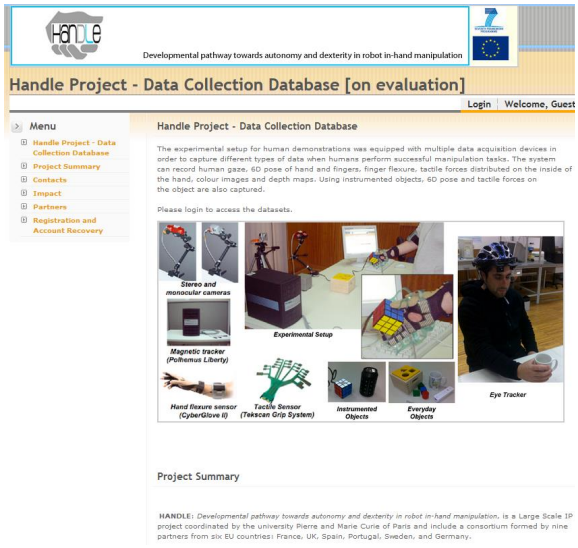


Figure 4.7: Homepage of the HANDLE project [HANDLE, 2009] online data repository.

4.3 Datasets

The datasets described in this section were recorded at the experimental area of *Artificial Perception for Intelligent Systems and Robots* AP4ISR laboratory during the PhD studies, under the scope of the HANDLE project. The datasets were acquired to fulfil the objectives and requests made by other partners. The publications and applications related to each of the datasets are listed in Table 4.1. All the datasets are available online at the HANDLE project web repository.

The participants were seated in a chair in front of the table presented in Figure 4.1. The data acquisition devices and objects used in each session are summarized in Table 4.1, and the configuration of the experimental area is demonstrated in the figures listed in the description of each dataset (section 4.3.1 to 4.3.8).

4.3.1 Dexterous manipulation of a laboratory pipette

The participants were instructed to move a liquid from a container to a second container, using a laboratory pipette. The experimental materials were dispensed as presented in Figure 4.8. The laboratory pipette was grasped and manipulated using only the dominant hand. Before starting each run, the pipette was placed on a pipette stand. The participants were instructed to place the pipette back on the stand or to leave the pipette directly on top of the table by the end of each run. A video demonstrating this task is available online `www.rmartins.net/phd-docs/ds01`.

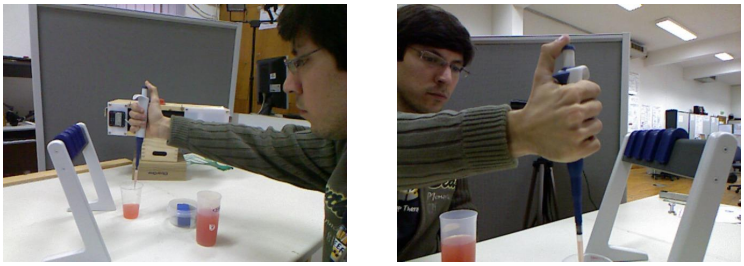
To stimulate and challenge the participants with unexpected situations, during some

Table 4.1: List of HANDLE datasets acquired during the PhD studies

Task	**Data Acquisition Devices**	**Participants**	**Runs**	**Application**
Dexterous manipulation of a laboratory pipette	2 *Microsoft Kinect*	12	240	final demonstration: HANDLE project
Thumb movement during manipulation tasks	1 *Microsoft Kinect* 7 *Polhemus Liberty* sensors 1 *Cyberglove II*	1	5	journal paper: HANDLE partner [Berglund et al., 2012]
Screwdriver in-hand rotation	1 *Microsoft Kinect* 1 *Cyberglove II* 1 *Tekscan Grip*	1	5	conference paper, PhD thesis: HANDLE partner [Cheng et al., 2012], [Cheng, 2013]
In-hand manipulation of toys	1 *Microsoft Kinect* 8 *Polhemus Liberty sensors* 1 *Videre* camera 1 *Tekscan Grip*	1	35	conference paper, PhD thesis: HANDLE partner [Cheng et al., 2012], [Cheng, 2013]
Grasp the Wii remote and press a button	1 *Cyberglove II* 1 *Unibrain* camera 6 *Polhemus Liberty* sensors 1 *Videre* camera 1 *Tekscan Grip*	3	18	conference papers: HANDLE partner [Hendrich et al., 2010], [Hendrich et al., 2012]
Fill a toy sorting box with objects	1 *Cyberglove II* 1 *Unibrain* camera 5 *Polhemus Liberty* sensors 1 *Videre* camera 1 *Tekscan Grip*	3	17	conference papers: HANDLE partner [Hendrich et al., 2010], [Hendrich et al., 2012]
Pick up a pen and write	1 *Unibrain* camera 8 *Polhemus Liberty* sensors 1 *Videre* camera 1 *Tekscan Grip*	3	10	conference papers: HANDLE partner [Hendrich et al., 2010], [Hendrich et al., 2012]
Pick an object and slide	1 Instrumented rubik cube 1 *Unibrain* camera 7 *Polhemus Liberty* sensors 1 *Videre* camera 1 *Tekscan Grip*	4	20	conference papers: HANDLE partner [Hendrich et al., 2010], [Hendrich et al., 2012]

runs, the manipulation of the pipette was done without using the thumb to press the pipette buttons. This variant of the protocol was used to record the ways in which the participants adapted the standard manipulation strategy to this constraint. The data acquired during the demonstrations was manually annotated (grasp and in-hand manipulation primitives) using the MATLAB tool developed for that purpose and presented in section 4.4.3.

The datasets were used by several partners of the HANDLE project to develop various applications concerning robotic learning by human demonstration. This approach was used to provide the SHADOW robotic platform with pipette manipulation skills. The results were demonstrated during the final review meeting of the HANDLE project. The HANDLE final demonstration was presented publicly in *Euronews Futuris*, disseminating science and technology. The video of the demonstration is available online at `http://www.youtube.com/watch?v=XSw5QVdzGW4`. Figure 4.9 summarizes some segments of the TV show. The datasets are available online at `http://www.rmartins.net/phd-docs/ds01/`.



(a) (b)

Figure 4.8: Participant demonstrating the task *Dexterous manipulation of a laboratory pipette*.

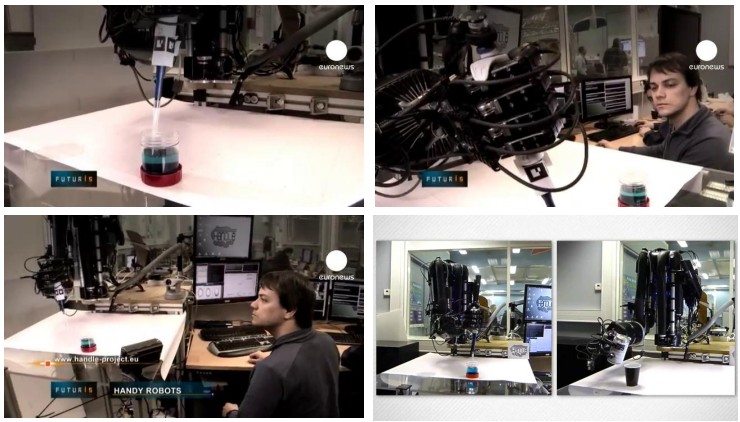
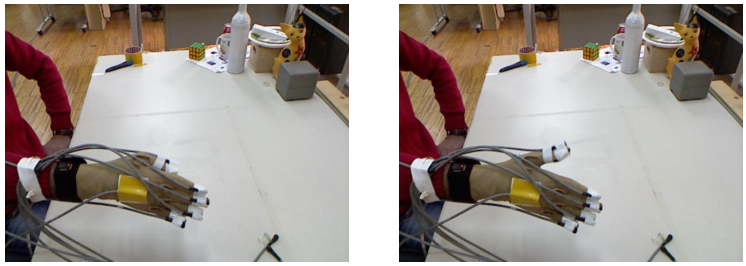


Figure 4.9: Final demonstration of HANDLE project presented during *Euronews* TV show *Futuris* (url: `https://www.youtube.com/watch?v=XSw5QVdzGW4`). The final demonstration consisted of the robotic dexterous manipulation of a laboratory pipette learned from the human demonstrations. The TV show also featured the instrumented Rubik cube data visualization tool presented in section 4.4.4 (url: `https://www.youtube.com/watch?v=KJybBorZjH0`).

4.3.2 Thumb movement during manipulation tasks

The participant started with the palm of the right hand flat on the table. The hand was lifted off the table, keeping the palm and index, middle, and little fingers straight (Figure 4.10). The thumb was then moved around. The participant went through all possible movements of the thumb several times, exploring all the degrees of freedom of that finger. A video demonstrating this task is available online at `www.rmartins.net/phd-docs/ds02/`.

The data recorded by the data glove, motion tracker, and camera was sent to the ORU partner (Orebro University - AASS Learning Systems Lab) of the HANDLE project. The data was used to train an algorithm that learns the mapping and correspondences between the kinematic structure of the human hand and Shadow robotic hand. The human thumb



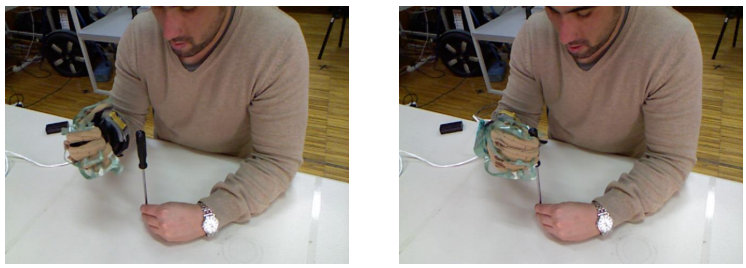
(a) (b)

Figure 4.10: Participant demonstrating the task *Thumb movement during manipulation tasks*.

has a complex mechanical structure, which is different from the kinematic structure of the robotic thumb. The data was used to improve the behaviour of the robot during robotic learning from human demonstrations or in tele-manipulation applications. The results of this work were published in an international journal [Berglund et al., 2012].

4.3.3 Screwdriver in-hand rotation

The participant is instructed to use the two hands to grasp a screw and use a screwdriver to insert and adjust the screw in a hole. The participant was seated in front of the table, where a screwdriver and screw were placed in their initial configurations (Figure 4.11).



(a) (b)

Figure 4.11: Participant demonstrating the task *Screwdriver in-hand rotation*.

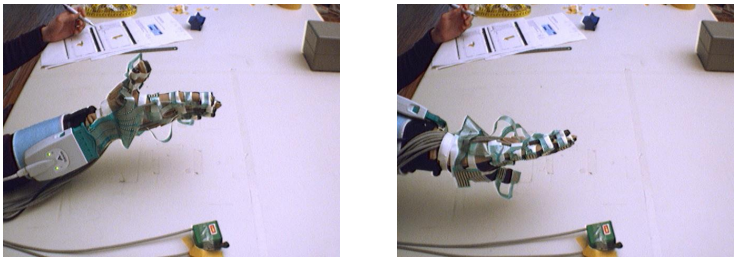
The left hand was not instrumented, and it was used to grasp the screw, insert it in the hole, and keep it stable as long as the task progressed. The right hand was used to perform consecutive in-hand manipulation movements (re-grasps), using the screwdriver, to insert the screw. By the end of each run, the participant placed the screwdriver on the table (initial configuration). A video demonstrating this task is available online `www.rmartins.net/phd-docs/ds03/`.

The data was delivered to the UHAM partner (University of Hamburg - Technical Aspects of Multimodal Systems (TAMS) Lab) of the HANDLE project. The data was

used to develop a software tool to automatically segment dexterous in-hand manipulation movements. The automatic segmentation was compared (benchmark) to the segmentation performed manually by a human operator. This work was published at an international conference [Cheng et al., 2012] and in a PhD thesis [Cheng, 2013].

4.3.4 In-hand manipulation of toys

The participant was instructed to grasp the toys (wood pieces) placed on top of the table and then perform several in-hand manipulation movements to rotate the toy around itself inside the hand. When the demonstration finished, the toy was placed back on the table.



(a) (b)

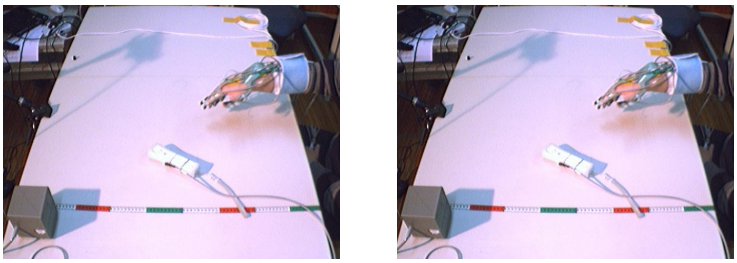
Figure 4.12: Participant demonstrating the task *In-hand manipulation of toys*.

The data was delivered to the UHAM partner (University of Hamburg - Technical Aspects of Multimodal Systems (TAMS) Lab) of HANDLE project. The objectives and applications of this data acquisition were the same as the objectives reported for the dataset described in section 4.3.4. A video demonstrating this task is available online `www.rmartins.net/phd-docs/ds04/`.

4.3.5 Grasp the Wii remote and press a button

The participant was seated in a chair in front of the table where the *Nintendo Wii-mote* was placed. Three different orientations of the *Nintendo Wii-mote* on the table were used as starting pose. As the data recording starts, the subject moves the right hand toward the *Nintendo Wii-mote*, grasping and picking-up the object (Figure 4.13). The subject performed some in-hand manipulation movements so that the *Nintendo Wii-mote* points toward the frontal camera recording the session. Participants used the thumb to press a button on the *Nintendo Wii-mote* and then put the *Nintendo Wii-mote* back on the table in the starting configuration.

This dataset was acquired and delivered to the UHAM partner (University of Hamburg - Technical Aspects of Multimodal Systems (TAMS) Lab) of the HANDLE project. The multi-modal dataset was used to develop a method to automatically segment the data



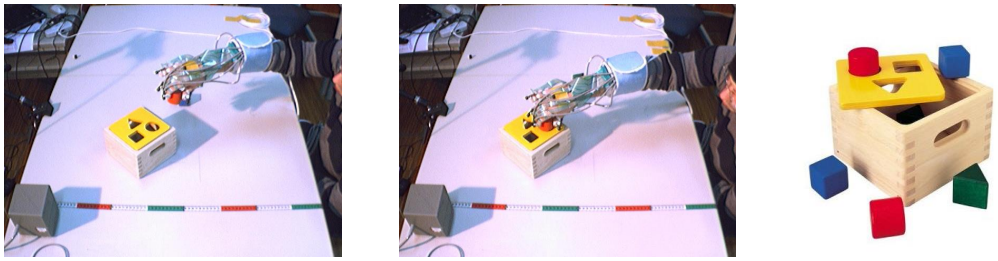
(a) (b)

Figure 4.13: Participant demonstrating the task *Grasp the Wii remote and press a button.*

from in-hand manipulation tasks. The primitives consist of a basic set of finger and hand movement patterns, which can be used to model manipulation tasks. The learned model was transferred to robotic platforms. The datasets were used in the experimental implementation of the work [Hendrich et al., 2010], which was published in the proceedings of an international conference. A video demonstrating this task is available online at `www.rmartins.net/phd-docs/ds05/`.

4.3.6 Fill a toy sorting box with objects

The participant was seated in front of a table where the toy sorting box and the toys were placed in their starting configurations, as shown in Figure 4.14.



(a) (b) (c)

Figure 4.14: a) - b) Participant demonstrating the task *Fill a toy sorting box with objects.* c) Toy sorting box and toys (wood pieces).

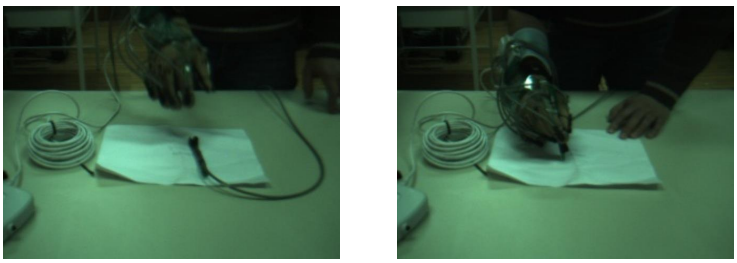
The participant was instructed to use the right hand to consecutively grasp each of the toys displayed on the table and fit each of them in the corresponding hole of the box. The positions of the toys were adjusted using only in-hand manipulation movements during the transport movement between the table and the toy sorting box. Each run ended when all the toys were placed inside the box.

This dataset was acquired and delivered to the UHAM partner (University of Hamburg - Technical Aspects of Multimodal Systems (TAMS) Lab) of the HANDLE project. The

motivation and applications of this dataset were the same as for the dataset described in section 4.3.5. A video demonstrating this task is available online at `www.rmartins.net/phd-docs/ds06/`.

4.3.7 Pick up a pen and write

The participant was seated in front of a table where a pen and a piece of paper were placed in their starting configurations (Figure 4.15). The participant manipulated a standard ball-point pen with typical active click mechanics on the back-end. The pen was placed in different starting poses. The pen was also instrumented with a motion tracking sensor.



(a) (b)

Figure 4.15: Participant demonstrating the task *Pick up a pen and write*.

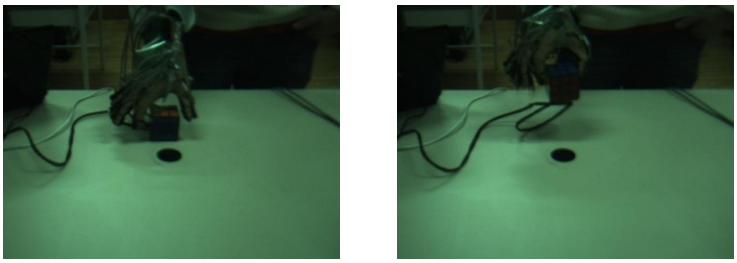
The participant started each run by picking up the pen. After grasping the pen, in-hand manipulation movements were performed to move the pen to a configuration suitable to activate the click mechanism with the thumb a few times. Then, the pen was in-hand manipulated to achieve a configuration for writing. The participant wrote *"HANDLE"* on a piece of paper. After that, the participant placed the pen back on the table in the starting configuration.

This dataset was acquired and delivered to the UHAM partner (University of Hamburg - Technical Aspects of Multimodal Systems (TAMS) Lab) of the HANDLE project. The motivation and applications of this dataset were the same as for the dataset described in section 4.3.5. A video demonstrating this task is available online at `www.rmartins.net/phd-docs/ds08/`.

4.3.8 Pick an object and slide

The participant was comfortably seated in front of the table where the objects were placed. The participants performed this task using two different objects: an instrumented Rubik cube and a box, as described in Figures 4.16 and 4.17, respectively.

The participant was instructed to grasp the object (box or instrumented Rubik cube) with the index finger and thumb. This precision grasp was performed not in the geometric

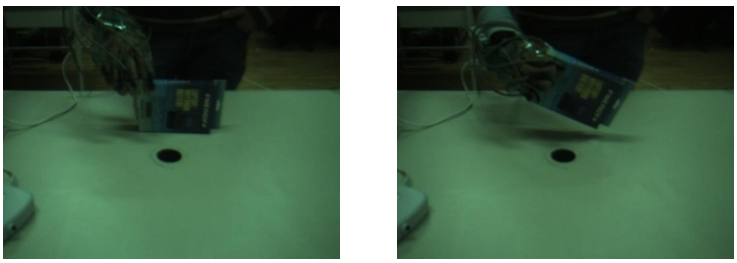


(a) (b)

Figure 4.16: Participant demonstrating the task *Pick an object and slide* with an instrumented Rubik cube.

center of the object, but near one end. The object was moved up using only the index and thumb, maintaining the orientation of the object. The index and thumb applied a pressure, which prevented the inertial momentum of the object from changing its pose. After reaching the maximum height of the movement, the participant reduced the pressure applied by the index and thumb fingers, allowing the object to slide and rotate around the index and thumb fingers, changing its orientation. The object was then moved back to its starting pose on the table.

This dataset was acquired and delivered to the UHAM partner (University of Hamburg - Technical Aspects of Multimodal Systems (TAMS) Lab) of the HANDLE project. The motivation and applications of this dataset were the same as for the dataset described in section 4.3.5. A video demonstrating this task is available online `www.rmartins.net/phd-docs/ds09/`.



(a) (b)

Figure 4.17: Participant demonstrating the task *Pick an object and slide* with box.

4.4 Software tools

4.4.1 Software clients for data acquisition devices

As described in section 4.1, each of the data acquisition devices has a dedicated software client. The software clients are responsible for the integration of the data acquisition devices in the distributed data acquisition architecture. They connect to a server software which coordinates the data acquisition session, and they receive the triggers to start and stop the data acquisition during each run.

Parameters such as data acquisition rate, image resolution, and calibration parameters are configured locally in each client prior to the start of the session. The data acquired by each device is formatted by the corresponding software client according to the XML format presented in the technical report *Protocol for the corpus of sensed grasp and handling data: storage of multi-modal datasets* [HANDLE-UC, 2009].

During the PhD studies reported in this thesis, I was responsible for the implementation of the software clients for the *Cybersystems Cyberglove II* [Martins, 2009b], instrumented Rubik cube [Martins, 2009c], instrumented sensing can [Martins, 2009d], *Tekscan Grip* system [Martins, 2009f], and an alternative software client for *Polhemus Liberty* [Martins, 2009e]. All software clients were developed using *C++* programming language for *Ubuntu* environment. The alternative software client for *Polhemus Liberty* was developed using *Python*.

The programming of these clients required the analysis of requisites for each of them and the understanding of the technologies: I/O communications (USB, RS-232, CAN-bus), TCP-IP socket communications, and XML data structure.

4.4.2 *importDatasetTB*: toolbox for integrating data in MATLAB

A MATLAB toolbox (*importDatasetTB*) was developed to promote the use of the datasets by the HANDLE partners, improve dissemination of the datasets, and facilitate the integration of the datasets in the research activities of the HANDLE partners (including this PhD study). The *importDatasetTB* MATLAB toolbox automatically imports the different types and configurations of datasets that were recorded using the XML storage scheme described in the technical report *Protocol for the corpus of sensed grasp and handling data: storage of multi-modal datasets* [HANDLE-UC, 2009].

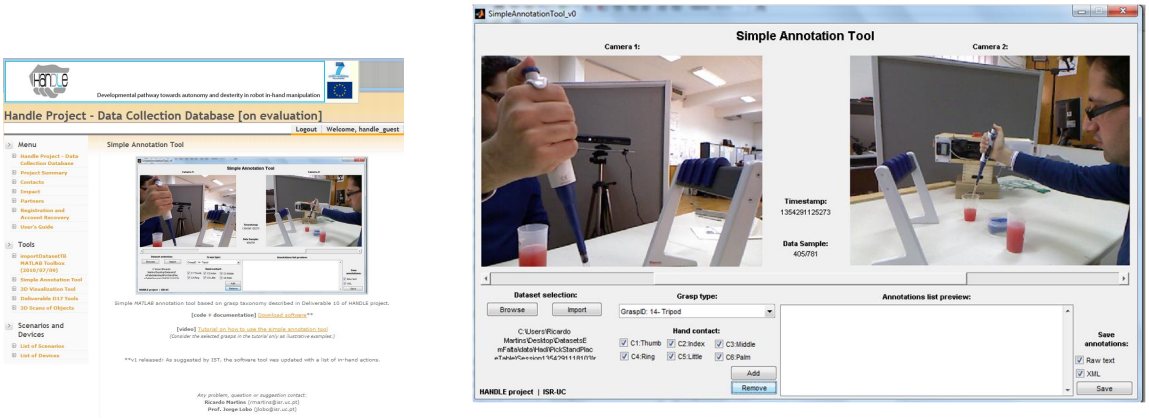
The toolbox allows an automatic and fast integration of large multi-modal datasets. Initially, the toolbox was compatible with datasets containing data structures of *Cybersystems Cyberglove II*, *Polhemus Liberty*, instrumented Rubik cube, instrumented sensing

Figure 4.18: Section of HANDLE project website [HANDLE, 2009] presenting and supporting the MATLAB toolbox *importDatasetTB*.

can, *Tekscan Grip* system, monocular RGB cameras, and stereo RGB cameras. Afterward, it was extended and made compatible with additional data acquisition devices of the *Artificial Perception for Intelligent Systems and Robots* AP4ISR laboratory: *Microsoft Kinect*, *Nintendo Wii-mote*, *NDI optotrak*, and *Xsens IMU*. The *importDatasetTB* MATLAB toolbox source code and documentation are described in the technical report *"importDatasetTB: a MATLAB toolbox to integrate multi-modal datasets"* [Martins, 2009g].

This MATLAB toolbox, *importDatasetTB*, was developed on top of a Java library (*jParserToolbox*) which was programmed from scratch for this tool. This Java library implements a XML *SAX* (Simple API for XML) parser for the XML dataset structures [HANDLE-UC, 2009]. This custom-designed Java library was used in the development of the MATLAB toolbox because MATLAB natively includes a Java Virtual Machine. Thus, the Java interpreter can be used via the MATLAB scripting shell. The *importDatasetTB* MATLAB toolbox imports datasets of hundreds of megabytes in a few seconds. This tool improved the low performance and bad memory management for large XML data structures of the XML parsing tools native with MATLAB; it is based in DOM (Document Object Model).

The source code and documentation with practical examples are available online at [Martins, 2009g] (url: `http://www.rmartins.net/phd-docs/st02/`).



(a) (b)

Figure 4.19: a) Section of HANDLE project website [HANDLE, 2009] presenting and supporting the MATLAB annotation tool for multi-modal datasets. b) Snap shot of the graphical interface of the MATLAB annotation tool for multi-modal datasets.

4.4.3 Annotation tool for multi-modal human grasping datasets

The dataset *Dexterous manipulation of a laboratory pipette* presented in section 4.3.1 was annotated offline by several partners of the HANDLE project. To facilitate this task, an annotation tool was developed using MATLAB. The MATLAB annotation tool presents to the user a graphical interface displaying side-by-side two synchronized images from the two *Microsoft Kinect* RGB cameras. The user performing the annotation employs the two perspectives of the manipulation strategy to annotate the images with the types of grasps and in-hand manipulation actions performed by the participant. The dictionary of possible labels is presented on the graphical interface and follows the categories of manipulation movements given in section 3.4. The MATLAB exports the annotations in plain text or XML format.

The annotation tool is demonstrated in Figure 4.19a. The MATLAB code, documentation, and video tutorials are available online at [Martins, 2012c] (url: `http://www.rmartins.net/phd-docs/st09/`).

4.4.4 Instrumented Rubik cube: touch data visualization tool

An interactive and intuitive 3D visualization tool was developed to demonstrate the design and touch sensing capabilities of the instrumented Rubik cube (presented in section 4.2.4). The 3D model of a Rubik cube changes the color of each cell, as long as the corresponding sensing element is pressed on the real cube. The visualization tool was developed in *Ubuntu*, using the *C++ OpenGL-GLUT* library.

The 3D visualization tool was presented during the final demonstration of the HAN-

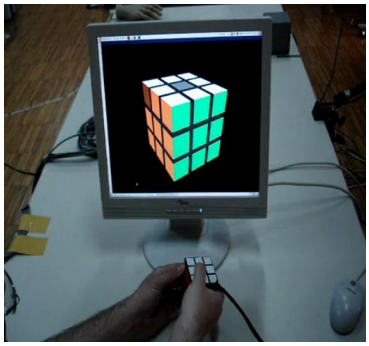


Figure 4.20: Demonstration of the instrumented Rubik cube visualization tool during a interactive session.

DLE project. The source code, tutorials for the installation of the *CAN-BUS* Linux drives, and *OpenGL-GLUT* libraries used on this Linux C++ software project are available online at [Martins, 2012a] (url: `http://www.rmartins.net/phd-docs/st08/`). A video demonstrating a typical live interaction with the instrumented Rubik cube is available as well [Martins, 2012a], and it is illustrated in Figure 4.20.

Chapter 5

Recognition of grasping primitives using tactile sensory data

5.1 Introduction

Over the past few years, several research fields (e.g. human-computer interface sciences, crowd behaviour, medical rehabilitation, robotics, surveillance, and sport performance analysis) have focused some of their attention on the understanding and analysis of human behaviour and human motions [Aggarwal and Ryoo, 2011]. In robotics, the analysis of human movements has been applied (among others) in research areas concerning the task of learning by imitation of human demonstrations [Billard et al., 2008]. This approach was motivated by principles described in several studies from human developmental sciences, which propose that humans learn most of their skills by observing and analysing others performing those tasks (observational learning) [Magill and Anderson, 2007].

Robot learning by human demonstration consists of using examples (successful and failed) of a task performed by humans, to extract key-points and other types of constraints (e.g. velocities, contact intensity, and trajectories) and statistics (e.g. causal dependencies, alternative redundancies, and contextual preferences). The data extracted from the demonstrations is used to estimate several parameters of the model of the task being learned.

The diversity of the demonstrations is essential to provide the robotic system with a robust model of the task. The robustness of the model establishes the capability to deal autonomously with (partially) new contexts (generalization capability). The approaches based on the robot learning by human demonstration try to implement generalizable models of the tasks, in contrast to the traditional approaches consisting of the full analytic formulation and specification of the models.

The work presented in this chapter of the thesis intends to contribute to the development of autonomous robotic hands by modelling the strategies used by humans to manipulate objects using the intrinsic movements of the hand (fingers, palm), which is

Table 5.1: Comparison between the contributions of this work and the related works

Study	**Task Model** [a]	**Approach** [b]	**Features** [c]	**Application** [d]
This Work	SL	P	T	HMA
[Delson and West, 1996]	TL	D	F, M	RPD
[Tso and Liu, 1996]	TL	P	M	RPD
[Calinon et al., 2007]	TL	P	M	RPD
[Kondo et al., 2011]	SL	D	T	RPD
[Bernardin et al., 2005]	SL	P	T, M	HMA
[Kruger et al., 2010]	SL	P	M	RPD

[a] SL- symbolic level; TL- trajectory level.
[b] P- probabilistic; D- deterministic.
[c] T- tactile based; F- force based; M- movement trajectories based;
[d] HMA- Human movement analysis; RPD- robotic platform development;

known as in-hand manipulation. This type of movements requires the complex coordinated action of the fingers and palm. The temporal characteristics and sequence of the contact between the object, fingers, and palm plays a crucial role in the stabilization of the object being manipulated and consequently in the success of the manipulation task.

This work intends to contribute with the definition of a set of primitives to represent in-hand manipulation movements, as well as the statistical relations between them, in order to model different tasks of this class performed by humans; this is termed generalization capability.

5.2 Related work

In the robotics research field, several approaches to solve the motion learning problem from human demonstrations have been proposed [Billard et al., 2008]. Typically, the proposed approaches can be grouped in two main categories.

One approach represents the movements at the trajectory level and generalizes the representation of the movements through the extraction of statistical regularities from several human demonstrations of the movements. Researchers [Tso and Liu, 1996] applied Hidden Markov Models to encode a training dataset built from a set of human demonstrations. Given a human demonstration as input, the system reproduces the trajectory of the training dataset with the highest likelihood. A simple approach was also presented by [Delson and West, 1996] . The authors simply made a statistical analysis of human demonstrations of a pick-and-place task and defined the range of Cartesian trajectories that can be performed to achieve that task. Calinon [Calinon et al., 2007] proposes to extract continuous constraints from a set of demonstrations, using different initial positions of the object. The Cartesian trajectories of these demonstrations are projected using Principal Component Analysis, and then the constraints are represented through Gaussian Mixture Models. To reproduce the task, the constraints are reprojected on the original data space, and the generalized version of the Cartesian trajectory is found

Figure 5.1: Schematic representation of the typical contact signatures of different grasps. Adapted from [Bernardin et al., 2005]. The boxes highlighted with orange border show different demonstrations of the grasp. The green border highlights the regions of the hand recruited to perform that type of grasp.

by estimating the trajectory that satisfies all the constraints. The approaches described previously propose the learning and encoding of movements at the trajectory level.

This work follows an alternative class of approaches defined using a symbolic learning and encoding of manipulation movements, performing the supervised segmentation and labelling of the primitives during the learning stage. Several works use Support Vector Machines (SVM) to extract sequences of primitives from human demonstrations. The output of the SVM, temporal sequences of labelled data, is combined with Hidden Markov Models (HMM), which provides the most probable temporal sequence of primitives [Vicente, 2007]. The HMM complements the initial sequence of primitives estimated by SVM, which does not consider the temporal relations and dependencies between the data elements.

[Kondo et al., 2011] proposes a method to describe in-hand manipulation movements by recognizing a sequence of contact state transitions between the human hand and the manipulated object. The recognition algorithm is based on a Dynamic Programming approach by comparing the similarity of the contact state transition between an input sequence and templates of manipulation primitives.

[Bernardin et al., 2005] describes a technique to recognize continuous human grasping sequences using HMM. Twelve different grasp primitives are recognized, combining data from palm tactile sensors and hand joint flexure levels from a data glove.

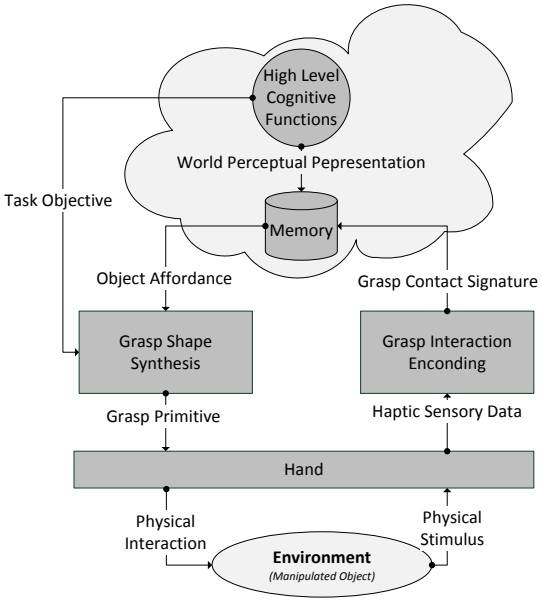


Figure 5.2: Modular representation of the processes involved in the planning and execution of a manipulation task. The representation is simplified to highlight the mechanisms (grasping primitives based on contact signatures) supporting the approach proposed in this chapter, Figure 5.3.

The work by [Kruger et al., 2010] presents the automatic extraction of action primitives (without the necessity of presegmentation and manual labelling) and the corresponding grammar from continuous movements of several human demonstrations of grasping tasks. The approach considers that all the actions can be described by a set of elementary building blocks (action primitives). A grammar (set of rules) defines how the action primitives can be combined. The action primitives are represented by parametric HMM (an extension of HMM). The extraction of the motion primitives from the movements also considers the changes in object state.

Matsuo proposed a segmentation method of human manipulation task that measures the contact force imposed by a human hand on the grasped object [Matsuo et al., 2009]. The work proposes a metric, whose values are used for segmenting a manipulation movement into primitives. The temporal evolution of the metric is calculated from the contact forces sensed at different regions of the hand, as long as the manipulation task progresses.

5.3 Approach overview

This work presents an approach to model the strategies underlying the in-hand manipulation tasks performed by humans. The main contributions of this chapter are summarized in Figure 5.3 and are detailed throughout the next sections of the manuscript.

Several studies [Johansson and Flanagan, 2009], [Castiello, 2005] concluded that a general human manipulation movement can be decomposed on different stages such as reach, load, lift, hold, replace, and unload. The manipulation movements can be segmented and represented as a sequence of primitives, which can be thought of as the elementary building blocks of the task model. The temporal transition between different primitives is made by specific events such as the variation of the intensity and extension of the hand-object contact areas, variation of grip aperture, and type of grasp.

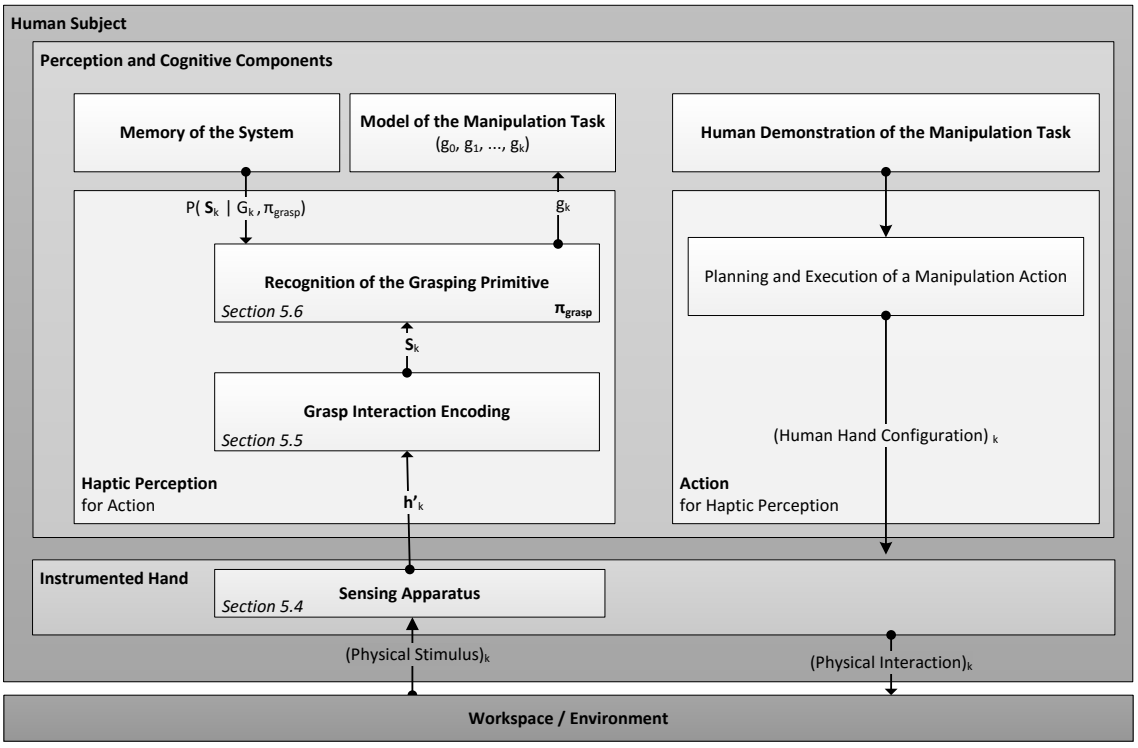


Figure 5.3: Global architecture of the approach proposed in this chapter. The main contributions (description of the sensing apparatus, grasp interaction encoding, and inference of grasping primitives) are highlighted in bold and presented in sections 5.4, 5.5, and 5.6, respectively. The variables representing the flow of the data are detailed in Table 5.2.

Table 5.2: Summary of the relevant variables used in this chapter

Variable	Description	Domain
k	Time iteration	$\mathbb{N}^0$
$\mathbf{h}_k$	Tactile sensing output of the instrumented hand. (360 elements)	$\mathbb{R}^{360}, \quad h^i \in [0, 255]$
$\mathbf{h}_k^{'i}$	Tactile sensing output of the instrumented hand (15 cluster regions)	$\mathbb{R}^{15}, \quad h^{'i} \in [0, 255]$
G_k	Category of the grasping primitive	$\{Primitive_1, \ldots, Primitive_7\}$
$\mathbf{S}_k$	Tactile activation descriptor of the instrumented hand.	$S^i \in \{"NotActive", "LowActive", "HighActive"\}$

The approach described in Figure 5.3 models the in-hand manipulation tasks by a

temporal sequence of primitives. The in-hand manipulation movements are performed to reorient and to repost the manipulated object, which require the change of the type of grasp applied to the object and a precise contact interaction between the object and the hand.

Each type of grasp can be characterized by a specific hand-object contact signature (Figure 5.1), resulting from the interaction of the hand and the object during that period of the task. Thus, the in-hand manipulation task can be described by a temporal sequence of the contact signatures corresponding to different types of hand configurations interacting with the object. The primitives used to model the in-hand manipulation task are defined on the tactile sensing domain.

For each primitive, the spatial configuration of the contact signatures is stored, as well as the force intensities for each region of the hand. These parameters characterizing each primitive can then be used as control states, described by the tactile intensity and hand locations, during the transfer of these skills to a robotic platform with manipulation capabilities.

The flow of the data is summarized in Figure 5.3. The human demonstrator performs an in-hand manipulation task using an instrumented data glove equipped with tactile sensors distributed on the hand palm and finger surface region. The haptic sensory data output is presented in section 5.4. The descriptor used to encode the grasping interaction is detailed in section 5.5. During the execution of the task, a sequence of the elementary primitives, selected among the set of pre-defined primitives, is extracted from the raw data provided by the data glove. The detection of primitives is performed by a Bayesian model detailed in section 5.6. The set of pre-defined primitives is shown in Figure 5.5. The set of task primitives is defined and learned *a-priori* from human demonstrations.

The diversity of the demonstrations promotes the exploration of the variability of the strategies used by humans to perform the same task. The essential primitives of those strategies will emerge as permanent elements. Then, it is possible to build the temporal and functional relations between those elements to find a canonical representation of those strategies. This canonical representation of different in-hand manipulation tasks and the learned parameters describing each of the primitives can be transferred to robotic platforms (not addressed in this thesis).

5.4 Haptic sensory data

This work considers that the hand of the participant manipulating the object is instrumented with a data glove, which is equipped with a distributed tactile sensing array throughout the hand surface (palm and fingers). The methods presented in this work

were formulated considering the *Tekscan Grip* (Tekscan Inc, Boston, MA, U.S.) tactile sensing array (Figure 5.4). However, the proposed approach can be adapted easily to other types of tactile sensing devices.

During a manipulation task, at each time interaction k , the instrumented hand interacts with the object, producing the haptic sensory output presented in equation 5.1.

$$\begin{aligned} \mathbf{h}_k &= (h^1, h^2, \ldots, h^{360}) \\ h^1, h^2, \ldots, h^N &\in [0, 255] \end{aligned} \tag{5.1}$$

The variables $h^1, h^2, \ldots, h^{360}$ represent the raw tactile sensing outputs of each of the 360 elements of the tactile sensing array. The output of each of the *Tekscan Grip* sensing elements is an eight-bit integer (equation 5.1).

The sensing outputs are used to encode (section 5.5) and categorize (section 5.6.2) different classes of grasping primitives.

Figure 5.4: Representation of the fifteen spatial segments $Region_i$ and their correspondence with the sensing elements of the instrumented glove.

5.5 Encoding of the grasping interaction

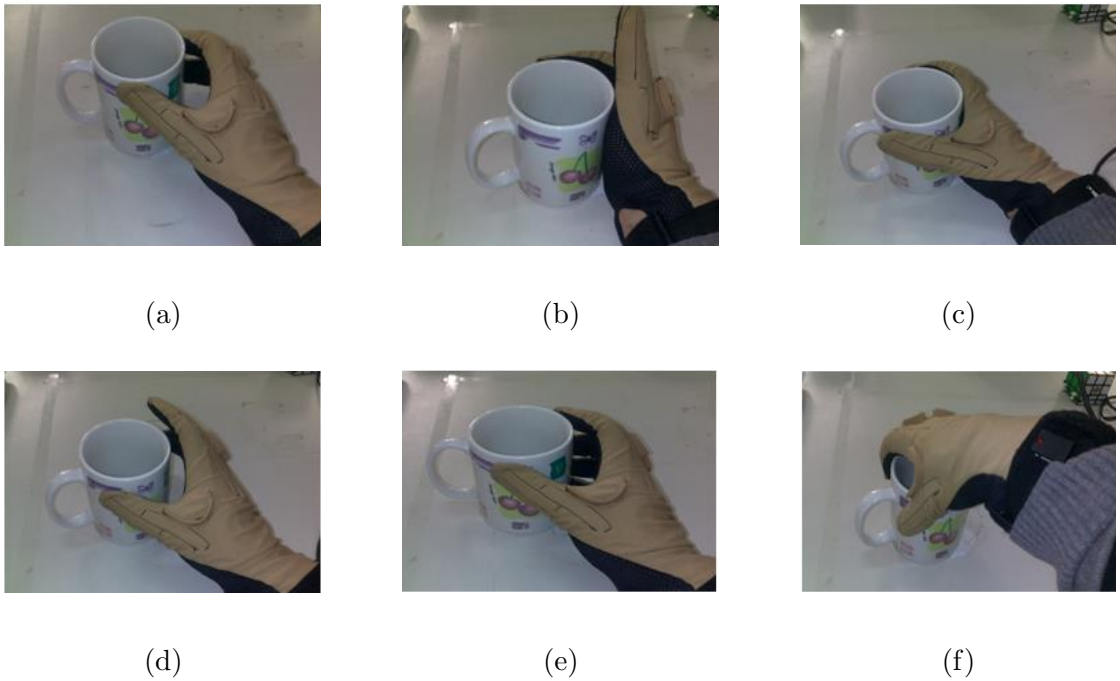
This section proposes the descriptor used to model the tactile sensing signatures produced during the interaction between the instrumented hand and the manipulated object. The tactile sensing device *Tekscan Grip* consists of 360 sensing elements distributed by the hand palm and finger surface, as presented in Figure 5.4. This work groups the 360 sensing elements in 15 regions (highlighted in red, Figure 5.4).

The contact sensing output of each of these 15 regions $Region_i$ of the hand is described by the variable presented in equation 5.2 .

$$\begin{aligned} \mathbf{h}_k^{'} &= (h^{'1}, h^{'2}, \ldots, h^{'15}) \\ h^{'1}, h^{'2}, \ldots, h^{'15} &\in [0, 255] \end{aligned} \tag{5.2}$$

The variable $h^{'i}$ (equation 5.3) represent the mean output of the tactile sensing elements h^j belonging to the $Region_i$ of the instrumented hand shown in Figure 5.4.

$$h^{'i} = mean(\quad \{\forall_{j \in Region_i} \quad h^j\} \quad) \tag{5.3}$$



(a) (b) (c)

(d) (e) (f)

Figure 5.5: Human demonstration of the grasping primitives: a) $Primitive_1$, b) $Primitive_2$, c) $Primitive_3$, d) $Primitive_4$, e) $Primitive_5$, f) $Primitive_6$. $Primitive_7$ corresponds to a grasp in which the hand does not contact the object.

5.6 Recognition of the grasping primitive

5.6.1 Random variables of the model

The Bayesian model π_{grasp} presented in this section is used to discriminate different types of grasp primitives during a manipulation task. A grasping primitive recognized by the system, at time iteration k, is represented by the discrete random variable G_k, described in equation 6.8.

$$\begin{aligned} G_k - &\text{"Category of the grasping primitive."} \\ G_k &\in \{"Primitive_1", \ldots, "Primitive_7"\} \end{aligned} \tag{5.4}$$

This work considers that the system is able to recognize seven different grasping primitives (equation 5.4). Six of these grasping primitives are demonstrated in Figure 5.5. The remaining one, "$Primitive_7$", corresponds to the situation when there is no contact between the hand and the object.

The subset of seven grasping primitives was selected from an extended set of grasping primitives presented in chapter 3. This subset was considered representative for the type of manipulation tasks proposed in this work.

During each time iteration k, the interaction of the instrumented hand equipped with the tactile sensing array and the manipulated object is described by the sensory output $\mathbf{h}_k^{'}$ presented in equation 5.2. The level of tactile activation of each of those 15 regions $h^{'i}$ during the manipulation task is described by the discrete random variable S^i presented in equation 5.5.

$$\begin{aligned} \mathbf{S}_k - &\text{"Tactile activation descriptor of the instrumented hand at instant } k\text{."} \\ \mathbf{S}_k &= (S^1, S^2, \ldots, S^{15}) \\ S_k^i &\in \{"NotActive", "LowActive", "HighActive"\} \end{aligned} \tag{5.5}$$

The tactile activation levels $NotActive$, $LowActive$, and $HighActive$ are defined as proposed in equation 5.6.

$$
\begin{aligned}
\text{"}NotActive\text{"} &: h^{'i} \in [0, 10] \\
\text{"}LowActive\text{"} &: h^{'i} \in [11, 190] \\
\text{"}HighActive\text{"} &: h^{'i} \in [191, 255]
\end{aligned} \tag{5.6}
$$

The 3 levels of discretization of the contact intensity are considered appropriate to characterize and distinguish the fundamental functional levels of mobilization of the different regions of the hand. The proposed contact activation levels are used to distinguish different stages of the interaction between the hand and object. Regions of the hand corresponding to *NotActive* are involved in pre-grasp segments of the manipulation task and in transitions between consecutive re-grasp. *LowActive* regions participate in initial contact with the object and are partially involved in a stage of the manipulation task. *HighActive* regions are highly involved in stabilization of the object.

5.6.2 Inference of the category of grasping primitive

The inference of the category of the grasping primitive G_k at each time iteration step k is performed by the Bayesian model π_{grasp} presented in Figure 5.6a.

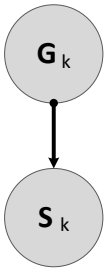
Based on the statistical independence relations between the random variables G_k and $\mathbf{S}_k$ described in Figure 5.6a, the joint probability distribution function $P(G_k, \mathbf{S}_k|\pi_{grasp})$ is decomposed as summarized in Figure 5.6b and presented in equation 5.7.

$$P(G_k, \mathbf{S}_k|\pi_{grasp}) = P(\mathbf{S}_k|G_k, \pi_{grasp})P(G_k|\pi_{grasp}) \tag{5.7}$$

The factor $P(\mathbf{S}_k|G_k, \pi_{grasp})$ expresses the likelihood of a specific grasping contact profile, given a category of grasping primitive. This probability distribution function is modelled by a histogram function as described in detail in section 5.6.3. The factor $P(G_k|\pi_{grasp})$ expresses the *a-priori* probability distribution function of the category of the grasping primitive. In this work, $P(G_k|\pi_{grasp})$ is modelled by a uniform probability distribution function.

The category of grasping primitive G_k is inferred by running the Bayesian program described in Figure 5.6b with the question proposed in equation 5.8.

$$P(G_k|\mathbf{s}_k, \pi_{grasp}) = \frac{P(\mathbf{s}_k|G_k, \pi_{grasp})P(G_k|\pi_{grasp})}{\sum\limits_{G_k} P(\mathbf{s}_k|G_k, \pi_{grasp})P(G_k|\pi_{grasp})} \tag{5.8}$$



(a)

Bayesian Program

- Description
  - Specification
    - Relevant variables:
    - $\mathbf{S}_k, G_k$
    - Decomposition:
    - $P(\mathbf{S}_k, G_k|\pi_{grasp}) = P(\mathbf{S}_k|G_k, \pi_{grasp}).P(G_k|\pi_{grasp})$
    - Parametric forms:
    - $P(\mathbf{S}_k|G_k, \pi_{grasp}) \equiv$ *Histogram*
    - $P(G_k|\pi_{grasp}) \equiv$ *Uniform*
  - Identification:
  - *Histogram learned during the training sessions: Human demonstrations.*
- Question:
- $P(G_k|\mathbf{S}_k, \pi_{grasp})$

(b)

Figure 5.6: Description of the Bayesian model π_{grasp} "Recognition of the grasping primitive". a) Graphical representation. b) Bayesian program.

The estimated category of the grasping primitive $\hat{g}_k$ is given by equation 5.9 via *Maximum a-Posteriori* decision rule (MAP).

$$\hat{g}_k = \arg\max_{G_k} P(G_k|\mathbf{s}_k, \pi_{grasp}) = \arg\max_{G_k} P(\mathbf{s}_k|G_k, \pi_{grasp})P(G_k|\pi_{grasp}) \tag{5.9}$$

5.6.3 Determination of $P(\mathbf{S}_k|G_k, \pi_{grasp})$

The parameters of the histogram function modelling the probability distribution function $P(\mathbf{S}_k|G_k, \pi_{grasp})$ are learned during a training period. Each of the grasping primitives G_k is demonstrated for a pre-defined number of training runs. For each training run, the corresponding contact signature of the instrumented hand $\mathbf{S}_k$ is acquired. After completing the training runs, the parameters of the histogram function are statistically estimated.

This methodology is demonstrated during the presentation of the experimental results (section 5.7.2).

5.7 Experimental results

5.7.1 Experimental setup

During the Human demonstrations of the in-hand manipulation tasks, the subject wears in the right hand an instrumented glove (*Cyberglove II*) with a tactile sensing array (*Tekscan Grip System*) attached to the palm and fingers.

The objects that are placed on the top of a table are manipulated only with one hand (right hand). The subject is seated during the demonstration of in-hand manipulation tasks. The data from the tactile sensing array is sampled at 500 Hz. The configuration of the tactile sensing array, as well as the typical configuration of the experimental area during the task demonstration, are shown in Figure 5.8.

5.7.2 Learning of the grasping primitives $P(\mathbf{S}_k|G_k, \pi_{grasp})$

During the training period, a participant performs five runs demonstrating each of the seven grasping primitives, illustrated previously in section 5.6. The data acquired from the demonstrations is used to estimate the parameters of the probability distribution function $P(\mathbf{S}_k|G_k, \pi_{grasp})$ for each primitive G_k. The results of the learning stage of the grasping primitives are shown in Figure 5.7.

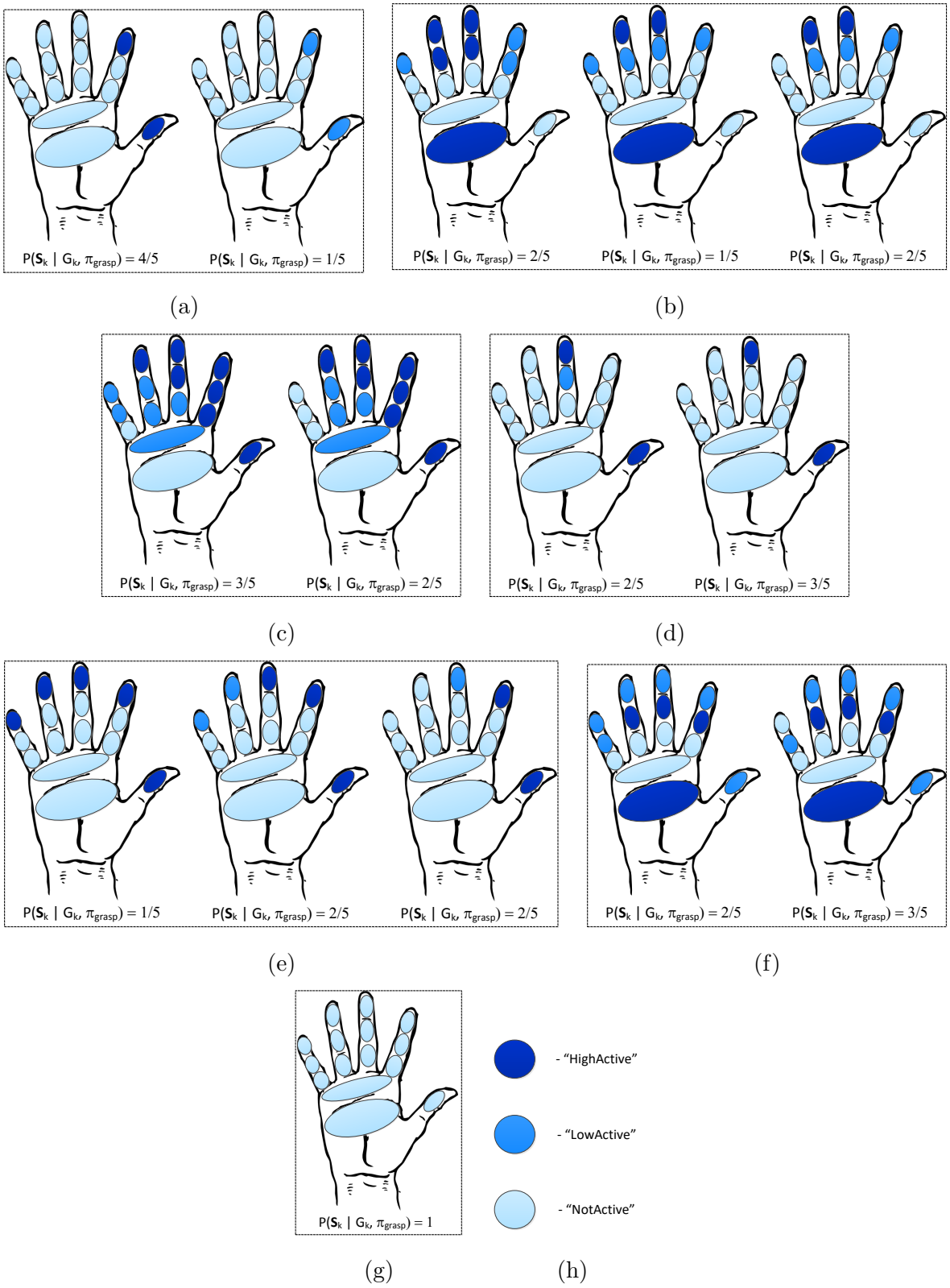


Figure 5.7: Illustration of the probability distribution function $P(\mathbf{S}_k|G_k, \pi_{grasp})$ learned from the human demonstration data (training period). a) $Primitive_1$. b)$Primitive_2$. c)$Primitive_3$. d)$Primitive_4$. e)$Primitive_5$. f)$Primitive_6$. g)$Primitive_7$. h) Colormap

5.7.3 Detection of grasp primitives in manipulation tasks

The approach proposed in section 5.6, to segment a human manipulation task as a sequence of grasping primitives, was tested for two different tasks, as shown in Figure 5.8.

(a)

(b)

Figure 5.8: Human demonstration of the tasks. a) *Task I: "Mug reorientation"*. b) *Task II: "Mug displacement/elevation"*.

During the execution of both tasks, the participant is seated comfortably in front of a table. A mug is placed on top of the table in its initial configuration.

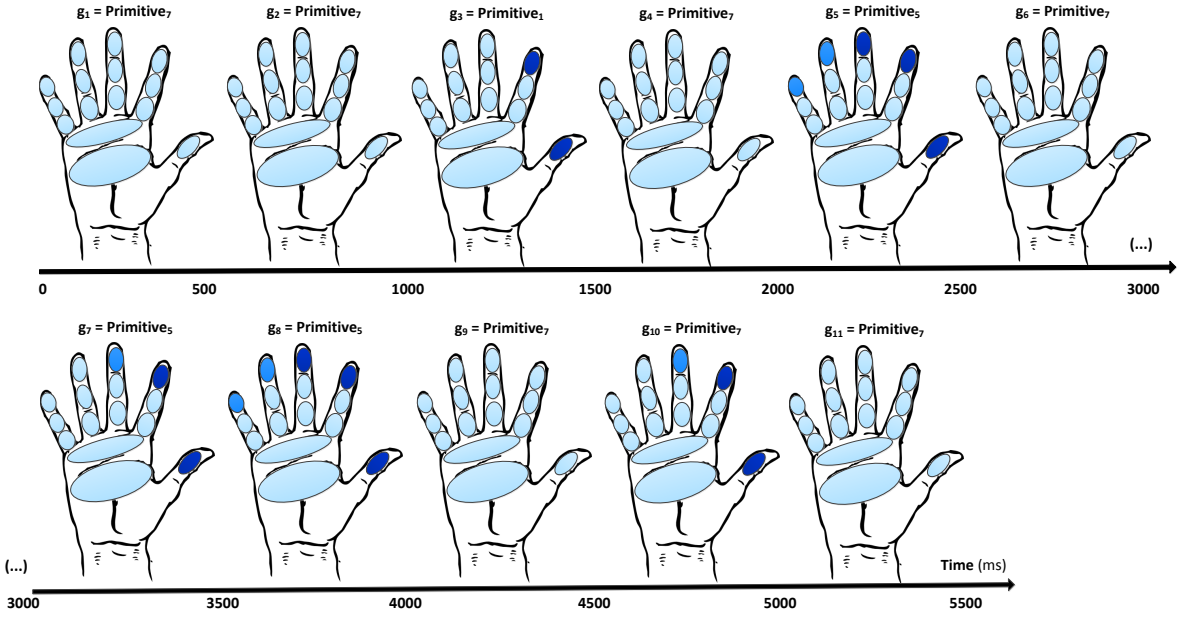
In *Task I: "Mug reorientation"*, the participant rotates the mug around the longitudinal axis. This rotation moves the the handle of the mug to a pose suitable to be grasped by the right hand of the participant (Figure 5.8a).

In *Task II: "Mug displacement/elevation"*, the participant grasps the mug and elevates it along the direction of the longitudinal axis. The participant finishes the task by placing the mug back on the table (Figure 5.8b).

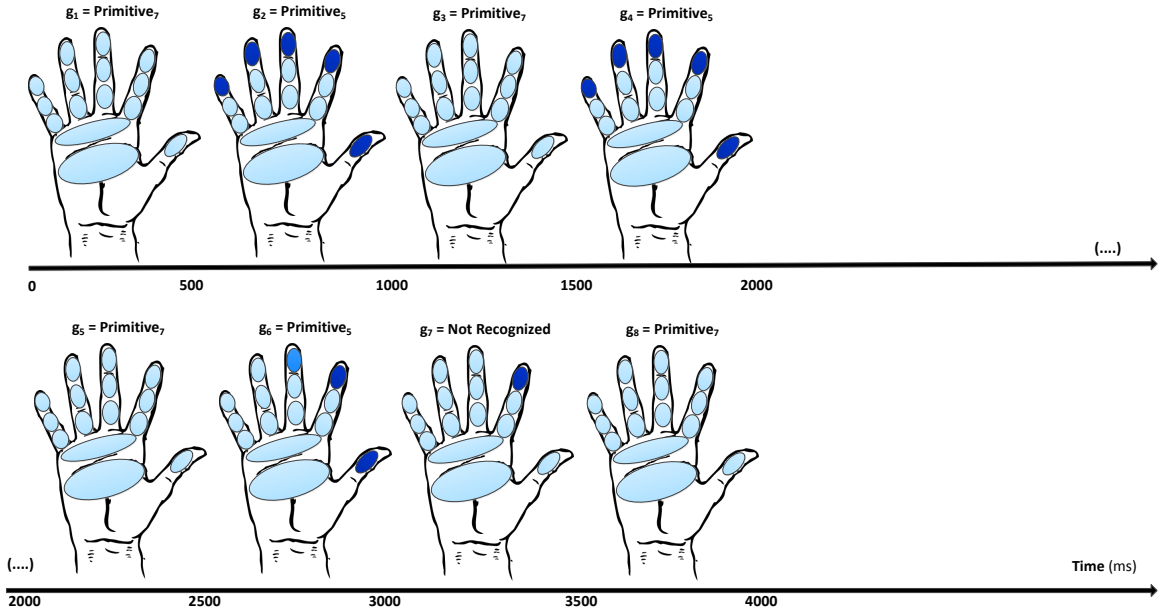
The data acquired during the demonstrations of the two tasks was segmented and time-averaged using a time window of 500 ms. Each time iteration k corresponds to a temporal segment of the sensory data.

The results of the segmentation of the data and recognition of grasping primitives are presented in Figure 5.9 (*Task I: "Mug reorientation"*) and Figure 5.10b (*Task II: "Mug displacement/elevation"*). In both tasks, the first segments are categorized as $Primitive_7$. The human hand is not contacting the object yet. The segments correspond to the reach-to-grasp stage, when the hand moves toward the object which will be manipulated.

The runs of *Task I: "Mug reorientation"* (figure 5.9) were segmented, by the Bayesian model π_{grasp}, on a cyclic sequence of grasping ($Primitive_1$, $Primitive_5$) and releasing ($Primitive_7$) the object. The sequence was used to reorient the mug placed on top of

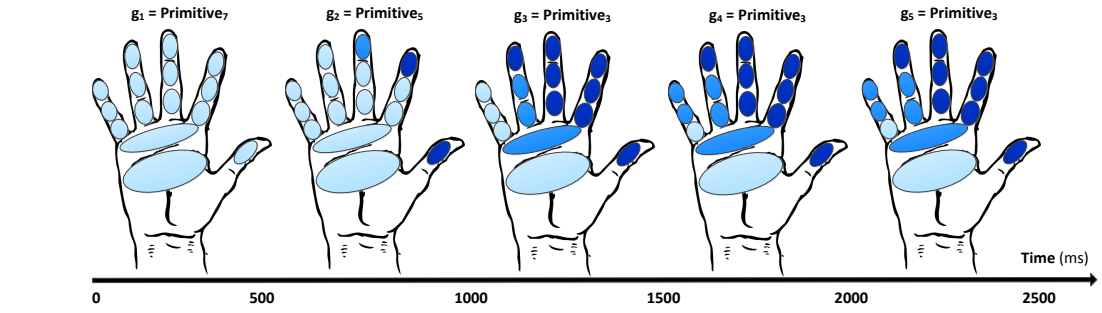


(a)

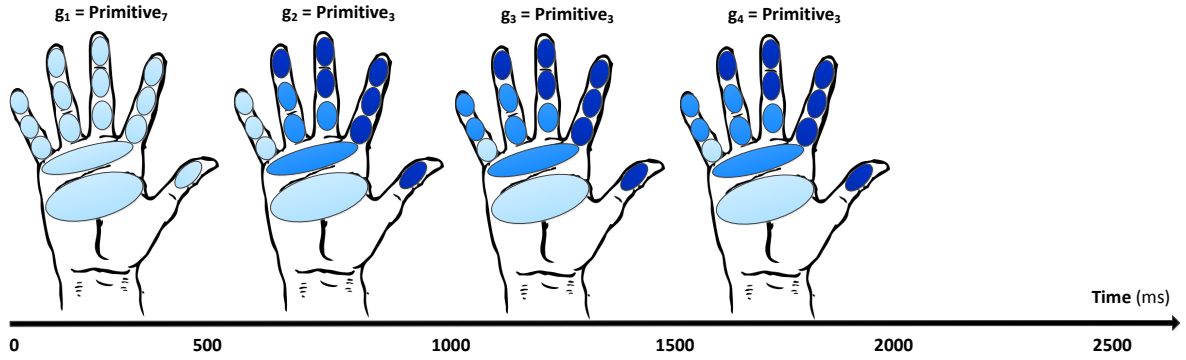


(b)

Figure 5.9: Grasping primitives $\hat{g}_k$ inferred from the data acquired during the execution of *Task I: "Mug reorientation"*. a) Run 1. b) Run 2. Colormap represented in figure 5.7h.



(a)



(b)

Figure 5.10: Grasping primitives $\hat{g}_k$ inferred from the data acquired during the execution of *Task II: "Mug displacement/elevation"*. a) Run 1. b) Run 2. Colormap represented in figure 5.7h.

the table. The sequence of grasp-release allows the participant to reposition the hand on the object, adapting the grasp configuration to the new pose of the object. This strategy promotes the maximization of the effect of the subsequent grasp primitive actuating the object. The regions of the fingers recruited during the reorientation of the mug are predominantly the distal segments of the index, thumb, and middle fingers. These grasping primitives are involved in actions requiring the fine and precise control of the movements of the object: precision grasps.

The runs of *Task II: "Mug displacement/elevation"* (Figure 5.10) were segmented, by the Bayesian model π_{grasp}, on a continuous sequence of grasping primitives ($Primitive_3$, $Primitive_5$). Due to the objective of the task (object displacement), the object was not reoriented. The grasping primitives modelling the strategy are characterized by the recruitment of larger extensions and with more intensity of the palm and surface of the fingers. These primitives provide powerful grasps, contributing to the stability of the execution of the task: power grasps.

5.8 Conclusions

This work presents an approach to modeling the mechanisms underlying the strategies performed by humans to perform manipulation tasks requiring the in-hand manipulation (reorientation and repositioning) of objects. The description and representation of the tasks is made symbolically by using a set of primitives defined on the tactile domain. Each primitive represents a specific spatial distribution of tactile force intensities across the palm and fingers.

The Bayesian model π_{grasp} was able to categorize different types of grasping primitives using as input only the hand-object contact interaction signature. The sequence of primitives modelling two different manipulation tasks was inferred by the Bayesian model π_{grasp}. *Task I* required regrasping and precision grasps. *Task II* was demonstrated using power grasps.

Chapter 6

Categorization of soft objects during haptic exploration tasks

6.1 Introduction

Due to the introduction of robotic platforms in new types of environments (chapter 1), the principles and demands guiding the implementation of robotic platforms are changing. The robotic systems need to interact autonomously with a wide variety of objects (size, shape, compliance, and texture) [Feix et al., 2014]. The work presented in this chapter is focused on the study of objects with different hardness-softness properties [Tiest, 2010].

Typically, the autonomous planning of a robotic manipulation task starts by the estimation of an initial model of the object, through the extraction of features from the vision data, as described in chapter 2 and [Bohg, 2011]. The initial estimation of distance to object, shape affordances, and other characteristics of the surface is based on previous perceptual experiences. It allows the robot to infer several parameters of the reach-to-grasp movement and initial grasp required to hold the object without slipping.

The contributions presented in this chapter are associated with the manipulation movements happening after this initial interaction. When the perceived model representing the object is not sufficiently informative to perform the required task, the system uses the robotic hand to explore the object progressively. The haptic exploration movements are used to perceive complementary properties of the object such as hardness, texture, weight, shape, and temperature (chapter 3).

This chapter presents a probabilistic spatial framework suitable to integrate multi-modal data (vision, tactile, and motion) acquired during the interaction with an object. The multi-modal data is used to build a perceptual model of the manipulated object and infer the category of material being explored. The categories of material considered in this chapter have different perceived hardness characteristics. In terms of hardness characteristics, this work considers the perception that the robotic hand receives when it interacts with objects made of compliant materials. The methods proposed in this

chapter do not achieve a full, precise, and universal characterization of the objects as studied in materials science (stiffness, Young's modulus, mass spring system, finite element methods). The approach presented in this chapter formulates a descriptor defined from features extracted during the interaction between an exploratory element and a surface, to discriminate different objects using its previous knowledge about a finite set of objects.

The description of the perceived hardness can contribute to the discrimination and recognition of objects and adapt the manipulation strategies accordingly.

6.2 Related works

The study of deformable objects has been a research field explored extensively in very different areas: computer vision [Dufour et al., 2011] (tracking, reconstruction, and recognition), computer graphics [Garre et al., 2011] [Ni et al., 2011] (primarily for rendering purposes), industrial materials sciences (very restricted and controlled tests and applications labs or clean rooms with high-accuracy measurement devices), medicine [Liu et al., 2010a] [Ni et al., 2011] (organ analysis and anatomical abnormality detection, virtual haptics simulation for medical training).

Table 6.1: Comparison between the contributions of this work and the related works

Study	**Apparatus**		**Haptic Perception**	
	Platform [a]	Sensing [b]	Approach [c]	Features [d]
This work	HH	HS	P	CI, CD
[Okamura et al., 2001]	RS	HS	D	C
[Oddo et al., 2011]	RS	HS	D	V, T
[Hongbin Liu, 2011]	RS	HS	P	F
[Fishel and Loeb, 2012]	RS	HS	P	CD, CI, V, H
[Hui and Kuchenbecker, 2014]	RS	HS	D	CI, CD
[Xu et al., 2013]	RS	HS	P	CI, V, T
[Faria et al., 2012]	HH	HS, VS	P	C, L
[Chitta et al., 2011]	RS	HS	D	CI
[Frank et al., 2010]	RS	HS, VS	D	L, CI, CD

[a] RS- robotic system; HH- human hand.
[b] HS- haptic sensing; VS- visual sensing.
[c] P- probabilistic; D- deterministic.
[d] T- texture; C- curvature; F- friction coefficient; L- RGBD cloud points; CI- contact intensity; CD- contact indentation; V- micro-vibration; H- heat flow.

Typically, soft objects can be represented by 3D computational models in the discrete (e.g. mass spring systems) or continuous (e.g. finite element methods) domain. This type of computational model provides an accurate description of the dynamics of soft objects. However, the elaboration of this type of model by an autonomous robotic platform faces several challenges and constraints.

Some experimental conditions of the material sciences laboratories are difficult to replicate in the context of autonomous robotic manipulation (restrictions on hardware design: hand dexterity, calibration accuracy of tactile and vision sensors, real-time computational

power, and energy efficiency). Additionally, autonomous robotic platforms must deal with the uncertainty associated with the dynamics of the environment and noise of the data measurements.

In the cognitive autonomous robotic manipulation field, different approaches have been proposed to improve the perceived representation of the object during the manipulation and exploration tasks. Some approaches are dedicated to the estimation of the surface characteristics of the object, such as texture and stickiness. [Hongbin Liu, 2011] proposes an algorithm to categorize objects using the surface friction properties. The friction coefficients of the surfaces are estimated from force and torque data sensed by a robotic finger.

The work by [Oddo et al., 2011] proposes the design of a robotic fingertip with an artificial haptic perception system for surface texture discrimination. [Okamura et al., 2001] proposes a method to identify different types (cusp, step, and bump) of surface features during the lateral sliding of a robotic finger.

Recently, several works have used the *Biotac* robotic fingertip to discriminate objects. This compact device replicates the anatomic characteristics of a human fingertip, being capable of sensing temperature, micro-vibration, and pressure. [Fishel and Loeb, 2012] applied lateral sliding movements of a metallic bar equipped with *Biotac* to recognize different textures. The work also studied the impact of velocity and pressure of lateral sliding movement to the performance of the system. The work by [Hui and Kuchenbecker, 2014] used the *Biotac* device integrated on the tip of a probe to explore simulated samples of tissues. Haptic palpation was used to recognize lumps on tissues by integrating pressure and indentation depth features. The work of [Xu et al., 2013] integrates *Biotac* fingertip on a *Shadow* robotic hand and performs two types of exploratory movements to extract three type of haptic features: texture (lateral sliding movement), compliance (palpation/press-and-release), and heat flow (palpation/press-and-release). These three features are integrated simultaneously to discriminate different classes of objects.

Other works are focused on representing the perceived shape of the object to find suitable regions for stable grasping. [Faria et al., 2012] builds a volumetric representation of the shape of the object as long as the object is progressively explored. The representation is decomposed in volumetric primitives, and the best elementary regions for stable grasping are identified based on human demonstrations. [Chitta et al., 2011] infers the internal state (empty, full, open, closed) and recognizes objects using deformation signatures during a haptic exploration task. The approach only uses tactile data.

The work of [Frank et al., 2010] tries to integrate information from a vision and force/torque sensing system to estimate the elasticity properties of objects during a haptic exploration task. The work establishes a relation between the deformation induced by the

robotic end-effector and the intensities of the sensed forces. Different objects have distinct deformation / force intensity signatures. The work does not use a dexterous robotic hand to interact with the object. A rectangular bar is used to explore the environment.

The motivation for this work is introduced in section 6.1. The proposed approach is summarized in section 6.3 and compared with related works in Table 6.1.

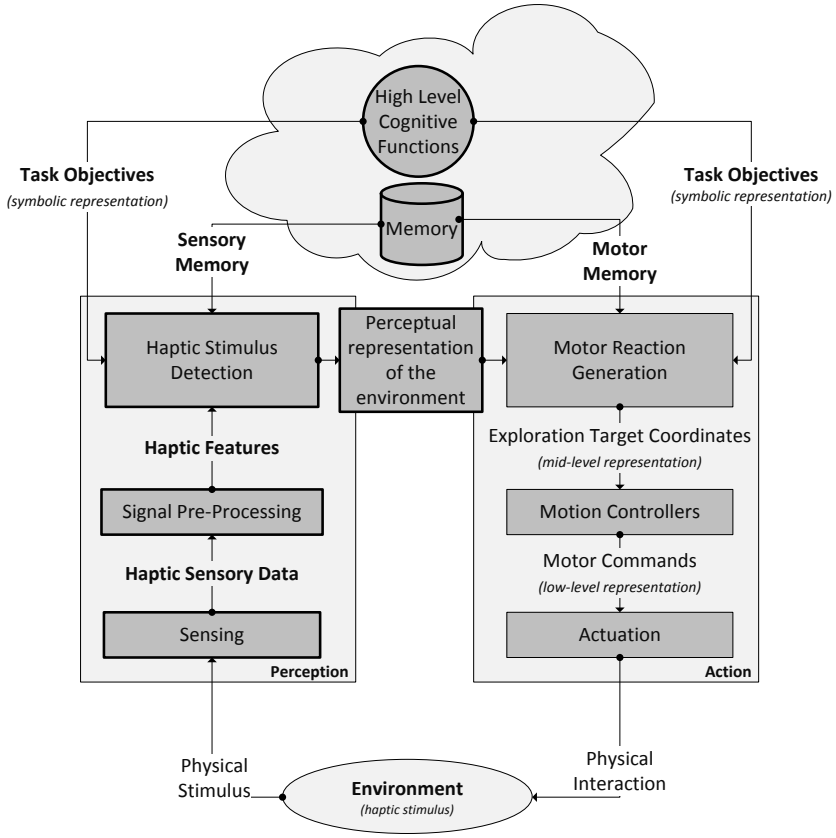


Figure 6.1: Schematic representation of the action-perception loop involved in the haptic exploration of surfaces. The elements highlighted with bold border are discussed in this chapter.

6.3 Approach overview

To endow the robotic systems with the capability to recognize and categorize distinct materials (different hardness-softness properties), the approach presented in this chapter analyses the principles and strategies used by humans to perform such types of tasks. The capability to discriminate the materials results from the integration of different types of haptic data.

During the haptic exploration tasks, the perception and discrimination of hardness-

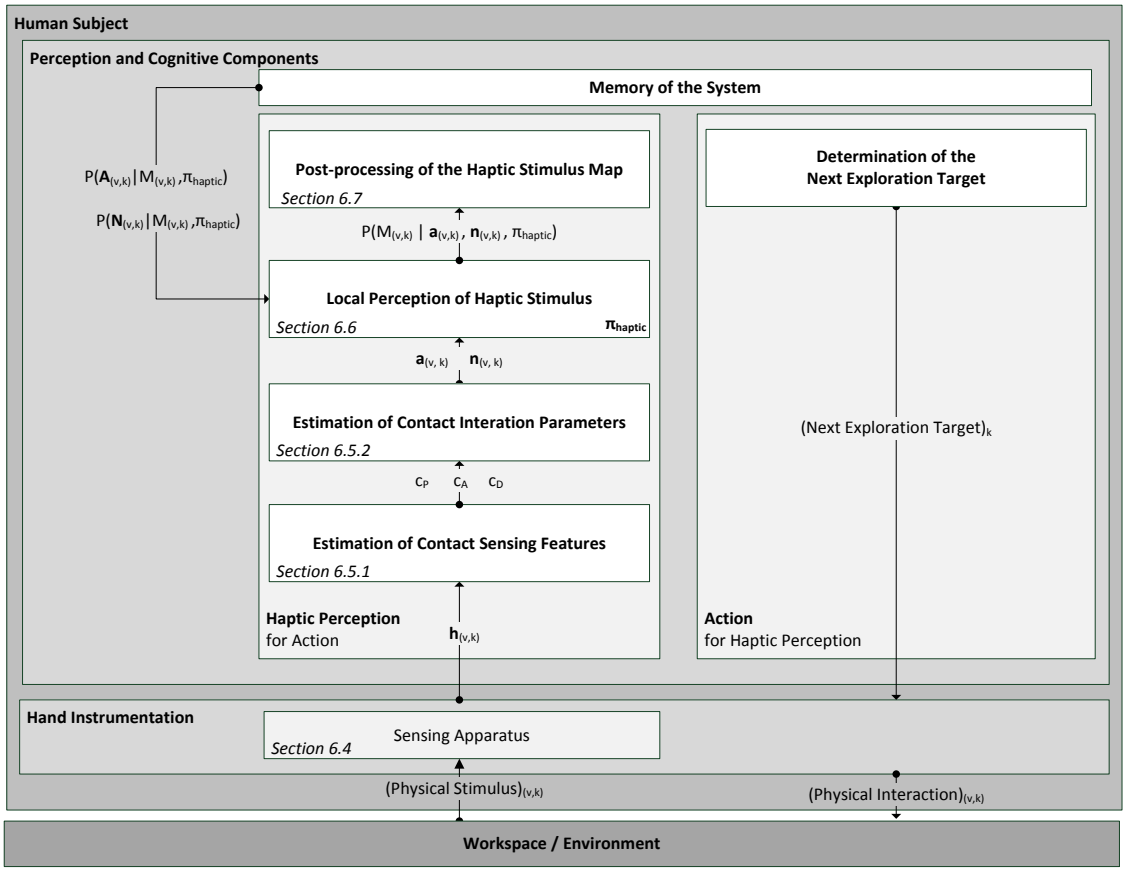


Figure 6.2: Global architecture of the approach presented in this chapter. The main contributions are identified and referenced in the scheme (description of the sensing apparatus, estimation of the contact interaction parameters, local perception of haptic stimulus, and post-processing of haptic stimulus map). The variables involved in the flow of data are summarized in Table 6.2.

Table 6.2: Summary of the relevant variables used in this chapter

Variable	Description	Domain
v	Cell of the workspace grid.	$\mathbb{R}^2$
k	Time / exploration iteration.	$\mathbb{N}^0$
$\mathbf{h}_{(v,k)}$	Raw haptic sensing data acquired on v.	$\mathbb{R}^n$ *
c_P	Contact sensing: intensity.	$\mathbb{R}_0^+$
c_A	Contact sensing: area.	$\mathbb{R}_0^+$
c_D	Contact sensing: indentation distance.	$\mathbb{R}_0^+$
$M_{(v,k)}$	Material category of v	$\{Material_1, \ldots, Material_3\}$
$A_{(v,k)}$	Cutaneous contact interaction parameters.	$\mathbb{R}^2$
$N_{(v,k)}$	Kinesthetic contact interaction parameters.	$\mathbb{R}^2$
$H(v)$	Entropy of the grid cell v.	$\mathbb{R}$

* In this work, the description of the sensory apparatus of the instrumented hand platform follows a generic formulation. In this context, n represents the dimensionality of the raw sensing data provided by the haptic sensory apparatus of the robot.

softness characteristics of objects depends on the simultaneous integration of cutaneous and kinesthetic information by performing "press-and-release" movements [Lederman and Klatzky, 1987] - active haptic perception. This haptic exploration strategy and integration of multi-modal data was demonstrated by psychophysical experiments performed in [Srinivasan and LaMotte, 1995] and [Lederman and Klatzky, 1987], respectively.

As described in detail in chapter 3 and shown using a simplified representation in Figure 6.1, several types of haptic features can be extracted from the cutaneous and kinesthetic data. In this chapter, the human hand is instrumented with a tactile sensing array and a motion tracking system, described in section 6.4.

The haptic data is subjected to a feature extraction processing pipeline formulated in chapters 6.5.1 (extraction of contact sensing features) and 6.5.2 (extraction of contact interaction parameters). The features extracted in the previous stages are integrated by the Bayesian model π_{haptic} *"Local Perception of the Haptic Stimulus"* (chapter 6.6).

The Bayesian model π_{haptic} infers and updates the category of material describing the region of the workspace which was explored. The workspace is represented by a planar 2D probabilistic (inference) grid (details about this representation framework are provided in chapter 2). The description of the perceived category of material is updated after each haptic exploration iteration step, integrating new sensory inputs on the haptic processing pipeline and on the Bayesian model π_{haptic}.

This representation framework is suitable to integrate noisy multi-modal sensory inputs from multiple exploratory elements (dexterous robotic hands). The grid also allows the representation of non-homogeneous objects/surfaces.

In this chapter, a exploration iteration step corresponds to a press-and-release exploration movement. This is the exploration pattern used to extract hardness-softness properties of objects (literature review presented in chapter 3). The perceived spatial representation of the object is improved and optimized using methods presented in section 6.7.

In the approach proposed in this chapter, the press-and-release movements are performed at pre-defined locations of the workspace. In the next chapter of this thesis (chapter 7), the the regions of the workspace which are going to be explored next are inferred online and autonomously as the exploration progresses (closes the action-perception loop).

6.4 Haptic sensory data

During the haptic exploration of an object, the sensory apparatus interacts with the object producing raw sensory outputs represented by the variable $\mathbf{h}_{(v,k)}$. The type of sensory

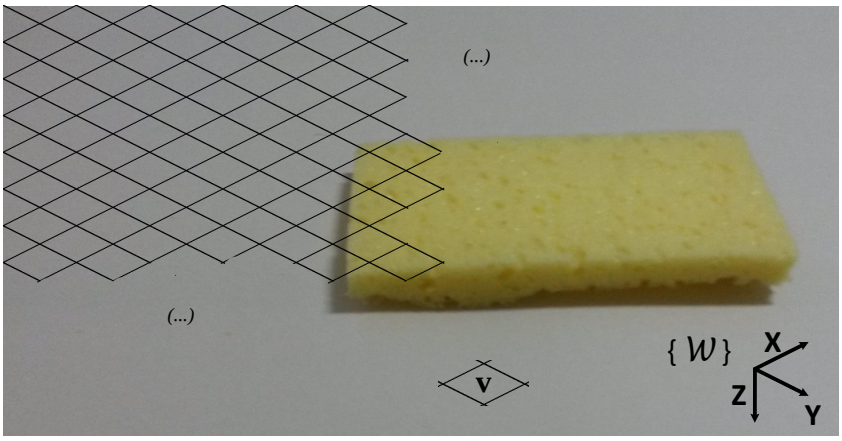


Figure 6.3: Partial representation of the bidimensional grid framework, which is used to describe the workspace region (e.g.: sponge object).

outputs $\mathbf{h}_{(v,k)}$ produced during the exploration process is dependent on the technology and design of the sensing apparatus (force, torque, tactile, temperature, vibration; single point, array), as well as the type of movement strategy used to perform the haptic exploration.

The methods proposed in this chapter are defined following a generic formulation. This work considers that the agent involved in the exploration process is equipped with a tactile sensing array and that the exploration strategy consists of a sequence of "press-and-release" (palpation) movements.

Thus, at each time iteration k, the haptic sensory output resulting from the interaction between the exploratory element and the material at the region v of the workspace is described by the variable presented in equation 6.1.

$$\begin{aligned} \mathbf{h}_{(v,k)} &= (h^1, h^2, \ldots, h^N, h^X, h^Y, h^Z) \\ h^1, h^2, \ldots, h^N \in \mathbb{R}_0^+, &\quad h^X, h^Y, h^Z \in \mathbb{R} \end{aligned} \tag{6.1}$$

The variables $h^1, h^2, \ldots, h^N$, represent the tactile sensing outputs of each of the N elements of the tactile sensing array. The outputs h^i express the contact intensity sensed by each element of the array. The Cartesian coordinates of the end-effector of the exploration system are expressed in the inertial reference frame $\{\mathcal{W}\}$ and represented by the variables h^X, h^Y, h^Z.

A cell v of the bidimensional grid and the inertial reference frame $\{\mathcal{W}\}$ are illustrated in Figure 6.3.

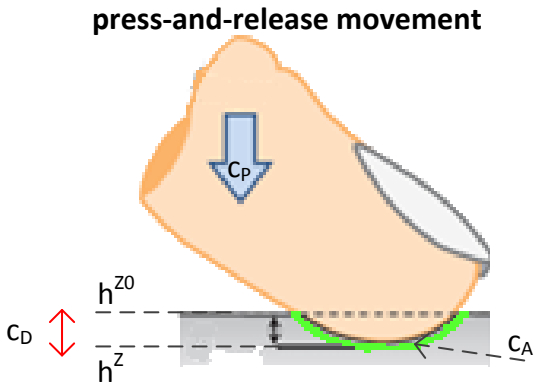


Figure 6.4: Schematic representation of the contact sensing variables extracted during the haptic exploration of a surface. Variables described in section 6.5. Image adapted from [van Kuilenburg et al., 2013].

6.5 Pre-processing of the haptic sensory data

6.5.1 Determination of the contact sensing features c_P, c_D, c_A

The sensory outputs $\mathbf{h}_{(v,k)}$ presented in equation 6.1 are processed to extract features modelling the contact interaction behaviour of the material. A description of the contact sensing features is presented in Figure 6.4.

The total intensity of the contact is described by the variable c_P presented in equation 6.2. This variable corresponds to the total sum of the individual outputs of the elements h^j of the tactile array.

$$c_P = \sum_{j=1}^{N} h^j \tag{6.2}$$

As long as the "press-and-release" exploration movement is performed, the area of the tactile sensing array contacting the surface of the object changes. This area is described by the variable c_A described in equation 6.3 .

$$
\begin{aligned}
c_A &= \vartheta h^a \\
\vartheta &= \#\{\forall_{h^j} : h^j > 0\} \\
c_A &\in [0, N h^a]
\end{aligned} \tag{6.3}
$$

The total area of contact interaction is determined by the number of active tactile sensing elements of the array, ϑ, and the area, h^a, of each element of the array, as described in equation 6.3. A sensing element of the tactile array is considered active if its individual output h^j is higher than zero (equation 6.3). This means that if a sensing elements is contacting the object, the sensing element is active.

In this work, the objects to be explored are made of soft materials. Thus, during a "press-and-release" exploration movement, the deformation of the surface of the object can be described by the variable c_D, as shown in equation 6.4 and illustrated in Figure 6.4.

$$
\begin{aligned}
c_D &= h^Z - h^{Z_0} \\
c_D &\in \mathbb{R}_0^+
\end{aligned} \tag{6.4}
$$

This work considers that the "press-and-release" exploration motion is performed uniaxially along the Z axis of the inertial reference $\{\mathcal{W}\}$. Considering this, c_D measures the indentation distance, h^Z, of the exploratory element relative to the point where it made the initial contact with the natural surface of the object, h^{Z_0}.

The contact sensing features c_A, c_P, c_D extracted from the haptic sensory data $\mathbf{h}_{(v,k)}$ are used as input to the estimation of the contact interaction parameters described in section 6.5.2.

6.5.2 Estimation of the cutaneous and kinesthetic interaction parameters

During a "press-and-release" exploration movement, several samples of the features c_A, c_P, c_D are acquired. The profile represented by the set of the corresponding data points (c_A, c_P, c_D) encodes relevant information about the haptic characteristics of the material being explored and can be used to discriminate different classes of materials.

The categorization of different materials based on their contact interaction signature can be performed by integrating two different sources of information. As demonstrated by [Scilingo, 2010], the simultaneous integration of information related to cutaneous tactile

sensing (relation between contact intensity and contact area) and information related to kinesthetic sensing (relation between contact intensity and contact indentation level) during a haptic exploration task, is essential to perceive and discriminate soft materials based on their haptic properties. This multi-modal integration is also essential to resolve some ambiguities which can occur if each of type of information is used separately.

Considering an exploratory element with a spherical design [Scilingo, 2010] (e.g.: human fingertip), the cutaneous component of the contact interaction signature is described by the relation presented in equation 6.5.

$$\begin{aligned} c_P &= a_1 c_A^{\frac{3}{2}} + a_2 \\ a_1, a_2 &\in \mathbb{R} \end{aligned} \tag{6.5}$$

The kinesthetic component in the contact interaction signature is described in equation 6.6 .

$$\begin{aligned} c_P &= n_1 c_D^{\frac{3}{2}} + n_2 \\ n_1, n_2 &\in \mathbb{R} \end{aligned} \tag{6.6}$$

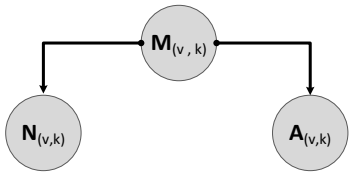
The constants a_1, a_2 and n_1, n_2 are the cutaneous and kinesthetic interaction parameters, respectively. These parameters are different for each class of materials. They encode the signature of the dynamic behaviour of the material during a "press-and-release" exploration movement. The behaviour is related to the haptic properties of that material. The parameters can be used as descriptors of the class of the materials to discriminate objects made of different materials.

Given some data c_A, c_P, c_D resulting from a "press-and-release" exploration movement, the parameters a_1, a_2 and n_1, n_2 are estimated using the method MLE (Maximum Likelihood Estimation). The general goal of this method is to identify the parameters of the models presented in equations 6.5 and 6.6 which are most likely to have generated the set of data points c_A, c_P, c_D.

The next section presents the Bayesian model proposed in this chapter to discriminate different classes of materials.

6.6 Perception of the haptic stimulus map

This section proposes a Bayesian model to infer the perceived category of haptic stimulus (class of material) from the contact and kinesthetic interaction parameters extracted



(a)

Bayesian Program

Description

Specification

Relevant variables:

$N_{(v,k)}, A_{(v,k)}, M_{(v,k)}$

Decomposition:

$P(N_{(v,k)}, A_{(v,k)}, M_{(v,k)}|\pi_{haptic}) =$

$P(N_{(v,k)}|M_{(v,k)}, \pi_{haptic}).P(A_{(v,k)}|M_{(v,k)}, \pi_{haptic}).P(M_{(v,k)}|\pi_{haptic})$

Parametric forms:

$P(N_{(v,k)}|M_{(v,k)}, \pi_{haptic}) \equiv \mathcal{N}_N(\mu_N(M), \sigma_N(M))$

$P(A_{(v,k)}|M_{(v,k)}, \pi_{haptic}) \equiv \mathcal{N}_A(\mu_A(M), \sigma_A(M))$

For $k = 0$:

Map initialization.

$P(M_{(v,0)}|\pi_{per}) \equiv$ *Uniform*

For $k > 0$:

Map update using previous state information.

$P(M_{(v,k)}|\pi_{per}) = P(M_{(v,k-1)}|n_{(v,k-1)}, a_{(v,k-1)}, \pi_{haptic})$

Identification:

$\mu_N(M), \Sigma_N(M), \mu_A(M), \Sigma_A(M)$

- *Constants defined during the learning stage.*

Question:

$P(M_{(v,k}|n_{(v,k)}, a_{(v,k)}, \pi_{haptic})$

(b)

Figure 6.5: Bayesian model π_{haptic} *Perception of the haptic stimulus map.* a) Graphical representation. b) Bayesian program.

from the sensory data $\mathbf{h}_{(v,k)}$. The Bayesian model π_{haptic} is presented in Figure 6.5a and described in detail in the next sections.

6.6.1 Random variables of the model

The perceived category of the haptic stimulus is modelled by the discrete random variable $M_{(v,k)}$, as described in equation 6.7. N is the total number of different materials which can be discriminated by the system.

$$\begin{aligned} M_{(v,k)} &- \text{"Material category of v"} \\ M_{(v,k)} &\in \{Material_1, \ldots, Material_N\} \end{aligned} \tag{6.7}$$

In this Bayesian model, the information about the cutaneous contact interaction is integrated by the continuous random variable $A_{(v,k)}$, detailed in equation 6.8.

$$\begin{aligned} A_{(v,k)} &- \text{"Cutaneous contact interaction parameters"} \\ A_{(v,k)} &= (a_1, a_2) \\ A_{(v,k)} &\in \mathbb{R}^2 \end{aligned} \tag{6.8}$$

Conversely, the information encoded by the kinesthetic interaction parameters is described by the continuous random variable $N_{(v,k)}$, presented in equation 6.9.

$$\begin{aligned} N_{(v,k)} &- \text{"Kinesthetic contact interaction parameters"} \\ N_{(v,k)} &= (n_1, n_2) \\ N_{(v,k)} &\in \mathbb{R}^2 \end{aligned} \tag{6.9}$$

The description and methods used to determine the parameters a_1, a_2, n_1, n_2 from the haptic sensory data $\mathbf{h}_{(v,k)}$ are presented in section 6.5.2.

6.6.2 Inference of the haptic stimulus category

Figure 6.5a shows a graphical representation of the Bayesian model proposed in this chapter. It describes the statistical independence relationship between the random variables $M_{(v,k)}$, $A_{(v,k)}$ and $N_{(v,k)}$ of the model. Considering the relations presented in Figure 6.5a, the joint probability distribution function $P(M_{(v,k)}, N_{(v,k)}, A_{(v,k)}|\pi_{haptic})$ can be decom-

posed as presented in Figure 6.5b and equation 6.10.

$$\begin{aligned} &P(M_{(v,k)}, N_{(v,k)}, A_{(v,k)}|\pi_{haptic}) = \\ &= P(N_{(v,k)}|M_{(v,k)}, \pi_{haptic})P(A_{(v,k)}|M_{(v,k)}, \pi_{haptic})P(M_{(v,k)}|\pi_{haptic}) \end{aligned} \tag{6.10}$$

The factors $P(N_{(v,k)}|M_{(v,k)}, \pi_{haptic})$ and $P(A_{(v,k)}|M_{(v,k)}, \pi_{haptic})$ express the likelihood of having specific measurements of kinestethic $n_{(v,k)}$ and cutaneous $a_{(v,k)}$ contact interaction parameters for a given type of material (haptic stimulus) $M_{(v,k)}$. $P(M_{(v,k)}|\pi_{haptic})$ is the *a-priori* probability of exploring a specific type of material.

The probability distribution function modelling the category of the haptic stimulus (type of material) is formulated from equation 6.10, as described in equation 6.11.

$$\begin{aligned} &P(M_{(v,k)}|n_{(v,k)}, a_{(v,k)}, \pi_{haptic}) = \\ &= \frac{P(n_{(v,k)}|M_{(v,k)}, \pi_{haptic})P(a_{(v,k)}|M_{(v,k)}, \pi_{haptic})P(M_{(v,k)}, \pi_{haptic})}{\sum\limits_{M_{(v,k)}} P(n_{(v,k)}|M_{(v,k)}, \pi_{haptic})P(a_{(v,k)}|M_{(v,k)}, \pi_{haptic})P(M_{(v,k)}, \pi_{haptic})} \end{aligned} \tag{6.11}$$

The category of the haptic stimulus is inferred using the *Maximum a-Posteriori* (MAP) decision rule, as described in equation 6.12.

$$\begin{aligned} \hat{m}_{(v,k)} &= \arg\max_{m_{(v,k)}} P(M_{(v,k)}|n_{(v,k)}, a_{(v,k)}, \pi_{haptic}) \\ \hat{m}_{(v,k)} &= \arg\max_{m_{(v,k)}} P(n_{(v,k)}|M_{(v,k)}, \pi_{haptic})P(a_{(v,k)}|M_{(v,k)}, \pi_{haptic})P(M_{(v,k)}|\pi_{haptic}) \end{aligned} \tag{6.12}$$

Each of the factors involved in the inference of $\hat{m}_{(v,k)}$ is modelled by a probability distribution function presented in Figure 6.5b and detailed in section 6.6.3.

6.6.3 Determination of $P(N_{(v,k)}|M_{(v,k)}, \pi_{haptic})$ and $P(A_{(v,k)}|M_{(v,k)}, \pi_{haptic})$

The probability distribution functions $P(N_{(v,k)}|M_{(v,k)}, \pi_{haptic})$ and $P(A_{(v,k)}|M_{(v,k)}, \pi_{haptic})$ are normal probability distribution functions. Thus, both are modelled by bi-dimensional Gaussian functions as described by equations 6.13 and 6.14.

$$P(N_{(v,k)}|M_{(v,k)}, \pi_{haptic}) \equiv \mathcal{N}_N(\mu(M), \Sigma(M)) \tag{6.13}$$

$$P(A_{(v,k)}|M_{(v,k)}, \pi_{haptic}) \equiv \mathcal{N}_A(\mu(M), \Sigma(M)) \tag{6.14}$$

For each of the reference materials $M_{(v,k)}$ recognized by the system, the mean μ and covariance matrix Σ of the functions $\mathcal{N}_N$ and $\mathcal{N}_A$ are learned during a training period.

In this training period, a sample of each reference material recognized by the system is explored during several "press-and-release" movements. The (c_P, c_A, c_D) data acquired during each of the "press-and-release" cycles are used to determine the contact interaction parameters (a_1, a_2) and (n_1, n_2) for that material, using the Maximum Likelihood Estimation (MLE) method.

After a pre-defined number of press-and-release cycles, the average μ values of (a_1, a_2) and (n_1, n_2) are determined, as well as the respective covariance matrix Σ.

6.7 Post-processing of haptic stimulus map

As mentioned in the previous section, the haptic properties of the objects explored in the workspace are represented using a probabilistic (inference) grid. The formulation of this type of representation framework (analogous to an occupancy grid) considers that the representation of each grid cell v is independent from the remaining cells v of the grid (see details in section 2). This assumption frequently originates from unexplored grid cells having high uncertainty, even if those regions are surrounded by cells which were explored, having an informative haptic description assigned to them and thus a lower uncertainty.

This section proposes a method which is applied to the final haptic stimulus map provided by the Bayesian model π_{haptic} presented in section 6.6. The methods proposed in this chapter consider some constraints derived from the physical world. The physical world (explored object) has spatial structure; thus, it is expected to find local spatial continuity in the haptic perceptual characterization of contiguous grid cells.

The methods proposed in this section improve the perceptual representation of the haptic properties of the object by estimating the representation of unexplored grid cells based on the representation of neighbour cells which were explored.

The set of unexplored grid cells is defined by the variable ϑ_{unexpl}. The haptic description of an unexplored cell is estimated as the weighted mean of the haptic description of the eight grid cells surrounding that cell, as described in 6.15.

$$\forall_v \in \vartheta_{unexpl}, \quad P(M_{(v,k)}|N_{(v,k)}, A_{(v,k)}, \pi_{haptic}) = \frac{\sum_{j=1}^{8} \frac{1}{H(vj)} P(M_{(vj,k)}|N_{(vj,k)}, A_{(vj,k)}, \pi_{haptic})}{\sum_{j=1}^{8} \frac{1}{H(vj)}} \tag{6.15}$$

$H(vj)$ is the entropy of the grid cell vj (more details in chapter 2). The weights $\frac{1}{H(vj)}$ of the extrapolation method assign a higher contribution to the representation of grid cells with lower entropy $H(vj)$ (less uncertainty). vj represents each of the eight neighbour cells of v. The determination of the entropy of a random variable is detailed in section 2.

6.8 Experimental results

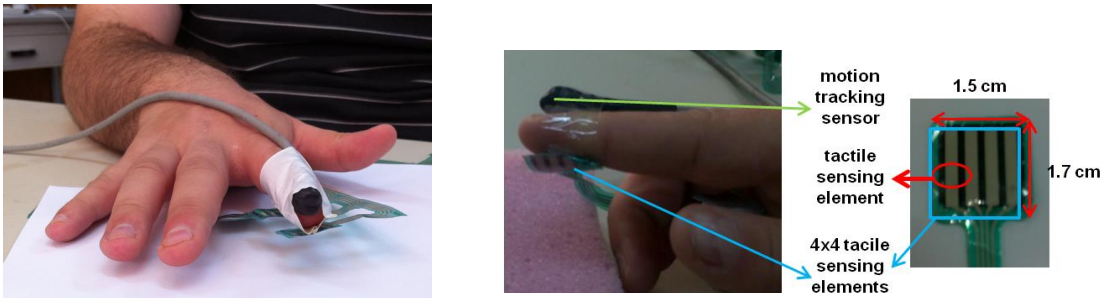
6.8.1 Experimental setup

The methods proposed in this work were tested using a human hand as exploration agent. The right hand of the participant was instrumented with a tactile sensing array *Tekscan Grip System* (Tekscan Inc, Boston, MA, U.S.) and a motion tracking system *Polhemus Liberty* (Polhemus Inc, Colchester, VT, U.S.). The instrumented hand is shown in Figure 6.6b.

The tactile sensing array is attached to the surface of the fingers using glue tape. It provides information about the spatial configuration of the regions of the hand contacting the object, the intensity, and the area of contact (cutaneous information). In this work, only one pad (array with $16 = 4 \times 4$ sensing elements) of the *Tekscan Grip System* is attached to the index fingertip (Figure 6.6b) (Figure 6.6a). The output of each sensing element of the array is an eight-bit integer. This work considers the raw outputs of the tactile sensing array without prior calibration. A method, called equilibration, proposed by the manufacturer of the system, is applied before each run of the data acquisition. The method is used to compensate for unexpected variations of the output between sensing elements. The data of each tactile sensing element is sampled at 50 Hz.

One sensor of the motion tracking device *Polhemus Liberty* was attached using glue tape to the same index fingertip of the tactile sensing array, as shown in Figure 6.6b. This motion tracking sensor provides information about the 6D pose (3D position, 3D orientation) of the fingertip and inherently of the tactile sensing array. This data provided the kinesthetic component of the interaction, related to the indentation depth of the fingertip in the natural shape of the object. The motion tracking sensor was sampled at approximately 30 Hz.

The multi-modal data samples were individually timestamped in millisecond *ms* reso-



(a) (b)

Figure 6.6: Instrumented finger involved in the haptic exploration of objects. a) global overview. b) detailed view of the integration of the *Tekscan Grip* system and the motion tracking sensor *Polhemus Liberty*.

lution by the software applications developed for each of the data acquisition devices (see chapter 4). The clocks of the different computers involved in the architecture of the data acquisition were synchronized using Network Time Protocol (NTP) (see chapter 4).

6.8.2 Learning of the contact interaction parameters of the reference materials

Experimental protocol

During the learning stage of the contact interaction parameters (a_1, a_2), and (n_1, n_2), physical samples of three reference materials $Material_1$, $Material_2$, and $Material_3$ are selected and placed in the experimental area.

The samples selected for this work are presented in Figure 6.7. They were selected to represent common objects of daily life and to have distinct haptic properties. $Material_1$, $Material_2$, and $Material_3$ were evaluated empirically by a human operator (exploration using a non-instrumented human hand) and have increasing levels of perceived hardness, respectively.

Each of the reference materials $Material_1$, $Material_2$, and $Material_3$ was submitted to ten press-and-release exploration cycles using the index fingertip of the instrumented hand, as presented in Figure 6.6b. All the press-and-release exploration cycles were performed in the same region of the surface of the object. Thus, the subject applies a unidirectional movement perpendicular to the natural shape of the physical samples of the reference materials (Figure 6.7d).

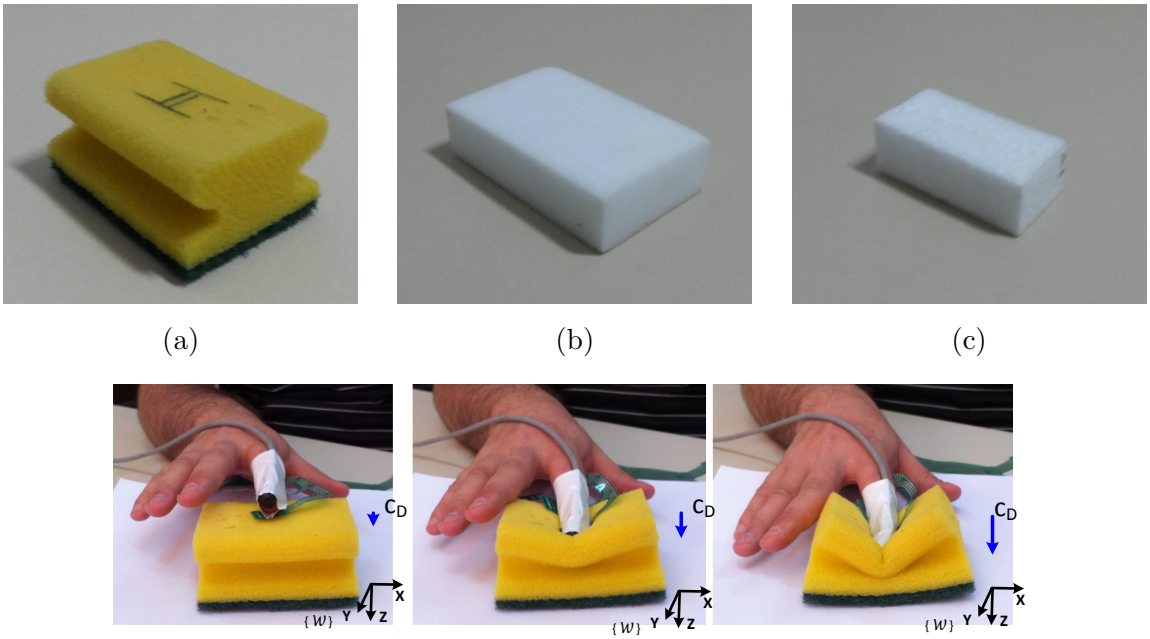


Figure 6.7: Reference materials a) $Material_1$. b) $Material_2$. c)$Material_3$. d) Demonstration of the press-and-release exploration pattern.

Estimation of $P(N_{(v,k)}|M_{(v,k)}, \pi_{haptic})$ and $P(A_{(v,k)}|M_{(v,k)}, \pi_{haptic})$

The contact sensing features c_P, c_A, and c_D extracted from the haptic sensory data $\mathbf{h}_{(v,k)}$ during all ten press-and-release exploration movements are presented in Figure 6.8, 6.9, and 6.10, for each of the reference materials $Material_1$, $Material_2$, and $Material_3$, respectively.

For all the reference materials, during each of the ten press-and-release movements, as the press segment of the movement progresses, the contact intensity c_P, contact area c_A, and contact indentation level c_D increase, as long as the press movements progress. During the release segment, they decrease. This work only uses segments from the press exploration movements. The contact sensing features c_P, c_A, c_D from the ten press segments of the press-and-release cycles for $Material_1$, $Material_2$, and $Material_3$ are represented in Figure 6.11.

However, although the global behaviour is analogous, each of the reference materials has its own characteristic profile. That unique profile is specified by the contact interaction cutaneous (a_1, a_2) and kinesthetic (n_1, n_2) parameters.

For each of the press segments of the press-and-release exploration movement, the parameters (a_1, a_2) and (n_1, n_2) were estimated using the *MATLAB Curve Fitting Toolbox* (The MathWorks Inc, MA, U.S.). The results are shown in Table 6.3, Table 6.4, and Table 6.5, for $Material_1$, $Material_2$, and $Material_3$, respectively. The curve fitting errors, sum

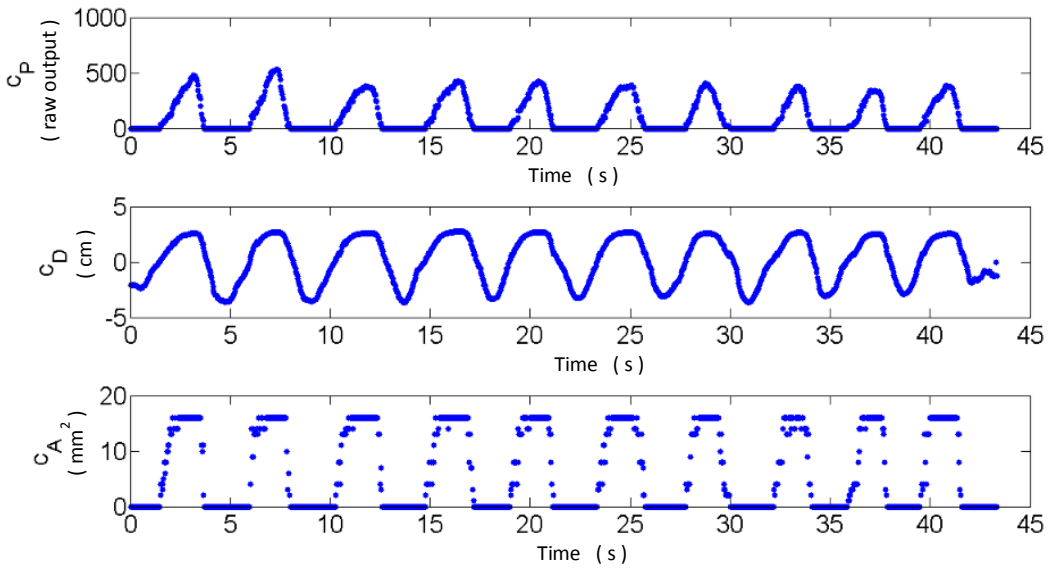


Figure 6.8: Typical temporal profile of the variables c_P, c_A, and c_D during ten press-and-release exploration movements of $Material_1$.

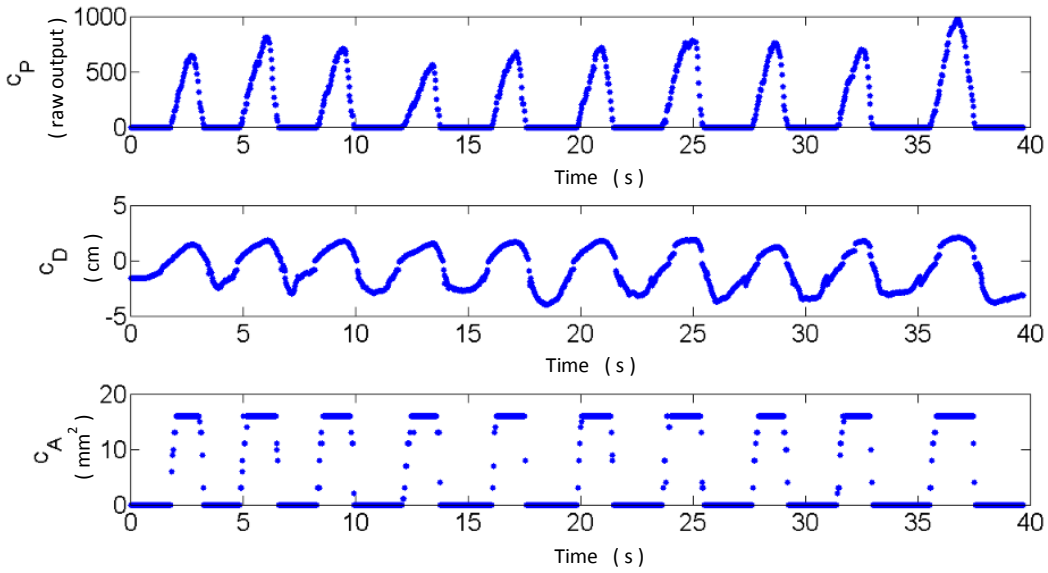


Figure 6.9: Typical temporal profile of the variables c_P, c_A, and c_D during ten press-and-release exploration movements of $Material_2$.

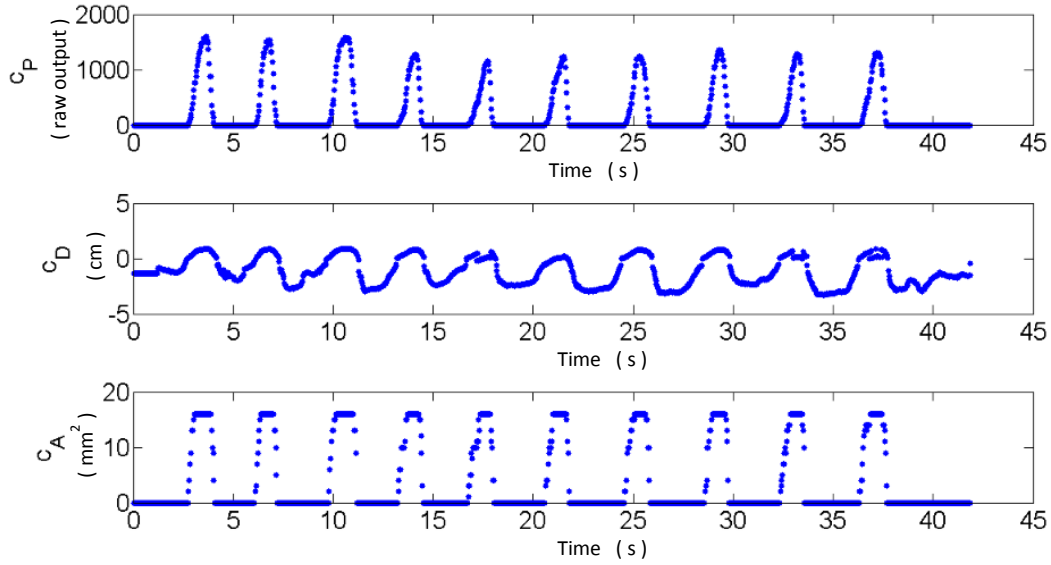


Figure 6.10: Typical temporal profile of the variables c_P, c_A, and c_D during ten press-and-release exploration movements of $Material_3$.

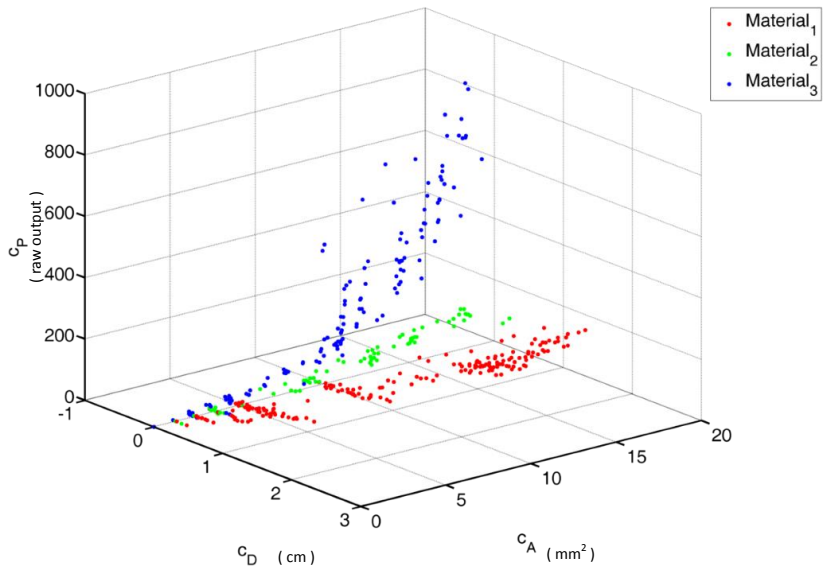


Figure 6.11: Typical contact interaction profile (c_P, c_A, c_D) of $Material_1$, $Material_2$, $Material_3$.

Table 6.3: Contact interaction parameters $N_{(v,k)}$ (kinesthetic) and $A_{(v,k)}$ (cutaneous) for $Material_1$

Exploration	$N_{(v,k)}$			$A_{(v,k)}$		
Run	n_1	n_2	Φ_N	a_1	a_2	Φ_A
1	83.1	15.6	7706.7	2.9	10.1	12997.9
2	66.7	18.7	9977.2	3.4	2.8	19108.7
3	70.1	19.8	4518.4	3.2	2.3	8505.6
4	74.2	11.8	6777.1	3.3	14.5	5485.2
5	66.4	10.5	4493.4	3.0	20.9	5285.1
6	67.4	10.8	4868.4	3.1	14.1	3389.3
7	66.7	18.8	6408.7	3.6	15.1	15691.7
8	102.4	19.1	3109.7	2.6	18.9	1773.5
9	82.8	-2.8	2793.3	2.8	17.4	3036.6
10	72.6	16.9	5570.4	2.3	18.5	3750.4

Table 6.4: Contact interaction parameters $N_{(v,k)}$ (kinesthetic) and $A_{(v,k)}$ (cutaneous) for $Material_2$

Exploration	$N_{(v,k)}$			$A_{(v,k)}$		
Run	n_1	n_2	Φ_N	a_1	a_2	Φ_A
1	493.6	30.2	1715.9	3.0	-0.7	415.4
2	467.9	-3.4	2147.5	2.8	12.1	5172.4
3	380.4	-7.7	1989.5	2.9	5.5	1182.3
4	436.0	-16.9	3571.0	2.9	0.1	1290.9
5	679.7	7.9	625.6	2.4	7.3	572.4
6	604.2	-1.3	450.5	2.8	0.0	1372.9
7	525.9	4.0	4496.3	2.1	14.4	446.1
8	177.2	38.9	9807.9	3.8	-9.7	2282.8
9	385.8	4.6	1723.4	3.1	6.4	389.9
10	175.5	40.0	8935.9	3.0	-0.2	818.8

Table 6.5: Contact interaction parameters $N_{(v,k)}$ (kinesthetic) and $A_{(v,k)}$ (cutaneous) for $Material_3$

Exploration	$N_{(v,k)}$			$A_{(v,k)}$		
Trial	n_1	n_2	Φ_N	a_1	a_2	Φ_A
1	2030.4	-31.2	5091.9	10.6	-64.7	70064.9
2	952.6	-112.6	69435.6	10.1	-48.1	35691.9
3	1858.4	-30.2	31394.8	12.9	-117.7	165029.3
4	2144.5	-48.8	26389.5	11.2	-91.4	132650.9
5	183.6	213.1	671769.1	10.5	-44.2	101726.5
6	2154.0	-17.1	13875.6	9.8	-37.2	36101.0
7	2128.4	-8.0	19706.1	12.2	-91.1	239131.8
8	2204.5	-14.4	8201.4	10.6	-69.5	104792.6
9	1818.9	-45.8	19224.5	10.9	-73.1	103725.3
10	1357.5	68.1	410851.5	10.3	-59.8	164797.7

Table 6.6: LearMean (μ) and co-variation matrix (Σ) parameters learned from the data in Table 6.3, Table 6.4, and Table 6.5

Parameters	$Material_1$	$Material_2$	$Material_3$
μ_N	$\begin{pmatrix}75.2 & 13.9\end{pmatrix}$	$\begin{pmatrix}432.6 & 9.6\end{pmatrix}$	$\begin{pmatrix}1638.3 & -2.7\end{pmatrix}$
Σ_N	$\begin{pmatrix}117.4 & -3.5\\ -3.5 & 42.5\end{pmatrix}$	$\begin{pmatrix}24042.6 & -1626.1\\ -1626.1 & 356.2\end{pmatrix}$	$\begin{pmatrix}396276.5 & -32354.7\\ -32354.7 & 6959.4\end{pmatrix}$
μ_A	$\begin{pmatrix}3.0 & 12.5\end{pmatrix}$	$\begin{pmatrix}2.9 & 3.5\end{pmatrix}$	$\begin{pmatrix}10.9 & -69.7\end{pmatrix}$
Σ_A	$\begin{pmatrix}0.1 & -1.1\\ -1.1 & 38.1\end{pmatrix}$	$\begin{pmatrix}0.2 & -2.4\\ -2.4 & 44.9\end{pmatrix}$	$\begin{pmatrix}0.8 & -20.2\\ -20.2 & 554.2\end{pmatrix}$

of squared 2-norm of the residuals, are represented by the variables Φ_N and Φ_A and reported in Table 6.3, Table 6.4, and Table 6.5. The determination of Φ_N and Φ_A is described in equations 6.16 and 6.17. S represents the number of sensory samples used in each trial as input to the curve fitting method.

$$\Phi_N = \sum_{i=1}^{S} (c_{P,i} - (\hat{n}_1 c_{D,i}^{\frac{3}{2}} + \hat{n}_2))^2 \tag{6.16}$$

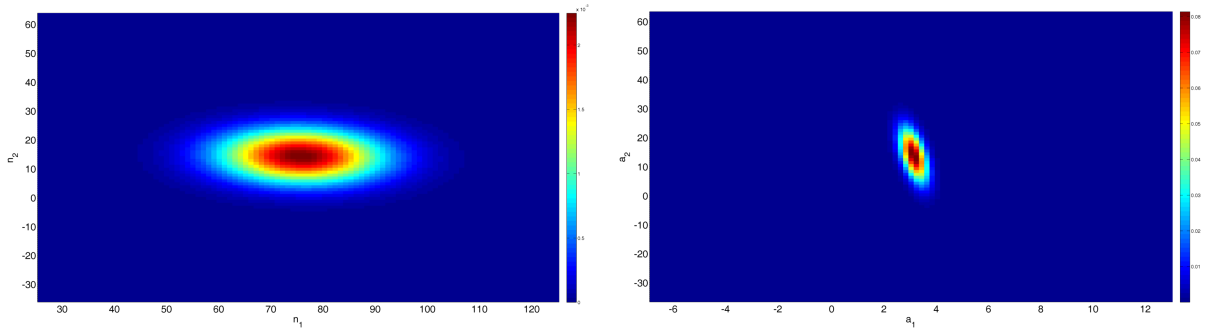
$$\Phi_A = \sum_{i=1}^{S} (c_{P,i} - (\hat{a}_1 c_{A,i}^{\frac{3}{2}} + \hat{a}_2))^2 \tag{6.17}$$

The results in Tables 6.3, 6.4, and 6.5 are used to determine the parameters average μ_A, μ_N and co-variance matrices Σ_A, Σ_N for each of the reference materials, as presented in Table 6.6. These parameters, which are learned from the training data, are used to specify the probability distribution functions $P(N_{(v,k)}|M_{(v,k)}, \pi_{haptic})$ and $P(A_{(v,k)}|M_{(v,k)}, \pi_{haptic})$, represented in Figure 6.12, Figure 6.13, and Figure 6.14.

The probability distribution functions are used in section 6.8.3 to infer the haptic properties of objects unknown to the system.

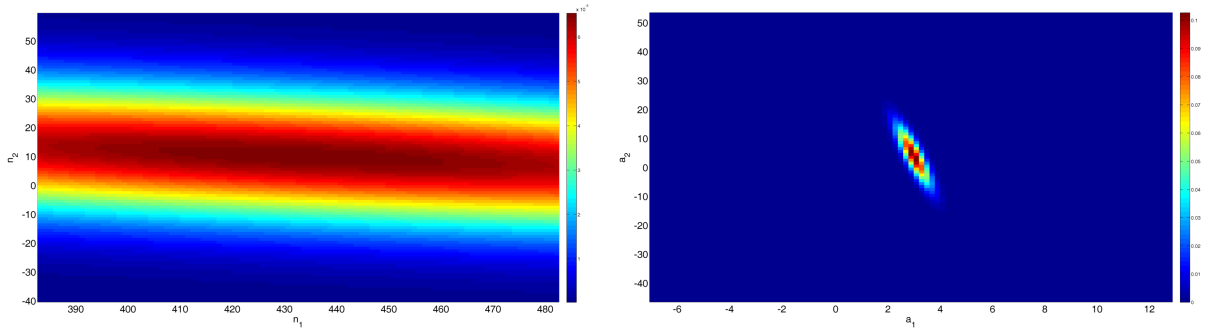
Evaluating the learned parameters of $P(N_{(v,k)}|M_{(v,k)}, \pi_{haptic})$ and $P(A_{(v,k)}|M_{(v,k)}, \pi_{haptic})$

The evaluation of the probability distribution functions $P(N_{(v,k)}|M_{(v,k)}, \pi_{haptic})$ and $P(A_{(v,k)}|M_{(v,k)}, \pi_{haptic})$ learned from the experimental data is performed following the cross-validation scheme leave-one-out. For each reference material, $P(N_{(v,k)}|M_{(v,k)}, \pi_{haptic})$ and $P(A_{(v,k)}|M_{(v,k)}, \pi_{haptic})$ are learned from nine of the ten press segments of the press-and-release exploration movements and tested for the remaining. The categorization of the sample (n_1, n_2), (a_1, a_2) is made according to equation 6.18.



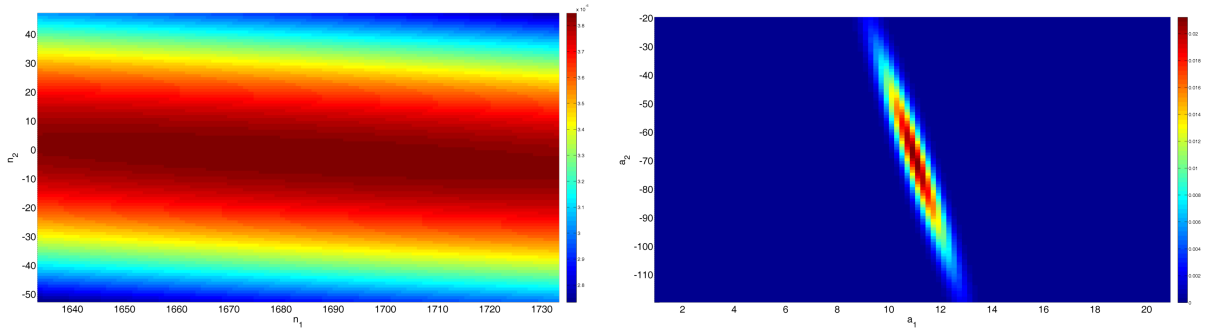
(a) (b)

Figure 6.12: Graphical representation of the probability distribution functions learned from $Material_1$ data. a) $P(N_{(v,k)}|M_{(v,k)}, \pi_{haptic})$. b) $P(A_{(v,k)}|M_{(v,k)}, \pi_{haptic})$.



(a) (b)

Figure 6.13: Graphical representation of the probability distribution functions learned from $Material_2$ data. a) $P(N_{(v,k)}|M_{(v,k)}, \pi_{haptic})$. b) $P(A_{(v,k)}|M_{(v,k)}, \pi_{haptic})$.

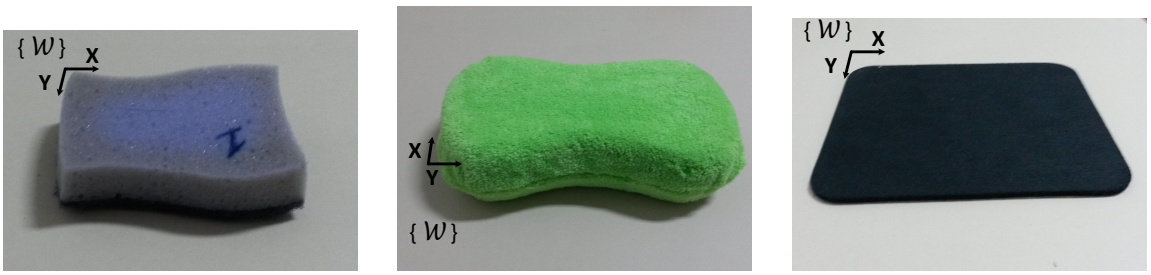


(a) (b)

Figure 6.14: Graphical representation of the probability distribution functions learned from $Material_3$ data. a) $P(N_{(v,k)}|M_{(v,k)}, \pi_{haptic})$. b) $P(A_{(v,k)}|M_{(v,k)}, \pi_{haptic})$.

Table 6.7: Confusion table for the categorization of $Material_i$ (ground truth) as $M.i$ (perceived category) by the Bayesian model π_{haptic}, using a leave-one-out cross-validation scheme

	$M.1$	$M.2$	$M.3$
$Material_1$	0.82	0.18	0.00
$Material_2$	0.00	1.00	0.00
$Material_3$	0.00	0.00	1.00



(a) (b) (c)

Figure 6.15: Unknown objects a) $Object_1$ b) $Object_2$ c)$Object_3$.

$$\hat{m}_{(v,k)} = \arg\max_{m_{(v,k)}} P(n_{(v,k)}|M_{(v,k)}, \pi_{haptic}) P(a_{(v,k)}|M_{(v,k)}, \pi_{haptic}) P(M_{(v,k)}|\pi_{haptic}) \quad (6.18)$$

Table 6.7 presents the confusion table resulting for the cross-validation of the Bayesian model π_{haptic} following a leave-one-out scheme.

The results presented in Table 6.7 show that the proposed Bayesian model π_{haptic} has a high capability to discriminate the proposed reference materials. The model shows only a minimal confusion in the discrimination of $Material_1$ and $Material_2$.

6.8.3 Haptic exploration of unknown objects

Experimental protocol

After the learning stage of the Bayesian model π_{haptic} for the reference materials $Material_1$, $Material_2$, and $Material_3$, the model π_{haptic} can be used to infer the similarity of haptic properties of new objects unknown to the system with the reference materials previously learned by the system. This work proposes the haptic exploration of three new objects, $Object_1$, $Object_2$, and $Object_3$, presented in Figures A, B, and C, respectively.

The workspace region is partitioned on a bi-dimensional inference grid. Each cell is a square of $1cm$ side. Each of the new objects is placed in this workspace region and explored in five pre-defined regions of the surface, using a total of ten press-and-release exploration movements (two press-and-release movements per region). The list of five

Table 6.8: Pre-defined coordinates of the cells on the grid where the press-and-release exploration movements are performed

	$Object_1$	$Object_2$	$Object_3$
$Region_1$	(3,3)	(5,5)	(3,3)
$Region_2$	(4,6)	(12,7)	(4,6)
$Region_3$	(3,8)	(5,8)	(3,8)
$Region_4$	(5,3)	(20,5)	(5,3)
$Region_5$	(5,8)	(20,8)	(5,8)

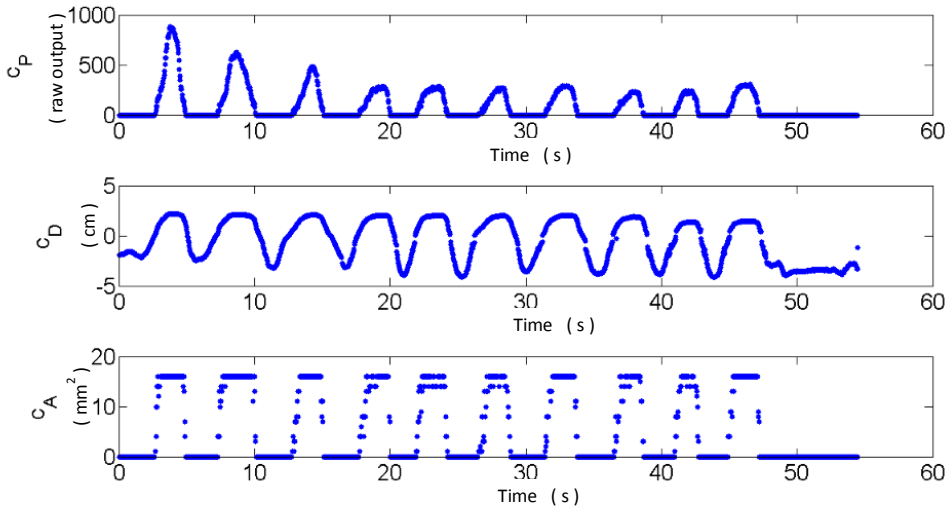


Figure 6.16: Typical temporal profile of the variables c_P, c_A and c_D during press-and-release exploration movements of $Object_1$.

pre-defined regions is proposed in Table 6.8.

The hand of the subject performing the haptic exploration of the new objects is instrumented with the same tactile sensing array *Tekscan Grip System* and motion tracking sensor *Polhemus Liberty*, as presented in Figure 6.6b.

Representation and update of the haptic stimulus map of unknown objects

The profile of the contact sensing features c_P, c_A, c_D (gathering the data from the ten exploration movements) of each of the new objects is presented in Figures 6.19, 6.20, and 6.21.

The temporal evolution of the haptic representation of each of the unknown objects $Object_1$, $Object_2$, and $Object_3$ is illustrated in Figures 6.23, 6.24, and 6.25 and detailed in Tables 6.9, 6.9, and 6.9. Initially, $(k = 0)$, and all the cells v of the workspace are described by a uniform probability distribution function $P(M_{(v,k)}|A_{(v,k)}, N_{(v,k)}, \pi_{haptic}) \equiv Uniform$. As long as the exploration of the unknown objects progresses $(k = 1, \ldots, k = 10)$ in the pre-defined regions listed in Table 6.8, the representation of those regions improves

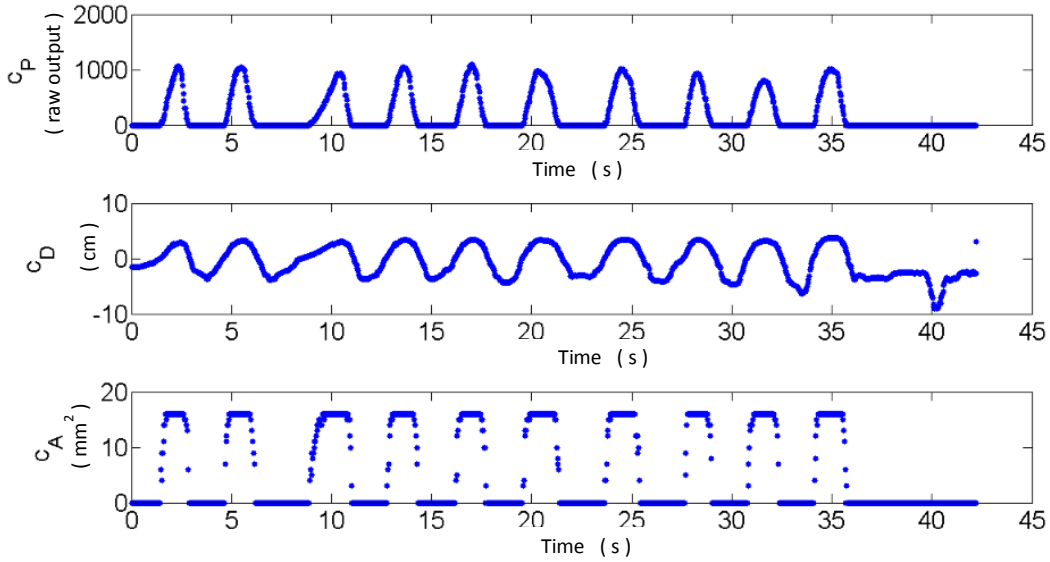


Figure 6.17: Typical temporal profile of the variables c_P, c_A and c_D during press-and-release exploration movements of $Object_2$.

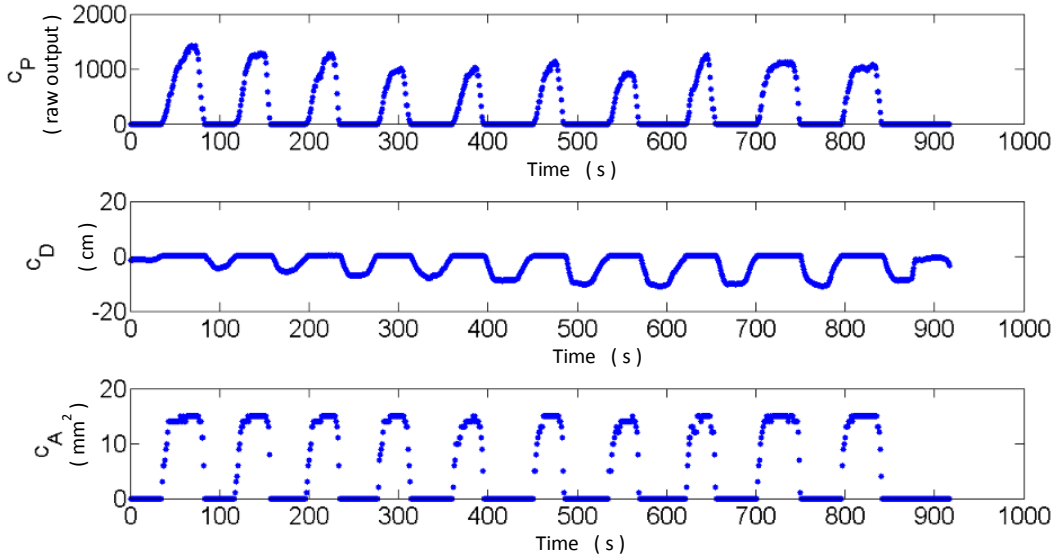


Figure 6.18: Typical temporal profile of the variables c_P, c_A and c_D during press-and-release exploration movements of $Object_3$.

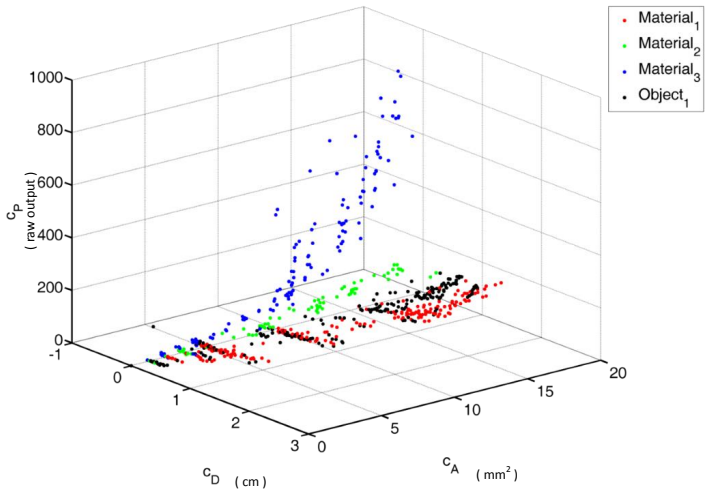


Figure 6.19: Typical contact interaction profile (c_P, c_A, c_D) of $Object_1$, compared with $Material_1$, $Material_2$, $Material_3$.

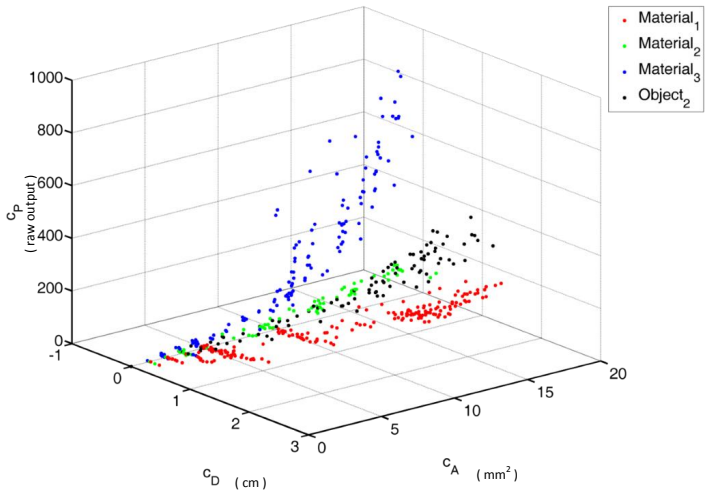


Figure 6.20: Typical contact interaction profile (c_P, c_A, c_D) of $Object_2$, compared with $Material_1$, $Material_2$, $Material_3$.

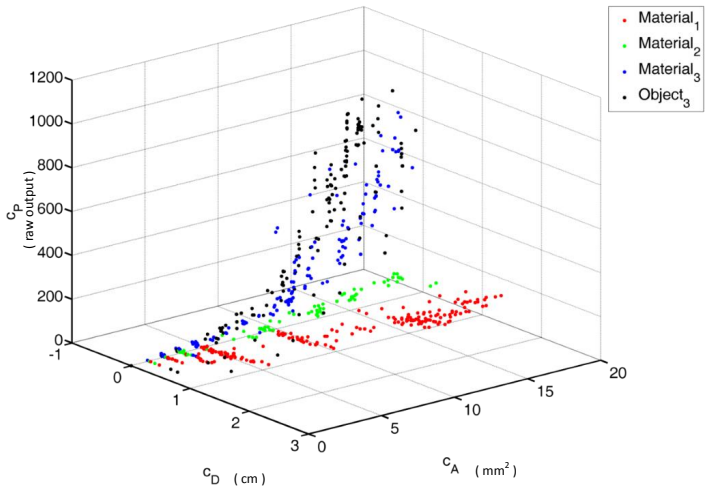


Figure 6.21: Typical contact interaction profile (c_P, c_A, c_D) of $Object_3$, compared with $Material_1$, $Material_2$, $Material_3$.

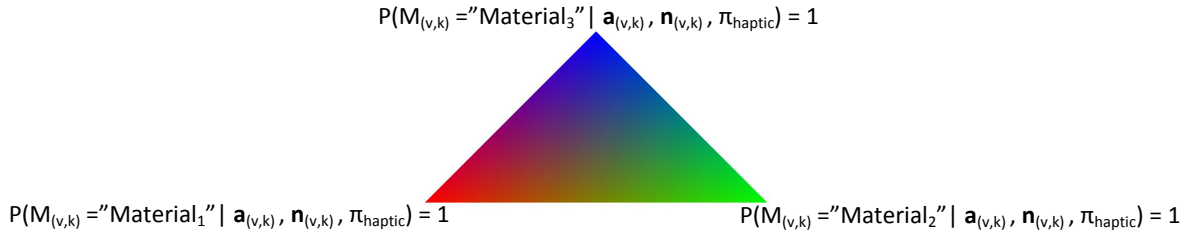


Figure 6.22: Colormap used to represent $P(M_{(v,k)}|N_{(v,k)}, A_{(v,k)}, \pi_{haptic})$.

remarkably.

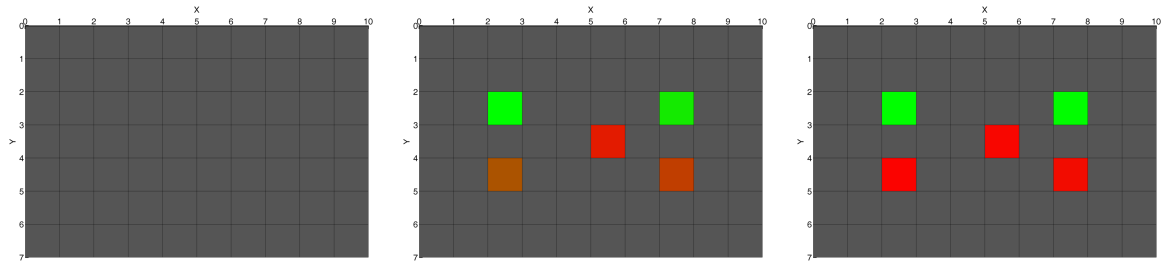
The regions of $Object_1$ are described as being similar (higher probability) to $Material_1$. Some of these regions improve their representation by integrating additional haptic data acquired in the exploration cycles. Two of the regions of $Object_1$ show a contact interaction behaviour recognized as $Material_2$.

Alternatively, all the explored regions of $Object_2$ and $Object_3$ are described as being made of materials with a contact interaction behaviour similar to $Material_2$ and $Material_3$, respectively. Since the initial exploration cycles, that tendency is evident.

Extrapolation of the representation of unexplored grid ells

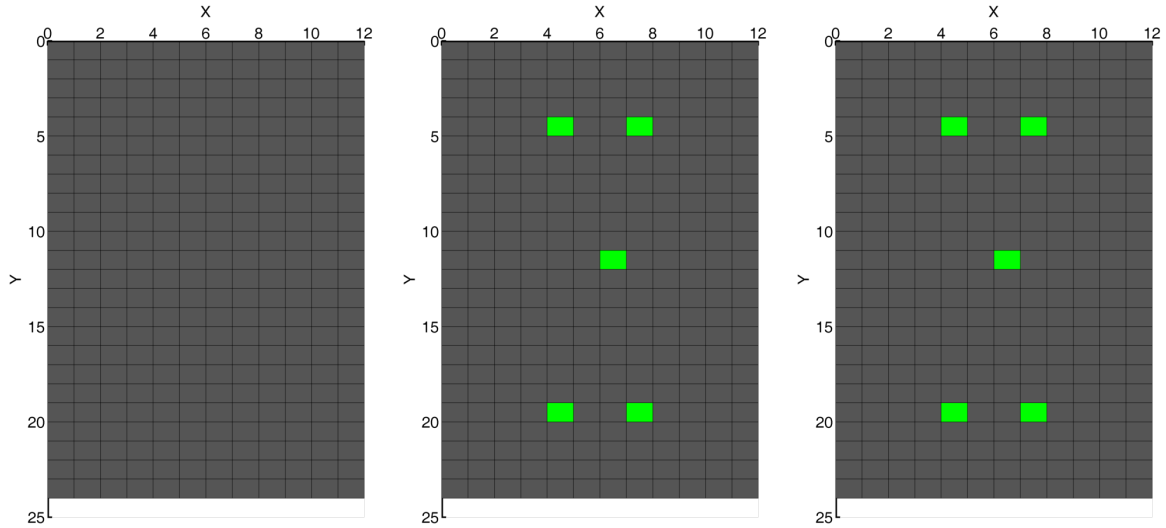
The final representation of the workspace of $Object_1$, $Object_2$, and $Object_3$ are processed in an effort to improve the representation of the unexplored cells. The method presented in section 6.7 intends to reduce the uncertainty of the representation of the workspace by reducing the entropy of the representation. The results are shown in Figure 6.26.

By comparing Figures 6.23c, 6.24c, 6.25c, and 6.26, the extension of uncertain regions is reduced. The extension of regions characterized as being similar to $Material_1$,



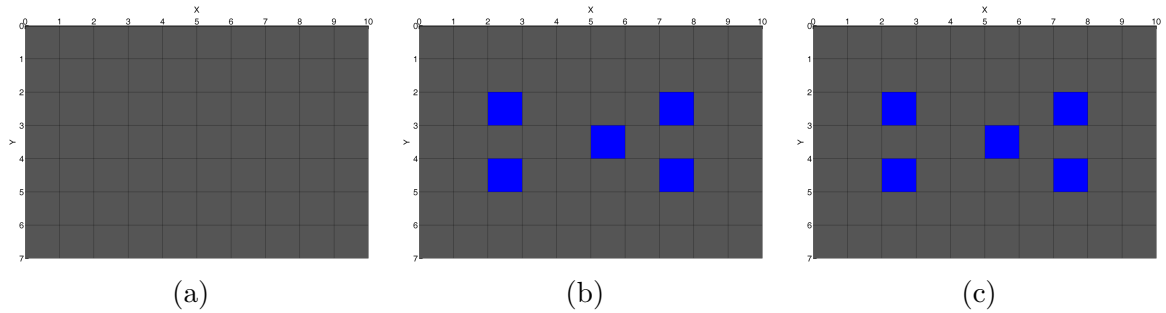
(a) (b) (c)

Figure 6.23: Probabilist representation of the result of the haptic exploration of $Object_1$. a) Initial representation ($k = 0$). b) Representation after five press-and-release movements ($k = 5$). c) Representation after five press-and-release movements ($k = 10$). Colormap described in Figure 6.22.



(a) (b) (c)

Figure 6.24: Probabilist representation of the result of the haptic exploration of $Object_1$. a) Initial representation ($k = 0$). b) Representation after five press-and-release movements ($k = 5$). c) Representation after five press-and-release movements ($k = 10$). Colormap described in Figure 6.22.



(a) (b) (c)

Figure 6.25: Probabilist representation of the result of the haptic exploration of $Object_1$. a) Initial representation ($k = 0$). b) Representation after five press-and-release movements ($k = 5$). c) Representation after five press-and-release movements ($k = 10$). Colormap described in Figure 6.22.

Table 6.9: Evolution of the probability distribution function $P(M_{(v,k)}|n_{(v,k)}, a_{(v,k)}, \pi_{haptic})$ during the exploration of $Object_1$

	Initial condition ($k = 0$)			1^{st} exploration cycle ($k = 1, \ldots, 5$)			2^{nd} exploration cycle ($k = 6, \ldots, 10$)		
	$Mat._1$	$Mat._2$	$Mat._3$	$Mat._1$	$Mat._2$	$Mat._3$	$Mat._1$	$Mat._2$	$Mat._3$
$Region_1$	0.33	0.33	0.33	0.01	0.99	0.00	0.00	1.00	0.00
$Region_2$	0.33	0.33	0.33	0.89	0.11	0.00	0.97	0.03	0.00
$Region_3$	0.33	0.33	0.33	0.08	0.92	0.00	0.01	0.99	0.00
$Region_4$	0.33	0.33	0.33	0.67	0.33	0.00	0.98	0.02	0.00
$Region_5$	0.33	0.33	0.33	0.75	0.25	0.00	0.95	0.05	0.00

Note: For simplicity and compactness of the notation, $Mat._1$, $Mat._2$, and $Mat._3$ were used to designate $Material_1$, $Material_2$, and $Material_3$, respectively.

Table 6.10: Evolution of the probability distribution function $P(M_{(v,k)}|n_{(v,k)}, a_{(v,k)}, \pi_{haptic})$ during the exploration of $Object_2$

	Initial condition ($k = 0$)			1^{st} exploration cycle ($k = 1, \ldots, 5$)			2^{nd} exploration cycle ($k = 6, \ldots, 10$)		
	$Mat._1$	$Mat._2$	$Mat._3$	$Mat._1$	$Mat._2$	$Mat._3$	$Mat._1$	$Mat._2$	$Mat._3$
$Region_1$	0.33	0.33	0.33	0.00	1.00	0.00	0.00	1.00	0.00
$Region_2$	0.33	0.33	0.33	0.00	1.00	0.00	0.00	1.00	0.00
$Region_3$	0.33	0.33	0.33	0.00	1.00	0.00	0.00	1.00	0.00
$Region_4$	0.33	0.33	0.33	0.00	1.00	0.00	0.00	1.00	0.00
$Region_5$	0.33	0.33	0.33	0.01	0.99	0.00	0.00	1.00	0.00

Note: For simplicity and compactness of the notation, $Mat._1$, $Mat._2$, and $Mat._3$ were used to designate $Material_1$, $Material_2$, and $Material_3$, respectively.

Table 6.11: Evolution of the probability distribution function $P(M_{(v,k)}|n_{(v,k)}, a_{(v,k)}, \pi_{haptic})$ during the exploration of $Object_3$

	Initial condition ($k = 0$)			1^{st} exploration cycle ($k = 1, \ldots, 5$)			2^{nd} exploration cycle ($k = 6, \ldots, 10$)		
	$Mat._1$	$Mat._2$	$Mat._3$	$Mat._1$	$Mat._2$	$Mat._3$	$Mat._1$	$Mat._2$	$Mat._3$
$Region_1$	0.33	0.33	0.33	0.00	0.00	1.00	0.00	0.00	1.00
$Region_2$	0.33	0.33	0.33	0.00	0.00	1.00	0.00	0.00	1.00
$Region_3$	0.33	0.33	0.33	0.00	0.00	1.00	0.00	0.00	1.00
$Region_4$	0.33	0.33	0.33	0.00	0.00	1.00	0.00	0.00	1.00
$Region_5$	0.33	0.33	0.33	0.00	0.00	1.00	0.00	0.00	1.00

Note: For simplicity and compactness of the notation, $Mat._1$, $Mat._2$, and $Mat._3$ were used to designate $Material_1$, $Material_2$, and $Material_3$, respectively.

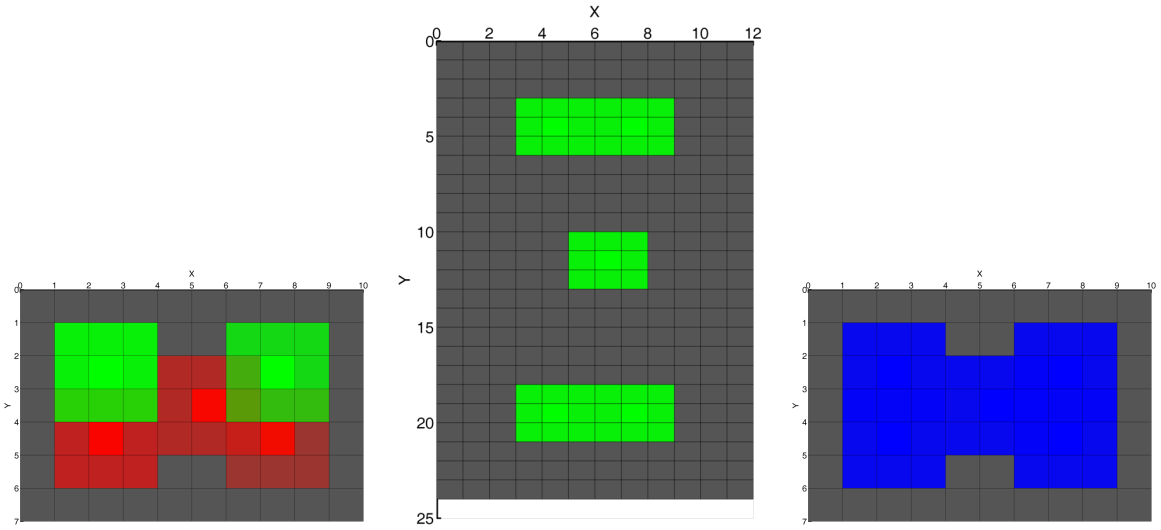


Figure 6.26: Probabilist representation of the result of the haptic exploration and post-processing of a) $Object_1$. b) $Object_2$. c) $Object_3$. Colormap described in Figure 6.22.

$Material_2$, or $Material_3$ increases. This methodology can be used, not only to improve the representation of the workspace, but also to speed up the exploration of the workspace toward an optimal representation.

6.9 Conclusions

The work presented in this chapter contributed to the development of autonomous dexterous robotic hand platforms by proposing a probabilistic inference grid to represent and discriminate the perceived hardness-softness characteristics extracted during the exploration of soft objects. The proposed approach follows some principles inspired by human strategies to perceive and estimate the haptic characteristics of objects in uncertain environments. The perceived hardness-softness characteristics of unknown objects are described by a probabilistic combination of previously known characteristics of a set of reference materials (haptic memory of the system).

This approach is designed to progressively receive haptic inputs (cutaneous and kinesthetic data). As long as the object exploration progresses, the total entropy of the representation is reduced, showing that the representation becomes less uncertain. The processing stages related to the integration of local context information have also contributed to the improvement of the representation.

Chapter 7

Active haptic exploration of surfaces using robotic hands

7.1 Introduction

Dexterous robotic hands are a key element for interacting with the surrounding environment because of their mechanical (high number of degrees of freedom) and sensory (tactile, force, torque, and heat) capabilities, which are becoming tendentiously analogous to human hands. They allow robotic platforms to perform precise manipulation of objects [Johansson and Flanagan, 2009] (reach, grasp, transport, and in-hand reorientation), as well as to perform the haptic exploration of surfaces using different patterns of movements (lateral motion, press-and-release, and static contact) promoting the extraction and integration of different haptic properties of the surfaces [Lederman, 1994] (contours, texture, compliance, and temperature).

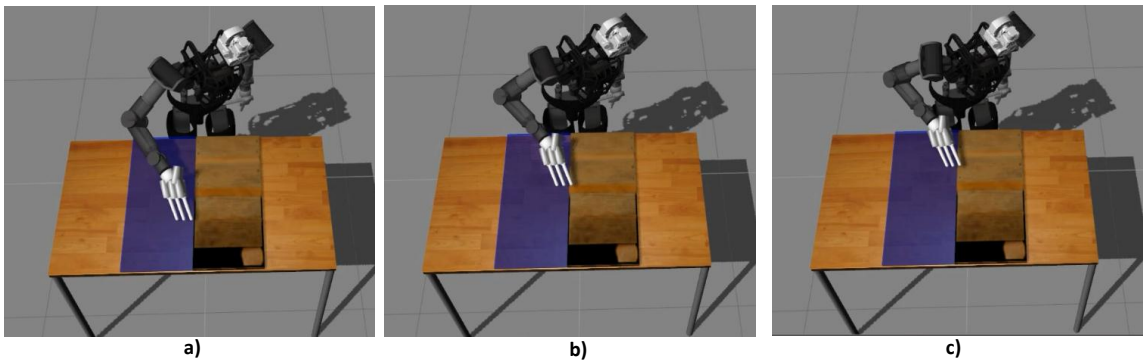


Figure 7.1: Results presented in [Martins et al., 2014], illustrating the typical exploration behaviour performed by a robotic system robot during a haptic discontinuity following task: straight line geometry. Three time instants a), b), and c) are represented.

The contributions presented in this work relate to the robotic haptic exploration of surfaces.The haptic exploration of surfaces plays a fundamental role in reduced visibility

scenarios (i.e. underwater robotic manipulation, smoky and foggy disaster environments, and partial or complete occlusion of elements in the scenario). Although this work only addresses the implementation of haptic exploration strategies, the proposed Bayesian framework allows the integration of additional sensory sources such as vision (depth, color) and laser to infer the robotic exploration path. The approach proposed in this work can be used to complement methods already available to explore surfaces using non-haptic sensory inputs exclusively [Meng et al., 2016] [Ban and Lee, 2006] [Gomes et al., 2013].

Figure 7.1 shows an example presented in [Martins et al., 2014]. A robotic platform follows a haptic discontinuity between two surfaces made of different materials (different haptic properties) placed on top of a table.

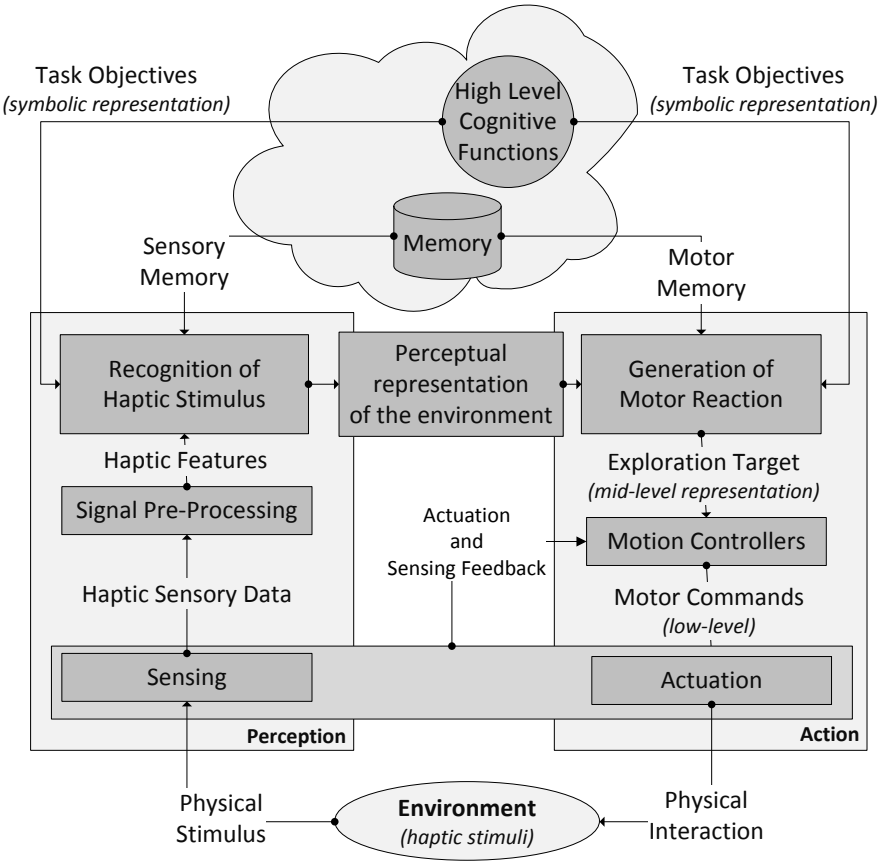


Figure 7.2: Conceptual representation of the action-perception loop [Ernst and Bulthoff, 2004] involved in the haptic exploration of surfaces [Wacker, 2011]. In this work, the objectives of the task and corresponding solution are represented in two levels: symbolic and mid-level.

Section 7.2 of this manuscript describes several works performing the haptic exploration of surfaces, using deterministic methods with a set of control rules defined *a-priori*.

This strategy demonstrates difficulties to overcome the high diversity of noisy sensory signals and to show a generalization capability when dealing with the uncertainty of the structure of the environments (see Table 7.1).

To deal with these characteristics of the application scenarios, this chapter implements an approach that follows a typical architecture of a cognitive robotic system, formalized using a Bayesian framework. The Bayesian models implementing the action-perception loop and the attention mechanisms integrated to increase the efficiency of those processes are presented in section 7.3 and detailed in sections 7.4, 7.5, and 7.6.

The experimental setup implemented in this chapter is described in section 7.7. The generalization capability of the proposed approach is tested experimentally in one scenario, which is explored autonomously by the robotic system. Section 7.8 presents the main conclusions of this chapter and key guidelines for future developments of this approach.

7.2 Related works

The robotic exploration of surfaces using haptic inputs has been a research topic pursued for a long time, with seminal works by [Harmon, 1982], [Klatzky et al., 1985], [Dario and De Rossi, 1985], and [Nicholls and Lee, 1989]. The complexity and variety of application fields of the proposed approaches followed the developments verified during the last years on force, torque, and tactile sensing manufacturing technologies [Lucarotti et al., 2013], [Yousef et al., 2011]: miniaturization, higher resolution and accuracy, improved mechanical resistance, variety of materials (e.g. compatible with MRI environments). The integration of this type of sensing devices on robotic platforms, especially in the new generation of dexterous robotic hands, has become standard and tendentiously mandatory, as presented in [Dahiya et al., 2013].

The literature describes a class of approaches performing the haptic exploration of surfaces with the objective of achieving a categorization of the surface or object. The exploration is performed locally, assuming that the explored region is homogeneous or uniform in terms of the haptic features under analysis.

Discrimination between the different classes of surfaces is performed by extracting different types of haptic features such as surface curvature [Okamura et al., 2001], texture [Oddo et al., 2011], [Xu et al., 2013], [Fishel and Loeb, 2012], [Chathuranga et al., 2013], compliance [Martins et al., 2012a], [Xu et al., 2013], stickiness [Liu et al., 2012], and thermal conductivity [Xu et al., 2013], [Castelli, 2002] from the haptic sensory signals. The formalization of the descriptors of each of those haptic features is dependent on the type of robotic platform and type of sensing apparatus involved in the exploration. The modelling of the contact interaction and the characteristics of the sensory signals produced

Table 7.1: Comparison between the contributions of this work and related works

Study	**Apparatus** [a]	**Local Haptic Perception**		**Global Exploration of the workspace**			
		Approach [b]	Features [c]	Approach [d]	Task [e]	Strategy [f]	Workspace [g]
This Work	HS	P	T, CO	P	M:E, F:E	AE	GD:2D
[Okamura et al., 2001]	HS	D	C	-	-	-	-
[Oddo et al., 2011]	HS	D	T	-	-	-	-
[Xu et al., 2013]	HS	P	CO, T, TC	-	-	-	-
[Fishel and Loeb, 2012]	HS	P	T	-	-	-	-
[Chathuranga et al., 2013]	HS	D	T	-	-	-	-
[Martins et al., 2012a]	HS	P	C	-	-	-	-
[Liu et al., 2012]	HS	P	S	-	-	-	-
[Castelli, 2002]	HS	D	C, TC	-	-	-	-
[Liu et al., 2010b]	HS	D	CO	D	M:C	PD	CS:2D
[Bologna et al., 2013]	HS	D	RO	P	F:T	PD	-
[Li et al., 2013]	HS	P	C	D	F:C	AE	-
[Bohg et al., 2010]	HS, VS	P	TO	P	M:E	AE	GD:2D
[Martinez et al., 2013]	HS	P	RO	D	M:E, F:E	AE	GD:2D
[Li et al., 2013]	HS	D	CI, CR	D	F	AE	CS:3D

[a] HS- haptic sensing; VS- visual sensing.
[b],[d] P- probabilistic; D- deterministic.
[c] T- texture; CO- compliance; C- curvature; TC- thermal conductivity; S- stickiness; RO- raw sensory output; CI- contact intensity; CR- contact orientation.
[e] M:E- mapping edge; M:C - mapping compliance; F- following; F:E- following edge; F:T- following texture; F:C: following curvature;
[f] AC- active exploration; PD- pre-defined exploration path.
[g] GD:2D - bi-dimensional grid; CS:2D- bi-dimensional Cartesian space points; CS:3D- tri-dimensional Cartesian space points;

during the interaction are specific to the different types of robotic setups. However, independently of the factors regarding the design of the sensory apparatus, the specific motor exploratory procedures, which are involved in the perception of each type of haptic feature, are the same in different works.

This work contributes to this class of approaches by presenting a Bayesian model to discriminate different categories of materials by integrating compliance and texture features. This work presents a formulation to perform the perceptual discrimination of different materials that abstracts the contact interaction models between the exploratory element and the surface.

A second group of works integrates sensing, perception, and local exploration mechanisms similar to the previous works. However, they expand the exploration strategy to large and heterogenous surfaces in terms of the haptic domain under analysis. The perceptual map of the surface can be constructed following different strategies. Some works proposed a fixed global exploration strategy and defined *a-priori*. In [Liu et al.,

2010b], the haptic exploration is performed using pre-defined exploration paths to build a stiffness map of biological tissues. As long as the perception of the haptic stimulus occurs, it does not influence the exploration movement. In [Bologna et al., 2013], Braille reading is performed real-time by a robotic platform. The Braille symbols are explored and recognized. The exploration velocity is adjusted depending on the recognition uncertainty. Nevertheless, the exploration path is also pre-defined.

Although restricted exploration strategies may be successful when substantial information about the structure of the environment is available, in most of the scenarios identified during the motivation of this work, the structure of the environment is initially unknown (partially or completely) or can change during the exploration task. Thus, the exploration strategy should have an (re)active behaviour to progressively integrate and analyse the local perceptual representation of the environment (perception for action) and estimate which should be the next global region to be explored and perceived (action for perception) [Dahiya et al., 2010].

[Barron-Gonzalez et al.,] presents a surface-following controller for active haptic exploration. The controller reactively adapts the orientation of the robotic hand according to the perceived variations of curvature. The active exploration of a scene represented by a occupancy grid was proposed in [Bohg et al., 2010]. An initial estimation of the scene structure is made using stereo vision data which is projected on a 2D occupancy grid. The exploration strategy is dependent on that initial representation, and haptic inputs (lateral contact/non-contact) are used to confirm and update the occupancy of the grid representing the scene.

In other works, the active exploration task is started without any knowledge about the structure of the scene. The work of [Martinez et al., 2013] proposes a method to perform the active contour following of objects by executing tap movements using a robotic fingertip equipped with a tactile array. The reaction of the system is formulated considering the interaction profile between the haptic stimuli with the tactile sensing array, integrating specific deterministic rules defined a-priori by a human operator. By the end of the exploration, the contour of the object, which is described by the full exploration path, is used to recognize the object. In [Herzog et al., 2013], the categorization of the structure of the object is applied to estimate the affordance of the object and synthesize the corresponding robotic grasps [Bohg et al., 2013]. In [Martinez et al., 2013], [Herzog et al., 2013] the structure of the exploration path did not influence the active exploration of the workspace. [Li et al., 2013] presents a generic formulation of a control framework for different types of tasks requiring tactile servoing (e.g. tracking a touched object and tactile object active exploration). The different behaviours are obtained by adjusting several matrix parameters and selecting the corresponding haptic primitives extracted

from a tactile array.

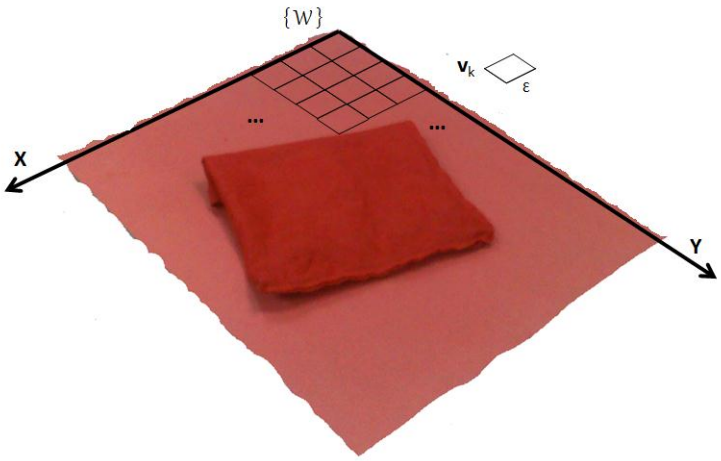


Figure 7.3: Illustration of a 2D isometric grid partitioning a real-world workspace area. Each cell $\mathbf{v}$ has a dimension ε and is described by position (x, y) expressed in $\{\mathcal{W}\}$.

This work contributes to this class of approaches by proposing a formulation of Bayesian models implementing touch attention mechanisms involved in the active haptic exploration of unknown surfaces by generic robotic hands and sensory apparatus. The formulation is independent of the contact interaction model between the surface and the robotic exploratory element. The definition of the architecture of the Bayesian models follows the principles for how humans manage uncertainty to make motor decisions from perceptions [Ernst and Bulthoff, 2004]. This formulation assumes that the workspace is unknown *a-priori* to the system (blind exploration). The exploration path is adapted actively by the touch attention mechanisms, as long as the exploration occurs. The system integrates information from the saliency of the perceived workspace (bottom-up), objectives of the task, structure of current exploration path, uncertainty of the perceptual map representing the workspace, and inhibition-of-return mechanisms (top-down). The implementation provides the ability to deal with ambiguous sensory signals corrupted with noise and to perform the active haptic exploration of surfaces with different geometries. It is expected that a generalized behaviour of the systems emerges from this formulation.

7.3 Approach overview

The approach proposed in this chapter assumes that the exploration task will be performed on top of a table (workspace defined by a planar surface) and using a generic robotic system with manipulation capability. The internal structure and configuration of the workspace is unknown *a priori* to the robotic system. Thus, the solution to the

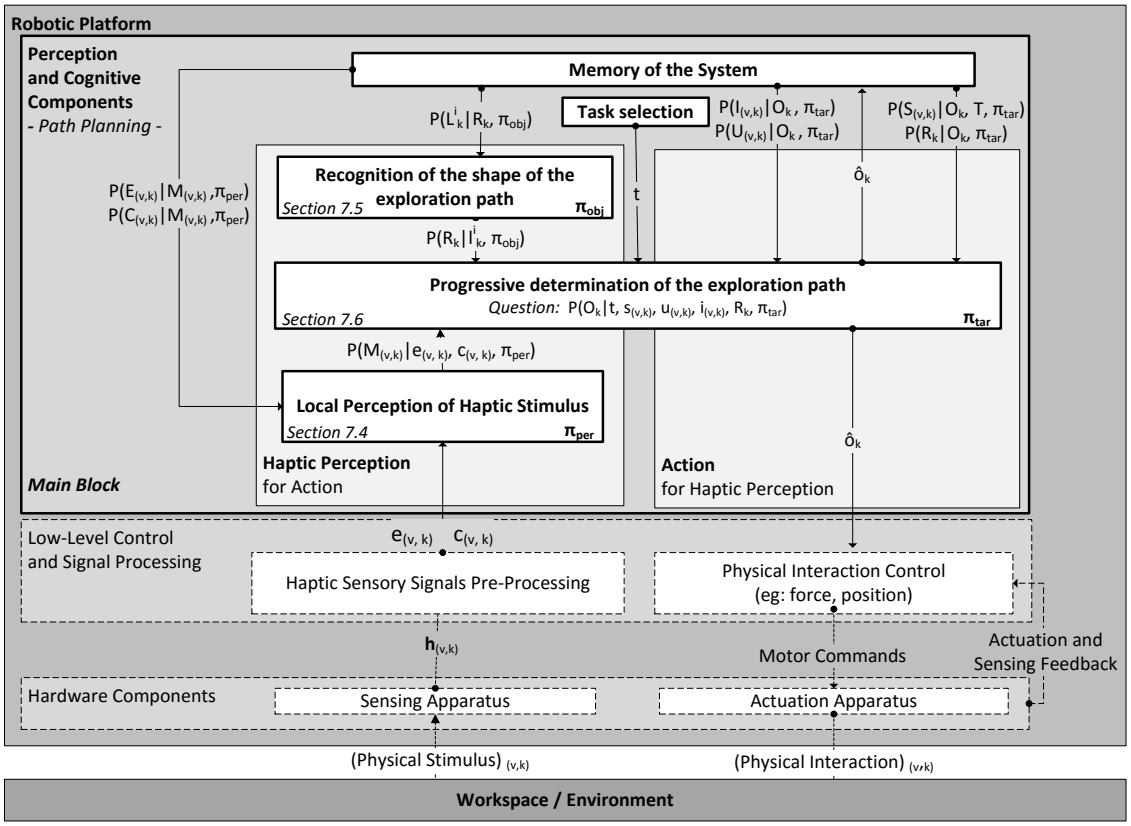


Figure 7.4: Detailed diagram of the architecture of the proposed system. The main contributions of this work are identified in the diagram as *main block* (local perception of haptic stimulus, recognition of the shapes of discontinuities, progressive determination of the exploration path). The variables of the system are summarised in Table 7.2.

haptic exploration task is described in the bidimensional Cartesian space by progressively determining the sequence of regions of the workspace that should be visited by the robotic platform during the task execution.

The 2D Cartesian space is partitioned using a planar isometric 2D grid (square cells), as represented in Figure 7.3 . Each cell $\mathbf{v}_k$ has a side dimension ε and is described by a 2D Cartesian location (x, y) expressed in the inertial world referential $\{\mathcal{W}\}$. This type of representation framework, which can be used as an inference grid, has been used extensively in robotics [Ferreira and Dias, 2014c] to integrate multi-modal sensing data (see chapter 2.2). The inference grid establishes the loop between the perception and action models.

The methods presented in this chapter follow the principles and architecture of the human somatosensory processing pipeline and cognition (reviewed in chapter 3.2.2). A simplified overview of the approach is presented in Figure 7.2.

During the local interaction of the robotic exploratory elements with the workspace

Table 7.2: Summary of the relevant variables used in this chapter

Variable	Description	Domain
v	Cell of the workspace grid.	$\mathbb{R}^2$
k	Time / exploration iteration.	$\mathbb{N}^0$
R_k	Category of structure of the exploration path.	$\{"Shape_1","Shape_2"\}$
$template_i$	Set of points defining the template of each category of structure.	$\mathbb{R}^2$
L_k^i	Matching error between the exploration path and $template_i$.	[0,1]
$M_{(v,k)}$	Material category of v	$\{Material_1, \ldots, Material_{10}\}$
$E_{(v,k)}$	Texture characterization of v.	$\mathbb{R}$
$C_{(v,k)}$	Compliance characterization of v.	$\mathbb{R}$
$\mathbf{h}_{(v,k)}$	Raw haptic sensing data acquired on v.	$\mathbb{R}^n$ *
O_k	Next workspace region to be explored.	v
$I_{(v,k)}$	Inhibition level for cell v.	$[0,1]$
$U_{(v,k)}$	Uncertainty level for cell v.	$[0,1]$
$S_{(v,k)}$	Saliency of the perceived haptic stimulus in region v.	$[0,1]$
T	Objective of the haptic exploration task.	$\{T_1, T_2, T_3\}$

* In this work, the description of the sensory apparatus of the robotic platform follows a generic formulation. In this context, n represents the dimensionality of the raw sensing data provided by the haptic sensory apparatus of the robot.

at the region $\mathbf{v}_k$, the haptic sensory inputs $\mathbf{h}_{(v,k)}$ are acquired. Haptic features such as texture, compliance, and temperature are extracted. The features are used to discriminate the different classes of materials found in the workspace.The local perception of haptic stimulus is modelled by the Bayesian model π_{per} presented in section 7.4.

Following the somatosensory processing pipeline, the robotic system uses the update perceptual representation of the workspace to select the next region of the workspace that should be explored. The mechanisms involved on this selection are described by the Bayesian model π_{tar} and presented in section 7.6. The model implements touch attention mechanisms (feature based) by integrating the haptic saliency and uncertainty of the perceptual representation of the workspace, inhibition-of-return mechanisms, and shape cues about the structure of the discontinuity extracted from the current exploration path (determined by the Bayesian model π_{obj}, section 7.5). The integration of the touch attention mechanisms on the action-perception loop architecture contributes to increasing the efficiency of the action-perception loop mechanisms by promoting the selective exploration of regions of the workspace likely to be useful to the task.

The detailed structure of the approach proposed in this chapter is presented in Figure 7.4, following an action-perception loop architecture.

7.3.1 Path planning of the global haptic exploration strategy

The framework conceptually represented in Figure 7.2 and detailed in Figure 7.4 implements a haptic exploration path planning method, which infers a series of global via-points in the workspace that should be probed by the robotic system.

This work does not address the low-level control loop involved in physical interaction of the fingers with the surface and the ability to move the fingers along the surface by

keeping contact. In other words, the low-level modelling and control of local contact interaction (e.g. force, impedance, and position control) and processing of haptic sensory data are not discussed by this work. These processes are implemented in Figure 7.4 by the module *Low-Level Control and Signal Processing* and inner loop labelled *Actuation and Sensing Feedback*.

Our solution assumes that algorithms (dependent on specific robotic devices and sensing apparatus) implemented by other works (e.g. [Xu et al., 2013]) extract different haptic features and control the local movements during the haptic exploration of a region $\mathbf{v}_k$. The integration between these lower-level control models (dashed boxes) and the global exploration path planning method (bold boxes) proposed by this work is detailed in Figure 7.4.

7.4 Local perception of haptic stimulus

7.4.1 Random variables of the model

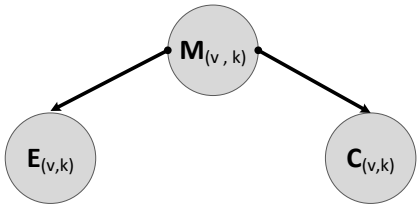
During the local exploration of the region v of the workspace at time iteration step k, the robotic system senses the haptic measurements $\mathbf{h}_{(v,k)}$. The haptic sensory inputs are integrated by the Bayesian model π_{per}, determining the perceived category of material describing the cell v of the workspace.

The type of material describing the workspace region v is represented by the discrete random variable $M_{(v,k)}$, defined in equation (7.1).

$$\begin{aligned} M_{(v,k)} &- \text{"Material category of v"} \\ M_{(v,k)} &\in \{Material_1, \ldots, Material_{10}\} \end{aligned} \tag{7.1}$$

The categories of materials are characterized by different properties of texture and compliance. The texture and compliance characteristics of the region v of the workspace are described by the continuous random variables $E_{(v,k)}$ and $C_{(v,k)}$, respectively, according to equations (7.2) and (7.3).

$$\begin{aligned} E_{(v,k)} &- \text{"Texture characterization of v"} \\ E_{(v,k)} &= f(\mathbf{h}_{(v,k)}), \quad E_{(v,k)} \in \mathbb{R} \end{aligned} \tag{7.2}$$



(a)

Bayesian Program

Description

Specification

Relevant variables:

$E_{(v,k)}, C_{(v,k)}, M_{(v,k)}$

Decomposition:

$P(E_{(v,k)}, C_{(v,k)}, M_{(v,k)}|\pi_{per}) =$

$P(E_{(v,k)}|M_{(v,k)}, \pi_{per}).P(C_{(v,k)}|M_{(v,k)}, \pi_{per}).P(M_{(v,k)}|\pi_{per})$

Parametric forms:

$P(E_{(v,k)}|M_{(v,k)}, \pi_{per}) \equiv \mathcal{N}(\mu_E(M), \sigma_E(M))$

$P(C_{(v,k)}|M_{(v,k)}, \pi_{per}) \equiv \mathcal{N}(\mu_C(M), \sigma_C(M))$

For $k = 0$:

Map initialization.

$P(M_{(v,0)}|\pi_{per}) \equiv$ *Uniform*

For $k > 0$:

Map update using previous state information.

$P(M_{(v,k)}|\pi_{per}) =$

$P(M_{(v,k-1)}|e_{(v,k-1)}, c_{(v,k-1)}, \pi_{per})$

Identification:

$\mu_E(M), \sigma_E(M), \mu_C(M), \sigma_C(M)$

- Constants defined during the learning stage.

Question:

$P(M_{(v,k}|e_{(v,k)}, c_{(v,k)}, \pi_{per})$

(b)

Figure 7.5: Bayesian model π_{per}:*"Local perception of haptic stimulus"*. a) Graphical representation. b) Description of the Bayesian program.

$$C_{(v,k)} - \text{"Compliance characterization of v"}$$
$$C_{(v,k)} = g(\mathbf{h}_{(v,k)}), \quad C_{(v,k)} \in \mathbb{R} \tag{7.3}$$

The parameter $\mathbf{h}_{(v,k)}$ represents haptic sensing measurements provided by the sensory apparatus. The function g transforms the haptic sensing measurements $\mathbf{h}_{(v,k)}$ in a compliance characterization of the explored surface, while f transforms $\mathbf{h}_{(v,k)}$ in a texture characterization of the surface. This work uses the data provided by [Xu et al., 2013]; thus it considers the same operator functions f and g of the work [Xu et al., 2013].

This work considers the same set of materials ($n = 10$) that were used in [Xu et al., 2013]: acrylic, brick, copper, damp sponge, feather, rough foam, plush toy, silicone, soft foam, and wood. In the datasets made available by [Xu et al., 2013], the different categories of materials are characterized by varying properties of texture, compliance, and thermal conductivity, extracted using *BioTac* biomimetic tactile sensor raw data (contact intensity, vibration, and heat flow).

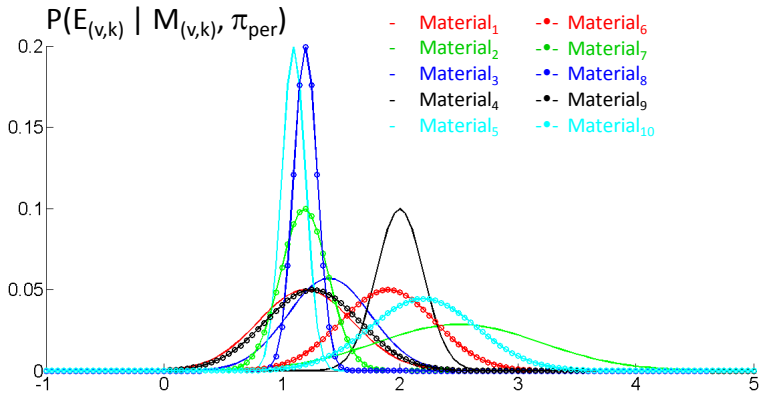
7.4.2 Inference of the haptic stimulus category

The statistical independence relations between the random variables $E_{(v,k)}, C_{(v,k)}, M_{(v,k)}$ are expressed in Figure 7.5 a). Based on these statistical assumptions, the joint probability distribution function $P(E_{(v,k)}, C_{(v,k)}, M_{(v,k)}, \pi_{per})$ is decomposed as described in Figure 7.5 b). Each of the factors follows a probability distribution function presented in Figure 7.5 b).

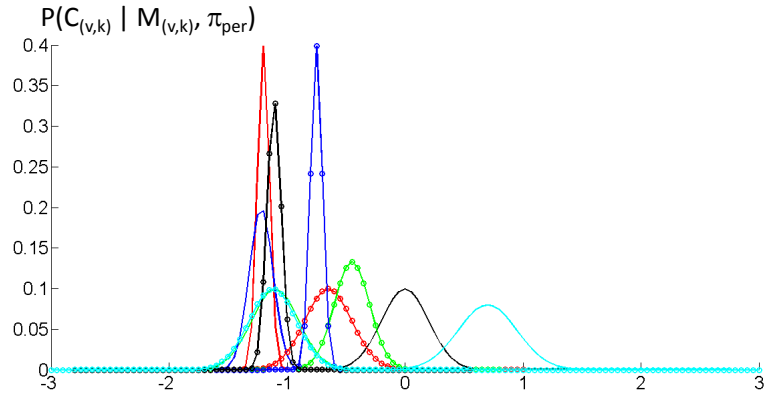
On each time iteration step, the probability distribution function $P(M_{(v,k)}|e_{(v,k)}, c_{(v,k)}, \pi_{per})$ describing the category of material of v is inferred using the observed data $e_{(v,k)}, c_{(v,k)}$ (extracted from the samples acquired by the sensory apparatus of the robotic system). Thus, the Bayesian program described in Figure 7.5 b) is run with the question presented in equation 7.4.

$$P(M_{(v,k)}|e_{(v,k)}, c_{(v,k)}, \pi_{per}) = \frac{P(e_{(v,k)}|M_{(v,k)}, \pi_{per})P(c_{(v,k)}|M_{(v,k)}, \pi_{per})P(M_{(v,k)}, \pi_{per})}{\sum_{M_{(v,k)}} P(e_{(v,k)}|M_{(v,k)}, \pi_{per})P(c_{(v,k)}|M_{(v,k)}, \pi_{per})P(M_{(v,k)}, \pi_{per})} \tag{7.4}$$

The determination of the free parameters $\mu_E(M)$, $\sigma_E(M)$, $\mu_C(M)$, $\sigma_C(M)$ of the Gaussian functions used to define the normal probability distributions $P(E_{(v,k)}|M_{(v,k)}, \pi_{per})$ and $P(C_{(v,k)}|M_{(v,k)}, \pi_{per})$ is detailed in section 7.4.3.



(a)



(b)

Figure 7.6: Representation of $P(E_{(vi,k)}|M_{(vi,k)}, \pi_{per})$ (a)) and $P(C_{(vi,k)}|M_{(vi,k)}, \pi_{per})$ (b)) learned for ten reference materials. Data extracted from [Xu et al., 2013].

7.4.3 Determination of $P(E_{(v,k)}|M_{(v,k)}, \pi_{per})$ and $P(C_{(v,k)}|M_{(v,k)}, \pi_{per})$

The free parameters $\mu_E(M)$, $\sigma_E(M)$, $\mu_C(M)$, $\sigma_C(M)$ of the Gaussian functions used to model the normal probability distributions $P(E_{(v,k)}|M_{(v,k)}, \pi_{per})$ and $P(C_{(v,k)}|M_{(v,k)}, \pi_{per})$ are estimated during experimental learning sessions. As described in [Xu et al., 2013], during the learning period, standard local exploration procedures are performed for each of the $n = 10$ reference materials.

After the pre-determined number of standard local explorations, the free parameters $\mu_E(M)$, $\sigma_E(M)$, $\mu_C(M)$, $\sigma_C(M)$ of the normal ($\mathcal{N}$) distributions are determined by calculating the averages μ and standard deviations σ of E and C for each reference material. The results are represented in Figures 7.6 a) and 7.6 b), extracting the data available from the manuscript of the work [Xu et al., 2013].

7.5 Recognition of the shape of the exploration path

7.5.1 Random variables of the model

In this work, the structure of the discontinuity is actively estimated as long as the haptic exploration performed by the robotic hand progresses. The exploration path and thus the structure of the haptic discontinuity is described in the bi-dimensional Cartesian space.

The Bayesian model π_{obj} implemented in this Bayesian program determines the category of the structure of the discontinuity being followed in the workspace. The category of the structure of the discontinuity is represented by the discrete random variable R_k, described in equation (7.5). The robotic system can recognize $\Theta = 2$ categories of structures of discontinuities.

$$R_k - \text{"Category of structure of the exploration path"}$$
$$R_k \in \{Shape_1, \ldots, Shape_\Theta\} \tag{7.5}$$

Each class of structure described by the discrete random variable R_k is associated with a template, which consists of a set of points representing the boundaries of each category of structure of the discontinuity, as described in equation (7.6).

$$\forall_{i \in \{1,\ldots,\Theta\}} \quad \langle \text{"}Shape_i\text{"}, template_i \rangle \tag{7.6}$$

The categories of structure recognized by the system consist of two elementary geometric shapes described by the edge points of a triangle and a rectangle. The templates are illustrated in Figure 7.7.

The sequence of regions explored by the robotic system until the time iteration $(k-1)$ is described by the set of workspace locations $(\hat{o}_0, \hat{o}_1, \ldots, \hat{o}_{k-1})$ (section 7.6). The categorization process consists of establishing a match between the points $(\hat{o}_0, \hat{o}_1, \ldots, \hat{o}_{k-1})$ explored by the robotic system until the iteration $(k-1)$ and each of the templates, $template_i$, representative of each category of structure of discontinuity. The normalized matching error between each template and the current exploration path is described by the continuous random variable L^i, described by the equation (7.7).

$$L^i - \text{"Normalized matching error between } (\hat{o}_0, \hat{o}_1, \ldots, \hat{o}_{k-1}) \text{ and } template_i\text{"}$$
$$[L^i, \Upsilon_i] = f_{ICP}((\hat{o}_0, \hat{o}_1, \ldots, \hat{o}_{k-1}), template_i)$$
$$L^i \in [0, 1] \qquad (7.7)$$

The matching between the two sets of points $(\hat{o}_0, \hat{o}_1, \ldots, \hat{o}_{k-1})$ and $template_i$ is determined using the method *Iterative Closest Point* (ICP) [Zhang, 1994], as described in equation 7.7.

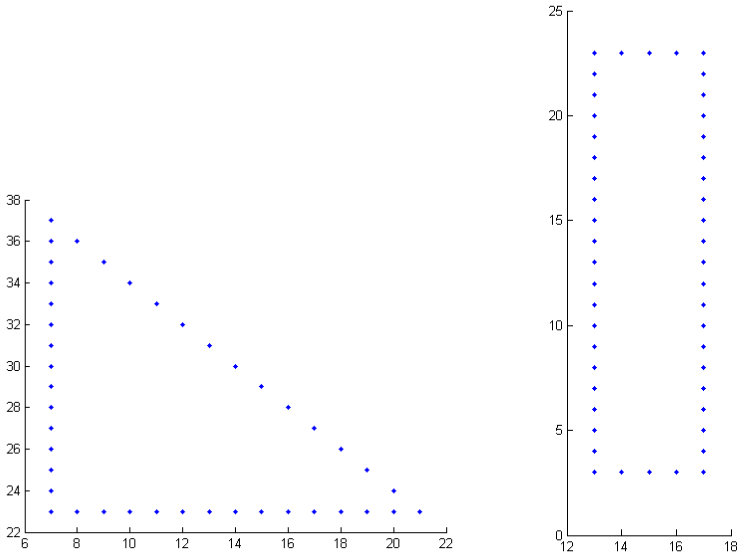
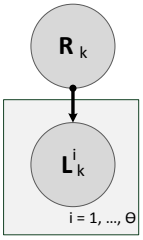


Figure 7.7: Datasets $\langle \text{"}Shape_i\text{"}, template_i \rangle$ for a) $R_k = \text{"}Template_1\text{"}$ b) $R_k = \text{"}template_2\text{"}$

The ICP function f_{ICP} also returns the estimation of the geometrical transformation Υ_i between the two sets of points. This transformation can be used to determine a new set of points $template_i^{'}$ which results from the registration of the points $template_i$ on the structure described by the set of points $(\hat{o}_0, \hat{o}_1, \ldots, \hat{o}_{k-1})$.

This relation can be described by the geometrical transformation represented in equation (7.8).

$$template_i^{'} = \Upsilon_i.template_i \qquad (7.8)$$



(a)

Bayesian Program

- Description
  - Specification
    - Relevant variables:
    - $R_k, L_k^1, \ldots, L_k^\Delta$
    - Decomposition:
    - $P(R_k, L_k^1, \ldots, L_k^\Delta | \pi_{obj}) =$
    - $P(L_k^1 | R_k, \pi_{obj}) \ldots . P(L_k^\Delta | R_k, \pi_{obj}) . P(R_k | \pi_{obj})$
    - Parametric forms:
    - $P(L_k^i | R_k, \pi_{obj}) \equiv \mathcal{B}(\alpha_{L^i}, \beta_{L^i})$
    - $P(R_k | \pi_{obj}) \equiv \mathit{Uniform}$
  - Identification:
  - $\alpha_{L^i}, \beta_{L^i}$
  - $-$ *Constants defined empirically.*
- Question:
- $P(R_k | l_k^1, \ldots, l_k^\Delta, \pi_{obj})$

(b)

Figure 7.8: Bayesian model π_{obj}:*"Recognition of the shapes of discontinuities."*. a) Graphical representation. b) Description of the Bayesian program.

7.5.2 Inference of the category of structure

The graphical representation of the Bayesian model π_{obj}, presented in Figure 7.8 a), expresses the statistical dependence relations between the random variables $L_k^1, \ldots, L_k^\Theta, R_k$. According to these dependence relations, the joint probability distribution function $P(L_k^1, \ldots, L_k^\Theta, R_k|\pi_{obj})$ can be factorized as presented in Figure 7.8 b).The probability distribution function followed by each of those factors is also presented in 7.8 b).

The inference of the category of the structure of the discontinuity is performed by running, at each time iteration step k, the Bayesian program presented in Figure 7.8 with the question $P(R_k|l_k^1, \ldots, l_k^\Theta, \pi_{obj})$, detailed in equation 7.9. This probability distribution function is determined by considering the observed normalized matching errors $l_k^1, \ldots, l_k^\Theta$.

$$P(R_k|l_k^1, \ldots, l_k^\Theta, \pi_{obj}) = \frac{P(l_k^1|R_k, \pi_{obj}).\ldots.P(l_k^\Theta|R_k, \pi_{obj}).P(R_k|\pi_{obj})}{\sum_{R_k} P(l_k^1|R_k, \pi_{obj}) \ldots P(l_k^\Theta|R_k, \pi_{obj})P(R_k|\pi_{obj})} \tag{7.9}$$

The determination of the probability distribution functions $P(l_k^i|R_k, \pi_{obj})$ is detailed in section 7.5.3.

7.5.3 Determination of $P(l_k^i|R_k, \pi_{obj})$

The probability distribution functions $P(l_k^i|R_k, \pi_{obj})$ are described by beta probability distribution functions $\mathcal{B}_L$ with the constant parameters $\alpha_L = 1.0$ and $\beta_L = 4.5$. We assume that all the Θ probability distribution functions are identical.

The typical profile of the probability distribution functions $P(l_k^i|R_k, \pi_{obj})$ is represented in Figure 7.9. The profile proposed for $P(l_k^i|R_k, \pi_{obj})$ attributes higher probabilities for lower levels of normalized matching errors l_k^i and lower probabilities to higher values of l_k^i. This promotes the selection of categories of the structure R_k that have a template similar to the current exploration path $(\hat{o}_0, \hat{o}_1, \ldots, \hat{o}_{k-1})$.

7.6 Integration of attention mechanisms in the inference of the exploration path

7.6.1 Random variables of the model

After the local exploration of the region v of the workspace is completed, the perceptual representation of the workspace is updated with the sensory measurements acquired at v (update mechanisms presented in section 7.4) and the robotic system has to determine

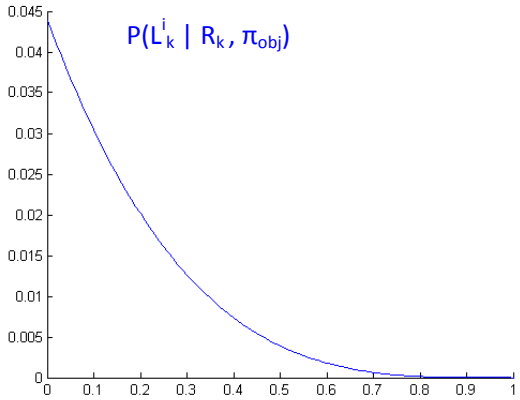


Figure 7.9: Graphical representation of $P(l_k^i|R_k, \pi_{obj})$.

which is going to be the next region v of the workspace grid to be explored.

The mechanisms involved in the selection of that region are described by the Bayesian model π_{tar}, presented in this section of the work. The next exploration target is represented by the discrete random variable O_k, presented in equation (7.10). θ is the total number of cells in the grid representing the workspace. v^i is a compact representation of the cell identifier.

$$O_k - \text{"Next workspace region to be explored"}$$
$$O_k \in \{v^1, v^2, v^3, \ldots, v^\theta\} \tag{7.10}$$

The sequence of regions on workspace $(\hat{o}_0, \hat{o}_1, \ldots, \hat{o}_{k-1})$ that were previously explored by the robotic system may provide cues about the spatial structure of the discontinuity being followed and indirectly influence the estimation of $\hat{o_k}$. The cues are provided by matching the current structure of the exploration path with representations of typical shapes stored in the memory of the robotic system. As presented in section 7.6.3, regions of the workspace that are coincident with the structure of the shape templates will be more likely to be explored.

The selection of O_k is also conditioned by inhibition-of-return mechanisms. The inhibition level imposed by the inhibition-of-return process involved in the touch attention mechanisms is implemented by the continuous random variable $I_{(v,k)}$, described in equation (7.11). The inhibition level assigned to each cell in the grid varies according to equation (7.11).

$$I_{(v,k)} - \text{"Inhibition level for cell v."}$$
$$I_{(v,k)} = 1 - \Theta d^{\alpha-1}(1-d)^{1-\beta}, \quad I_{(v,k)} \in [0,1] \tag{7.11}$$

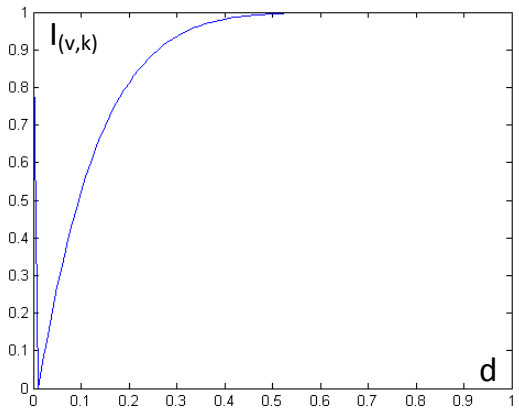


Figure 7.10: Graphical representation of $I_{(v,k)}$.

Due to the characteristics of the haptic exploration procedures presented in section 7.1, the inhibition-of-return process promotes, at time iteration $k+1$, the exploration of regions of the workspace different from the current position of the end-effector of the robotic system ($\hat{o}_{k-1}$). However, simultaneously, the inhibition-of-return process inhibits the exploration of regions too distant from $\hat{o}_{k-1}$, to avoid breaks during the following exploration tasks. The inhibition levels $I_{(v,k)}$ for each cell v can be described by the equation 7.11, considering $\alpha = 1.01$ and $\beta = 9$ (profile represented in Figure 7.10). The parameter d is determined by $d = d_k/d_{max}$. The parameter d_k expresses the Euclidean distance between o_k and $\hat{o}_{k-1}$, and d_{max} is a constant representing the maximum possible distance between o_k and $\hat{o}_{k-1}$ for the workspace dimensions. Θ is a normalization constant. The values of $I_{(v,k)}(d)$ range between 0 and 1. $I_{(v,k)} = 0$ indicates that the inhibition-of-return mechanism applies no inhibition to cell v, whereas $I_{(v,k)} = 1$ indicates a full inhibition of v.

The selection of the region O_k is also dependent on mechanisms to avoid the return to regions already explored and perceived with low uncertainty. These mechanisms are formulated to promote the curiosity of the system and are represented by the continuous random variable $U_{(v,k)}$, described in equation (7.12). The operator $\mathcal{H}$ determines the information entropy [Shannon, 2001b] of the discrete random variable $M_{(v,k)}$.

$$U_{(v,k)} - \text{"Uncertainty level for cell v."}$$
$$U_{(v,k)} = \frac{\mathcal{H}(M_{(v,k)})}{max(\mathcal{H}(M_{(v,k)}))}, \quad U_{(v,k)} \in [0,1] \tag{7.12}$$

Another factor conditioning the determination of O_k is the saliency of the haptic stimulus perceived in region v of the workspace and in its surroundings. In addition to depending on the perceived haptic stimulus $M_{(v,k)}$ map, the formulation of the saliency of those haptic stimulus is also dependent on the current objectives of the exploration task. The objectives of the task being executed by the robotic platform are represented by the discrete random variable $T =$ "Task objective.", given that $T \in \{T_1, \ldots, T_\Phi\}$. During an experimental trial, the value of $T = t$ is considered constant in time k. Φ expresses the total number of tasks that can be executed by the robotic platform.

Based on these considerations, the saliency of the haptic stimulus perceived in the surroundings of v can be formulated by the continuous random variable $S_{(v,k)}$, presented in equation (7.13).

$$S_{(v,k)} - \text{"Saliency of the perceived haptic stimulus in region v."} \tag{7.13}$$

This work defines $S_{(v,k)}$ for a class of tasks T="*Search and follow of discontinuities between regions of surfaces with* $Material_a$ *and* $Material_b$*.*", as presented in equation 7.14. $S_{(v,k)}$ is related by a soft evidence relation with the perceived haptic stimulus $M_{(v,k)}$ characterization of the workspace.

$$S_{(v,k)} = \frac{\max(|s_x|, |s_y|)}{s_{norm}}$$
$$S_{(v,k)} \in [0,1] \tag{7.14}$$

The parameters $s_x = \mathcal{G}_{sobel_x}(\mathbf{d})$, $s_y = \mathcal{G}_{sobel_y}(\mathbf{d})$, and $s_z = \mathcal{G}_{sobel_z}(\mathbf{d})$ are determined using the edge detector $\mathbf{G}_{sobel}$ following an approach analogous to the operator proposed approach in [Bhattacharya and Wild, 1996]. Considering that the exploratory element is located at region v of the workspace, a $type-8$ neighbourhood can be defined around that location. For a given neighbourhood $v1, \ldots, v8$, we can define $\mathbf{d} = (\Omega_{(v1,k)}, \ldots, \Omega_{(v8,k)})$ as the set of values of $\Omega_{(vi,k)}$. We consider that the haptic stimulus perceived at each of the cells $(v1, \ldots, v8)$ of a neighbourhood can be described by a probability distribution function $P(M_{(vi,k)}|e_{(vi,k)}, c_{(vi,k)}, \pi_{per})$, respectively (details in section 7.4). Considering

each region vi, a constant $\Omega_{(vi,k)}$ is defined, expressing the similarity of the perceived material category of a region with $Material_a$ or $Material_b$. The constant $\Omega_{(vi,k)} \in [0,1]$ is determined by equation 7.15.

$$\Omega_{(vi,k)} = \frac{1 - P(M_{(vi,k)} = Mat._b | e_{(vi,k)}, c_{(vi,k)}, \pi_{per}) + P(M_{(vi,k)} = Mat._a | e_{(vi,k)}, c_{(vi,k)}, \pi_{per})}{2} \tag{7.15}$$

7.6.2 Inference of the next exploration target

Based on the statistical independence relations between the random variables O_k, $I_{(v,k)}$, $U_{(v,k)}$, R_k $S_{(v,k)}$, T, presented in Figure 7.11 a), the joint probability distribution function $P(O_k, T, S_{(v,k)}, U_{(v,k)}, I_{(v,k)}, R_k | \pi_{tar})$ for the Bayesian model π_{tar} can be decomposed as summarized in Figure 7.11 b). Each factor is described by a probability distribution function presented in Figure 7.11 b). The final estimate for the next exploration target $\hat{o}_k$ is given via a *Maximum a Posteriori* (MAP) decision rule, as expressed in equation 7.16 given a specific task $T = t$.

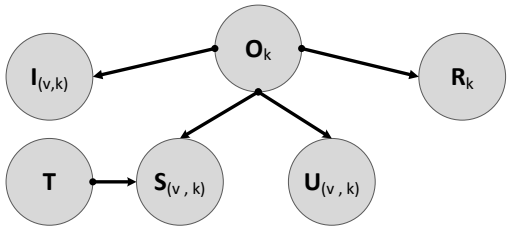
$$\begin{aligned} \hat{o}_k &= \arg\max_{o_k} P(O_k | t, s_{(v,k)}, i_{(v,k)}, u_{(v,k)}, R_k, \pi_{tar}) \\ \hat{o}_k &= \arg\max_{o_k} \sum_{R_k} \begin{pmatrix} P(t|\pi_{tar}) P(i_{(v,k)}|O_k, \pi_{tar}). \\ P(O_k|\pi_{tar}) P(s_{(v,k)}|O_k, t, \pi_{tar}). \\ P(u_{(v,k)}|O_k, \pi_{tar}) P(R_k|O_k, \pi_{tar}) \end{pmatrix} \end{aligned} \tag{7.16}$$

The determination of the probability distribution functions $P(S_{(v,k)}|O_k, T, \pi_{tar})$, $P(I_{(v,k)}|O_k, \pi_{tar})$, $P(U_{(v,k)}|O_k, \pi_{tar})$, $P(R_k|O_k, \pi_{tar})$ involved in equation (7.16) is described in detail in section 7.6.3.

7.6.3 Determination of $P(S_{(v,k)}|O_k, T, \pi_{tar})$, $P(I_{(v,k)}|O_k, \pi_{tar})$, $P(U_{(v,k)}|O_k, \pi_{tar})$, $P(R_k|O_k, \pi_{tar})$

The probability distribution function $P(R_k|O_k, \pi_{tar})$ is characterized by a Gaussian Mixture Model (GMM), as presented in equation 7.17.

$$P(R_k = "Object_j" | O_k, \pi_{tar}) = \sum_{i \in template'_j} w_i . g_i(O_k | \mu_i, \Sigma) \tag{7.17}$$



(a)

Bayesian Program

Description

Specification

Relevant variables:

$O_k, T, S_{(v,k)}, U_{(v,k)}, I_{(v,k)}, R_k$

Decomposition:

$$P(O_k, T, S_{(v,k)}, U_{(v,k)}, I_{(v,k)}, R_k|\pi_{tar}) =$$
$$\left(\begin{array}{c} P(O_k|\pi_{tar}).P(T|\pi_{tar}).P(I_{(v,k)}|O_k,\pi_{tar}). \\ P(U_{(v,k)}|O_k,\pi_{tar}).P(R_k|O_k,\pi_{tar}).P(S_{(v,k)}|O_k,T,\pi_{tar}) \end{array}\right)$$

Parametric forms:

$P(O_k|\pi_{tar}) \equiv Uniform$

$P(T|\pi_{tar}) \equiv Uniform$

$P(I_{(v,k)}|O_k,\pi_{tar}) \equiv \mathcal{B}(\alpha_I, \beta_I)$

$P(U_{(v,k)}|O_k,\pi_{tar}) \equiv \mathcal{B}(\alpha_U, \beta_U)$

$P(R_k|O_k,\pi_{tar}) \equiv G.M.M.$

$P(S_{(v,k)}|O_k,T,\pi_{tar}) \equiv \mathcal{B}(\alpha_S, \beta_S)$

Identification:

$\alpha_I, \beta_I, \alpha_U, \beta_U, \alpha_S, \beta_S$

$-$ *Constants defined empirically.*

Question:

$P(O_k|t, s_{(v,k)}, u_{(v,k)}, i_{(v,k)}, R_k, \pi_{tar})$

Decision criteria: *MAP - Maximum a posteriori.*

$\hat{o}_k = \arg\max_{o_k} P(O_k|t, s_{(v,k)}, u_{(v,k)}, i_{(v,k)}, R_k, \pi_{tar})$

(b)

Figure 7.11: Bayesianl model π_{tar}:*"Selection of the next exploration target"*. a) Graphical representation. b) Description of the Bayesian program.

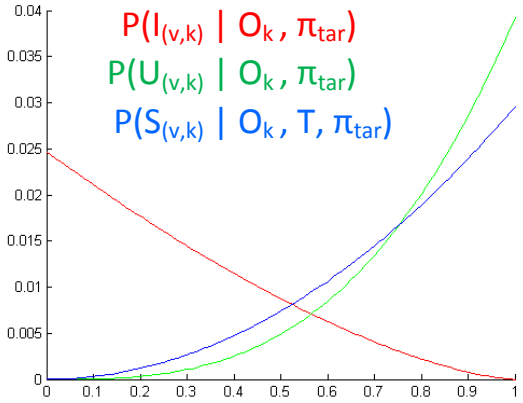


Figure 7.12: Graphical representation of $P(I_{(v,k)}|O_k, \pi_{tar})$, $P(U_{(v,k)}|O_k, \pi_{tar})$, $P(S_{(v,k)}|O_k, T, \pi_{tar})$.

The Gaussians g_i of the Gaussian Mixture Model are centred at the locations μ_i of the workspace, with a covariance matrix Σ. Assuming a 2-D structure of the workspace, each of the Gaussian function g_i is represented by equation 7.18.

$$g_i(O_k|\mu_i, \Sigma) = \frac{1}{2\pi \mid \Sigma \mid^{1/2}} \exp^{-\frac{1}{2}(O_k-\mu_i)^T \Sigma^{-1}(O_k-\mu_i)} \tag{7.18}$$

The centres μ_i of the Gaussians correspond to the points belonging to the set $Template_j^{'}$, determined as presented in detail in section 7.5. The exploration of regions predicted as belonging to stereotyped structures of exploration paths is promoted by assigning higher probabilities to them.

As presented in Figure 7.11 b), $P(I_{(v,k)}|O_k, \pi_{tar})$ is described by a beta probability distribution function $\mathcal{B}_I$ characterized by the constants $\alpha_I = 1$ and $\beta_I = 2.5$. The profile of the probability distribution function $P(I_{(v,k)}|O_k, \pi_{tar})$ is represented in Figure 7.12. The selected profile for $P(I_{(v,k)}|O_k, \pi_{tar})$ attributes higher probabilities for lower levels of $I_{(v,k)}$ and lower probabilities to higher values of $I_{(v,k)}$ to promote the selection of regions of the workspace with low values of inhibition level.

Following an analogous approach, $P(U_{(v,k)}|O_k, \pi_{tar})$ is described by a beta probability distribution function $\mathcal{B}_U$ (Figure 7.12) with the constant parameters $\alpha_U = 4$ and $\beta_U = 1$. $P(U_{(v,k)}|O_k, \pi_{tar})$ attributes higher probability values to regions of the workspace perceived with higher uncertainty $U_{(v,k)}$, promoting the curiosity of the system.

$P(S_{(v,k)}|O_k, T, \pi_{tar})$ is described by a beta probability distribution function $\mathcal{B}_R$ defined by $\alpha_R = 3$ and $\beta_R = 1$ (Figure 7.12), assigning higher probability values to workspace regions v with higher values of saliency $S_{(v,k)}$, promoting the exploration of regions of the

workspace with relevant haptic stimulus for the task under execution.

7.7 Experimental results

7.7.1 Computational simulation environment

The path planning method proposed by this work, supporting the global haptic exploration strategy, was simulated in a computational environment. As detailed previously in section 7.3.1, this work does not address the low-level (motor and sensing) control loop involved in physical interaction between robotic fingers and surface.

The simulation environment consists of a planar 2D probabilistic grid representing the workspace placed in front of a robotic platform, as represented in Figure 7.13. This work considers that all the regions of the workspace are reachable by the robotic exploratory element. In the scenario illustrated in Figure 7.13, three different materials were used: wood ($Material_{10}$, brown cells), silicone ($Material_8$, blue cells), and flush ($Material_7$, green cells). The spatial distribution of the three materials is employed to simulate a hypothetical real-world scenario, as shown in Figure 7.13.

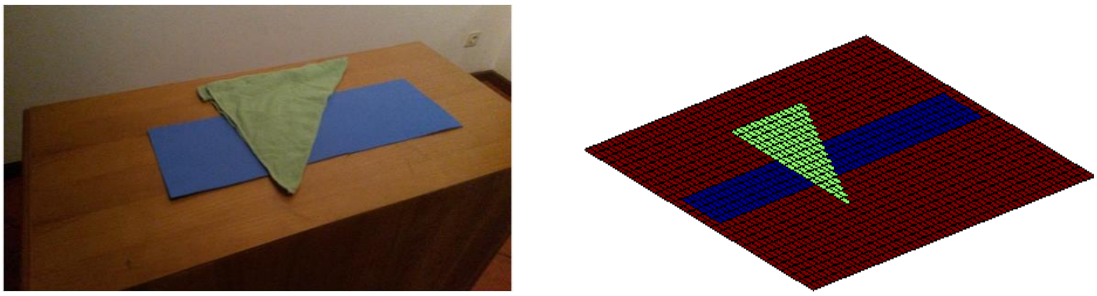
In the computational simulation, the sensory features modelling the haptic properties, texture ($E_{(v,k)}$) and compliance ($C_{(v,k)}$), of materials $Material_7$, $Material_8$, $Material_{10}$, were extracted from a previous work [Xu et al., 2013], as detailed in section 7.4.

The workspace region is integrated on top of the table. The workspace region is partitioned in a bidimensional grid, as suggested in Figure 7.3. In this work, the workspace grid has the following lower and upper dimensions, respectively: $X_l^W = 0m$, $X_u^W = 0.30m$, $Y_l^W = 0m$, $Y_u^W = 0.60m$. Each cell (square) has a side dimension of $\varepsilon = 0.01m$.

Scenario

Different configurations (homogeneous surfaces, heteronegous surfaces) and behaviours of haptic stimulus can be specified in the simulation environment. This work considers haptic stimulus consisting of heterogeneous surfaces. The haptic exploration tasks were tested on surfaces made of three different materials.

The scenario proposed in this work is presented in Figure 7.13 a). Three different materials are used: wood (brown), silicone (blue), and flush (green). This scenario (spatial configuration of the haptic stimulus) is mounted on the top of the table presented in Figure 7.13 b)-c) and explored autonomously by the robotic platform. The objective of the exploration tasks tested in this scenario are described in section 7.7.3, and the different perspectives about the experimental results are discussed as well.



(a) (b)

Figure 7.13: a) Real-world representation of the scenario. b) Schematic representation of configuration of the haptic stimulus placed in the workspace. The materials wood ($Material_{10}$), silicone ($Material_8$) and flush ($Material_7$) are represented in brown, blue, and green, respectively.

7.7.2 Evaluation of the haptic stimulus perception model

As mentioned previously, this work extracts the parameters $\mu_E(M)$, $\sigma_E(M)$, $\mu_C(M)$, $\sigma_C(M)$ from the work of [Xu et al., 2013] . However, the Bayesian program proposed in this work to categorize the haptic stimulus $Material_1, \ldots, Material_n$ implements a different approach than in [Xu et al., 2013].

Following an approach analogous to several previous works (e.g. [Xu et al., 2013], [Bologna et al., 2013], [Liu et al., 2012], [Pape et al., 2012]), the performance of the Bayesian model π_{per} proposed in section 7.4 was evaluated by performing a numerical simulation of 400 runs consisting of the local haptic exploration of samples of each of the reference materials $Material_1, \ldots, Material_{10}$. On each run, the local haptic exploration of the reference materials was simulated by generating random samples $e'_{(v,k)}$ and $c'_{(v,k)}$, obtained from $e_{(v,k)}$ and $c_{(v,k)}$, corrupted with additive white Gaussian noise (q_C and q_E), according to the formulations $c'_{(v,k)} = c_{(v,k)} + q_C$ and $e'_{(vi,k)} = e_{(v,k)} + q_E$, respectively.

As presented previously, $C_{(v,k)} \sim \mathcal{N}(\mu_C(M), \sigma_C(M))$ and $E_{(v,k)} \sim \mathcal{N}(\mu_E(M), \sigma_E(M))$ and in this experimental setup the additive white Gaussian noise is described by $Q_C \sim \mathcal{N}(0, \frac{\mu_C(M)}{2})$ and $Q_E \sim \mathcal{N}(0, \frac{\mu_E(M)}{2})$. For each reference material $Material_i$, the classification performance was evaluated for a initial exploration of that region v ($k = 0$) and for progressive exploration of that region of the workspace v during ($k = 1, \ldots, 4$) exploration iterations. The categorization $\hat{m}_{(v,k)}$) of the perceived haptic stimulus is determined by *MAP - Maximum a Posteriori* following equation 7.19.

$$\hat{m}_{(v,k)} = \underset{m_{(v,k)})}{\arg\max}\, P(M_{(v,k)} | e'_{(v,k)}, c'_{(v,k)}, \pi_{per}) \tag{7.19}$$

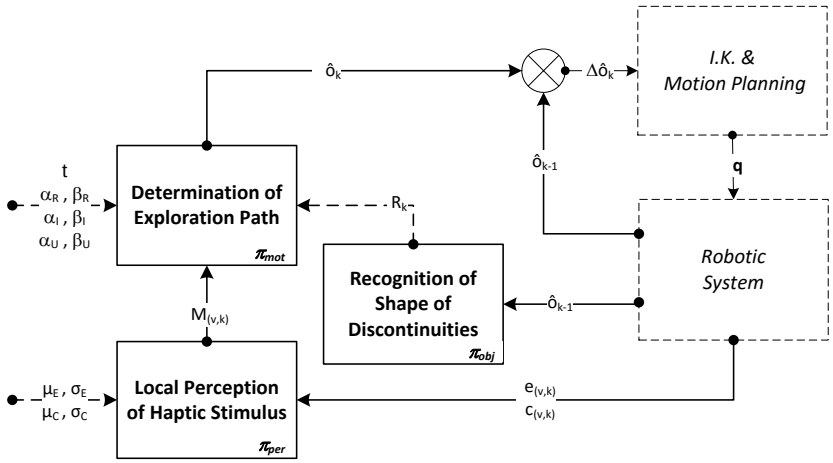


Figure 7.14: The modules and variables involved in the determination of the reference signal $\hat{o}_k$ are represented with a solid line.

Table 7.3: Confusion table for the categorization of $Material_i$ (ground truth) as $M.i$ (perceived category) by the Bayesian model π_{per}, using only one exploration sample $k = 0$ (400 runs).

	M.1	*M.2*	*M.3*	*M.4*	*M.5*	*M.6*	*M.7*	*M.8*	*M.9*	*M.10*
$Material_1$	243	0	60	0	0	0	0	0	77	20
$Material_2$	17	176	20	0	0	45	1	4	22	115
$Material_3$	108	2	146	0	0	0	0	0	97	47
$Material_4$	0	0	0	371	0	25	4	0	0	0
$Material_5$	0	0	0	0	397	0	3	0	0	0
$Material_6$	0	11	1	24	0	257	38	29	14	26
$Material_7$	0	1	0	6	1	19	340	32	1	0
$Material_8$	0	0	0	0	0	10	9	381	0	0
$Material_9$	78	1	9	0	0	0	0	0	270	42
$Material_{10}$	15	65	32	0	0	42	1	6	38	201

The evaluation of the performance of the Bayesian model π_{per} proposed in section 7.4 is presented in the confusion tables using one exploration iterations ($k = 0$, Table 7.3) in location v and using five exploration iterations ($k = 4$, Table 7.4).

The results presented in Table 7.3 show that, globally, the Bayesian model π_{per} has a good capability to discriminate and categorize the perceived haptic stimulus with the correct category of reference materials $Material_i$. The Bayesian model π_{per} shows a worst classification performance for haptic stimulus $Material_2$, $Material_3$, and $Material_{10}$. By integrating a higher number of sensory samples ($k = 4$, Table 7.4), the global performance of the Bayesian model π_{per} increases, including $Material_2$, $Material_3$, and $Material_{10}$ materials. The integration of five sensory samples ($k = 4$) allows the system to improve the erroneous effect introduced by the uncertainties of the measurements and by the additive white Gaussian noise. As in other works [Xu et al., 2013], [Bologna et al., 2013], [Liu et al., 2012] and [Pape et al., 2012], the different materials are correctly discriminated

Table 7.4: Confusion table for the categorization of $Material_i$ (ground truth) as $M.i$ (perceived category) by the Bayesian model π_{per}, using five exploration samples $k = 4$ (400 runs).

	M.1	*M.2*	*M.3*	*M.4*	*M.5*	*M.6*	*M.7*	*M.8*	*M.9*	*M.10*
$Material_1$	331	1	50	0	0	0	0	0	18	0
$Material_2$	0	316	2	0	0	0	0	0	0	82
$Material_3$	34	0	337	0	0	0	0	0	25	4
$Material_4$	0	0	0	399	0	1	0	0	0	0
$Material_5$	0	0	0	0	400	0	0	0	0	0
$Material_6$	0	0	0	0	0	397	1	0	0	2
$Material_7$	0	0	0	0	0	0	400	0	0	0
$Material_8$	0	0	0	0	0	0	0	400	0	0
$Material_9$	17	1	19	0	0	0	0	0	362	1
$Material_{10}$	0	92	2	0	0	1	0	0	0	305

with a high performance (average recognition rate higher than 90%).

This work also studied how the classification performance of the Bayesian model π_{per} can be affected by increasing levels of additive white noise Q_C and Q_E. The increasing levels of additive white Gaussian noise were simulated by increasing the standard deviation of the distributions of Q_C and Q_E, as presented in Figure 7.15. By increasing the magnitude of the standard deviation of the distributions of Q_C and Q_E, the classification performance of the Bayesian model π_{per} decreases. This effect is attenuated by the consecutive integration of several sensory samples ($k = 4$). This demonstrates the relevancy of implementing an active haptic exloration strategy in order to promote the exploration of uncertainty regions of the workspace to improve the current perceptual representation.

7.7.3 Autonomous exploration of the workspace

This work assumes that at each time iteration step k of the system, illustrated in Figure 7.4, an exploratory element of a robotic hand probes a workspace region v. The sensory samples modelling texture $e_{(v,k)}$ and compliance $c_{(v,k)}$ are artificially synthesised from the respective probability distribution functions $P(E_{(v,k)}|m_{(v,k)}, \pi_{per})$ and $P(C_{(v,k)}|m_{(v,k)}, \pi_{per})$, given the known ground truth material $m_{(v,k)}$ for that region of the workspace, as defined in Figure 7.13. Following the architecture of the sensory processing pipeline represented in Figure 7.4, the sensory feature samples $e_{(v,k)}$, $c_{(v,k)}$ are integrated by the Bayesian models to infer the next region (via point) of the workspace that should be probed by a robotic system.

In this scenario, the exploratory element of the robotic system is initialized ($k = 0$) at random locations of the 2D grid representing the workspace. The full list of initialization locations for the 100 runs is available online at `http://www.rmartins.net/phd-docs/ae/`. Unlike the previous work [Martins et al., 2014], these cells of the grid are not only located on a haptic discontinuity between the different materials of the scenario, they can be located on homogeneous regions. This provides a completely blind and unbiased

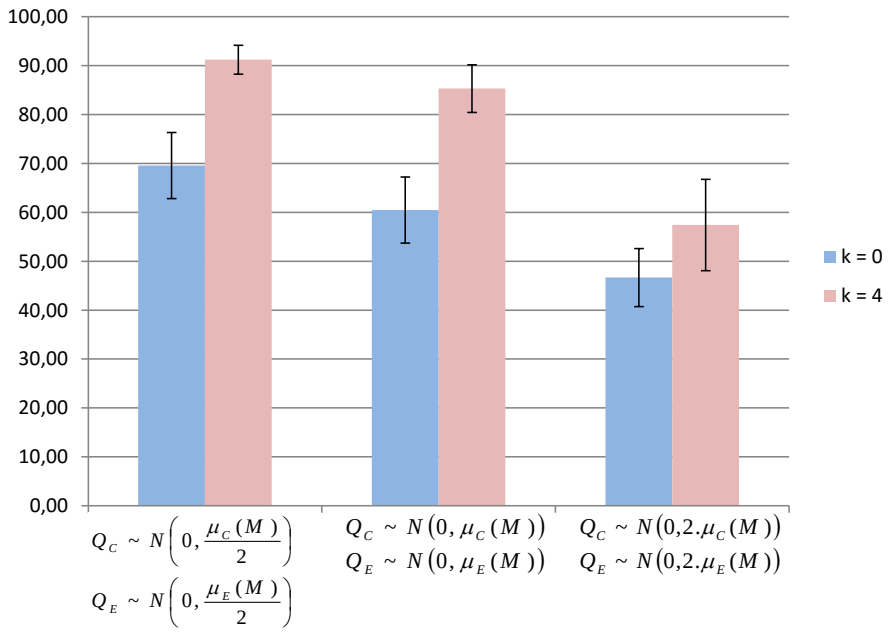


Figure 7.15: Classification performance (average for ten materials) of the Bayesian model π_{per}, using sensory samples corrupted with three different levels of additive white noise. The performance is evaluated integrating 1 ($k = 0$) and 5 ($k = 4$) successive sensory samples. The error bars represent the SEM (standard error of mean).

initialization of the exploration task for each exploration run.

For each of the exploration tasks described next, the workspace presented in Figure 7.13 was explored during 100 runs (100 different initial locations of the exploratory element). For each run, the exploration procedure lasts $k = 100$ time iterations.

Exploration tasks

To evaluate the specificity and robustness of the Bayesian models implementing the touch attention mechanisms proposed in this work, the autonomous exploration of the workspace was performed using three different tasks (T_1, T_2, and T_3). The objectives of T_1, T_2, and T_3 are the following:

- T_1="search and follow of discontinuities between $Material_7$ and other materials";
- T_2="search and follow of discontinuities between $Material_8$ and other materials";
- T_3="search and follow of discontinuities between $Material_{10}$ and other materials";

The ground-truth exploration path for each of the tasks T_1, T_2, and T_3 is defined in Figure 7.17. The ground-truth exploration paths are used for benchmarking purposes dur-

ing the evaluation of the general behaviour of the system and analysis of the contribution of the different components of the Bayesian models.

Performance metric

Although the internal structure and configuration of the haptic stimulus encountered in the workspace is unknown *a-priori* to the robotic system, the ground truth describing the target locations (cells) of the workspace that should be visited by the robotic platform during the task execution are defined by a human operator for benchmark purposes and represented by $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \ldots, \mathbf{b}_l\}$, $\mathbf{b}_i = (x, y) \in \mathbb{R}^2$. The set of workspace regions visited by the robotic platform during the task execution can be represented by $\mathcal{V} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_k\}$, $\mathbf{v}_i = (x, y) \in \mathbb{R}^2$.

The performance of the autonomous execution of the task by the robotic platform during an experimental run can be evaluated by an error metric defined in equation 7.20.

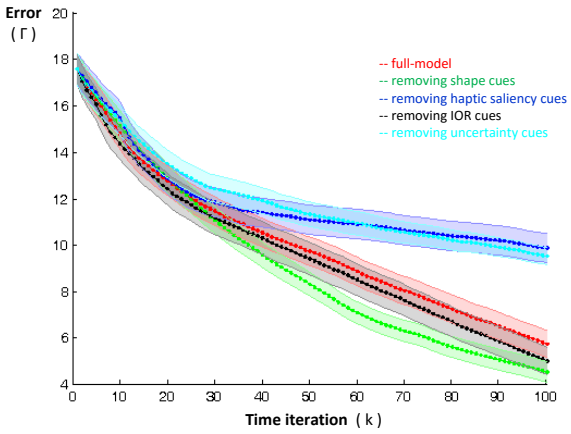
$$\Gamma = \sum_{i=1}^{l} \|\mathbf{b}_i - \mathbf{v}_{nearest}\|, \quad \text{given that}$$
$$\forall_{\mathbf{v}_i \in \mathcal{V}} \quad \exists_{\mathbf{v}_{nearest}} : \|\mathbf{b}_i - \mathbf{v}_{nearest}\| \leq \|\mathbf{b}_i - \mathbf{v}_i\| \tag{7.20}$$

$\|\ldots\|$ represents the Euclidean distance operator. Better autonomous exploration strategies provide lower values of Γ. This metric determines the total Euclidean distance between each ground truth point and the nearest point belonging to the exploration path executed by the robotic platform.

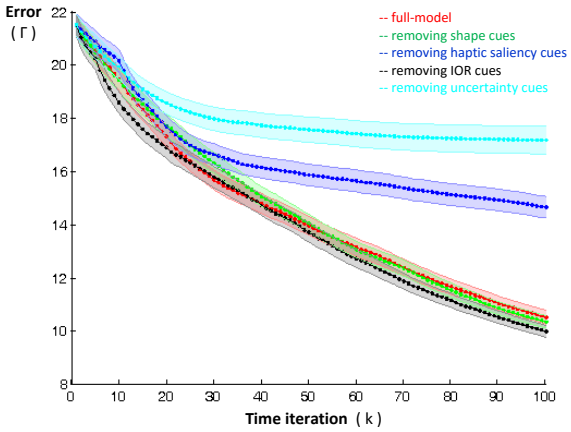
The impact of the different components (discontinuity shape cues, uncertainty, haptic saliency, and inhibition-of-return) of the Bayesian models implementing the touch attention mechanisms was also evaluated by comparing the exploration performance after discarding specific components of the Bayesian model π_{tar}: shape cues R_k, haptic saliency $S_{(v,k)}$, inhibition-of-return mechanisms $I_{(v,k)}$, and uncertainty cues $U_{(v,k)}$. The influence of those components was discarded by assuming that each of those random variables is described by a uniform probability distribution during all the time iterations of those experimental runs.

Discussion of the experimental results

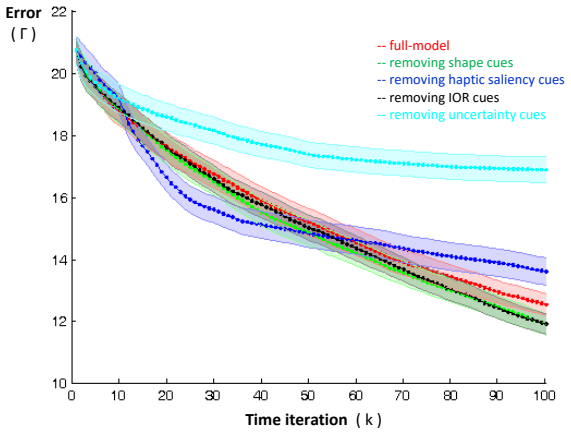
The ground truth exploration paths for the objectives of the exploration tasks T_1, T_2, and T_3 are represented in Figures 7.17 I.a), II.a), and III.a) representing the borders of the $Material_7$, $Material_8$, and $Material_{10}$ with the remaining materials in the workspace, respectively.



(a)



(b)



(c)

Figure 7.16: Temporal evolution (from $k = 0$ to $k = 100$) of mean value (average for the 100 runs; shaded colors represent SEM: standard error of mean) of performance metric Γ by integrating different configurations of Bayesian model π_{tar}: full-model, removing shape cues R_k, removing haptic saliency $S_{(v,k)}$, removing inhibition-of-return mechanisms $I_{(v,k)}$, removing uncertainty cues $U_{(v,k)}$. a) Task T_1. b) Task T_2. c) Task T_3.

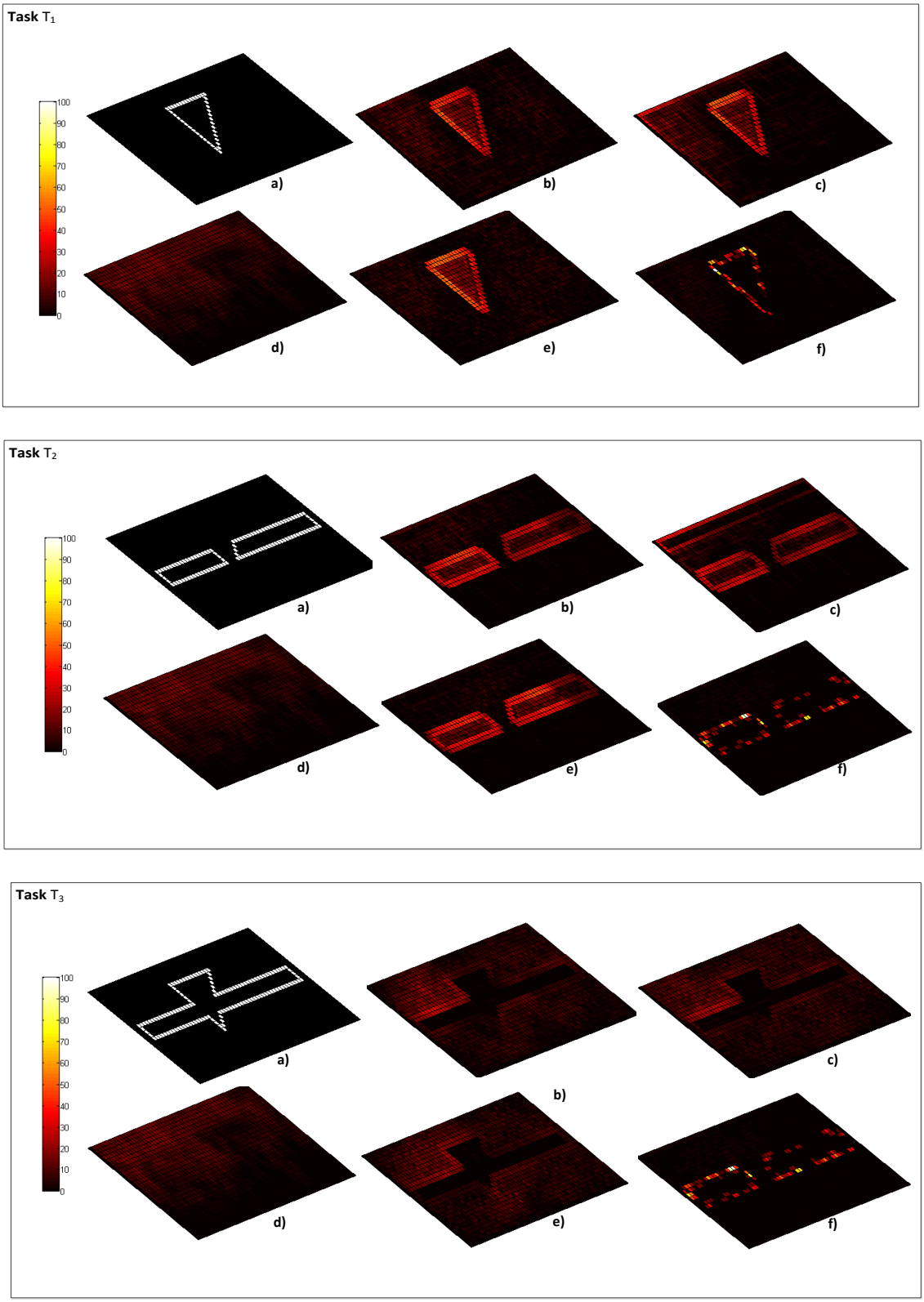


Figure 7.17: I.a) Ground truth exploration path for task T_1. II.a) Ground truth exploration path for task T_2. III.a) Ground truth exploration path for task T_3. I-III b)-f) Heat map of the exploration paths after 100 exploration runs with a duration of 100 time iterations each. Different exploration behaviours by integrating different configurations of the Bayesian model π_{tar}: b) full-model c) removing shape cues R_k d) removing haptic saliency $S_{(v,k)}$ e) removing inhibition-of-return mechanisms $I_{(v,k)}$ f) removing uncertainty cues $U_{(v,k)}$.

By performing an empirical comparison between the ground truth exploration paths and the heat maps resulting from the exploration behaviour inferred by the full Bayesian model π_{tar} in Figures 7.17 I.a)-b), 7.17 II.a)-b), and 7.17 III.a)-b), there is a high correspondence between the spatial structure of the most explored regions and the spatial structure of the ground truth exploration paths. The performance metric presented in Figure 7.16 also shows that the full-model always provides a good result. The touch attention mechanisms implemented by the Bayesian model π_{tar} have promoted the exploration of regions corresponding to the discontinuities described in the objectives of the tasks T_1, T_2, and T_3, ignoring other types of haptic discontinuities in the scenario.

The structural correspondence is better for the exploration tasks T_1 and T_2, which involve the search and follow of discontinuities involving $Material_7$ and $Material_8$, respectively, and the remaining materials. This better performance is justified by the better perceptual discrimination capability of this system concerning $Material_7$ and $Material_8$ relative to $Material_{10}$. The perceptual discrimination capability of this system was evaluated extensively in section 7.7.2. The confusion of $Material_{10}$ with other materials causes the haptic sensory data acquired during the exploration of regions made of $Material_{10}$ to be perceived erroneously as other categories of materials. During the execution of the exploration task T_3, this phenomenon has induced a high magnitude of haptic saliency during the exploration of homogeneous regions of the workspace, dispersing the system during the exploration task. However, it is evident from Figure 7.17 III.b) that the exploration paths follow the borders between $Material_{10}$ and the remaining materials in the workspace.

The analysis of the results of discarding the influence of specific components of the Bayesian model π_{tar} (Figure 7.17) shows that the degradation of performance of the exploration behaviour is big (Figure 7.16) when the effect of the haptic saliency $S_{(v,k)}$ is not considered. This causes the system to explore the workspace randomly, considering only criteria related to the regions where it was previously and not any information about relevancy for the task regarding the perceptual characteristics of the sensed haptic stimulus.

By neutralizing the integration of the information about the uncertainty of the perceptual representation of the workspace (Figure 7.17), the robotic system fails to have an exploration strategy that produces results similar to the ground truth. Although the Bayesian model π_{tar} implements inhibition-of-return mechanisms, their effect is transient on time, and after some time iterations, the system tends to return to the same regions of the workspace that were explored before and were perceived with low uncertainty; thus it can provide high haptic saliency features. The plot of the performance metric Γ for those conditions shows that the degradation of performance of the exploration behaviour

is considerable. Figure 7.16 describes this poor behaviour of the robotic system.

Conversely, by disabling the integration of the effects of the inhibition-of-return mechanisms (Figure 7.17), the performance of the execution of the exploration tasks is not as degraded. The plots of the metric Γ, presented in Figure 7.16, support this evidence by showing a performance of the system at the same level as the full-model condition. The removal of the transient effect of the inhibition-of-return mechanisms is compensated by the integration of information of the mechanisms related with the uncertainty of the perceptual representation of the workspace. Once the explored regions in this work tend to have low uncertainty (due to the high capability of the system to perceive and discriminate the materials used in this work), these regions tend to be avoided by the system, even without the influence of the inhibition-of-return mechanisms. However, the inhibition-of-return mechanisms can play a more relevant role in scenarios made of materials that the system can only discriminate with higher uncertainty. This topic will be evaluated in future works.

Discarding the effects provided by the integration of shape cues (Figure 7.17) does not have a strong influence on the performance of the exploration behaviour of the system (Figure 7.16). The weak contribution of the shape cues of the discontinuity to the improvement of the performance of the robotic system was caused by the low number of shape primitives recognized by the robotic system (only two: rectangle and triangle) and by the high number of points that were used to describe each of the shape templates (around 50 points). In future developments of this work, elementary shape primitives should be recognized by the system, and alternative methods to ICP should be tested so the system can recognize earlier tendencies about the shape of the discontinuity and match the current exploration path with the shape templates more robustly (noise, scale, and orientation).

7.8 Conclusions and future work

The integration of the touch attention mechanisms during the exploration of surfaces by robotic hands have proved to be effective to search and follow of haptic discontinuities in unknown and noisy environments. The updated perceptual representation of the workspace, provided by the Bayesian model π_{per}, together with shape cues about the structure of the discontinuity being followed, provided by the Bayesian model π_{obj} (extension of previous work [Martins et al., 2014]), are then used by the Bayesian model π_{tar} to infer the region that should be explored during the next time iteration.

The Bayesian models were tested in a scenario made of three different materials during three different haptic exploration tasks. The results presented in section 7.7.3 have

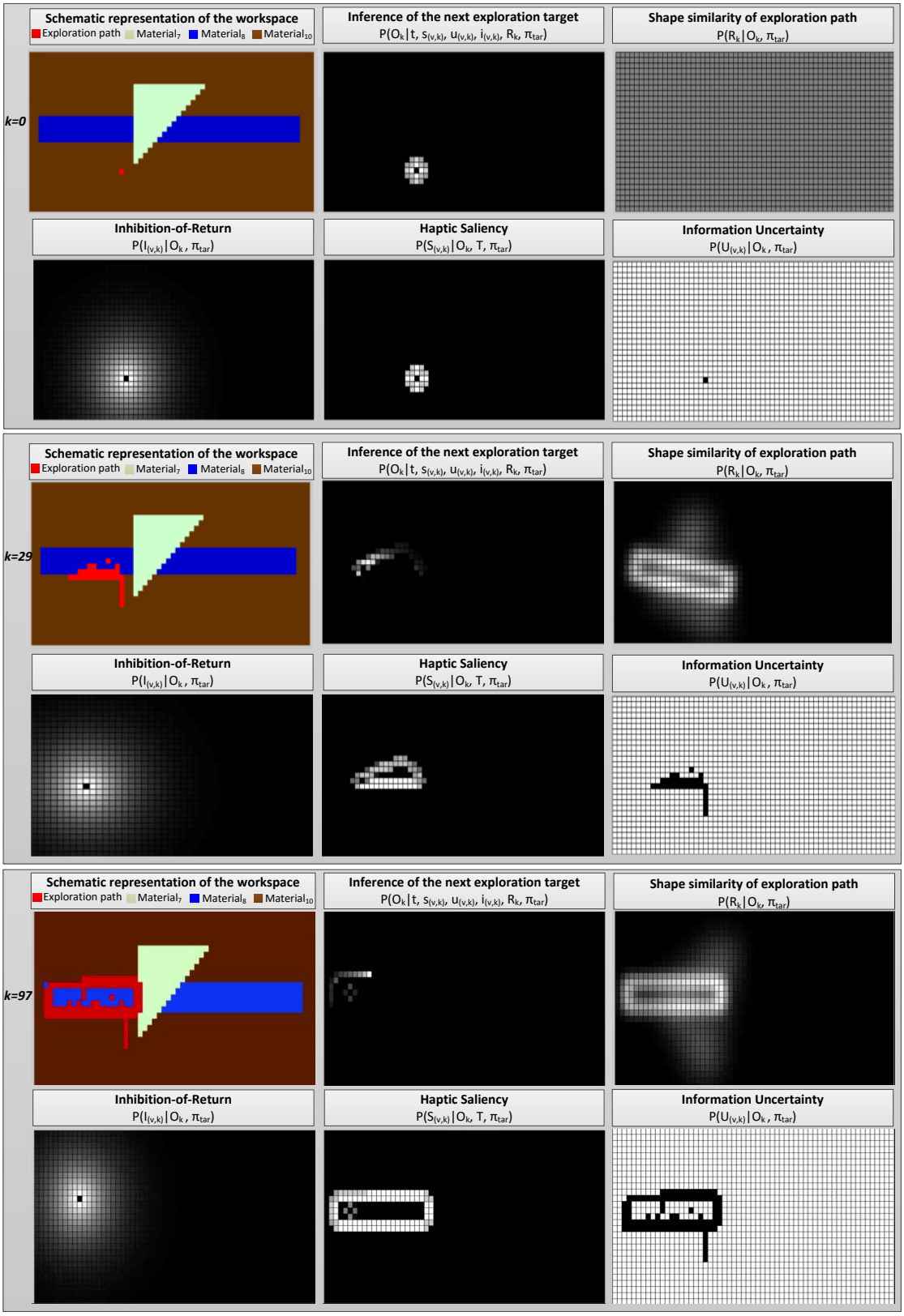


Figure 7.18: Representation of the $P(I_{(v,k)}|O_k, \pi_{tar})$, $P(S_{(v,k)}|O_k, T, \pi_{tar})$, $P(U_{(v,k)}|O_k, \pi_{tar})$, $P(R_k|O_k, \pi_{tar})$, $P(O_k|t, r_{(v,k)}, i_{(v,k)}, u_{(v,k)}, r_k, \pi_{tar})$ probability distribution functions and the exploration behaviour during the execution of the task T_2 *search and follow of discontinuities between* $Material_8$ *and remaining materials*, run 18. Dark colors represent lower values. Light colors represent higher values. Animated versions of this type of representations for autonomous exploration tasks T_1,T_2 and T_3, are available online `http://www.rmartins.net/phd-docs/ae/`.

demonstrated that the proposed approach provides to the robotic system a good framework to define and generalize the exploration behaviours. As in [Martinez et al., 2013], the system was able to handle severe changes in the slope of the discontinuities. In all the tasks, the robotic system was able to follow haptic discontinuities with progressive inversions in the slope of the discontinuity, which clearly demonstrates the generalization capability of the proposed approach. This emergent behaviour of the system presents an improvement of the results presented in [Martinez et al., 2013]. The test of the system with other slope variations in discontinuities than right angles (90 degrees) was suggested by [Martinez et al., 2013] for future work.

According to the results presented in section 7.7.3, the performance of the robotic system during the haptic exploration tasks is heavily dependent on the integration by the Bayesian model π_{tar} of information about the haptic saliency $S_{(v,k)}$ and uncertainty $U_{(v,k)}$ of the perceptual representation of the workspace.

The formulation of the contributions of the inhibition-of-return mechanisms $I_{(v,k)}$ and shape cues of the haptic discontinuities R_k will be studied extensively in future works to improve and optimize the contributions of these components of the Bayesian model π_{tar}.

The future developments of this work will also investigate the implementation of the automatic computational optimization of the parameters defining the profile of Beta distribution functions. Currently, the selection of parameters is made empirically, testing different sets of values and analysing the behaviour of the system.

List of Figures

List of Tables

Bibliography

[Adams et al., 2009] Adams, J., Hodges, K., Kujawa, J., and Metcalf, C. (2009). Test-retest reliability of the southampton hand assessment procedure. *International Journal of Rehabilitation Research*, 32:S18.

[Aggarwal and Ryoo, 2011] Aggarwal, J. and Ryoo, M. (2011). Human activity analysis: A review. *ACM Comput. Surv.*, 43(3):16:1–16:43.

[Amigoni et al., 2015] Amigoni, F., Bastianelli, E., Berghofer, J., Bonarini, A., Fontana, G., Hochgeschwender, N., Iocchi, L., Kraetzschmar, G., Lima, P., Matteucci, M., Miraldo, P., Nardi, D., and Schiaffonati, V. (2015). Competitions for benchmarking: Task and functionality scoring complete performance assessment. *Robotics Automation Magazine, IEEE*, 22(3):53–61.

[Amso and Scerif, 2015] Amso, D. and Scerif, G. (2015). The attentive brain: insights from developmental cognitive neuroscience. *Nature Reviews Neuroscience*, 16(10):606–619.

[Arbib et al., 2008] Arbib, M. A., Metta, G., and van der Smagt, P. P. (2008). Neurorobotics: From vision to action. In *Springer Handbook of Robotics*, pages 1453–1480.

[Argall et al., 2009] Argall, B. D., Chernova, S., Veloso, M., and Browning, B. (2009). A survey of robot learning from demonstration. *Robotics and Autonomous Systems*, 57(5):469 – 483.

[ATLAS, 2013] ATLAS (2013). Atlas - the agile anthropomorphic robot. URL: `http://www.theroboticschallenge.org/aboutrobots.aspx`.

[Ban and Lee, 2006] Ban, S.-W. and Lee, M. (2006). Selective attention-based novelty scene detection in dynamic environments. *Neurocomputing*, 69(1315):1723 – 1727. Blind Source Separation and Independent Component AnalysisSelected papers from the {ICA} 2004 meeting, Granada, SpainBlind Source Separation and Independent Component Analysis.

[Barron-Gonzalez et al.,] Barron-Gonzalez, H., Porrill, J., Lepora, N., Chinellato, E., Metta, G., and Prescott, T. Cerebellum-based adaptation for fine haptic control over the space of uncertain surfaces. In *World Haptics 2013*, pages 353–358.

[Berglund et al., 2012] Berglund, E., Iliev, B., Palm, R., Krug, R., Charusta, K., and Dimitrov, D. (2012). Mapping between different kinematic structures without absolute positioning during operation. *Electronics Letters*, 48(18):1110–1112.

[Bernardin et al., 2005] Bernardin, K., Ogawara, K., Ikeuchi, K., and Dillmann, R. (2005). A sensor fusion approach for recognizing continuous human grasping sequences using hidden markov models. *Robotics, IEEE Transactions on*, 21(1):47 – 57.

[Bhattacharya and Wild, 1996] Bhattacharya, P. and Wild, D. (1996). A new edge detector for gray volumetric data. *Computers in Biology and Medicine*, 26(4).

[Billard et al., 2008] Billard, A., Calinon, S., Dillmann, R., and Schaal, S. (2008). Robot programming by demonstration. In *Springer Handbook of Robotics*, pages 1371–1394.

[Bohg, 2011] Bohg, J. (2011). *Multi-Modal Scene Understanding for Robotic Grasping*. PhD thesis, KTH, Computer Vision and Active Perception, CVAP, Centre for Autonomous Systems, CAS.

[Bohg et al., 2010] Bohg, J., Johnson-Roberson, M., Bjoandrkman, M., and Kragic, D. (2010). Strategies for multi-modal scene exploration. In *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pages 4509 –4515.

[Bohg et al., 2013] Bohg, J., Morales, A., Asfour, T., and Kragic, D. (2013). Data-driven grasp synthesis - a survey. *Robotics, IEEE Transactions on*.

[Bologna et al., 2013] Bologna, L. L., Pinoteau, J., Passot, J.-B., Garrido, J. A., Vogel, J., Vidal, E. R., and Arleo, A. (2013). A closed-loop neurobotic system for fine touch sensing. *Journal of Neural Engineering*, 10(4).

[Bonsignorio and del Pobil, 2015] Bonsignorio, F. and del Pobil, A. (2015). Toward replicable and measurable robotics research [from the guest editors]. *Robotics Automation Magazine, IEEE*, 22(3):32–35.

[Calinon et al., 2007] Calinon, S., Guenter, F., and Billard, A. (2007). On learning, representing, and generalizing a task in a humanoid robot. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 37(2):286 –298.

[Calli et al., 2015] Calli, B., Walsman, A., Singh, A., Srinivasa, S., Abbeel, P., and Dollar, A. (2015). Benchmarking in manipulation research: Using the yale-cmu-berkeley object and model set. *Robotics Automation Magazine, IEEE*, 22(3):36–52.

[Case-Smith, 2000] Case-Smith, J. (2000). Effects of occupational therapy services on fine motor and functional performance in preschool children. *American Journal of Occupational Therapy*, 54(4):372–380.

[Castelli, 2002] Castelli, F. (2002). An integrated tactile-thermal robot sensor with capacitive tactile array. *IEEE T. Industry App.*, 38(1):85–90.

[Castiello, 2005] Castiello, U. (2005). The neuroscience of grasping. *Nature Reviews Neuroscience*, 6(9):726–736.

[Chapman, 2009] Chapman, C. (2009). Tactile attention. In Binder, M., Hirokawa, N., and Windhorst, U., editors, *Encyclopedia of Neuroscience*, pages 3990–3995. Springer Berlin Heidelberg.

[Chathuranga et al., 2013] Chathuranga, D., Ho, V., and Hirai, S. (2013). Investigation of a biomimetic fingertip's ability to discriminate fabrics based on surface textures. In *Int. Conf. Advanced Intelligent Mechatronics*, pages 1667–1674.

[Cheng, 2013] Cheng, G. (2013). *State-Action Gist based In-hand Manipulation Learning from Human Demonstration*. PhD thesis, Universitt Hamburg, Von-Melle-Park 3, 20146 Hamburg.

[Cheng, 2014] Cheng, G. (2014). *Humanoid Robotics and Neuroscience: Science, Engineering and Society*. CRC Press.

[Cheng et al., 2012] Cheng, G., Hendrich, N., and Zhang, J. (2012). Action gist based automatic segmentation for periodic in-hand manipulation movement learning. In *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, pages 4768–4775.

[Chitta et al., 2011] Chitta, S., Sturm, J., Piccoli, M., and Burgard, W. (2011). Tactile sensing for mobile manipulation. *Robotics, IEEE Transactions on*, 27(3):558 –568.

[Chung and AitSahlia, 2012] Chung, K. L. and AitSahlia, F. (2012). *Elementary probability theory: with stochastic processes and an introduction to mathematical finance*. Springer Science & Business Media.

[Corbetta and Shulman, 2002] Corbetta, M. and Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature reviews neuroscience*, 3(3):201–215.

[Cutkosky, 1989] Cutkosky, M. (1989). On grasp choice, grasp models, and the design of hands for manufacturing tasks. *Robotics and Automation, IEEE Transactions on*, 5(3):269 –279.

[CyberGloveII, 2008] CyberGloveII (2008). Cyberglove ii - wireless instrumented glove. URL: `http://www.cyberglovesystems.com/cyberglove-ii/`.

[Dahiya et al., 2013] Dahiya, R., Mittendorfer, P., Valle, M., Cheng, G., and Lumelsky, V. (2013). Directions toward effective utilization of tactile skin: A review. *Sensors Journal, IEEE*, 13(11):4121–4138.

[Dahiya et al., 2010] Dahiya, R. S., Metta, G., Valle, M., and Sandini, G. (2010). Tactile sensing - from humans to humanoids. *Robotics, IEEE Transactions on*, 26(1):1–20.

[Dario and De Rossi, 1985] Dario, P. and De Rossi, D. (1985). Tactile sensors and the gripping challenge: Increasing the performance of sensors over a wide range of force is a first step toward robotry that can hold and manipulate objects as humans do. *Spectrum, IEEE*, 22(8):46–53.

[del Pobil et al., 2014] del Pobil, A., Bonsignorio, F., and Messina, E. (2014). Fostering progress in performance evaluation and benchmarking of robotic and automation systems [tc spotlight]. *Robotics Automation Magazine, IEEE*, 21(1):22–25.

[Delson and West, 1996] Delson, N. and West, H. (1996). Robot programming by human demonstration: adaptation and inconsistency in constrained motion. In *Robotics and Automation, 1996. Proceedings., 1996 IEEE International Conference on*, volume 1, pages 30–36 vol.1.

[DEXMART, 2008] DEXMART (2008). project dexmart - dexterous and autonomous dual-arm/hand robotic manipulation with smart sensory-motor skills: A bridge from natural to artificial cognitio. URL:`http://www.dexmart.eu/`.

[DEXMART, 2009] DEXMART (2009). Dexmart project - deliverable d6.1 specification of benchmarks. `http://www.dexmart.eu/fileadmin/dexmart/public_website/downloads/216239_D6-1_Specification_of_Benchmarks.pdf`.

[Diego Faria, 2012] Diego Faria, Ricardo Martins, J. L. J. D. (2012). A probabilistic framework to detect suitable grasping regions on objects. In *SYROCO 2012*, pages 247–252.

[Dijkerman and Haan, 2007] Dijkerman, H. C. and Haan, E. H. F. D. (2007). Somatosensory processes subserving perception and action. *Behavioral and Brain Sciences*, 30(2):189 –239.

[Dipietro et al., 2008] Dipietro, L., Sabatini, A., and Dario, P. (2008). A survey of glove-based systems and their applications. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 38(4):461 –482.

[Dufour et al., 2011] Dufour, A., Thibeaux, R., Labruyere, E., Guillen, N., and Olivo-Marin, J.-C. (2011). 3-d active meshes: Fast discrete deformable models for cell tracking in 3-d time-lapse microscopy. *Image Processing, IEEE Transactions on*, 20(7):1925 – 1937.

[Elfes, 1989] Elfes, A. (1989). Using occupancy grids for mobile robot perception and navigation. *Computer*, 22(6):46–57.

[Elliott and Connolly, 1984] Elliott, J. M. and Connolly, K. (1984). A classification of manipulative hand movements. *Developmental Medicine & Child Neurology*, 26(3):283–296.

[Ernst and Bulthoff, 2004] Ernst, M. O. and Bulthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in cognitive sciences*, 8:162–169.

[EURON, 2008] EURON (2008). Euron 2006 - survey and inventory of current efforts in comparative robotics research - manipulation and grasping. URL: `http://www.robot.uji.es/EURON/en/manipulation.htm`.

[Exner, 1992] Exner, C. E. (1992). *Development of hand skills in the child*, chapter In-hand manipulation skills, pages 35–45. J. Case-Smith & C. Pehoski.

[Fagg and Arbib, 1998] Fagg, A. H. and Arbib, M. A. (1998). Modeling parietalpremotor interactions in primate control of grasping. *Neural Networks*, 11(78):1277 – 1303.

[Faix et al., 2015] Faix, M., Mazer, E., Laurent, R., Othman Abdallah, M., Le Hy, R., and Lobo, J. (2015). Cognitive computation: A bayesian machine case study. In *Cognitive Informatics Cognitive Computing (ICCI*CC), 2015 IEEE 14th International Conference on*, pages 67–75.

[Faria et al., 2010a] Faria, D., Martins, R., Lobo, J., and Dias, J. (2010a). Probabilistic representation of 3d object shape by in-hand exploration. In *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pages 1560 –1565.

[Faria et al., 2009] Faria, D. R., Martins, R., and Dias, J. (2009). Human reach-to-grasp generalization strategies: a bayesian approach. In *Robotics: Science and Systems 2009, Workshop: "Understanding the Human Hand for Advancing Robotic Manipulation"*.

[Faria et al., 2010b] Faria, D. R., Martins, R., and Dias, J. (2010b). *Grasp Exploration for 3D Object Shape Representation Using Probabilistic Map*, pages 215–222. Springer Berlin Heidelberg, Berlin, Heidelberg.

[Faria et al., 2010c] Faria, D. R., Martins, R., and Dias, J. (2010c). Learning motion patterns from multiple observations along the actions phases of manipulative tasks. In *Workshop on Grasping Planning and Task Learning by Imitation: IEEE/RSJ IROS*.

[Faria et al., 2011] Faria, D. R., Martins, R., Lobo, J., and Dias, J. (2011). *Manipulative Tasks Identification by Learning and Generalizing Hand Motions*, pages 173–180. Springer Berlin Heidelberg, Berlin, Heidelberg.

[Faria et al., 2012] Faria, D. R., Martins, R., Lobo, J., and Dias, J. (2012). Extracting data from human manipulation of objects towards improving autonomous robotic grasping. *Robotics and Autonomous Systems*, 60(3):396 – 410.

[Feix et al., 2014] Feix, T., Bullock, I., and Dollar, A. (2014). Analysis of human grasping behavior: Object characteristics and grasp type. *Haptics, IEEE Transactions on*, 7(3):311–323.

[Feix et al., 2013] Feix, T., Romero, J., Ek, C., Schmiedmayer, H., and Kragic, D. (2013). A metric for comparing the anthropomorphic motion capability of artificial hands. *Robotics, IEEE Transactions on*, 29(1):82–93.

[Ferreira and Dias, 2014a] Ferreira, J. and Dias, J. (2014a). Bayesian decision theory and the action-perception loop. In *Probabilistic Approaches to Robotic Perception*, volume 91 of *Springer Tracts in Advanced Robotics*, pages 121–145. Springer International Publishing.

[Ferreira and Dias, 2014b] Ferreira, J. and Dias, J. (2014b). Probabilistic learning. In *Probabilistic Approaches to Robotic Perception*, volume 91 of *Springer Tracts in Advanced Robotics*, pages 147–167. Springer International Publishing.

[Ferreira and Dias, 2014c] Ferreira, J. F. and Dias, J. (2014c). *Probabilistic Approaches to Robotic Perception*, volume 91 of *Springer Tracts in Advanced Robotics*. Springer.

[Fishel and Loeb, 2012] Fishel, J. A. and Loeb, G. E. (2012). Bayesian exploration for intelligent identification of textures. *Frontiers Neurorobotics*, 6.

[Flanagan et al., 2006] Flanagan, J. R., Bowman, M. C., and Johansson, R. S. (2006). Control strategies in object manipulation tasks. *Curr Opin Neurobiol*, 16(6):650–659.

[Frank et al., 2010] Frank, B., Schmedding, R., Stachniss, C., Teschner, M., and Burgard, W. (2010). Learning the elasticity parameters of deformable objects with a manipulation robot. In *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pages 1877 –1883.

[Garre et al., 2011] Garre, C., Hernández, F., Gracia, A., and Otaduy, M. A. (2011). Interactive simulation of a deformable hand for haptic rendering. In *Proc. of World Haptics Conference*. IEEE.

[Gelfand and Yaglom, 1993] Gelfand, I. and Yaglom, A. (1993). Amount of information and entropy for continuous distributions. In Shiryayev, A., editor, *Selected Works of A. N. Kolmogorov*, volume 27 of *Mathematics and Its Applications*, pages 33–56. Springer Netherlands.

[GeRT, 2011] GeRT (2011). Project gert - generalizing robot manipulation tasks. URL: `http://www.gert-project.eu`.

[Gibson, 1962] Gibson, J. J. (1962). Observations on active touch. *Psychol Rev*, 69:477–491.

[Goldstein, 2002] Goldstein, E. B. (2002). *Sensation and Perception*. Wadsworth, Belmont, CA.

[Gomes et al., 2013] Gomes, R. B., de Carvalho, B. M., and Gonalves, L. M. G. (2013). Visual attention guided features selection with foveated images. *Neurocomputing*, 120:34 – 44. Image Feature Detection and Description.

[Goodale and Milner, 1992] Goodale, M. A. and Milner, A. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, 15(1):20 – 25.

[GRASP, 2008] GRASP (2008). project grasp - emergence of cognitive grasping through introspection, emulation and surprise. URL: `http://www.csc.kth.se/grasp/`.

[Guizzo and Ackerman, 2015] Guizzo, E. and Ackerman, E. (2015). The hard lessons of darpa's robotics challenge [news]. *Spectrum, IEEE*, 52(8):11–13.

[HANDLE, 2009] HANDLE (2009). project handle - developmental pathway towards autonomy and dexterity in robot in-hand manipulation. URL: `http://www.handle-project.eu/`.

[HANDLE-SHADOW, 2009] HANDLE-SHADOW (2009). Instrumented rubik cube - augmented sensing object. *HANDLE FP7 project: Developmental pathway towards autonomy and dexterity in robot in-hand manipulation*. URL: `http://www.rmartins.net/phd-docs/irc/`.

[HANDLE-SHADOW, 2010] HANDLE-SHADOW (2010). The can sensing object. *HANDLE FP7 project: Developmental pathway towards autonomy and dexterity in robot in-hand manipulation*. URL:`http://www.rmartins.net/phd-docs/isc/`.

[HANDLE-UC, 2009] HANDLE-UC (2009). Protocol for the corpus of sensed grasp and handling data: storage of multi-modal datasets. *HANDLE FP7 project: Developmental pathway towards autonomy and dexterity in robot in-hand manipulation*. `http://www.rmartins.net/phd-docs/tr02/`.

[Harmon, 1982] Harmon, L. D. (1982). Automated tactile sensing. *The International Journal of Robotics Research*, 1(2):3–32.

[Hendrich et al., 2010] Hendrich, N., Klimentjew, D., and Zhang, J. (2010). Multi-sensor based segmentation of human manipulation tasks. In *Multisensor Fusion and Integration for Intelligent Systems (MFI), 2010 IEEE Conference on*, pages 223–229.

[Hendrich et al., 2012] Hendrich, N., Klimentjew, D., and Zhang, J. (2012). 5th international conference on cognitive systems, cogsys2012, vienna.

[Hertzberg and Chatila, 2008] Hertzberg, J. and Chatila, R. (2008). Ai reasoning methods for robotics. In Siciliano, B. and Khatib, O., editors, *Springer Handbook of Robotics*, pages 207–223. Springer Berlin Heidelberg.

[Herzog et al., 2013] Herzog, A., Pastor, P., Righetti, L., Bohg, J., Asfour, T., and Schaal, S. (2013). Learning of grasp selection based on shape-templates.

[Hongbin Liu, 2011] Hongbin Liu, Xiaojing Song, T. N. K. A. L. S. (2011). Friction estimation based object surface classification for intelligent manipulation. In *Proceddings of IEEE International Conference on Robotics and Automation 2011 (ICRA 2011) - Workshop on Autonomous Grasping*.

[Hui and Kuchenbecker, 2014] Hui, J. C. T. and Kuchenbecker, K. J. (2014). *Evaluating the BioTac's Ability to Detect and Characterize Lumps in Simulated Tissue*, pages 295–302. Springer Berlin Heidelberg, Berlin, Heidelberg.

[Humphry et al., 1995] Humphry, R., Jewell, K., and Rosenberger, R. C. (1995). Development of in-hand manipulation and relationship with activities. *American Journal of Occupational Therapy*, 49(8):763–771.

[Jansen et al., 2013] Jansen, S. E., Tiest, W. M. B., and Kappers, A. M. L. (2013). Identifying haptic exploratory procedures by analyzing hand dynamics and contact force. *IEEE Transactions on Haptics*, 99:1.

[Johansson and Flanagan, 2009] Johansson, R. S. and Flanagan, J. R. (2009). Coding and use of tactile signals from the fingertips in object manipulation tasks. *Nat Rev Neurosci*, 10:345–359.

[Jones and Forster, 2014] Jones, A. and Forster, B. (2014). Neural correlates of endogenous attention, exogenous attention and inhibition of return in touch. *European Journal of Neuroscience*, 40(2):2389–2398.

[Jones, 2011] Jones, J. A. (2011). *Neural Correlates of Tactile Attention: Behavioural measures and event-related brain potentials of inhibition of return, exogenous and endogenous attention in touch.* PhD thesis, City University London.

[Jones and Lederman, 2006] Jones, L. A. and Lederman, S. J. (2006). *Human Hand Function.* Oxford University Press, USA, 1 edition.

[Kalsi-Ryan et al., 2012] Kalsi-Ryan, S., Curt, A., Verrier, M. C., and Fehlings, M. G. (2012). Development of the graded redefined assessment of strength, sensibility and prehension (grassp): reviewing measurement specific to the upper limb in tetraplegia. *Journal of Neurosurgery: Spine*, 17(Suppl1):65–76. PMID: 22985372.

[Kamakura et al., 1980] Kamakura, N., Ohmura, M., Ishii, H., Mitsubosi, F., and Miura, Y. (1980). Patterns of static prehension in normal hands. *The American Journal of Occupational Therapy*, 34(7):437–445.

[Klatzky et al., 1985] Klatzky, R. L., Lederman, S. J., and Metzger, V. A. (1985). Identifying objects by touch: An expert system. *Perception & Psychophysics*, 37(4):299–302.

[Klein, 2000] Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Sciences*, 4(4):138 – 147.

[Knill and Richards, 1996] Knill, D. C. and Richards, W., editors (1996). *Perception as Bayesian inference.* Cambridge University Press, New York, NY, USA.

[Kondo et al., 2011] Kondo, M., Ueda, J., and Ogasawara, T. (2011). Recognition of in-hand manipulation by observing contact state transition for robot hand control. In Kaneko, M. and Nakamura, Y., editors, *Robotics Research*, volume 66 of *Springer Tracts in Advanced Robotics*, pages 349–360. Springer Berlin / Heidelberg.

[Kruger et al., 2010] Kruger, V., Herzog, D., Baby, S., Ude, A., and Kragic, D. (2010). Learning actions from observations. *Robotics Automation Magazine, IEEE*, 17(2):30 –43.

[Lacey and Sathian, 2014] Lacey, S. and Sathian, K. (2014). Visuo-haptic multisensory object recognition, categorization, and representation. *Frontiers in psychology*, 5.

[Lebeltel et al., 2004] Lebeltel, O., Bessire, P., Diard, J., and Mazer, E. (2004). Bayesian robot programming. *Autonomous Robots*, 16(1):49–79.

[Lederman, 1994] Lederman, S. J. (1994). The intelligent hand: An experimental approach to human object recognition and implications for robotics and ai. *AI Magazine*, 15:774–785.

[Lederman and Klatzky, 1987] Lederman, S. J. and Klatzky, R. L. (1987). Hand movements: A window into haptic object recognition. *Cognitive Psychology*, 19(3):342 – 368.

[Lee, 2008] Lee, J. (2008). Hacking the nintendo wii remote. *Pervasive Computing, IEEE*, 7(3):39 –45.

[Lee and Kunii, 1995] Lee, J. and Kunii, T. (1995). Model-based analysis of hand posture. *Computer Graphics and Applications, IEEE*, 15(5):77 –86.

[Li et al., 2010] Li, K., Chen, I.-M., and Yeo, S. H. (2010). Design and validation of a multi-finger sensing device based on optical linear encoder. In *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, pages 3629 –3634.

[Li et al., 2013] Li, Q., Schrmann, C., Haschke, R., and Ritter, H. (2013). A control framework for tactile servoing. In *Proceedings of RSS*.

[Liu et al., 2010a] Liu, H., Noonan, D., Challacombe, B., Dasgupta, P., Seneviratne, L., and Althoefer, K. (2010a). Rolling mechanical imaging for tissue abnormality localization during minimally invasive surgery. *Biomedical Engineering, IEEE Transactions on*, 57(2):404 –414.

[Liu et al., 2010b] Liu, H., Noonan, D. P., Challacombe, B. J., Dasgupta, P., Seneviratne, L. D., and Althoefer, K. (2010b). Rolling mechanical imaging for tissue abnormality localization during minimally invasive surgery. *Biomedical Engineering, IEEE Transactions on*, 57(2):404–414.

[Liu et al., 2012] Liu, H., Song, X., Bimbo, J., Seneviratne, L., and Althoefer, K. (2012). Surface material recognition through haptic exploration using an intelligent contact sensing finger. In *IROS 2012*, pages 52–57. IEEE.

[Lobo et al., 2011] Lobo, J., Trindade, P., and Dias, J. (2011). Observing hand grasp type and contact points using hand distributed accelerometers and instrumented objects. In *Proc. of IEEE/ICRA 2011 Workshop on Autonomous Grasping, Shanghai, China*.

[Lucarotti et al., 2013] Lucarotti, C., Oddo, C. M., Vitiello, N., and Carrozza, M. C. (2013). Synthetic and bio-artificial tactile sensing: A review. *Sensors*, 13(2):1435–1466.

[Magill and Anderson, 2007] Magill, R. A. and Anderson, D. (2007). *Motor learning and control: Concepts and applications*, volume 11. McGraw-Hill New York.

[Martinez et al., 2013] Martinez, U., Dodd, T., Natale, L., Metta, G., Prescott, T., and Lepora, N. (2013). Active contour following to explore object shape with robot touch. In *World Haptics 2013*, pages 341–346.

[Martins, 2008] Martins, R. (2008). Modelling the human body and hand: kinematic structure, degrees-of-freedom. *PROMETHEUS FP7 project: PRediction and interpretatiOn of huMan behaviour based on probabilistic sTructures and HEterogeneoUs Sensors*. `http://www.rmartins.net/phd-docs/tr01/`.

[Martins, 2009a] Martins, R. (2009a). 3d interactive demonstrator of human body and hand: kinematic structure, degrees-of-freedom. *PROMETHEUS FP7 project: PRediction and interpretatiOn of huMan behaviour based on probabilistic sTructures and HEterogeneoUs Sensor*. URL: `http://www.rmartins.net/phd-docs/st01/`.

[Martins, 2009b] Martins, R. (2009b). Distributed data acquisition architecture: software client for cyberglove (data glove). *HANDLE FP7 project: Developmental pathway towards autonomy and dexterity in robot in-hand manipulation*. URL: `http://www.rmartins.net/phd-docs/st03/`.

[Martins, 2009c] Martins, R. (2009c). Distributed data acquisition architecture: software client for instrumented rubik cube (instrumented object). *HANDLE FP7 project: Developmental pathway towards autonomy and dexterity in robot in-hand manipulation*. URL: `http://www.rmartins.net/phd-docs/st04/`.

[Martins, 2009d] Martins, R. (2009d). Distributed data acquisition architecture: software client for instrumented sensing can (instrumented object). *HANDLE FP7 project:*

Developmental pathway towards autonomy and dexterity in robot in-hand manipulation. URL: http://www.rmartins.net/phd-docs/st05/.

[Martins, 2009e] Martins, R. (2009e). Distributed data acquisition architecture: software client for polhemus liberty system (6d motion tracking). *HANDLE FP7 project: Developmental pathway towards autonomy and dexterity in robot in-hand manipulation.* URL: http://www.rmartins.net/phd-docs/st07/.

[Martins, 2009f] Martins, R. (2009f). Distributed data acquisition architecture: software client for tekscan grip system (tactile sensing array). *HANDLE FP7 project: Developmental pathway towards autonomy and dexterity in robot in-hand manipulation.* URL: http://www.rmartins.net/phd-docs/st06/.

[Martins, 2009g] Martins, R. (2009g). importdatasettb: a matlab toolbox to integrate multi-modal datasets. *HANDLE FP7 project: Developmental pathway towards autonomy and dexterity in robot in-hand manipulation.* URL: http://www.rmartins.net/phd-docs/st02/.

[Martins, 2010] Martins, R. (2010). Distributed synchronization of multi-modal data acquisition devices using ntp (network time protocol). *HANDLE FP7 project: Developmental pathway towards autonomy and dexterity in robot in-hand manipulation.* URL: http://www.rmartins.net/phd-docs/tr03/.

[Martins, 2012a] Martins, R. (2012a). 3d visualization tool of instrumented rubik cube touch data. *HANDLE FP7 project: Developmental pathway towards autonomy and dexterity in robot in-hand manipulation.* http://www.rmartins.net/phd-docs/st08/.

[Martins, 2012b] Martins, R. (2012b). Experimental evaluation and calibration protocol of tekscan grip system. *HANDLE FP7 project: Developmental pathway towards autonomy and dexterity in robot in-hand manipulation.* URL: http://www.rmartins.net/phd-docs/tr05/.

[Martins, 2012c] Martins, R. (2012c). Matlab annotation tool for the dataset "dexterous manipulation of a laboratory pipette". *HANDLE FP7 project: Developmental pathway towards autonomy and dexterity in robot in-hand manipulation.* URL:http://www.rmartins.net/phd-docs/st09/.

[Martins, 2013] Martins, R. (2013). Installing controller area network (can-bus) drivers and compiling code on ubuntu. *HANDLE FP7 project: Developmental pathway towards autonomy and dexterity in robot in-hand manipulation.* URL: http://www.rmartins.net/phd-docs/tr04/.

[Martins et al., 2012a] Martins, R., Faria, D., and Dias, J. (2012a). Representation framework of perceived object softness characteristics for active robotic hand exploration. In *HRI2012*, USA.

[Martins et al., 2010] Martins, R., Faria, D. R., and Dias, J. (2010). Symbolic level generalization of in-hand manipulation tasks from human demonstrations using tactile data information. In *Workshop on Grasping Planning and Task Learning by Imitation, I2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) Taipei, Taiwan (October 2010)*.

[Martins et al., 2017] Martins, R., Ferreira, J. F., Castelo-Branco, M., and Dias, J. (2017). Integration of touch attention mechanisms to improve the robotic haptic exploration of surfaces. *Neurocomputing*, 222:204 – 216.

[Martins et al., 2012b] Martins, R., Ferreira, J. F., and Dias, J. (2012b). Touch attention bayesian models for object feature extraction in robotic blind manipulation. In *MaxEnt 2012 - Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*. Munich.

[Martins et al., 2013] Martins, R., Ferreira, J. F., and Dias, J. (2013). Touch attention bayesian models for robotic active haptic exploration. In *2nd Workshop on Recognition and Action for Scene Understanding (REACTS)*. UK.

[Martins et al., 2014] Martins, R., Ferreira, J. F., and Dias, J. (2014). Touch attention bayesian models for robotic active haptic exploration of heterogeneous surfaces. In *Proceedings of 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2014)*, pages 1208–1215, Chicago, USA. IEEE.

[Matheus and Dollar, 2010] Matheus, K. and Dollar, A. (2010). Benchmarking grasping and manipulation: Properties of the objects of daily living. In *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pages 5020–5027.

[Matsuo et al., 2009] Matsuo, K., Murakami, K., Hasegawa, T., Tahara, K., and Kurazume, R. (2009). Segmentation method of human manipulation task based on measurement of force imposed by a human hand on a grasped object. In *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*, pages 1767–1772. IEEE.

[Meng et al., 2016] Meng, F., Guo, B., Song, M., and Zhang, X. (2016). Image fusion with saliency map and interest points. *Neurocomputing*, 177:1 – 8.

[Microft-Kinect, 2009] Microft-Kinect (2009). Microsoft kinect - controller-free gaming and entertainment experience. URL: `https://dev.windows.com/en-us/kinect`.

[Miles Breslin and Exner, 1999] Miles Breslin, D. M. and Exner, C. E. (1999). Construct validity of the in-hand manipulation test: A discriminant analysis with children without disability and children with spastic diplegia. *American Journal of Occupational Therapy*, 53(4):381–386.

[Mller and Giabbiconi, 2008] Mller, M. M. and Giabbiconi, C.-M. (2008). Attention in sense of touch. In *Human Haptic Perception*. Birkhuser.

[Napier, 1956] Napier, J. R. (1956). The prehensile movements of the human hand. *J Bone Joint Surg Br*, 38-B(4):902–913.

[Ni et al., 2011] Ni, D., Chan, W. Y., Qin, J., Chui, Y.-P., Qu, I., Ho, S., and Heng, P.-A. (2011). A virtual reality simulator for ultrasound-guided biopsy training. *Computer Graphics and Applications, IEEE*, 31(2):36 –48.

[Nicholls and Lee, 1989] Nicholls, H. R. and Lee, M. H. (1989). A survey of robot tactile sensing technology. *The International Journal of Robotics Research*, 8(3):3–30.

[Nintendo, 2006] Nintendo (2006). Nintendo wii remote - the primary controller for nintendo wii console. URL: `http://www.nintendo.com/wii/console/controllers`.

[Nof, 1999] Nof, S. Y. (1999). *Handbook of Industrial Robotics*. John Wiley & Sons, Inc., New York, NY, USA, 2nd edition.

[Oddo et al., 2011] Oddo, C. M., Controzzi, M., Beccai, L., Cipriani, C., and Carrozza, M. C. (2011). Roughness encoding for discrimination of surfaces in artificial active-touch. *IEEE Trans. Robotics*, 27(3):522–533.

[Okamura et al., 2001] Okamura, A. M., Cutkosky, M. R., and I. (2001). Feature detection for haptic exploration with robotic fingers. *The International Journal of Robotics Research*, 20(12):925–938.

[Oztop and Arbib, 2002] Oztop, E. and Arbib, M. A. (2002). Schema design and implementation of the grasp-related mirror neuron system. *Biological Cybernetics*, 87:116–140. 10.1007/s00422-002-0318-1.

[Pape et al., 2012] Pape, L., Oddo, C. M., Controzzi, M., Cipriani, C., Frster, A., Carrozza, M. C., and Schmidhuber, J. (2012). Learning tactile skills through curious exploration. *Frontiers in Neurorobotics*, 6(6).

[Petersen and Posner, 2012] Petersen, S. E. and Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annual Review of Neuroscience*, 35(1):73–89. PMID: 22524787.

[Polhemus-Liberty, 2008] Polhemus-Liberty (2008). Polhemus liberty - electromagnetic motion 6d tracker. URL: `http://www.polhemus.com/?page=Motion_Liberty`.

[Pont et al., 2009] Pont, K., Wallen, M., and Bundy, A. (2009). Conceptualising a modified system for classification of in-hand manipulation. *Australian Occupational Therapy Journal*, 56(1):2–15.

[Pont et al., 2008] Pont, K., Wallen, M., Bundy, A., and Case-Smith, J. (2008). Reliability and validity of the test of in-hand manipulation in children ages 5 to 6 years. *American Journal of Occupational Therapy*, 62(4):384–392.

[Posner and Cohen, 1984] Posner, M. I. and Cohen, Y. (1984). Components of visual orienting. *Attention and performance X: Control of language processes*, 32:531–556.

[Rocha et al., 2005] Rocha, R., Dias, J., and Carvalho, A. (2005). Cooperative multi-robot systems: A study of vision-based 3-d mapping using information theory. *Robotics and Autonomous Systems*, 53(3-4):282–311.

[Romano et al., 2011] Romano, J., Hsiao, K., Niemeyer, G., Chitta, S., and Kuchenbecker, K. (2011). Human-inspired robotic grasp control with tactile sensing. *Robotics, IEEE Transactions on*, 27(6):1067–1079.

[Saal and Bensmaia, 2014] Saal, H. P. and Bensmaia, S. J. (2014). Touch is a team effort: interplay of submodalities in cutaneous sensibility. *Trends in Neurosciences*, 37(12):689 – 697.

[Saudabayev and Varol, 2015] Saudabayev, A. and Varol, H. (2015). Sensors for robotic hands: A survey of state of the art. *Access, IEEE*, 3:1765–1782.

[Scilingo, 2010] Scilingo, E.P.; Bianchi, M. G. G. B. A. (2010). Rendering softness- integration of kinesthetic and cutaneous information in a haptic device. *IEEE Transactions on Haptics*, 3:109 – 118.

[Shannon, 2001a] Shannon, C. E. (2001a). A mathematical theory of communication. *SIGMOBILE Mob. Comput. Commun. Rev.*, 5(1):3–55.

[Shannon, 2001b] Shannon, C. E. (2001b). A mathematical theory of communication. *SIGMOBILE Mob. Comput. Commun. Rev.*, 5(1):3–55.

[Shipp, 2004] Shipp, S. (2004). The brain circuitry of attention. *Trends in Cognitive Sciences*, 8(5):223 – 230.

[Smisek et al., 2013] Smisek, J., Jancosek, M., and Pajdla, T. (2013). 3d with kinect. In Fossati, A., Gall, J., Grabner, H., Ren, X., and Konolige, K., editors, *Consumer Depth Cameras for Computer Vision*, Advances in Computer Vision and Pattern Recognition, pages 3–25. Springer London.

[Srinivasan and LaMotte, 1995] Srinivasan, M. A. and LaMotte, R. H. (1995). Tactual discrimination of softness: Abilities and mechanisms. *J Neurophysiol*, 73:88–101.

[Stone and Gonzalez, 2015] Stone, K. D. and Gonzalez, C. L. R. (2015). The contributions of vision and haptics to reaching and grasping. *Frontiers in Psychology*, 6(1403).

[Tekscan-Grip, 2010] Tekscan-Grip (2010). Tekscan-grip system - tactile grip force and pressure sensing. URL: `http://www.tekscan.com/grip-pressure-analysis-system`.

[THE, 2010] THE (2010). project the - the hand embodied. URL: `http://www.thehandembodied.eu/`.

[Thrun et al., 2005] Thrun, S., Burgard, W., and Fox, D. (2005). *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents)*. The MIT Press.

[Tiest, 2010] Tiest, W. M. B. (2010). Tactual perception of material properties. *Vision Research*, 50(24):2775 – 2782. Perception and Action: Part I.

[Tso and Liu, 1996] Tso, S. and Liu, K. (1996). Hidden markov model for intelligent extraction of robot trajectory command from demonstrated trajectories. In *Industrial Technology, 1996. (ICIT '96), Proceedings of The IEEE International Conference on*, pages 294 –298.

[Unibrain, 2007] Unibrain (2007). Unibrain fire-i digital camera. URL: `http://www.unibrain.com/products/fire-i-digital-camera/`.

[van Kuilenburg et al., 2013] van Kuilenburg, J., Masen, M. A., and van der Heide, E. (2013). A review of fingerpad contact mechanics and friction and how this affects tactile perception. *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology*.

[Vicente, 2007] Vicente, I. S. (2007). Action recognition and understanding through motor primitives. *Advanced Robotics*, 21(15):1687–1707.

[Videre, 2006] Videre (2006). Videre sth-md cs stereo camera. URL: `http://www.rmartins.net/phd-docs/vsc`.

[Wacker, 2011] Wacker, E. (2011). *Tactile Feature Processing and Attentional Modulation in the Human Somatosensory System*. PhD thesis, TU Berlin.

[Westheimer, 2008] Westheimer, G. (2008). Was helmholtz a bayesian? *Perception*, 37(5):642–650.

[Wikipedia, 2014] Wikipedia (2014). Human hand - bones of the human hand. URL: `http://en.wikipedia.org/wiki/Hand`.

[Wurman and Romano, 2015] Wurman, P. and Romano, J. (2015). The amazon picking challenge 2015 [competitions]. *Robotics Automation Magazine, IEEE*, 22(3):10–12.

[Xu et al., 2013] Xu, D., Loeb, G. E., and Fishel, J. A. (2013). Tactile identification of objects using bayesian exploration. In *ICRA 2013*.

[Yancosek and Howell, 2009] Yancosek, K. E. and Howell, D. (2009). A narrative review of dexterity assessments. *Journal of Hand Therapy*, 22(3):258 – 270.

[Yousef et al., 2011] Yousef, H., Boukallel, M., and Althoefer, K. (2011). Tactile sensing for dexterous in-hand manipulation in robotics - a review. *Sensors and Actuators A*, 167(2):171 – 187.

[Yozbatiran et al., 2008] Yozbatiran, N., Der-Yeghiaian, L., and Cramer, S. C. (2008). A standardized approach to performing the action research arm test. *Neurorehabilitation and Neural Repair*, 22(1):78–90.

[Zhang, 1994] Zhang, Z. (1994). Iterative point matching for registration of free-form curves and surfaces. *International journal of computer vision*.

[Zhang, 2012] Zhang, Z. (2012). Microsoft kinect sensor and its effect. *MultiMedia, IEEE*, 19(2):4–10.

[Zollo et al., 2013] Zollo, L., Wada, K., and Van der Loos, H. (2013). Special issue on assistive robotics [from the guest editors]. *Robotics Automation Magazine, IEEE*, 20(1):16–19.