Title
Tactile Perception, Haptic Exploration, and Map Rendering for Robots that Operate within Granular Materials

Author
Shengxin, Jia

Publication Date
2022

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Tactile Perception, Haptic Exploration, and Map Rendering for Robots that Operate within Granular Materials

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Mechanical Engineering

by

Shengxin Jia

2022

ABSTRACT OF THE DISSERTATION

Tactile Perception, Haptic Exploration, and Map Rendering
for Robots that Operate within Granular Materials

by

Shengxin Jia
Doctor of Philosophy in Mechanical Engineering
University of California, Los Angeles, 2022
Professor Veronica Santos, Chair

Robots are expected to operate autonomously in unstructured, real-world environments. For effective physical interaction with the world, robots must build and refine their understanding of the environment through sensory feedback. However, tactile feedback has been used primarily in open-air environments and not within granular materials. When robots operate within opaque granular materials, tactile and proprioceptive feedback can be more informative than visual feedback. Our long-term objective is to leverage tactile sensors to develop efficient algorithms that enable robots to infer environmental conditions and to plan exploratory movements that reduce uncertainty in their models of the world. Motivated by the need to keep humans out of harm's way in search and rescue or other field environments, we address the challenge of using tactile feedback to locate objects buried in granular materials.

In study #1, we designed a tactile perception pipeline for sensorized robot fingertips

that directly interact with granular materials in teleoperated systems. We proposed an architecture called the Sparse-Fusion Recurrent Neural Network (SF-RNN) to detect contact with an object buried within granular materials. We leveraged multimodal tactile sensor data in order to classify contact states within five different granular materials. We also constructed a belief map that combines probabilistic contact state estimates and fingertip location.

In study #2, we developed a framework for tactile perception, mapping, and haptic exploration for the autonomous localization of objects buried in granular materials. The haptic exploration task was performed within densely packed sand mixtures using sensor models that account for granular material characteristics and aid in the interpretation of interaction forces between the robot and its environment. The haptic exploration strategy was designed to efficiently locate and refine the outline of a buried object while simultaneously minimizing potentially damaging physical interactions with the object. Continuous occupancy maps were generated that fused local, sparse tactile information into global maps.

In summary, we developed tactile-based frameworks for perception, planning, and mapping for the challenging task of localizing objects buried within granular materials. Our work can serve as a foundation for more complex, autonomous robotic behaviors such as the excavation and bimanual retrieval of fragile, buried objects.

The dissertation of Shengxin Jia is approved.

Song-Chun Zhu

Dennis W. Hong

Tetsuya Iwasaki

Veronica Santos, Committee Chair

University of California, Los Angeles

2022

To mom, dad, and my treasured grandma . . .

for their endless love and supporting, and let me going at my own pace.

TABLE OF CONTENTS

1 Introduction **1**

1.1 Motivation 1

1.2 Background 2

1.2.1 Granular Material Properties 2

1.2.2 Prior Robotics Research with Granular Materials 5

1.3 Contributions 6

2 Tactile Perception for Teleoperated Robotic Exploration within Granular Media **8**

2.1 Abstract 8

2.2 Introduction 9

2.3 Related Work 12

2.3.1 Complex Behaviors of Granular Media 12

2.3.2 Robot Interaction with Granular Media 13

2.3.3 Feedback to Human Teleoperator 14

2.4 Multimodal Model for Tactile Perception 15

2.4.1 Pre-processing of Tactile Data 18

2.4.2 RNN with GRUs for Binary Classification 21

2.5 Belief Map Representation of Tactile Information 22

2.6 Experimental Procedure: Robot Data Collection 24

2.6.1 Robot Testbed 24

2.6.2 Dataset for Model Training 26

2.6.3 Dataset for Model Testing 27

2.7 Experimental Procedure: Model Evaluation 28

2.7.1 Study of Model Architecture 29

2.7.2 Study of Model Robustness 36

2.7.3 Belief Map Representation of Tactile Information 37

2.8 Discussion 38

2.8.1 Notable Trends in the Tactile Data 38

2.8.2 Real-Time Tactile Perception and Visualization for Human Control of Remote Robotic Systems 45

2.9 Conclusion 47

2.9.1 Summary of Contributions 47

2.9.2 Limitations and Future Work 48

3 Tactile SLAM for Autonomous Haptic Localization and Mapping of Objects Buried in Granular Materials 50

3.1 Abstract 50

3.2 Introduction 51

3.3 Related Work 54

3.3.1 Soil-tool Interaction: Plate Drag 54

3.3.2 Tactile SLAM 55

3.4 Framework for Tactile SLAM within Granular Materials 57
3.4.1 Sensor Measurements and Assumptions 57
3.4.2 Tactile Perception 60
3.4.3 Continuous Occupancy Mapping 61
3.4.4 Haptic Exploration Strategy 66
3.5 Experimental Procedure and Evaluation 69
3.5.1 Granular Material Model 69
3.5.2 Haptic Exploration in Simulation 70
3.5.3 Haptic Exploration with a Real Robot 73
3.6 Conclusion 78

4 Summary and Conclusion 79
4.1 Contributions 79
4.2 Future Work 81
4.2.1 Enhance the Generalizability and Robustness of the Sensor Model . . 81
4.2.2 Extend the Continuous Occupancy Map 82
4.2.3 Employ Adaptive Behaviors for Haptic Exploration within Granular Materials 83

A Multimodal Tactile Perception Model 84
A.1 Empirical Mode Decomposition 84
A.2 Sparse Coding 85

A.3 RNN with GRUs 86

A.4 Belief Map Representation 87

A.5 Test of SF-RNN Model Performance 89

References **90**

LIST OF FIGURES

2.1 A block diagram is shown for a notional real-time execution of the action, perception, and visualization processes. 11

2.2 Sparse-Fusion RNN (SF-RNN) model architecture for multimodal tactile perception in granular media. (a) Vibration, internal fluid pressure, and 19 electrodes' impedance data were collected from BioTac sensors. (b) Vibration and internal fluid pressure data were filtered by EMD. (c) Feature tensors were composed for each modality. (d) Sparse coding was used as an unsupervised pre-training step. (e) An RNN mapped a 400D input space to a 2D output space. (f) The RNN sequentially labeled tactile inputs according to the estimated fingertip contact state. 16

2.3 (a) Raw vibration data and (b-i) squared deviations of its first 8 $IMFs$. Each deviation was calculated as *norm[IMF - mean(IMF)]*. 19

2.4 (a) Experimental set-up used to collect multimodal tactile sensor training data by sweeping sensorized fingertips through granular media (cottonwood beads) and across a buried cylinder. (b) Blue arrows indicate the linear trajectory of two fingertips through granular media (not shown for clarity). (c) In the presence of a buried cylinder, one fingertip was swept briefly across the cylinder (finger-object contact is indicated in red). (d) For a robustness study, different granular media types were used (plastic pellets are shown) and a 3D-printed adapter was used to change the orientation of the fingertip to simulate 20° of finger extension relative to the pose in (a-c). 23

2.5 (a) Vibration, (b) internal fluid pressure, (c) electrode impedance, and (d) fingertip contact state estimates from an SF-RNN model are shown for a representative 2 sec trial. "First contact" and "last contact" events were identified from open air "ground truth" trials (screenshots, orange dotted lines). For clarity, impedance values from an open air trial are shown for two electrodes only (min, max traces). 25

2.6 Using tSNE for visualization, one can observe the effects of different stages of the tactile perception pipeline. (a) Visualized in 3D space, the feature tensors from Fig. 2.2c are entangled. (b) The outputs from the sparse coding step in Fig. 2.2d are spread out in a high-dimensional feature space. (c) The feature space from the last layer of the model in Fig. 2.2e shows that the data are mostly separated into distinct regions according to contact state. 35

2.7 As the fingertip moves across a buried object (top row), a belief map (bottom row) is updated using fingertip contact state estimates from the SF-RNN model. Snapshots are shown as the fingertip (a) makes first contact with the buried object, (b) passes across the top of the object, and (c) makes last contact with the object and returns to free movement within the granular media. 39

2.8 (a) Vibration amplitude and (b) internal fluid pressure are shown as median values with 1st and 3rd quartile bands for both fingertip contact states during all trials with and without a buried object. 41

2.9 Electrode impedance is shown as median values with 1st and 3rd quartile intervals. (a) Trends for all 19 electrodes are shown for both fingertip contact states during all trials with and without a buried object for the same time window (theoretical contact period). (b) Trends for three electrodes at the fingertip (#1, 7, 11) are shown for all trials with a buried object for the entire 2 sec period. . . . 44

3.1 Localization and mapping in granular materials: (a) A robot haptically explores a sand mixture and updates an (b) occupancy map in log odds using fingertip forces interpreted by a sensor model that accounts for granular material characteristics. (c) Local, sparse tactile data are fused into a global Bayesian Hilbert Map to localize a buried object and augment (a) colorized point cloud data from a head-mounted sensor. 52

3.2 Visualization of the kernel-based sensor model described by Eq. 3.3 for three environmental interaction states: (a) free movement through granular materials, (b) granular material jamming, and (c) direct contact with a buried object. The sensor model was designed based on the theoretical failure zone. Per Wick and Perumpral [1], measurement range r and side crescent angle η of the failure zone are calculated as $r = d(cot\alpha + cot\beta)$ and $\eta = sin^{-1}(\frac{6.03+0.460r+0.0904\alpha}{r})$, respectively. 56

3.3 Framework for tactile perception, mapping, and haptic exploration for the autonomous localization of objects buried in granular materials. 59

3.4 Environmental interaction states were labeled according to their coordinates in a 2-D space comprised of fingertip force and force-rate. 61

3.5 Simulation results for the haptic localization of randomly placed objects having circular (top row) and elliptical (bottom row) outlines in the x-y plane: Occupancy maps in log odds are shown for Phase 1 (a,f) and Phase 2 (b,g). Optimal kernel shapes and placements were learned (c,h). Final Bayesian Hilbert Maps are shown for occupancy (d,i) and uncertainty (e,j). 67

3.6 Fingertip force for each of 7 EMs: Since fingertip forces fluctuate during interactions with granular materials, dashed and solid lines indicate mean values for free movement through granular materials and during granular material jamming, respectively. Colored bands indicate values that are one standard deviation from the mean. For visual comparison, the onset of granular material jamming is aligned at 0 sec (black dashed line) for EMs #1, 3, 4, 6, and 7. No direct contact with the object was perceived. 76

3.7 Final Bayesian Hilbert Map for a haptic search trial that autonomously ended after 7 EMs (magenta). The last fingertip force for each EM is shown in green. The estimated outline of the buried object for an occupancy level of 0.6 is shown with a black dashed line. 77

LIST OF TABLES

2.1 Granular Media Types: To test model robustness, we selected five distinct granular media types to represent the five aggregate types ranging from coarse sand to coarse gravel according to the Wentworth scale [2]. 28

2.2 Ablation Study with **Unimodal Models**: Model accuracy is reported as mean ± standard deviation in %. Best performing models are shown in bold. The models were trained and evaluated with cottonwood beads. The (b)-(d) annotations refer to the perception pipeline steps shown in Figure 2.2. 31

2.3 Ablation Study with **Multimodal Models**: Model accuracy is reported as mean ± standard deviation in %. Best performing models are shown in bold. The models were trained and evaluated with cottonwood beads. The (b)-(d) annotations refer to the perception pipeline steps shown in Figure 2.2. 32

2.4 Robustness Study: Model accuracy is reported as mean ± standard deviation in %. The models were trained on cottonwood beads, but evaluated on the expanded test dataset collected with different granular media types and finger-object contact conditions. 37

3.1 Haptic Exploration Parameters 71

3.2 Haptic Exploration Simulation Results 73

A.1 Effects of Multimodality on SF-RNN Model Performance: Model accuracy is reported as mean in %. Best performing models are shown in bold. The models were trained and evaluated with cottonwood beads. 89

ACKNOWLEDGMENTS

This work has been accomplished with the support of many inspiring individuals. First, I would like to express my heartfelt gratitude and respect to my advisor, Professor Veronica J. Santos. She has been providing patient guidance and tremendous support over the past years. I am very fortunate to have a graduate advisor who appreciates my potential and strengths. She shows me how to be an independent researcher from scratch and keeps encouraging me all the time. I am also grateful for the advice she gave on perfecting my writing and presentations.

Additionally, I would like to express my sincere gratitude to my committee members Professor Tetsuya Iwasaki, Professor Dennis Hong, and Professor Song-Chun Zhu. I have learned a lot from their inspiring lectures, presentations, and discussions with their graduate students.

For Chapter 2, I would like to thank Professor Chad Jenkins and Professor Rod Grupen for discussions on occupancy maps and belief maps, Dr. Randall Hellman and Dr. Kenneth Gutierrez for insightful discussions and guidance, and Dr. Eunsuk Chong and Lionel Zhang for assistance with editing. For Chapter 3, thanks to Professor Gaurav Sant for assistance with the preparation of the granular materials, Kenneth Hudson, and Professor Scott Brandenberg for assistance with the triaxial shear tests. Special thanks to Lionel Zhang, Huajing Zhao, and Jonathan Bopp for assistance with the real robot experiments.

The material presented in Chapters 2 and 3 were supported in part by the Office of Naval Research Award #N00014-16-1-2468 and Advanced Robotics for Manufacturing Institute Award #ARM-TEC-18-01-F-14. Any opinions, findings, conclusions, or recommendations are those of the authors and do not necessarily represent the official views, opinions, or policies of the funding agencies.

I enjoyed the company and friendship of all the current and past members of the UCLA Biomechatronics Laboratory. We shared happiness and frustration and built trust through project reviews and paper deadlines. A special thanks to Dr. Alexis E. Block for her valuable suggestions in preparing my defense slides and feedback on paper drafts. Thanks to Dr. Eunsuk Chong and Lionel Zhang for helping me assemble the robot hardware and building the control system from scratch.

VITA

2016–Pres.	Graduate Student Researcher, UCLA Biomechatronics Lab, Mechanical and Aerospace Engineering Department, UCLA, Los Angeles, California
2016	M.S. Mechanical Engineering, UCLA, Los Angeles, California
2015	Teaching Assistant, Physics Department, UCLA, Los Angeles, California
2014	Research Assistant, Institute of Automation, Chinese Academy of Science, China
2013	B.S. Mechanical Engineering, Harbin Institute of Technology, Harbin, China
2013	B.S. Mechanical Engineering with Honors, University of Bath, UK

PUBLICATIONS

Jia, S. and Santos, V.J. "Tactile Perception for Teleoperated Robotic Exploration within Granular Media," *ACM Transactions on Human-Robot Interaction,* vol. 10, no. 4, pp. 1-27, 2021.

Jia, S. and Santos, V.J. "Multimodal Haptic Perception within Granular Media via Recurrent Neural Networks," *Proceedings of Robotics: Science and Systems Workshop on "Tactile Sensing for Manipulation: Hardware, Modeling, and Learning,"* Boston, MA, USA, July 15, 2017.

CHAPTER 1

Introduction

1.1 Motivation

Tactile perception within granular materials is an important ability. The task of reaching into granular materials and searching for an object is an essential human survival skill for many applications such as harvesting root vegetables and clam digging below the surface of the tidal sand flats on beaches. Tactile perception is especially necessary for tasks that involve fine perception and manipulation of fragile and/or high consequence materials such as in archaeological excavation or manually searching for dangerous buried devices. However, the use of sensorized robot hands within granular material has not been extensively explored. Typically, experiments with sensorized robot hands are performed in uncluttered environments in open air, where contact and no contact conditions are easily distinguished from tactile sensor data.

This work is motivated by the need to advance capabilities for robots that are expected to locate and interact with dangerous objects in dusty, sandy environments. To keep humans out of harm's way, it is desirable to enable robots and their remote operators to haptically locate, excavate, retrieve, and manipulate dangerous devices that are buried within soil. Contact with unknown objects should be handled carefully. However, state-of-the-art robotic hands do not typically include tactile sensors and cannot match the tactile or manipulative abilities of human hands.

As highly capable tools, human hands can be used to distinguish between objects having different sizes, shapes, materials, and textures, thanks in part to an exquisite multimodal sensory system, the processing power of the central nervous system, and learning through unnumerable novel experiences. When maneuvering within granular materials, robot systems will require haptic intelligence in order to carefully move and interact with potentially fragile buried objects. While the field of computer vision has progressed much more rapidly in its technology and algorithm development, haptics has played a smaller role in the feedback control of robotic manipulators. By formulating models and leveraging machine learning techniques, multimodal tactile perception can be used to advance the ability of robots to perceive, understand, and control physical interactions with their environment.

1.2 Background

1.2.1 Granular Material Properties

Granular materials, such as snow, sand, grains, salts, and powders, are ubiquitous in nature and in industrial processes, and yet their complex behaviors are in many ways different from other substances. At non-extreme temperatures, temperature plays no role in granular material behavior, and the interactions between grains create a dissipative dynamical system because of static friction and the inelasticity of collisions. It is inappropriate to classify their properties as solid-like or liquid-like, as granular materials cut across these boundaries [3].

When in a resting state, the forces within a pile of granular materials appear to be heterogeneous and are conducted along "force chains" in which stress is largely transmitted by relatively rigid chains of particles forming a sparse network of larger contact forces [4].

Physicists modeled force chains by modeling each particle or grain as a mass placed on a lattice that rests its weight on the layer below unevenly and randomly [3]. When under external stress, force chains carry most of the external load due to complex interparticle forces and the patterns of pressure applied to the granular material.

A fundamental feature of granular material is its packing fraction (or density) ϕ, defined as the ratio of grains to the total volume occupied by the packed particles [5]. The packing fraction for randomly packed uniform spheres is typically in the 0.56 - 0.64 range [6]. The geometrical packing determines the rigidity and flow characteristics of granular materials. For example, dilatancy is an important property of granular materials where the overall volume of packed grains expands in order to flow or deform. This phenomenon is induced by shear stresses that force crowded, packed particles to spread apart [6]. If expansion is constrained, a hydrostatic stress develops on the constraining walls as the particles are sheared. The onset of dilatancy occurs when expansion is required to allow internal shear. For low packing fractions, the dilatancy effect is lessened.

Granular materials can flow like liquids when energy is added to the system, as through external shear forces. When granular materials flow, the velocity of each particle is a function of the average particle velocity in the direction of the current and a velocity due to random collisions with other particles [7]. Besides, unlike the classical assumption from kinetic gas theory, particle collisions are generally inelastic, which leads to clustering since two particles that collide inelastically will remain close to one another once energy has dissipated.

The phenomenon of granular material "jamming" is of particular interest for tactile perception and haptic exploration within granular materials. Subject to a mechanical perturbation, granular materials display a critical slowdown in dynamics, from a fluid state to a glassy state [8]. Jamming occurs when there is an increase packing density, which causes the granular material to behave in a stiff, solid-like manner. To distinguish jamming from its flowing counterpart, physicists have quantified the inhomogeneous distribution of stresses during jamming by measuring the distribution of interparticle normal forces $P(F)$ in experiments and simulations. The distribution $P(F)$ develops a plateau or small peak near transitions between jammed and non-jammed states that is hypothesized to induce yield stresses. The height of this peak increases with greater packing density and decreases with greater temperature and/or shear stress [9]. The time-averaged distribution of jammed granular material decays exponentially at large forces. By contrast, the shape of $P(F)$ in flowing systems, where shear stresses exceed yield stress, is described well by an equilibrium hertzian model at effective temperature [10].

Specific to our application, the flow and force dynamics of granular materials in the vicinity of sensorized robot fingertips are of greatest interest. The rigid robot fingertip could be treated as a flat plate. The raking movement of the robot fingertip through granular materials induces plate drag and results in a wedge-shaped region or "failure zone" ahead of the fingertip. Materials outside of the wedge-shaped flow region remain nearly at rest. Physical interactions between dry granular materials and plates, have been studied in simulation [11] and on physical testbeds over the past several decades [1, 12].

Gravish et al. reported that packing density greatly influences the behavior of granular materials around a plate "intruder" (a robot finger in our application) [13]. In general, granular materials tend to reach a critical packing fraction ϕ_c that can be used to divide granular materials into three classes of behaviors. Loose materials will compact to reach a "critical state," dense materials will dilate to reach a critical state, and materials already at the critical state will remain at a constant volume [14, 15]. When an intruder forces its way into granular materials, mean drag forces increase approximately linearly with ϕ. As ϕ

approaches ϕ_c, the packing fraction at a critical state, the onset of shear dilatancy induces flow and a bifurcation of drag forces. When ϕ exceeds ϕ_c, fluctuations in drag force grow as ϕ increases and the flowing region becomes increasingly stationary. While at low ϕ, as for loosely packed materials, drag forces fluctuate independently of ϕ, resulting in smooth surface deformation and intermittent flow. Gravish et al. reported that responses similar to those obtained using homogeneous glass beads were observed for other granular materials including poppy seeds and heterogeneous beach sand [13]. In our work, we assume this similarity in response across different types of granular materials.

Using a 3-D discrete element method simulation, Kobayakawa et al. [11] reproduced ϕ_0-dependent force oscillations observed experimentally by Gravish et al [12]. In the models, the drag force is determined by considering the equilibrium of forces acting on the failure zone formed ahead of the plate. By using a 3-D wedge model developed by Wick and Perumpral [1], Kobayakawa et al. also modeled relationships between the plate angle with respect to the granular materials surface, the "shear band" that separates flowing and stationary particles, and the drag force on the plate [11].

1.2.2 Prior Robotics Research with Granular Materials

The majority of prior work on robots that operate in granular material environments has focused on locomotion as opposed to grasp and manipulation. The performance of robots that locomote across and through granular materials has advanced through the careful study of granular materials and the control of robot interactions with granular materials. For example, Aguilar et al. [16] measured the kinematics of granular flow with high-speed videos as a spring-mass robot jumped across dry granular materials (poppy seeds). Particle image velocimetry analyses of particle flow showed that a cone of solidified particles rapidly developed beneath the robot's foot during locomotion. Aguilar et al. created a geometric model of the cone's development under the load of an "intruder" (flat, circular robot foot) in order to estimate the dynamics of impulsive interactions between the robot foot and the granular materials. The dynamic model was then explored in order to enable the robot to jump higher from and sink less into the granular materials. A sand-swimming robot [17] was optimized to swim limblessly within granular materials with the aid of a granular material model based on a multi-particle Discrete Element Method (DEM) [18] and Resistive Force Theory (RFT) [19]. DEM was used to compute robot-particle and particle-particle interaction forces, and RFT was used to predict the performance of the sand-swimming robot for locomotion through granular materials. DEM simulations also enabled the lightweight DynaRoAch robot [20] to run across granular materials by exploiting the fluidic properties of granular materials.

1.3 Contributions

In this dissertation, we address the underexplored topic of tactile perception and haptic exploration within granular materials as a foundation for future work on robotic excavation and retrieval of buried objects. Unlike prior work that used unsensorized robot feet to interact with granular materials, we leveraged sensorized robot fingertips to interact directly with granular materials and to draw inferences about local contact state and buried objects.

Additionally, prior works on robot interactions with granular materials have assumed that the granular materials were homogeneous in nature. In this work, we consider both homogeneous granular materials as well as heterogeneous soil mixtures. The presence of a foreign object buried within granular materials adds to the heterogeneity of the granular material environment and further affects tactile percepts.

Chapter 2 presents an approach for tactile perception when sensorized robot fingertips are used to directly interact with homogeneous granular materials. We established a multimodal learning architecture based on Recurrent Neural Networks with Long Short-Term Memory units that incorporates multiple tactile sensing modalities (vibration, internal fluid pressure, and electrode impedance as a proxy for fingerpad deformation). Our SF-RNN model architecture learned to fuse sparse representations of multimodal tactile sensor data and performed robustly on novel test data that included variations in granular material type and particle size, fingertip orientation, fingertip speed, and object location along the linear fingertip trajectory. We also introduced a belief map representation to visualize tactile information from an SF-RNN model for a human teleoperator. During real-time implementation, the tactile perception model and belief map were updated every 150 ms and every 100-300 ms, respectively.

Chapter 3 presents a framework for tactile perception, mapping, and haptic exploration for the autonomous localization of objects buried in heterogeneous granular materials, a densely packed sand mixture. The haptic exploration task was performed using a sensor model that accounts for granular material characteristics and aids in the interpretation of interaction forces between the robot and its environment. The task was designed to efficiently locate a buried object and refine its outline while simultaneously minimizing potentially damaging physical interactions with the buried object. A continuous occupancy map was generated that fused local, sparse tactile information into a global map. We demonstrated our framework in simulation and with a real robot with granular materials.

Chapter 4 summarizes the contributions of this doctoral dissertation and presents potential enhancements for haptic exploration tasks within granular materials as future work.

CHAPTER 2

Tactile Perception for Teleoperated Robotic Exploration within Granular Media

This chapter was based on work published in the journal ACM Transactions on Human-Robot Interaction [21].

2.1 Abstract

The sense of touch is essential for locating buried objects when vision-based approaches are limited. We present an approach for tactile perception when sensorized robot fingertips are used to directly interact with granular media particles in teleoperated systems. We evaluate the effects of linear and nonlinear classifier model architectures and three tactile sensor modalities (vibration, internal fluid pressure, fingerpad deformation) on the accuracy of estimates of fingertip contact state. We propose an architecture called the Sparse-Fusion Recurrent Neural Network (SF-RNN) in which sparse features are autonomously extracted prior to fusing multimodal tactile data in a fully connected RNN input layer. The multimodal SF-RNN model achieved 98.7% test accuracy and was robust to modest variations in granular media type and particle size, fingertip orientation, fingertip speed, and object location. Fingerpad deformation was the most informative modality for haptic exploration within

granular media while vibration and internal fluid pressure provided additional information with appropriate signal processing. We introduce a real-time visualization of tactile percepts for remote exploration by constructing a belief map that combines probabilistic contact state estimates and fingertip location. The belief map visualizes the probability of an object being buried in the search region and could be used for planning.

2.2 Introduction

The task of reaching into a granular material and searching for an object is an essential human skill for applications that include harvesting root vegetables, excavating archaeological artifacts, rescuing victims of landslides and avalanches, and humanitarian demining. Unique to such tasks is the importance of the sense of touch. This work is motivated by the need to keep humans out of harm's way by enabling robots and their teleoperators to haptically locate, excavate, retrieve, and manipulate dangerous devices that are buried within granular media, such as sand, snow, and rubble. Contact with unknown, potentially delicate and/or dangerous objects should be done with care. However, state-of-the-art robotic hands cannot match the haptic perception abilities of human hands. The challenge of remotely controlling a remote end-effector within granular media is compounded by issues that include a lack of direct, natural sensory feedback and compromised bidirectional communication between the teleoperator and robot. The cognitive burden on the teleoperators of robot manipulators could be reduced if tactile sensors were used in concert with visual feedback during interactions with granular media and buried objects. Traditionally, experiments with robot hands are performed in open air environments, where contact and no contact conditions are easily distinguished using tactile sensors. To our knowledge, the direct use of robot hands within granular media has not yet been explored. In this work, we address tactile perception challenges that arise when sensorized robot fingertips are used in granular media environments. The objective of this work is to develop the fundamental capability of tactile perception within granular media for semi-autonomous robotic systems that are supervised by human teleoperators. Specifically, the goal is to develop the tactile ability to distinguish between contact with an object embedded in granular media and ubiquitous contact with surrounding granular media.

While granular media has been manipulated by hand in virtual reality [22, 23], this work represents the first use of tactile sensors for actual robot hand movements in direct contact with granular media. The rapid development of tactile devices and skins over the last couple of years has created new opportunities for tactile information being used for a myriad of robotic applications [24, 25, 26, 27, 28] enabled by the development of novel tactile sensors [29, 30, 31]. However, unlike the rapid progress in the field of computer vision with respect to its technology and algorithm development, tactile feedback has played a smaller role in robot planning and feedback control.

While open-source tactile datasets are becoming available [32], there are still no datasets that are commonly used for evaluating tactile sensing algorithms in the way that ImageNet [33] and RoboNet [34] are used for evaluating computer vision algorithms. In contrast to image datasets, tactile datasets are expensive to collect and it is difficult to obtain ground truth for tactile perception tasks. Once tactile data are collected, a suitable computational method for deciphering the encoded percepts becomes especially important.

In this work, we present a multimodal tactile perception model that can be used for remote robotic tactile exploration within granular media. Extending beyond a prior workshop paper [35], the model performs tactile perception by combining signal processing, sparse

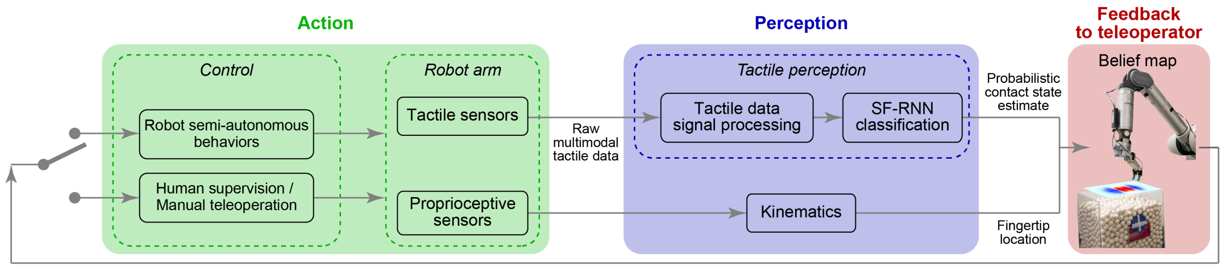


Figure 2.1: A block diagram is shown for a notional real-time execution of the action, perception, and visualization processes.

coding, and a recurrent neural network to integrate multiple tactile sensing modalities (vibration, pressure, fingerpad deformation). We refer to our model as a Sparse-Fusion RNN (SF-RNN in Figure 2.1) since sparse features are extracted in a pre-processing step that precedes a fusion layer of an RNN. The tactile percepts are visualized as a spatial probabilistic distribution using a "belief map" [36]. Figure 2.1 shows a high-level diagram of a notional computational flow during real-time execution. This work focuses primarily on the "Tactile perception" and "Feedback to teleoperator" portions of the block diagram shown in blue and red, respectively.

This article is organized as follows. Section 2 introduces related work. Section 3 describes the multimodal tactile perception algorithm and Section 4 describes the belief map representation of tactile information. Sections 5 and 6 detail the experimental procedure for robot data collection and model evaluation, respectively. Section 7 discusses notable trends in the tactile data and real-time tactile perception and visualization for human control of remote robotic systems. Section 8 concludes the work with a summary of contributions, limitations, and future directions.

2.3 Related Work

2.3.1 Complex Behaviors of Granular Media

Granular media, such as sand, grain, and powder are ubiquitous in nature and in industrial processes. Yet, the modeling of the complex behaviors of granular media remains an active area of study. It is inappropriate to classify granular material properties as strictly solid- or liquid-like [3], and there is no universal continuum model to predict the behavior of granular media in an arbitrary geometry under arbitrary loading. Affected by its packing condition, the rigidity and flow characteristics of granular media [5] exhibit complex behaviors, such as jamming [8], dilatancy [6], flow bifurcation in the vicinity of a "intruder" [13], and force fluctuations around the intruder. Models have been developed to estimate soil properties such as internal friction, interface friction, and soil density (e.g. [37] [38]) from large interaction forces. Grain-scale numerical models such as the Discrete Element Method [18] and Resistive Force Theory [19] have been applied to simulate behaviors of granular particles.

An important feature of granular media is the "force chain", which is a hallmark feature of physical interactions with granular media. Under external stress, the force chain carries most of the load via complex inter-particle forces [39]. When external forces are too small to cause particle rearrangements, granular media becomes an amorphous solid and the distribution of interparticle normal forces develops a peak near the jamming transition [9]. When an intruder moves through granular media, as when a robot finger is raked through sand, the granular media packing fraction will change and particle flow and force fluctuations will occur around the intruder. When the intruder is a sensorized finger, the finger-particle interactions directly affect the tactile measurements. The countless instances of making and breaking contact between the fingertips and surrounding granular media particles effectively add noise to tactile sensing modalities that might otherwise be useful for object detection, the perception of geometric features (e.g. [40]), and the detection of key haptic events.

2.3.2 Robot Interaction with Granular Media

Exploiting the sense of touch is essential for robots operating according to physical interaction cues and is often associated with computational and representation challenges [41]. As with biological systems, artificial tactile perception is an active process for sensing properties of one's environment, especially in the absence of vision [42]. The task of locating a buried rigid object has been previously studied using a leader-follower teleoperated robot system in which a custom haptic device displayed forces and torques directly to the operator's hand [43]. In that work, a human teleoperator manually controlled a robot that was sensorized with a force/torque sensor that remained outside of the granular media. Here, we aim to develop tactile perception capabilities for a semi-autonomous robot whose contact states are visually displayed for a human telesupervisor. We use multimodal tactile sensors that make direct contact with granular media and an object buried within the granular media.

Prior research has been conducted on robot interactions with granular media, primarily with a focus on improving robot locomotion. For example, studies have been performed on legged [20] and flipper-based locomotion [44], sidewinding [45] and sand-swimming [17]. Numerical simulation approaches have enabled high locomotion performance [16] by predicting forces and flow fields of the granular media, which are difficult to measure during experiments without sensorized feet. In addition, particle image velocimetry analyses via high speed video have enabled researchers to measure grain kinematics and predict quasistatic resistive forces to enable robots to jump [46] and walk [47] on flat dry granular surfaces. In general, however, legged robots do not incorporate local sensing on feet because sensors lack the robustness to repeated impact and wear.

In addition to locomotion, researchers have studied robot manipulation of granular media. For household applications, the granular media are poured from containers, such as a box [48] or powdered coffee creamer bottle [49], or manipulated with a tool, such as a spoon [50]. Unlike works in simulated environments [51, 52], it is difficult to observe the dynamics of granular media in real applications. As a result, researchers have relied heavily upon computer vision techniques. To detect the flow of granular media during pouring, Yamaguchi and Atkeson employed a stereo camera to reconstruct the 3D point cloud [49] of the granular media. Schenck et al. evaluated predictive models with depth images in order to infer the dynamics of scooping and pouring actions [50]. In these prior works, vision was the primary sensory modality and researchers did not touch the granular media directly using sensorized robot fingers.

2.3.3 Feedback to Human Teleoperator

To assist a remote human operator to haptically locate a buried object within granular media, a robotic system needs to provide accurate tactile information promptly, convey the properties of the environment, and communicate recommended future actions. Until immersive telepresence set-ups become the norm, current common methods for feedback include audio, visual, and tactile displays. Tactile feedback, which can be subdivided into kinesthetic and cutaneous feedback, relies on the design of tactile display hardware and the ability or willingness of the human teleoperator to don the hardware. Kinesthetic feedback can be achieved by coupling a human teleoperator to devices, such as exoskeletal-based robotic arms [53], that apply forces and torques to the limbs. Kinesthetic feedback devices tend to be heavy and expensive and they additionally require custom haptic devices unique to a particular task. Cutaneous feedback can be achieved by vibrating [54, 55], stretching [56], or applying pressure/force [57, 43] or changes in temperature [58] to a teleoperator's skin. The feedback device can be placed on a body part that mirrors the sensor location on a robot surrogate (as on fingertips) or on a body part that is available (as on the upper arm of a transradial amputee).

In many cases, some feedback methods are impractical for field-specific applications. For instance, audio feedback is not an option in loud environments. In another example, a teleoperator may be wearing gear that precludes the donning of bulky kinesthetic or cutaneous feedback devices. In this work, we have elected to convey the learned contact state information from our tactile perception model in a visual manner, as a proof-of-concept. The visualization can include spatial, temporal and probabilistic information, for example. For our example application, the belief map represents a top-down view of the exploration area and uses a color map to reflect tactile sensor-driven probabilities of an object being buried in a particular region (Figure 2.1). Such a visualization could serve as a source of actionable information, enabling teleoperators to safely gain situational awareness and enhancing the quality and efficiency of human-robot interaction.

2.4 Multimodal Model for Tactile Perception

The goal of the tactile perception algorithm is to enable a robot to distinguish a buried object from its surrounding granular media through tactile exploration. The two fingertip contact states of interest are "contact with a buried object" or "free movement through granular media." In order to address the tactile perception problem more generally, we propose

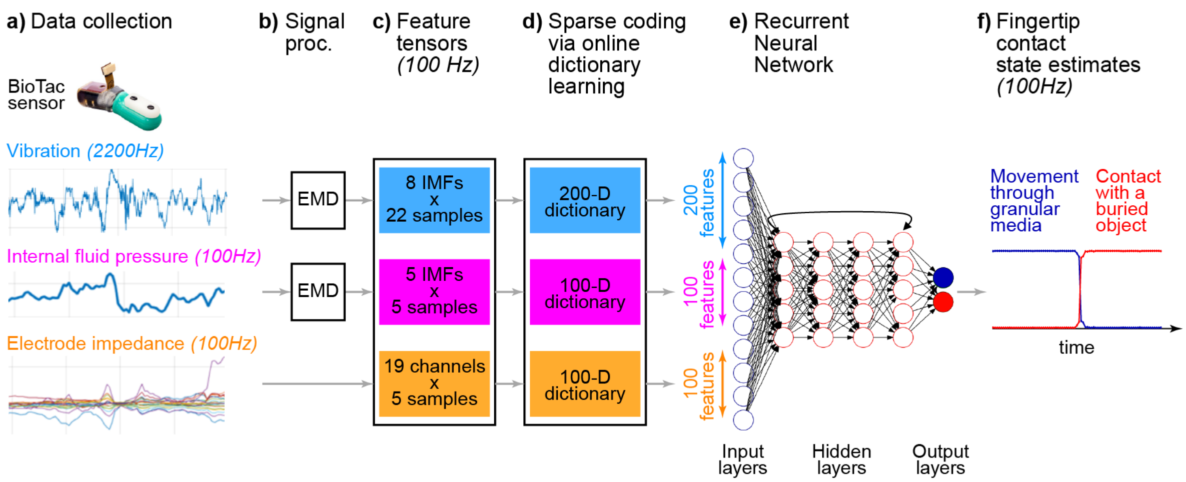


Figure 2.2: Sparse-Fusion RNN (SF-RNN) model architecture for multimodal tactile perception in granular media. (a) Vibration, internal fluid pressure, and 19 electrodes' impedance data were collected from BioTac sensors. (b) Vibration and internal fluid pressure data were filtered by EMD. (c) Feature tensors were composed for each modality. (d) Sparse coding was used as an unsupervised pre-training step. (e) An RNN mapped a 400D input space to a 2D output space. (f) The RNN sequentially labeled tactile inputs according to the estimated fingertip contact state.

a model architecture that we refer to as a Sparse-Fusion RNN (SF-RNN, Figure 2.2) since sparse features are extracted via an unsupervised learning pre-training step prior to fusion through a fully connected input layer of an RNN. The model incorporates multiple sensing modalities from the BioTac tactile sensor (SynTouch LLC, Los Angeles, CA): vibration, internal fluid pressure, and electrode impedance as a proxy for fingerpad deformation (Figure 2.2a).

Raw tactile sensor signals were first pre-processed in order to filter out noise. Features of the vibration and internal fluid pressure data were distilled using Empirical Mode Decomposition (EMD) (Figure 2.2b) prior to the creation of a data structure that combined the three tactile sensing modalities (Figure 2.2c). Per well-established practices of machine learning libraries such as Tensorflow [59] and pytorch [60], we referred to the combined features from different heterogeneous sensor data as feature tensors in our SF-RNN model architecture. Sparse coding via online dictionary learning was used as an unsupervised pre-training step (Figure 2.2d), which yielded a learned feature vector for use as an input to an RNN classifier (Figure 2.2e). The use of sparse coding enabled the efficient storage of the 400D feature vector prior to the RNN mapping to a 2D output space that reflected the perceived fingertip contact state (Figure 2.2f). The accuracy of the fingertip contact state estimate is important, as a probabilistic belief map will be created to visualize the data for a human teleoperator (Section 2.5) under the assumption that the contact state estimate is correct.

2.4.1 Pre-processing of Tactile Data

2.4.1.1 Tactile Sensor Signal Processing

Empirical Mode Decomposition (EMD) [61] (Figure 2.2b) is a powerful technique for analyzing nonlinear, non-stationary, natural signals because EMD decomposes a time series signal into a series of intrinsic mode functions $IMFs$ that preserve characteristics of the original signal, such as variations in frequency (Appendix A.1). Physical meaning can often be associated with EMD-filtered components, as in the analysis of electroencephalogram [62], water wave [61], and earthquake time series data [63].

The original tactile signal $x(t)$ is thus decomposed in the time domain with a small enough residue r_n.

$$x(t) = \sum_{j=1}^{n} IMF_j + r_n, \tag{2.1}$$

As shown in Figure 2.2b, EMD was applied to the BioTac's vibration and internal fluid pressure data. Per [61], we set the residue threshold at 10% of the norm of IMF_{n-1} for each the vibration data and internal fluid pressure data. We selected the first $n = 8$ $IMFs$ for the vibration data and the first $n = 5$ $IMFs$ for the internal fluid pressure data, leaving only monotonic final residues. Figure 2.3 shows the original vibration data for 2 sec of a representative trial along with deviations of its first 8 $IMFs$ from the mean IMF.

2.4.1.2 Reconciliation of Heterogeneous Datastreams

Our model incorporated features from vibration, internal fluid pressure and electrode impedance data from a multimodal tactile sensor. These tactile sensor data were heterogeneous due to the different transduction mechanisms, dimensionality of sensing channels, and

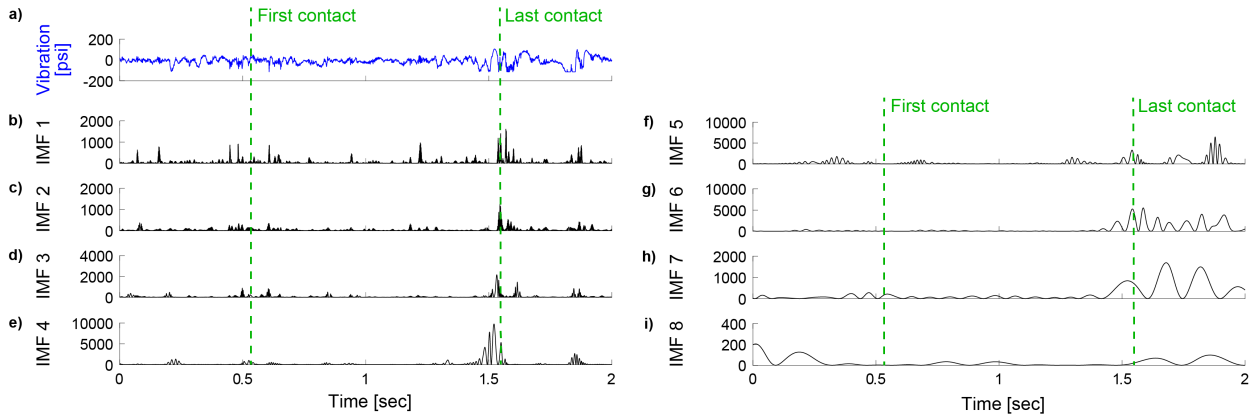


Figure 2.3: (a) Raw vibration data and (b-i) squared deviations of its first 8 $IMFs$. Each deviation was calculated as *norm[IMF - mean(IMF)]*.

sampling frequencies unique to each modality. One obvious way to fuse heterogeneous sensor data is by concatenating the datastreams. However, we found that models that utilized this simple concatenation approach performed poorly, as will be described in Section 2.7.1 and shown in columns 1 and 4 of Tables 2.2 and 2.3.

Ultimately, we used a sensor fusion architecture in order to leverage well-established computer vision techniques while reconciling the differences in the heterogeneous datastreams. We selected the update frequency of the feature tensors outputted from step "(c)" in Figure 2.2 to be 100 Hz. For the vibration data sampled at 2200 Hz and then filtered using EMD, the most recent 22 samples were extracted for each of 8 $IMFs$. For the internal fluid pressure data sampled at 100 Hz and then filtered using EMD, the most recent 5 samples were extracted for each of 5 $IMFs$. Finally, for the electrode impedance data sampled at 100 Hz, the most recent 5 samples were extracted. In this way, all three modalities would be represented by a spatiotemporal series of data. With an update frequency of 100 Hz, the following feature tensors were updated every 10 ms: 8 $IMFs$ x 22 samples for vibration data, 5 $IMFs$ x 5 samples for internal fluid pressure data, and 19 channels x 5 samples for electrode impedance data (Figure 2.2c).

2.4.1.3 Sparse Representation and Online Dictionary Learning

Upon creating feature tensors from three tactile sensing modalities, we used sparse coding (Figure 2.2d) to represent the multimodal tactile data as linear combinations of basis functions that capture higher-level features (Appendix A.2). These features were identified autonomously and adaptively using an unsupervised online optimization algorithm for dictionary learning [64]. This unsupervised, pre-training step was introduced specifically to improve the generalization of the multimodal tactile perception model [65].

Using sparse coding, we compactly represented the overcomplete model of the tactile sensor data. By solving a matrix factorization problem with a sparsity penalty (eq. A.2 in Appendix A.2), the algorithm learns a dictionary $D = [d_1, \dots d_k] \in R^{m \times k}$ and a sparse matrix of coefficients $\alpha = [\alpha_1, \dots \alpha_t] \in R^{k \times t}$ that can be used to approximate a signal $X = [x_1, \dots x_t] \in R^{m \times t}$ in a linearized manner as $\hat{X}$.

$$\hat{X} = D\alpha \tag{2.2}$$

Each dictionary size k, also referred to as feature vector length, was chosen heuristically as $k = 200$, 100, and 100 for vibration, internal fluid pressure, and electrode impedance modalities, respectively. According to [64], the sparsity penalty was set with a scaling parameter related to the dimension of the original signal. The feature tensors from step "(c)" in Figure 2.2 were flattened and passed through the sparse coding algorithm in parallel.

2.4.2 RNN with GRUs for Binary Classification

The final stage of our multimodal tactile perception pipeline was comprised of a binary RNN classifier. Through the use of an RNN architecture, we aim to develop a model capable of learning rich representations of tactile signals within granular media and capturing temporal dependencies. The inputs to the RNN were the sparse, overcomplete 200D, 100D, and 100D feature vectors for the vibration, internal fluid pressure, and electrode impedance data, respectively. These feature vectors were fused within the RNN's fully connected input layer. The output of the RNN was one of two fingertip contact states: "contact with a buried object" or "free movement through granular media."

The RNN was comprised of four hidden layers with 50 neurons per layer with parameters selected as described below. Model training and sequence labeling were performed with a Long Short Term Memory (LSTM) variation called Gated Recurrent Units (GRUs) [66] (Appendix A.3), which perform well for sequence learning and translation applications [67]. By combining RNNs with GRUs, an RNN can learn a nonlinear mapping between multimodal tactile feature inputs and contact state outputs. The long term memory duration was designed as 500 ms with a short term memory duration of 10 ms, per the 100 Hz update rate selected for step "(c)" of Figure 2.2. RNN weights were updated using Adaptive Momentum Estimation (Adam) Optimization, which computes adaptive learning rates for each parameter [68]. The cost function was based on cross-entropy loss and also included a penalty for heavy weights via L2-regulation. Dropout was applied with a rate of 0.4 in order to prevent overfitting and to enhance robustness of the model to the loss of input features, which will be evaluated using an ablation study in Section 2.7.1.

2.5 Belief Map Representation of Tactile Information

To convey the ever-changing fingertip state from our multimodal tactile perception pipeline to a human teleoperator, we lay the foundation for the visualization of an unknown object buried in granular media using a "belief map" [36]. The belief map combines spatial information related to fingertip location (derived from known robot kinematics) with probabilistic fingertip contact state estimates provided by our SF-RNN model architecture (Figure 2.2). A belief map can be used by a robot for semi-autonomous trajectory planning to further reduce uncertainty in the map or by a human teleoperator for high-level decision-making.

As a robot explores a region of interest, it samples from an unknown distribution and obtains new information. Specific to our application, a robotic hand with tactile sensors was used to explore and estimate the location of a buried object. As the robotic hand moved through space, it sampled the underlying spatial distribution by estimating the fingertip contact state and simultaneously associating the contact state with the fingertip location in space. The probability values for the entire spatial distribution were updated at each timestep with each new observation through a probabilistic classification method, specifically Support Vector Classification (SVC) with a radial basis function (RBF) kernel (Appendix A.4). Spatial information was stored in an RBF kernel that maps the influence of contact state at its spatial location into an infinite dimensional feature space. The spatial reach, or sphere of influence, of the radial basis function was customized for best performance for the size of our testbed, tactile sensor, and buried object. Since the SVC score function sums the learned mapping from all previous timesteps, the belief map is capable of tracking the states it has previously sampled. Finally, the SVC scores were calibrated using Platt Scaling [69] to produce a probability of whether an object is buried at a specific location. This probability

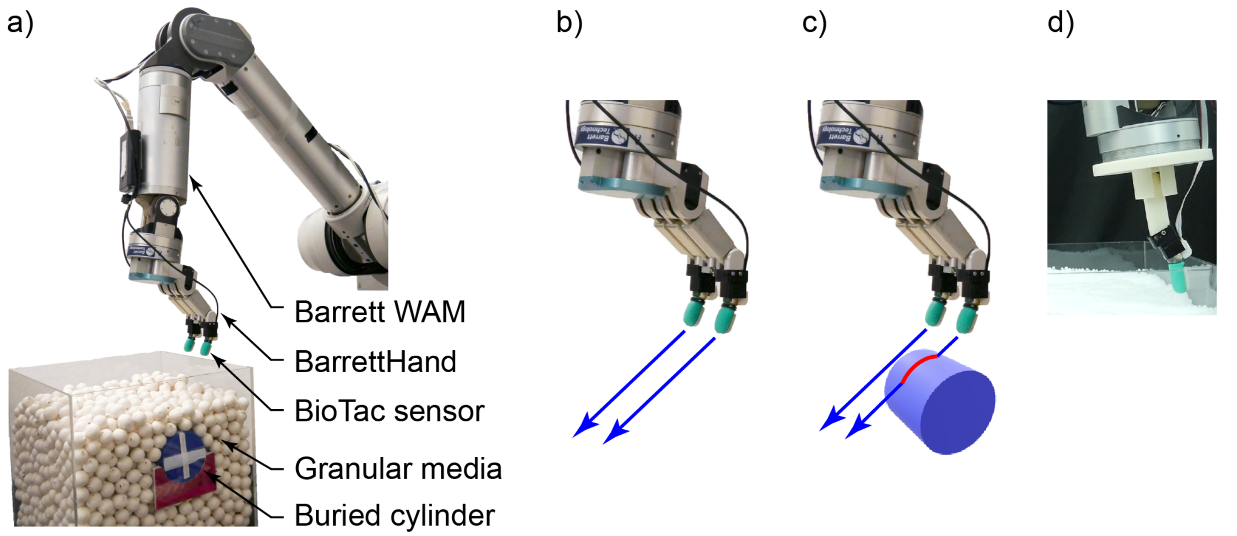


Figure 2.4: (a) Experimental set-up used to collect multimodal tactile sensor training data by sweeping sensorized fingertips through granular media (cottonwood beads) and across a buried cylinder. (b) Blue arrows indicate the linear trajectory of two fingertips through granular media (not shown for clarity). (c) In the presence of a buried cylinder, one fingertip was swept briefly across the cylinder (finger-object contact is indicated in red). (d) For a robustness study, different granular media types were used (plastic pellets are shown) and a 3D-printed adapter was used to change the orientation of the fingertip to simulate 20° of finger extension relative to the pose in (a-c).

integrates information about estimated fingertip contact state and known fingertip location. In this work, the Python Scikit-Learn package [70] was used for optimization and calibration. The distribution that represents the probability of fingertip contact with a buried object is updated for every 10 iterations of the RNN, a belief map update rate that was limited by our current computational power and is discussed further in Section 2.8.2.

2.6 Experimental Procedure: Robot Data Collection

In order to equip robots with the ability to locate objects buried within granular media, we focus on perceiving fingertip contact states for a preliminary set of fingertip-object-media interactions. In this section, we describe our robot testbed and datasets for model training and model testing.

2.6.1 Robot Testbed

Our robot testbed consisted of a 7 degree-of-freedom (DOF) Barrett Whole Arm Manipulator (WAM) with a 4 DOF BarrettHand (Barrett Technology, Cambridge, MA) (Figure 2.4). Two BioTac sensors (SynTouch LLC, Los Angeles, CA) were mounted on the BarrettHand fingertips. The multimodal BioTac sensor measures vibration at 2200 Hz, internal fluid pressure at 100 Hz, and fingerpad deformation data at 100 Hz via an array of 19 spatially distributed electrodes that sense impedance [29].

The sensorized robot fingertips were commanded, using position control, to perform linear exploratory movements within a rigid, 30 cm cube containing granular media. A 9.1 cm diameter, rigid plastic cylinder was buried within the granular media. The cylinder was fixed in place such that its long axis was perpendicular to the linear exploratory movements. Interactions with buried objects that can move with or within granular media are beyond the scope of this work. The top, curved surface of the cylinder was set at a depth of 4 cm below the top surface of the granular media. The robot fingertips were swept at a constant commanded velocity through the granular media at a depth of 5 cm to ensure finger-object contact during some portion of the exploratory movements.

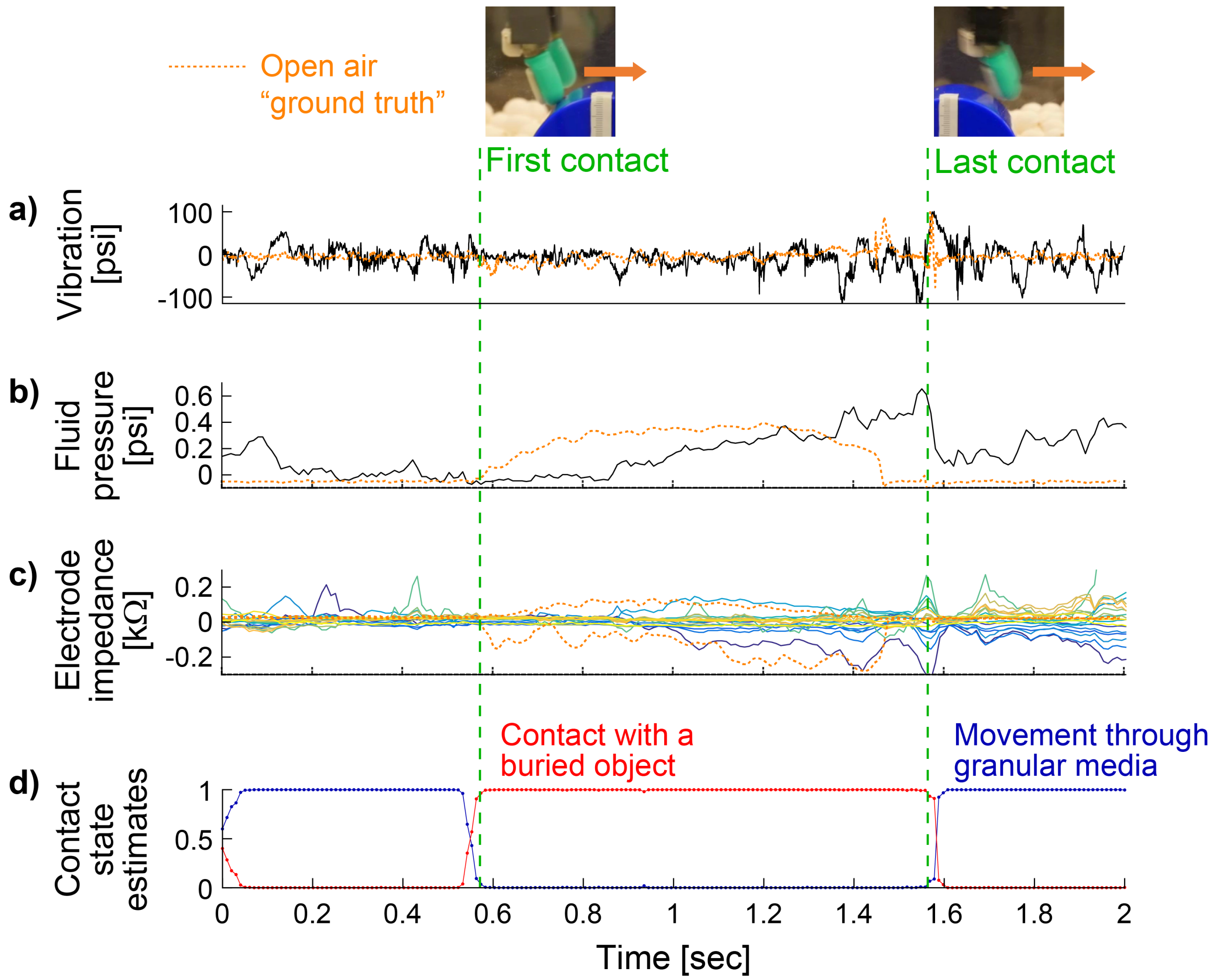


Figure 2.5: (a) Vibration, (b) internal fluid pressure, (c) electrode impedance, and (d) fingertip contact state estimates from an SF-RNN model are shown for a representative 2 sec trial. “First contact” and “last contact” events were identified from open air “ground truth” trials (screenshots, orange dotted lines). For clarity, impedance values from an open air trial are shown for two electrodes only (min, max traces).

2.6.2 Dataset for Model Training

As a precursor to future experiments in soil, we used 20 mm diameter wooden beads made of lightweight cottonwood as granular media for model training. The robot digits were swept through the granular media at a constant commanded speed of 4 cm/s. For one block of 250 trials, both digits were moved through the granular media without contacting the cylinder (Figure 2.4b). Pooling trials across both digits for this case resulted in a total of 500 trials in which the fingertips did not contact a buried object. As a first step toward multi-finger tactile perception, the commanded trajectory was such that only one digit contacted the cylinder while the other digit moved through the granular media (Figure 2.4c) for another block of 250 trials. This scenario generated a total of 250 trials in which a single fingertip made contact with a buried object. In summary, a total of 750 trials were collected (250 with contact and 500 without contact with a buried object).

In order to train and test the model, a 2 sec window of data was extracted from each trial in which the sensorized fingertip moved freely through granular media, made contact and broke contact with a buried object, and again moved freely through granular media (Figure 2.5). The making and breaking of contact with the buried cylinder will be referred to as "first contact" and "last contact" events, respectively. Trials without the buried cylinder featured only free fingertip movement through granular media. To mark the timing of the "first contact" and "last contact" events, trials were conducted with the cylinder exposed (in open air) for use as a "ground truth" (Figure 2.5).

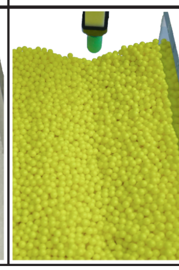
2.6.3 Dataset for Model Testing

In order to directly test the robustness of our models to variation in the granular media, we collected additional data for model testing that included a wide range of granular media types. We selected granular media that were representative of aggregates ranging from coarse sand to coarse gravel according to the Wentworth scale [2], which is the most frequently used geological grade scale for classifying the size of sediment particles. As shown in Table 2.1, we used 1 mm diameter poppyseed, 3 mm diameter plastic pellets for injection molding, 6 mm diameter plastic BBs for pellet guns, 16 mm diameter wooden beads made of pine, and 20 mm diameter wooden beads made of cottonwood.

We also tested the robustness of our models to variations in finger-object contact conditions. First, we used a 3D-printed adapter to change the orientation of the fingertip and simulate 20° of extension (Figure 2.4d) relative to the original pose of the BioTac when collecting data for model training (Figure 2.4a-c). Second, we swept the robot digit at a slower speed of 2 cm/s as compared to the 4 cm/s used for the training dataset. Finally, we moved the buried object 3.5 cm downstream from the original location such that the robot digit would not make contact until later in the exploratory movement. In doing so, we could directly test whether the spatiotemporal model trained with cottonwood beads was overfit.

For each of the five granular media types, we collected 50 trials in which a single fingertip contacted the buried object and 25 trials in which neither digit contacted the buried object. Since fingertip speed was halved for the test dataset, the window of data that was extracted from each test trial was doubled (i.e. 4 sec of data). Data labeling proceeded as described in Section 2.6.2.

Table 2.1: Granular Media Types: To test model robustness, we selected five distinct granular media types to represent the five aggregate types ranging from coarse sand to coarse gravel according to the Wentworth scale [2].

Aggregate type	Coarse sand	Very fine gravel	Fine gravel	Medium gravel	Coarse gravel
Particle diam.	1-2 mm	2-4 mm	4-8 mm	8-16 mm	16-32 mm
Selected media	Poppyseed	Plastic pellets	Plastic BBs	Wooden beads (pine)	Wooden beads (cottonwood)
Particle diam.	1 mm	3 mm	6 mm	16 mm	20 mm
Testbed					

2.7 Experimental Procedure: Model Evaluation

We extensively evaluated the SF-RNN on a tactile dataset with 750 trials of exploration within granular media. The dataset was randomized, then separated into 60% for training, 20% for validation, and 20% for testing. The validation dataset was used for tuning of model hyperparameters while the test dataset was used to evaluate the performance of the trained and validated model. We trained a number of classifier models with varying complexity following steps in the tactile perception pipeline (Figure 2.2). Models were trained for a sequence labeling task to distinguish between two fingertip states: "contact with a buried object" or "free movement through granular media." The "contact with a buried object" class includes cases in which the fingertip touches the object directly and when the fingertip encounters the object indirectly through a force chain of granular particles.

We evaluated tactile perception models with both linear support vector machine (SVM) [71] and recurrent neural network (RNN) classifiers on the following input architectures:

- “(b)”: $IMFs$ extracted via EMD from vibration and internal fluid pressure data plus raw electrode impedance data (Figure 2.2b).

- “(c)”: Feature tensors constructed from EMD-processed vibration and internal fluid pressure data and raw electrode impedance data (Figure 2.2c).

- “(d)”: Sparse features identified autonomously and adaptively using an unsupervised online optimization algorithm for dictionary learning (Figure 2.2d).

For the linear SVM classifier models, feature tensors were flattened and concatenated. The linear classifiers were implemented in MATLAB (MathWorks, Natick, MA) while RNN architectures were implemented in Python using TensorFlow [59] with two NVIDIA GeForce GTX 970 graphics cards.

2.7.1 Study of Model Architecture

For an ablation study, all model architectures were evaluated as unimodal models (vibration, internal fluid pressure, or electrode impedance inputs only) and as multimodal models (with pairs of modalities, or all three modalities simultaneously) using training, validation, and test datasets based on cottonwood beads. When the linear SVM classifier was trained as a multimodal model, input streams from each modality were simply concatenated after data alignment to account for different sampling frequencies. Since this sort of simple concatenation performs poorly for an RNN architecture, data from different tactile sensor modalities were processed by separate RNNs and then passed through a fully connected (fusion) layer, which was referred to as Fusion RNN (F-RNN) by Jain et al. [72]. For our proposed multimodal tactile perception SF-RNN model architecture, tactile sensor modalities were passed through a sparse coding process as an unsupervised pre-training step, which preceded a fully connected fusion layer of RNN (Figure 2.2e).

The data augmentation technique of oversampling was used to address our unbalanced dataset, increase the diversity of the training data, and minimize overfitting of the trained model. For linear classifier models, observations from the minority class ("contact with a buried object" class) were randomly selected and added to the training set during each iteration during training. For RNN models, since the long term memory duration was designed as 500 ms, the data were oversampled using a sliding window paradigm. The 500 ms tactile data streams were extracted from the selected 2 sec window with different increments. Trials featuring free movement through granular media were oversampled by shifting a 500 ms sliding window in 10-frame (100 ms) increments. Since there were half as many trials featuring contact with a buried object, a 5-frame increment was used to shift the sliding window in those cases. Tables 2.2 and 2.3 summarize the results of the ablation studies on the unimodal and multimodal models, respectively.

2.7.1.1 Effect of Model Inputs and Sparse Coding

We built multiple models to assess the usefulness of different model input combinations for both model training and testing: one tactile sensing modality, two modalities, and three modalities. As expected, models trained with two or three modalities (Table 2.3) achieved higher test accuracy than those trained with only one modality (Table 2.2).

In Table 2.2, we present the results of the ablation study for all six model architectures for models that used unimodal tactile inputs only. When the (d)+Linear SVM architecture was used, fluid pressure was the most informative tactile sensing modality, yielding a test accuracy of 95.8%. However, when the SF-RNN architecture was used, vibration and electrode impedance became much more informative, yielding test accuracies of 95.7% and

Table 2.2: Ablation Study with **Unimodal Models**: Model accuracy is reported as mean ± standard deviation in %. Best performing models are shown in bold. The models were trained and evaluated with cottonwood beads. The (b)-(d) annotations refer to the perception pipeline steps shown in Figure 2.2.

Modality	Category	Method					
		(b)+Linear SVM	(c)+Linear SVM	(d)+Linear SVM	(b)+RNN	(c)+RNN	(d)+RNN (SF-RNN)
vibration	validation (%)	51.1±0.7	62.2±0.6	73.5±0.5	65.91±0.5	66.0±0.5	95.9±0.2
	test (%)	48.7±0.7	60.4±0.6	72.0±0.5	65.3±0.5	65.6±0.5	**95.7**±0.2
	process time (ms)	1.68	2.28	2.47	1.5	0.35	2.58
	input	8 $IMFs$	8 $IMFs$ x 22 samples	200 sparse features	8 $IMFs$	8 $IMFs$ x 22 samples	200 sparse features
fluid pressure	validation (%)	52.7±0.6	73.3±0.5	92.5±0.3	67.5±0.5	67.1±0.5	85.4±0.4
	test (%)	49.0±0.7	69.0±0.6	**95.8**±0.2	67.4±0.5	67.2±0.5	85.8±0.4
	process time (ms)	0.06	0.05	0.08	0.19	0.09	0.26
	input	5 $IMFs$	5 $IMFs$ x 5 samples	100 sparse features	5 $IMFs$	5 $IMFs$ x 5 samples	100 sparse features
electrode impedance	validation (%)	82.2±0.4	84.7±0.4	93.4±0.3	90.3±0.3	92.1±0.3	97.8±0.2
	test (%)	78.6±0.5	78.7±0.5	85.4±0.4	88.3±0.4	90.1±0.3	**97.6**±0.2
	process time (ms)	0.22	0.17	0.24	0.23	0.17	0.4
	input	19 channels	19 channels x 5 samples	100 sparse features	19 channels	19 channels x 5 samples	100 sparse features
avg test accuracy (%)		58.76	69.38	84.42	73.68	79.06	**93.06**

Table 2.3: Ablation Study with **Multimodal Models**: Model accuracy is reported as mean ± standard deviation in %. Best performing models are shown in bold. The models were trained and evaluated with cottonwood beads. The (b)-(d) annotations refer to the perception pipeline steps shown in Figure 2.2.

Modality	Category	Method					
		(b)+Linear SVM	(c)+Linear SVM	(d)+Linear SVM	(b)+ F-RNN	(c)+ F-RNN	(d)+RNN (SF-RNN)
vibration, fluid pressure	validation (%)	51.8±0.7	78.1±0.5	99.7±0.1	67.1±0.5	73.0±0.5	96.5±0.2
	test (%)	46.6±0.7	72.2±.5	**99.5**±0.1	64.4±0.6	67.0±.5	96.5±0.2
	process time (ms)	2.62	1.87	2.52	7.34	0.75	2.72
	input	8 $IMFs$, 5 $IMFs$	8x22, 5x5 feature tensors	200D, 100D sparse features	8 $IMFs$, 5 $IMFs$	8x22, 5x5 feature tensors	200D, 100D sparse features
vibration, electrode impedance	validation (%)	84.7±0.4	86.4±0.4	93.8±0.3	85.1±0.4	90.4±0.3	98.3±0.2
	test (%)	78.3±0.5	84.9±0.4	91.7±0.3	83.9±0.4	86.4±0.4	**97.8**±0.2
	process time (ms)	1.68	2.24	2.73	7.41	0.77	2.83
	input	8 $IMFs$, 19 channels	8x22, 19x5 feature tensors	200D, 100D sparse features	8 $IMFs$, 19 channels	8x22, 19x5 feature tensors	200D, 100D sparse features
fluid pressure, electrode impedance	validation (%)	83.4±0.4	88.3±0.3	98.7±0.1	89.4±0.4	92.1±0.3	98.1±0.2
	test (%)	81.7±0.4	85.8±0.4	**98.9**±0.1	89.0±.4	91.0±0.3	97.5±0.2
	process time (ms)	0.27	0.19	0.302	0.43	0.54	0.51
	input	5 $IMFs$, 19 channels	5x5, 19x5 feature tensors	100D, 100D sparse features	5 $IMFs$, 19 channels	5x5, 19x5 feature tensors	100D, 100D sparse features
all modalities	validation (%)	85.0±0.4	89.6±0.3	99.9±0.1	84.7±0.4	92.0±0.3	98.4±0.1
	test (%)	80.8±0.4	89.2±0.3	**99.7**±0.1	84.1±0.4	84.3±0.4	98.7±0.1
	process time (ms)	8.08	2.39	2.82	8.06	0.85	2.96
	input	8 $IMFs$, 5 $IMFs$, 19 channels	8x22, 5x5, 19x5 feature tensors	200D, 100D, 100D sparse features	8 $IMFs$, 5 $IMFs$, 19 channels	8x22, 5x5, 19x5 feature tensors	200D, 100D, 100D sparse features
avg test accuracy (%)		68.8	81	96.7	79.1	81.5	**97.3**

97.6%, respectively. Use of electrode impedance data resulted in the highest test accuracy of 97.6% for all unimodal models, suggesting that electrode impedance data, generated by fingerpad deformation, was the single most informative tactile sensing modality for distinguishing between the two fingertip contact states (Table 2.2).

Of the unimodal models (Table 2.2), classifiers trained on overcomplete sparse representations of tactile data ("(d)+" cases) had obvious advantages as compared to other models. Notably, only the proposed SF-RNN ("(d)+RNN") model architecture was capable of achieving at least an 85% test accuracy for each of the three unimodal input cases. While the "(c)+RNN" method achieved 90.1% test accuracy for unimodal models trained on electrode impedance, test accuracy decreased to 65.6% and 67.2% for models trained only on vibration data or fluid pressure data, respectively.

In Table 2.3, we present the results of the ablation study for all six model architectures for models that used multimodal tactile inputs (with pairs of modalities, or all three modalities simultaneously). Both the linear SVM and RNN multimodal models trained on sparse features ("(d)+" cases) produced high test accuracy results. The average test accuracy was 96.7% and 97.3% for the (d)+Linear SVM and SF-RNN model architectures, respectively. Although the (d)+Linear SVM model performed well in Table 2.3, the (d)+Linear SVM model did not perform well during the robustness tests, which will be discussed in Section 2.7.2.

Training on three modalities did not provide much of an advantage over training on two modalities. For example, the SF-RNN model trained and tested on three modalities had a test accuracy of 98.7%, which was only slightly better than the levels of performance for the two-modality SF-RNN models, which ranged from 96.5% to 97.8% accuracy. One advantage of training on three tactile sensing modalities appears to be a robustness to incomplete input data. Without having to retrain the model, one could provide data from any of the tactile sensing modalities, or combinations thereof, and achieve comparable performance to models trained more specifically on fewer modalities (results of the SF-RNN model tested with incomplete inputs are shown in Table A.1 in Appendix A.5). This could be useful if one or more modalities becomes temporarily unavailable during tactile exploration. Considering the complexity and harsh working environment of interacting with and maneuvering within granular media, we recommend a multimodal SF-RNN structure instead of relying on a single modality for haptically perceiving contact state within granular media.

2.7.1.2 Effect of Sensor Data Representation

In order to check the sensor data representation and model training, we visualized the data and feature spaces throughout the multimodal modeling pipeline (Figure 2.2) using a dimensionality reduction algorithm called t-Distributed Stochastic Neighbor Embedding (tSNE) [73]. Feature spaces were visualized in Figure 2.6, in which each dot represented a labeled multimodal tactile sensor feature tensor that was updated at 100 Hz. This visualization provided insight as to how our proposed architecture dealt with dimensionality, learned features to represent multimodal tactile sensor data, and classified features with high accuracy.

First, the feature tensors from Fig. 2.2c (8 $IMFs$ x 22 samples for vibration data, 5 $IMFs$ x 5 samples for fluid pressure data, and 19 channels x 5 samples for electrode impedance data) were visualized in 3D space as a jumbled cluster of feature tensors with mixed fingertip contact states (Figure 2.6a). Using sparse coding, these entangled feature tensors were fully separated, with feature tensors classified as fingertip contact with a buried object now pushed into the outer regions of an overcomplete 400-dimensional sparse feature

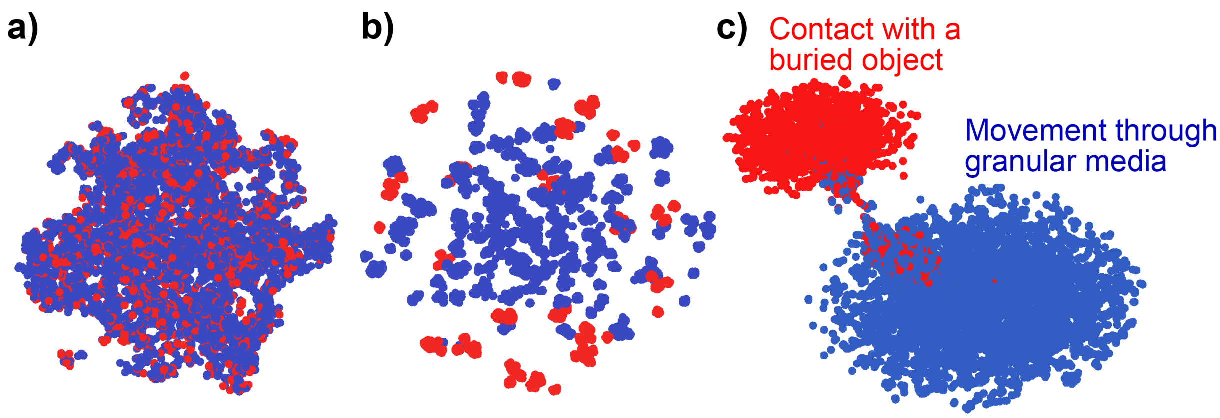


Figure 2.6: Using tSNE for visualization, one can observe the effects of different stages of the tactile perception pipeline. (a) Visualized in 3D space, the feature tensors from Fig. 2.2c are entangled. (b) The outputs from the sparse coding step in Fig. 2.2d are spread out in a high-dimensional feature space. (c) The feature space from the last layer of the model in Fig. 2.2e shows that the data are mostly separated into distinct regions according to contact state.

space (Figure 2.6b). Finally, the output of the 50-dimensional final layer of the RNN was visualized as two distinct clusters (one cluster per state) with a transitional bridge between the two clusters consisting of a mixture of feature tensors from both classes. As the RNN with GRUs reduced the feature space from 400 dimensions to 50 dimensions, redundant features that were not critical to detecting fingertip contact with a buried object were automatically excluded.

In Tables 2.2 and 2.3, we showed how accuracy improves with each pre-processing step in our proposed pipeline. In particular, we illustrated that the use of sparse coding prior to sensor fusion within a fully connected RNN input layer improves the performance of both linear and nonlinear models ("(d)+" cases). The motivation for introducing this unsupervised pre-training step was to improve generalization capabilities and to capture intricate dependencies between parameters. Sparse coding selects model hyperparameters that separate entangled tactile features (Figure 2.6a) such that the features that distinguish different fingertip contact states are as distinct as possible (Figure 2.6b). Given the low off-sample error of the SF-RNN for the test data (Table 2.3), the unsupervised pre-training step appears to have selected features that are near an optimal solution for high classification accuracy with reduced tendency for overfitting. Furthermore, the process of tuning hyperparameters is greatly simplified when sparse features are employed. The main disadvantage of unsupervised pre-training is that it is time consuming, especially for vibration data (Table 2.3). If offline pre-training time is not a major concern, then unsupervised pre-training is recommended.

2.7.2 Study of Model Robustness

As described in Section 2.6.3, for the purposes of a robustness study, we collected additional test data that intentionally varied the granular media types and finger-object contact conditions. We took multimodal models that were originally trained and tested with the cottonwood beads using all three modalities simultaneously and we tested them directly (without any hyperparameter tuning) on the new test dataset. The results of this robustness study are shown in Table 2.4. The robustness study was only performed on multimodal models that demonstrated decent performance with the original cottonwood bead dataset (i.e. "(d)+" cases that used sparse coding and all three modalities simultaneously in Table 2.3).

As shown in Table 2.3, the multimodal models that used sparse coding and all three modalities achieved 99.7% and 98.7% test accuracy for the linear SVM and SF-RNN methods, respectively. Due to the purposeful changes in the granular media types and finger-object

Table 2.4: Robustness Study: Model accuracy is reported as mean ± standard deviation in %. The models were trained on cottonwood beads, but evaluated on the expanded test dataset collected with different granular media types and finger-object contact conditions.

Material	Poppyseeds	Plastic pellets	Plastic BBs	Wooden beads (pine)	Wooden beads (cottonwood)
(d)+Linear SVM	46.0±0.7	53.7±0.7	58.5±0.6	52.4 ±0.7	52.2±0.7
(d)+RNN (SF-RNN)	74.8±0.5	78.7±0.5	88.4±0.4	87.0±0.4	86.2±0.4

contact conditions, the test accuracy of the SF-RNN model decreased, as expected, to as low as 74.8% for poppyseeds and only to 88.4% for plastic BBs. However, the test accuracy of the linear SVM model plummeted, only achieving as high as 58.5% accuracy for plastic BBs (Table 2.4). This is likely due to the inability of linear models to capture useful nonlinear relationships between rich, multimodal tactile signals and fingertip contact states.

On the other hand, nonlinear SF-RNN models were less influenced by variability in the input data. Test accuracy exceeded 74% for all granular media types and exceeded 86% for all granular media having particle diameters greater than or equal to 6 mm. This suggests that the model trained with cottonwood beads was capable of encoding useful features from a range of granular media types. SF-RNN models trained with cottonwood beads data collected with 4 cm/s exploratory movements (Tables 2.2 and 2.3) were robust to changes in granular media type and also fingertip speed, which was decreased to 2 cm/s for the test dataset (Table 2.4).

2.7.3 Belief Map Representation of Tactile Information

The state estimates from the SF-RNN model are simultaneously fed into a belief map as samples of the spatial distribution within the exploration space. Figure 2.7 shows the evolution of a belief map for a 2D, top-down view of the exploration area as the robot finger moves during a 2 sec trial. Upon initial contact with the buried object, the belief map reflects an increased probability of contact with a buried object (red region appears in Figure 2.7a). As the perceived contact persists during the finger movement, the red region of high probability increases in area (Figure 2.7b). Upon a return to free fingertip movement through the granular media, the belief map has been refined to show a region having high probability of contact with a buried object (dark red regions in Figure 2.7c) as well as adjacent regions that are believed to consist of granular media only (dark blue regions in Figure 2.7c).

Upon completion of the first exploratory movement, the width of the dark red region is approximately 10 cm (Figure 2.7c). The cylinder diameter is 9.1 cm and the chord length from first contact to last contact was approximately 5.4 cm. While the belief map could use refinement from additional exploratory movements of the fingertip, the dark red region is of the proper order of magnitude in size and could be used to plan future exploratory movements that reduce uncertainty in the shape and size of the visualization.

2.8 Discussion

2.8.1 Notable Trends in the Tactile Data

To further analyze the results in Table 2.2 and 2.3, we compared the response of each tactile sensor modality for each contact state (Figures 2.8, 2.9). In Figure 2.8, average trends in vibration amplitude and internal fluid pressure are shown for 500 trials without a buried object (green) and 250 trials with a buried object (blue during free movement through granular media and red during contact with a buried object). Since the tactile sensor data are positively skewed, we report median values along with 1st and 3rd quartile bands in Figures 2.8 and 2.9. The interquartile spread reflects inter-trial variability.

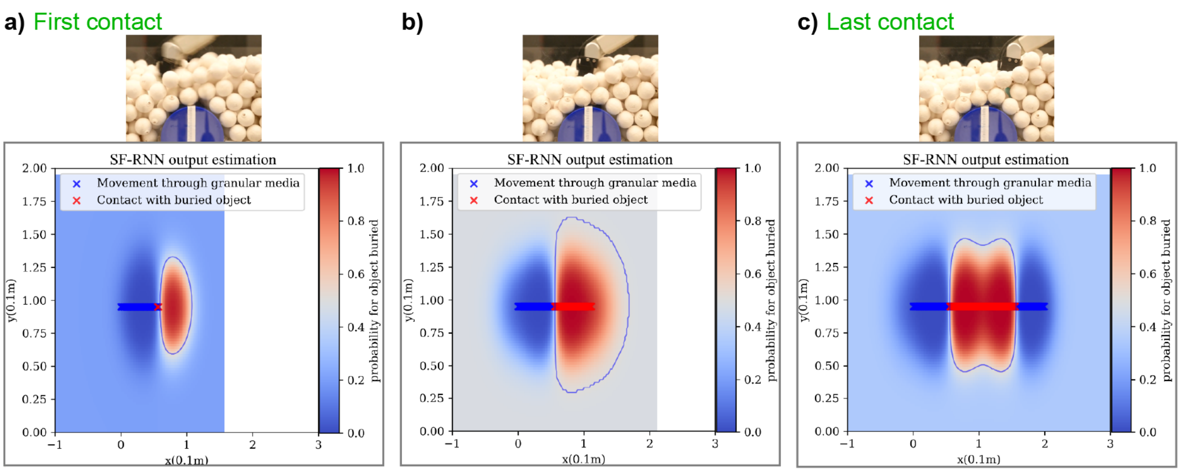


Figure 2.7: As the fingertip moves across a buried object (top row), a belief map (bottom row) is updated using fingertip contact state estimates from the SF-RNN model. Snapshots are shown as the fingertip (a) makes first contact with the buried object, (b) passes across the top of the object, and (c) makes last contact with the object and returns to free movement within the granular media.

2.8.1.1 Susceptibility of Different Tactile Modalities to Noise Induced by Movement within Granular Media

The BioTac sensor uses a fluidic mechanotransduction medium that enables the measurement of vibration and internal fluid pressure via high-pass and low-pass filtered signals, respectively, from a hydro-acoustic pressure sensor. For most tasks conducted in open air environments, the sensitivity of the BioTac's "fast-adapting" hydrophone receptor to stimuli such as contact transients and vibration is especially useful.

As the robotic finger rakes through the granular media, the BioTac sensor sensitively records vibrations transmitted through the fluid-filled fingerpad caused by repeated interactions with granular media particles. As a result, changes in hydrophone data that might be indicative of key contact events with a buried object get drowned out by noise caused by motion relative to granular media (Figure 2.8). In addition, the spread between the 1st and 3rd quartiles of the vibration data increased by a factor of 5 when the fingertip was moved through granular media as opposed to open air. Perhaps with advanced signal processing techniques, changes in vibration magnitudes observed near "first contact" and "last contact" events could be extracted for improved tactile perception.

The BioTac's "slow-adapting" receptors were also susceptible to noise induced by interactions with granular media. The interquartile spread increased by a factor of approximately 5, 3, and 6 for electrodes #1, 7, and 11 at the fingertip, respectively (Figure 2.9). The interquartile spread for the internal fluid pressure modality increased the most (by a factor of 15) when the fingertip was moved through granular media as opposed to open air. Nonetheless, larger sustained fingerpad deformations were clearly observed in the "slow-adapting" receptor data when the finger was in contact with the buried object (Figures 2.8b, 2.9). This

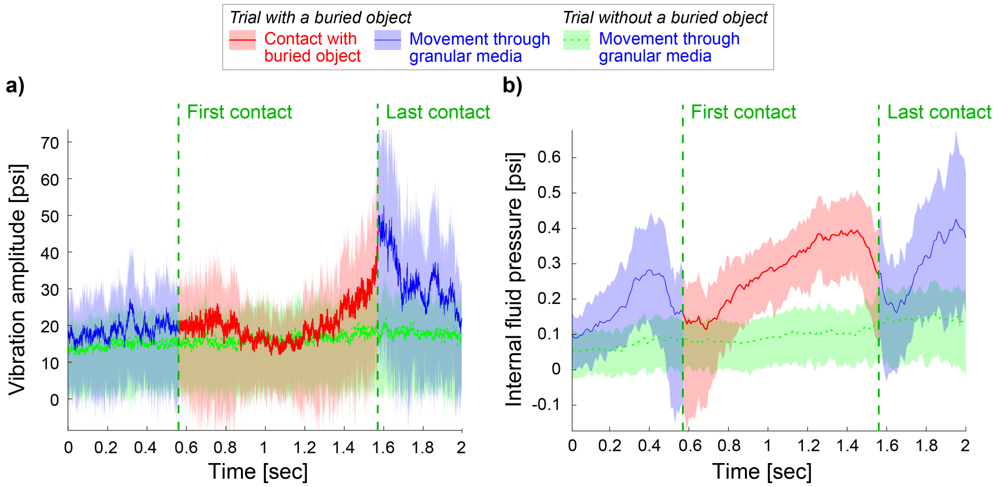


Figure 2.8: (a) Vibration amplitude and (b) internal fluid pressure are shown as median values with 1st and 3rd quartile bands for both fingertip contact states during all trials with and without a buried object.

was especially true for electrodes at the tip of the finger that would most likely have direct contact with a buried object (electrodes #1, 7, and 11 in Figure 2.9b).

2.8.1.2 Information Encoded by Different Tactile Modalities During Movement within Granular Media

Based on studies of human touch, it is believed that coarse textures and local shape are primarily perceived through spatial cues while fine textures are perceived through vibratory cues [74, 75]. Thus, we hypothesize that spatial information from fingerpad deformation is especially useful for distinguishing between sustained forces induced by a buried object and repetitive contacts with much smaller granular media particles. While changes in fluid pressure and electrode impedance both reflect fingerpad deformation, electrode impedance provides an array of spatial information that contains more information than a single scalar fluid pressure signal for the entire fingertip.

Despite the movement of granular media particles relative to the fingertip, the electrodes are capable of encoding sustained forces, whether due to direct contact with a buried object or granular media jamming preceding direct contact. In our experiments, the maximum height difference between the depth of the object and the depth of the exploratory movement was 1 cm. Thus, the distal aspect of the fingertip made primary contact with the buried object. When contacting the buried object, the fingerpad compresses at the fingertip, as reflected by the increase in impedance of electrode #7 (Figure 2.9b). This results in a bulging of the fingerpad that is reflected by the decrease in impedance of neighboring electrodes #1 and #11 on the sides of the fingertip. Such information could enable advanced behaviors within granular media such as contour-following, obstacle avoidance, and dynamic trajectory planning. Additionally, prior efforts to perceive the size and shape of local geometries using electrode impedance data could potentially be extended from open air scenarios [40] to granular media scenarios.

The "fast-adapting" hydrophone receptor appears to be sensitive to transients, especially the punctuated "last contact" event when the fingertip slips off of the buried object (Figure 2.8a). Videos of the experiment suggest that the fingertip, moving at a constant rate in the testbed, vibrates after suddenly losing its sustained contact with the buried object and getting flung into the surrounding granular media particles. Although further investigation is required, we surmise that the sudden release of the fingertip from contact with the buried object caused the temporary peak in vibration amplitude and short-lived increase in internal fluid pressure that were observed shortly after the "last contact" event (Figure 2.8). Vibration data may be useful for identifying transitions in fingertip contact states, although significant changes in vibration magnitude alone do not seem to occur with the "first contact" event.

2.8.1.3 Haptic Sensitivity to Granular Media Jamming

For transitions of fingertip contact state from "movement through granular media" to "contact with a buried object," the SF-RNN model estimated the transitions prematurely. The transition to "contact with a buried object" occurred approximately four timesteps (40 ms) prior to the true "first contact" event (Figure 2.5d). These errors are likely caused by our strict labeling policy during the training process. Fingertip contact state labels were assigned according to *direct contact* with a buried object in the absence of granular media (Figures 2.4c, 2.5). As a result, the "contact with a buried object" labels would not account for indirect contact with a buried object via particle force chains caused by granular media jamming [10]. This premature change in the fingertip contact state estimate just prior to direct "first contact" (theorized from "ground truth" trials in open air) makes intuitive sense when considering granular media jamming effects. In contrast, the change in fingertip contact state occurs in a timely manner at the "last contact" event. Approximately three timesteps (30 ms) were needed for the SF-RNN model to clearly transition away from a state of "contact with a buried object," but the drop in the state label's probability was immediate (Figure 2.5d).

Granular media jamming may have contributed to the increase in impedance for electrode #7 at the fingertip just prior to "first contact" (Figure 2.9b), but further investigation is required. Granular media jamming effectively injects uncertainty into the contact state estimate and the resulting spatial belief map. For instance, a buried object might be visualized as larger than its actual size due to premature perception of contact due to granular media jamming. One could reduce this uncertainty by accounting for the effects of jamming,

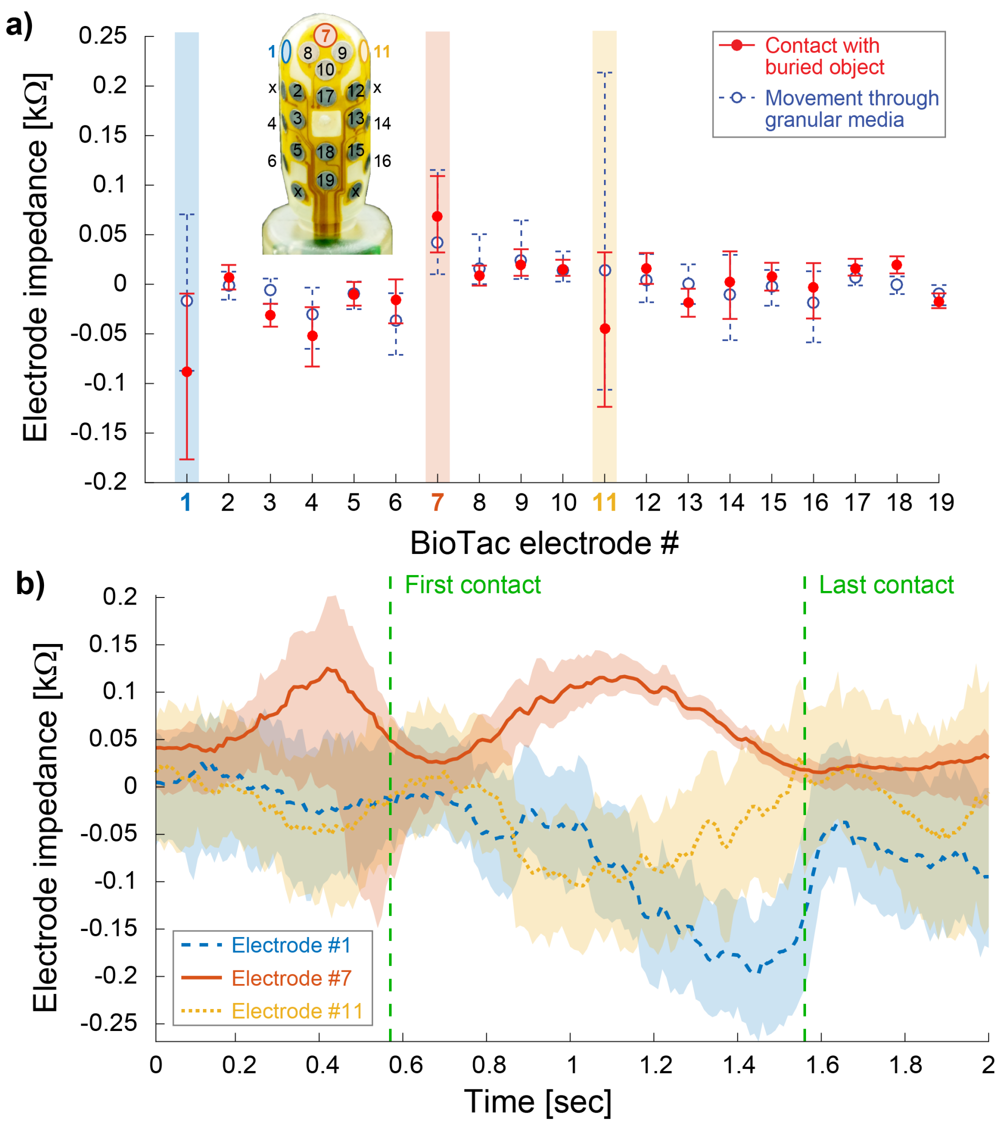


Figure 2.9: Electrode impedance is shown as median values with 1st and 3rd quartile intervals. (a) Trends for all 19 electrodes are shown for both fingertip contact states during all trials with and without a buried object for the same time window (theoretical contact period). (b) Trends for three electrodes at the fingertip (#1, 7, 11) are shown for all trials with a buried object for the entire 2 sec period.

especially how granular media characteristics and fingertip depth may prematurely trigger a transition in the fingertip contact state. The effect of granular media jamming is the subject of a separate investigation and is outside the scope of this work.

2.8.2 Real-Time Tactile Perception and Visualization for Human Control of Remote Robotic Systems

The supplemental video includes a real-time demonstration of the belief map update process, for which the graphic rendering was updated every 100-300 ms. During real-time execution, we combine tactile perception outputs from the SF-RNN model (fingertip contact state) with dynamic robot kinematic information (fingertip location) to produce a continuously updated 2D belief map. Since all prior haptic information is already encoded into the spatial distribution, the updates of the underlying spatial statistics of the belief map can be treated as improvements in the robot's belief with the benefit of new haptic experiences. The long term memory duration of the belief map can be adjusted according to the needs of the teleoperator.

A belief map conveys environment characteristics as visualizations of spatial probability distributions across a search space, from which a human teleoperator could infer the location, size, and orientation of a buried object. Information from multiple sensorized, semi-autonomous robots that are performing haptic search tasks in parallel could be combined into a single map. This could enable increased productivity for a single human telesupervisor as compared to a single human teleoperator of a leader-follower system. For safety critical tasks, the human may need to manually teleoperate the robot, as in surgical applications [76, 77] or the handling of high-consequence materials [78]. Custom haptic interfaces could be used to control the robot and display contact force information to a teleoperator's hand

[43]. In such an instantiation, the belief map would be used to convey remote environment characteristics and little autonomy would be given to the robot.

To increase the level of robot autonomy, the belief framework could be used to construct and update a map of an unknown environment covered by granular media. The ability to perceive within granular media could be used to extend existing haptic [79], and contact SLAM approaches [80]. Moreover, a belief map could be used to generate future exploration trajectories via optimal control [81] or ergodic control [82], which could enable belief maps to be updated more efficiently.

In the shared control or semi-autonomous condition, task execution and control could be shared between humans and robots. For instance, utilizing impedance control, a robot may follow a trajectory supervised by a human teleoperator while still responding naturally to external perturbations [83]. In another example, the robot might be responsible for controlling the orientation of the end-effector while the human controls the position [84]. Alternatively, a human may take the leader role for high-level, large-scale movements toward a region of interest while the robot manages the low-level, precise movements of the end-effector that rely critically upon local sensing [85, 86].

To support human-robot coordination and better utilize the robot, Gombolay et al. recommended to equip robots with as much autonomy as possible and greater robot authority over task allocation [87], such as telling the human operator when each sub-exploration could be started and finished. We posit that trajectory planning methods and algorithms to assess confidence based on the belief map will be necessary to improve the quality of human-robot interaction during exploration within granular media. However, it is important to enact appropriate arbitration between the human leader and robot follower during shared control, as the human domain expert must always retain final authority for safety purposes [83].

2.9 Conclusion

Haptic exploration within granular media is an essential task for robots designed for interacting with delicate and/or dangerous buried objects. Locating and retrieving buried devices requires haptic intelligence in order to carefully maneuver within the granular media, detect, and interact with buried objects. As such, we focused on the fundamental capability of haptically perceiving contact with a buried object and distinguishing such an event from ubiquitous contact with surrounding granular media particles.

2.9.1 Summary of Contributions

In this work, we established a multimodal learning architecture based on Recurrent Neural Networks with Long Short-Term Memory units that incorporates multiple tactile sensing modalities (vibration, internal fluid pressure, and electrode impedance as a proxy for fingerpad deformation). Our SF-RNN model architecture learned to fuse sparse representations of multimodal tactile sensor data and performed robustly on novel test data that included variations in granular media type and particle size, fingertip orientation, fingertip speed, and object location along the linear fingertip trajectory. We also introduced a belief map representation to visualize tactile information from an SF-RNN model for a human teleoperator. During real-time implementation, the tactile perception model and belief map were updated every 150 ms and every 100-300 ms, respectively.

Analyses of the different tactile modalities showed that fingerpad deformation was the most informative tactile sensor modality for distinguishing between free movement of the fingertip through granular media and contact with an object buried in granular media. With signal processing, filtered vibration and internal fluid pressure responded to first and last contact sensitively and improved perception accuracy beyond that achieved using electrode impedance data alone.

2.9.2 Limitations and Future Work

Here, we lay the foundation for more complex, semi-autonomous robotic behaviors such as exposure, excavation, and retrieval of buried objects. To improve upon this foundation, we highlight some limitations of the current work. For instance, our robot exploration movements were limited in speed and distance because of the size of our granular media testbed. In order to protect the exposed joints of the robot gripper, we used a 3D-printed adapter when performing exploratory movements amongst small diameter particles and only plunged the gripper into a testbed of large wooden beads.

Further investigation is needed to assess the effects of experimental variables such as object shape, object rigidity, depth of burial, and more complex, multi-digit exploratory movements on tactile perception. Moreover, accounting for tactile sensitivity to granular media jamming could help to reduce uncertainty in the belief map and better inform teleoperators. In this work, a robot fingertip was lightly raked across a buried object using a preplanned trajectory. Ideally, robots with haptic intelligence could be programmed to conservatively stop and autonomously replan their exploratory movements as soon as contact or granular media jamming is haptically perceived.

Further upgrades of the belief map updating algorithm (Appendix A.4) could include the use of local length-scale kernels with optimizations such as unsupervised, automatic kernel selections and placement [88, 89]. Fast kernel approximation methods [36] could be implemented to increase computational speed, thereby enabling the visualization of large 3D scenes. With regards to screen-based control interfaces, an attitude-based controller [90] could be integrated to allow teleoperators to use their natural movements and gestures to change the viewing angle of the belief map and command the robot.

CHAPTER 3

Tactile SLAM for Autonomous Haptic Localization and Mapping of Objects Buried in Granular Materials

3.1 Abstract

Robots are expected to operate autonomously in unstructured, real-world environments, for tasks such as locating buried objects in search and rescue applications. When robots operate within opaque granular materials, tactile and proprioceptive feedback can be more informative than visual feedback. However, since tactile measurements are local and sparse, it can be difficult to efficiently build a global, tactile-based model of a search area. In this work, we developed a framework for tactile perception, mapping, and haptic exploration for the autonomous localization of objects buried in granular materials. The haptic exploration task was performed within a densely packed sand mixture using a sensor model that accounts for granular material characteristics and aids in the interpretation of interaction forces between the robot and its environment. The haptic exploration strategy was designed to efficiently locate a buried object and refine its outline while simultaneously minimizing potentially damaging physical interactions with the buried object. Coverage path planning techniques were used to select haptic exploration movements from candidates that aimed to reduce map uncertainty. A continuous occupancy map was generated that fused local,

sparse tactile information into a global map. We demonstrated our framework in simulation and with a real robot with granular materials.

3.2 Introduction

Granular materials, such as sand, grain, salt, snow, and powder, are ubiquitous in nature. Humans can easily locate and identify objects buried within granular materials using tactile and proprioceptive feedback rather than visual feedback. Yet, it remains a challenge to transfer these human capabilities to robots for delicate operations such as search and rescue, archaeology, seafloor exploration, and demining.

Research in robotic grasp and manipulation has traditionally focused on rigid objects more than deformable objects or granular materials. One reason is due to the difficulty of modeling the complex dynamics of deformable objects and granular materials. Another reason is that tactile-based inference is often limited by sensing (hardware) and perception (algorithm) challenges. For instance, tactile sensors are typically unimodal in nature and one must carefully control finger-object contacts in order to build accurate models of objects and the environment. Additionally, tactile-based inference is limited by the narrow "field of view" of a fingertip. While a single camera image can provide an informative global estimate, a single tactile "image" is highly localized. Tactile information from multiple exploratory movements must be fused into a global model.

Jia and Santos presented the first use of tactile sensors on real robot fingertips making direct contact with granular materials to infer the contact state of the fingertips [21]. Building upon that prior work, we now develop *a tactile-based framework for tactile perception, mapping, and haptic exploration in granular materials.* We consider the effects of

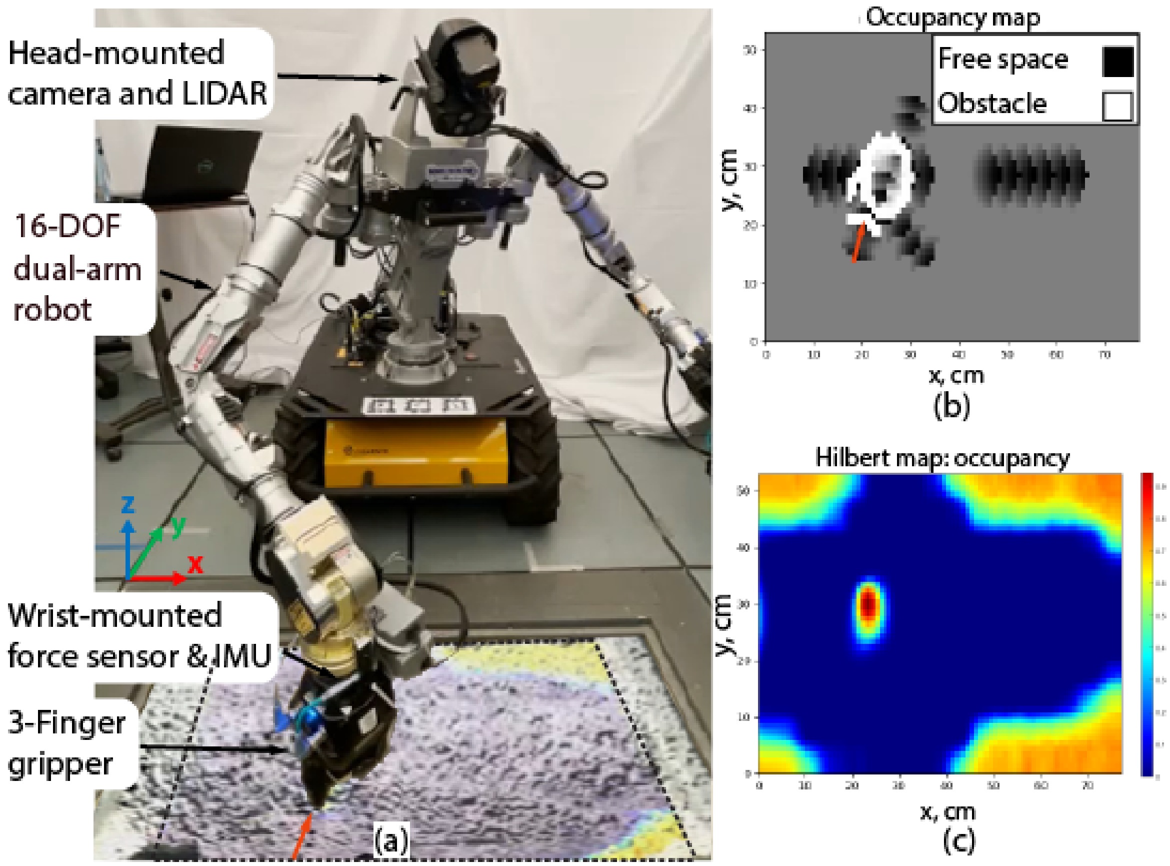


Figure 3.1: Localization and mapping in granular materials: (a) A robot haptically explores a sand mixture and updates an (b) occupancy map in log odds using fingertip forces interpreted by a sensor model that accounts for granular material characteristics. (c) Local, sparse tactile data are fused into a global Bayesian Hilbert Map to localize a buried object and augment (a) colorized point cloud data from a head-mounted sensor.

granular material jamming on tactile perception, introduce advancements to map development for granular material environments, and propose a path planning algorithm for haptic exploration in granular materials.

The research problem is depicted in Figure 3.1. A sensorized robot is used to haptically search for an object buried in granular materials and estimate the location, orientation, and footprint of the buried object via occupancy mapping and tactile measurements. The contributions of our work that extend tactile simultaneous localization and mapping (SLAM) to haptic exploration within granular materials are:

1. The design of a sensor model for tactile measurements that accounts for different environmental interaction states between the robot fingertip and granular materials as well as the local flow characteristics of granular materials.

2. The design of a continuous occupancy map algorithm, inspired by models of dry granular materials in response to "plate drag," that integrates localized, sparse tactile spatial information into a global map.

3. The design of a haptic exploration strategy for use within granular materials that maximizes localization efficiency from tactile measurements while minimizing forceful interactions with buried objects.

4. The demonstration of our haptic exploration framework on a haptic exploration task in simulated and real settings.

This paper is organized as follows: Section 3.3 provides background information on soil-tool interactions and tactile SLAM. Section 3.4 describes our tactile-based framework for perception, mapping, and haptic exploration in granular materials. Section 3.5 describes experiments in simulation and with a real robot. Section 3.6 addresses future work.

3.3 Related Work

We provide an overview of "plate drag" simulations and physical experiments that are relevant for our work, in which a rigid, rectangular robot fingertip is dragged through granular materials while searching for a buried object. We also summarize the state-of-the-art in tactile SLAM techniques, which do not consider granular material environments.

3.3.1 Soil-tool Interaction: Plate Drag

Soil-tool interactions, specifically physical interactions between dry granular materials and plates, have been studied in simulation [11] and on physical testbeds over the past several decades [1, 12]. As a plate advances across the surface of granular materials, flowing particles are confined to a wedge-shaped region or "failure zone" in front of the plate. Due to the advancement of the plate, the particles move forward in the direction of plate movement and upward toward the free surface of the granular materials. Materials outside of the wedge-shaped flow region remain nearly at rest.

As proposed by Wick and Perumpral [1], the 3-D analytical model for the granular material failure zone ahead of the plate is comprised of a central wedge whose width is equal to that of the plate and which is bound by a crescent shape on each side. The failure zone is assumed to be symmetric about the axis along which the plate is moved. Separating the flowing and stable regions, a "shear (force) band" originates from the bottom edge of the plate and extends upwards to the free surface of the granular materials. After the shear band fully develops, both the angle of the shear band and the drag force increase as the plate advances. When the drag force reaches the force required to disturb the more stable surrounding materials, a new shear band nucleates with a small wedge angle that grows as the plate advances. The shear band nucleation cycle repeats as the plate rakes through the granular materials [11]. A visualization of the failure zone concept applied to our research problem is shown in Figure 3.2.

3.3.2 Tactile SLAM

Recently, tactile sensors have attracted increased attention from the SLAM community [91, 92, 93] since the quality of sensory feedback from camera- and proximity-based sensors degrades under extreme lighting conditions and when aimed at highly reflective or absorptive surfaces [94]. For example, mobile robots have been outfitted with arrays of biomimetic whiskers to improve pose estimates [91]. A whisker-based approach to tactile SLAM called Whisker-RatSLAM [92] advanced navigation through the implementation of bio-inspired strategies via RatSLAM, a SLAM algorithm inspired by a model of a rodent hippocampus. ViTa-SLAM [93] introduced an extension to Whisker-RatSLAM that harnesses both visual and tactile information for performing SLAM.

Apart from the use of tactile SLAM for mobile robots described above, contact-based SLAM has also been applied to robot arms for object shape estimation. Zhang and Trinkle used particle filters for object shape estimation during object grasping [95]. Abraham et al. applied ergodic exploration for shape estimation with a low-resolution binary contact sensor [82]. Using a robot hand equipped with tactile arrays, Strub et al. demonstrated that the haptic learning of object shape could be approached as a tactile SLAM problem [96]. Based on batch-SLAM, Suresh et al. formulated a method for estimating the shape and pose of a planar object from a stream of tactile measurements [97]. While tactile SLAM has been used

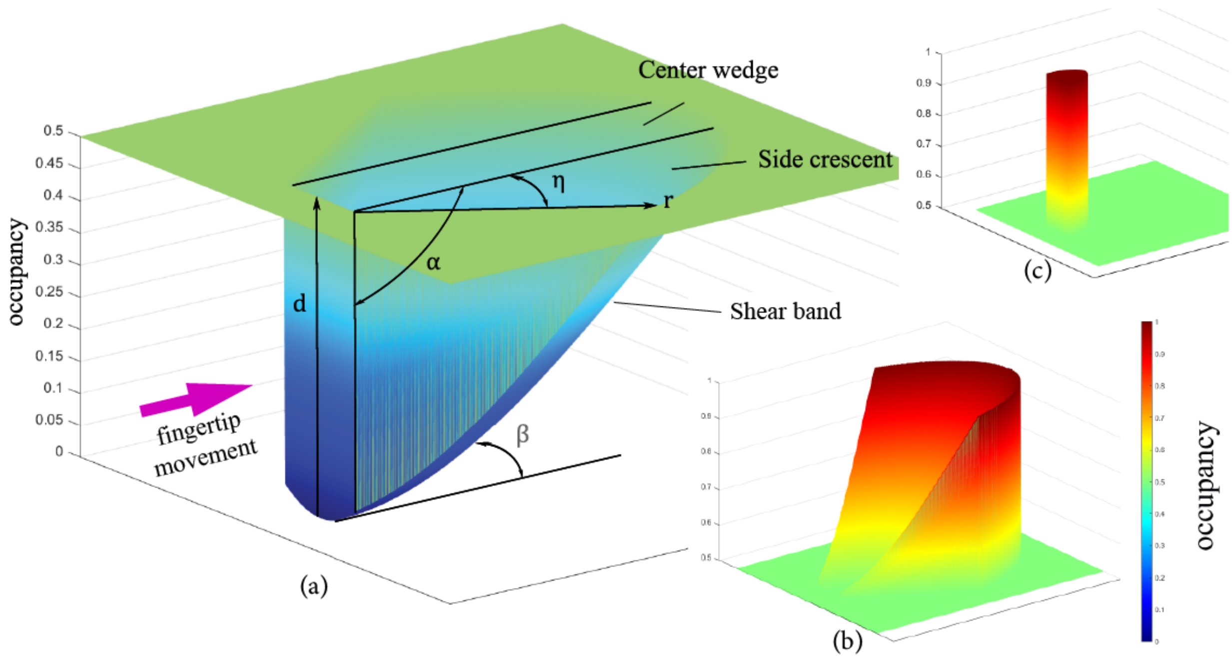


Figure 3.2: Visualization of the kernel-based sensor model described by Eq. 3.3 for three environmental interaction states: (a) free movement through granular materials, (b) granular material jamming, and (c) direct contact with a buried object. The sensor model was designed based on the theoretical failure zone. Per Wick and Perumpral [1], measurement range r and side crescent angle η of the failure zone are calculated as $r = d(cot\alpha + cot\beta)$ and $\eta = sin^{-1}(\frac{6.03+0.460r+0.0904\alpha}{r})$, respectively.

for mapping with mobile robots and object shape estimation with robot grippers and arms in open air environments, *tactile SLAM has not yet been applied within granular material environments.*

3.4 Framework for Tactile SLAM within Granular Materials

Here, we establish our assumptions, tactile perception methods, and mapping principles for a novel framework that extends tactile SLAM to haptic exploration within granular materials (Figure 3.3).

3.4.1 Sensor Measurements and Assumptions

In order to locate an object buried in granular materials, we use sensory feedback from a robot end-effector as its fingertip performs exploratory movements (EMs). Given streams of tactile measurements, we build an occupancy map, estimate the 2-D outline of the buried object, and estimate the location of the object's centroid in real-time.

At every timestep t we collect a set of sensor measurements $\mathbf{z}_t$ that are comprised of the 2-D coordinates of the fingertip location $\mathbf{x}_t$ in the x-y plane (Figure 3.1), the 2-D environmental interaction force on the fingertip $\mathbf{f}_t$ in the x-y plane, and the environmental interaction state Θ_t at the fingertip. There are three possible interaction states for the fingertip: free movement through granular materials $\theta_{t,0}$, granular material jamming $\theta_{t,1}$, and direct contact with a buried object $\theta_{t,2}$. The interaction states will be further described in Section 3.4.3.1.

$$\mathbf{z}_t = \{ \underbrace{\mathbf{x}_t \in \mathbb{R}^2}_{\text{Fingertip location}}, \underbrace{\mathbf{f}_t \in \mathbb{R}^2}_{\text{Fingertip force}}, \underbrace{\Theta_t = \{\theta_{t,0}, \theta_{t,1}, \theta_{t,2}\}}_{\text{Environ. interaction state}} \} \tag{3.1}$$

The robot's fingertip location is obtained directly from the robot's kinematic state. The environmental interaction force vector is estimated by a force observer model that fuses low-pass filtered inertial measurement unit (IMU) data and force/torque sensor data from the robot's wrist. However, information from fingertip sensors could also be used. The force observer model, specific to our experimental procedure, is described in Section 3.5.3.2. The environmental interaction state for the fingertip is estimated using rates of change in fingertip force, as described in Section 3.4.2.

For simplicity, we assumed the following:

1. Due to the shape of the rigid, rectangular robot fingertip and the densely packed nature of the heterogeneous sand mixture, plate drag and granular material failure zones will be similar to those observed in the soil-tool interaction literature [12].

2. For slow movements of the fingertip, inertial effects on the force observer model will be negligible. As such, force observer measurements will be attributed entirely to environmental interaction forces on the fingertip.

3. For a haptic exploration strategy that minimizes forceful interactions with the buried object, the object will remain stationary relative to the surrounding densely packed granular materials during haptic exploration.

4. Since the fingertip will be moved at a constant fingertip depth (z-value; Figure 3.1a) through the granular materials, the tactile perception, mapping, and path planning methods will be described in the 2-D, x-y space.

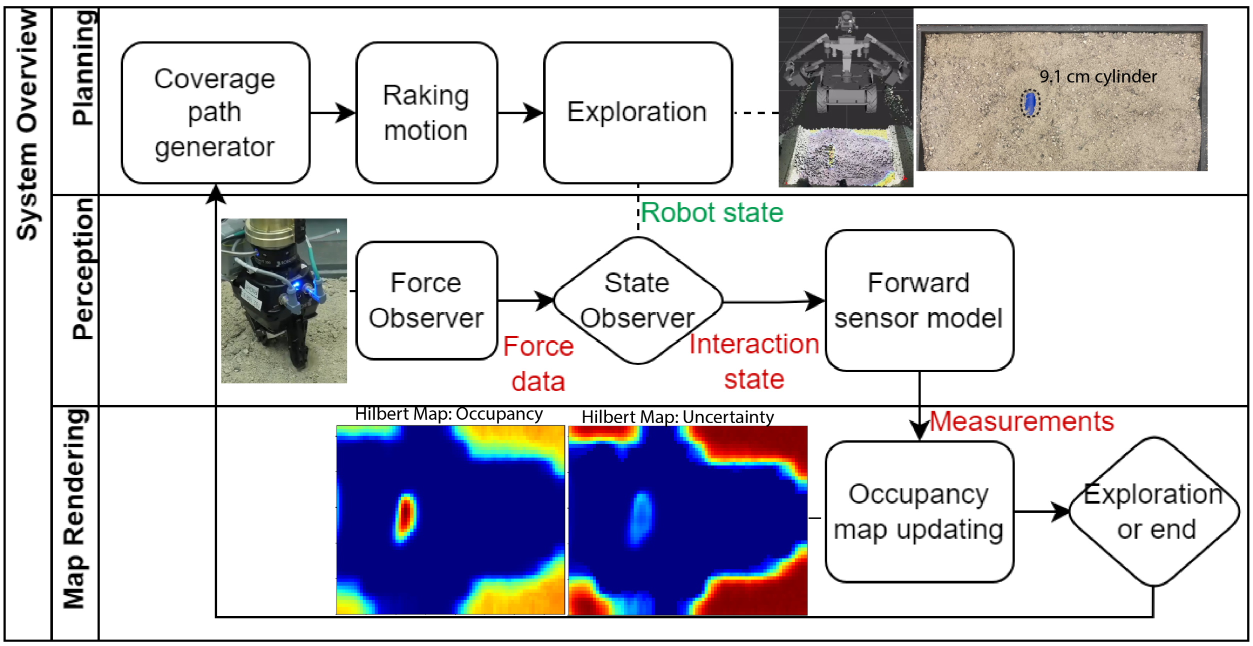


Figure 3.3: Framework for tactile perception, mapping, and haptic exploration for the autonomous localization of objects buried in granular materials.

3.4.2 Tactile Perception

In a prior study by Jia and Santos [21], the tactile perception pipeline considered only two environmental interaction states: free movement through granular materials and direct contact with a buried object. Here, we consider a third environmental interaction state: granular material jamming. This additional interaction state is necessary in order to avoid a premature, false detection of direct contact that would lead to inaccurate localization of the buried object.

As a plate is dragged through granular materials, a wedge-shaped failure zone forms based on the minimum lateral force required to displace the granular materials [98]. Assuming that all moving granular particles are part of a network of "force chains" [4], in which chain-like groups of particles are held together by mutual compressive forces, the failure zone can also be considered as a large network of force chains capable of transferring forces across the failure zone.

As a fingertip approaches a buried object, the wedge of displaced granular media ahead of the fingertip will contact the buried object before the fingertip makes direct contact. Due to granular material "jamming" caused by the network of force chains, the contact between the failure zone and buried object causes a surge in force on the fingertip even though the fingertip has not directly contacted the buried object. In a sense, granular material jamming endows tactile sensors with a proximity sensor-like capability when deployed within granular materials.

Given the challenge of modeling the formation of force chains during granular material jamming, we designed experimental data-driven criteria in order to detect the onset of granular material jamming. In a preliminary study, we measured environmental interaction forces

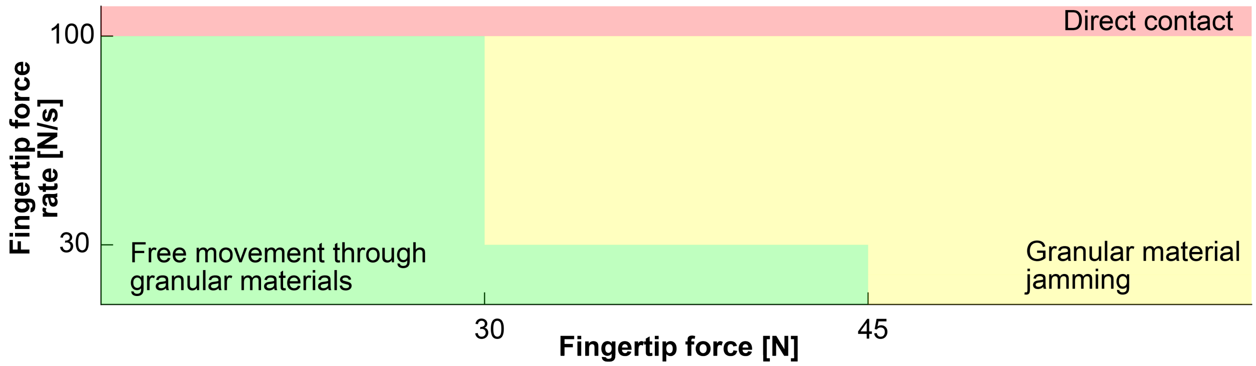


Figure 3.4: Environmental interaction states were labeled according to their coordinates in a 2-D space comprised of fingertip force and force-rate.

when a real robot fingertip was raked through granular materials. Figure 3.4 shows how environmental interactions were labeled in real-time according to their coordinates in a 2-D space comprised of fingertip force and force-rate. Interactions progressed from "free movement through granular materials" to "granular material jamming" and, finally, to "direct contact with a buried object" as fingertip force and force rates increased.

3.4.3 Continuous Occupancy Mapping

As with traditional SLAM applications for mobile robots, we assume that the robot arm collects tactile measurements while moving through the environment. However, the robot moves through an environment that is comprised of granular materials rather than air or water. Typically, occupancy grid maps are used for modeling the occupancy state of the environment by dividing the world into a fixed-size grid and applying a Bayes filter to individual cells [99]. The occupancy state of each discrete cell is indicated with a 1 for "occupied" and 0 for "obstacle-free" states.

Unlike proximity-based sensors, tactile sensors provide local measurements that are comparatively sparse. The sparsity of the tactile sensor data makes tactile-based map development quite inefficient. In order to maximize the spatial, state information contained in the sparse tactile measurements, we propose the use of continuous, kernel-based sensor models, inspired by models of dry granular materials in response to plate drag [1, 11]. The use of kernel-based sensor models enables us to calculate the likelihood of the tactile sensor measurements across the occupancy map. In this way, occupancy mapping becomes an optimization problem where the goal is to determine the occupancy map that maximizes the likelihood of the tactile sensor data.

3.4.3.1 Kernel-based Sensor Model

Through the use of a kernel-based sensor model, we fuse sparse tactile measurements $\mathbf{z}_t$ to the local neighborhood of the robot fingertip on the occupancy map. The occupancy map m has a width w along the x-axis and height h along the y-axis. The state of each occupancy map cell is a value representing the probability of the occupancy of that cell ($m \in [0,1]^{w,h}$), where 1 indicates the presence of a buried object and 0 indicates the presence of granular materials only.

The position of the robot fingertip $\mathbf{x}_t$ and the environmental interaction state Θ_t are used to model the granular material failure zone ahead of the moving fingertip. The environmental interaction state Θ_t is a vector of binary variables, where 1 indicates the current state and the remaining elements are set to 0. If $\theta_{t,0} = 1$, no buried object has been detected. This state of free movement through granular materials is analogous to the obstacle-free state in traditional SLAM performed by mobile robots. If $\theta_{t,1} = 1$, granular material jamming has been detected. It is assumed that contact between the failure wedge and the buried object results in increased forces at the fingertip due to granular material jamming. Although not used in our demos, the environmental interaction state $\theta_{t,2} = 1$ can be used to indicate direct contact between the fingertip and buried object.

Inspired by terminology used for proximity sensors, we define "measurement ranges" $r_t = \{r_{t,0}, r_{t,1}, r_{t,2}\}$ that apply an appropriate spatial range of influence for each environmental interaction state element $\{\theta_{t,0}, \theta_{t,1}, \theta_{t,2}\}$. When direct contact with a buried object has been detected ($\theta_{t,2} = 1$), the corresponding measurement range variable is set to zero ($r_{t,2} = 0$) to reflect the highly localized state information at the current fingertip location. Otherwise, the measurement range is defined according to the "rupture distance," the distance ahead of the fingertip at which the shear band intersects the free surface of the granular materials [98]. Based on Wick and Perumpral [1], we model the granular material failure zone as a central wedge bound by two crescents, where the 2-D zone is symmetric about the line of fingertip motion (Figure 3.2). The vertex of the central wedge intersects with the disto-proximal axis of the robot fingertip.

Let $\mathcal{B}(\mathbf{x}_t, r_{t,*})$ represent the neighborhood ahead of the current fingertip location $\mathbf{x}_t$ that corresponds to the granular material failure zone,

$$\mathcal{B}(\cdot) = \{\mathbf{x} \in \mathbf{X} \colon dist(\mathbf{x}, \mathbf{x}_t) \leq r_{t,*}, |\angle(\mathbf{x}, \mathbf{x}_t)| \leq \eta\}, \tag{3.2}$$

where η is the angle subtended by the side crescent adjacent to the central wedge model (Figure 3.2). The angle η and the measurement range $r_{t,*}$ are calculated based on empirical equations reported by Kobayakawa et al. [11] and Wick and Perumpral [1]. The empirical equations are based on plate drag experiments in simulation and on physical testbeds that account for different combinations of plate width, depth d, and angle α with respect to the free surface of the granular materials. Figure 3.2 shows how plate depth d and angle α are related to the angle of the shear band β and the angle η of the side crescent.

For points in the occupancy map $\mathbf{x}$ that are within the neighborhood $\mathcal{B}$ of the current fingertip position $\mathbf{x}_t$, the occupancy probability distribution $p(m_{\mathbf{x}}, \theta_{t,k})$ is approximated using the Gaussian kernel,

$$\begin{aligned} &k = 0, 1, 2, \forall \mathbf{x} \in \mathcal{B}(\mathbf{x}_t, r_{t,k}), \\ &p(m_{\mathbf{x}} = 1, \theta_{t,k} = 1) = a_k - exp(-\gamma_r \|\mathbf{x} - \mathbf{x}_t\|^2), \end{aligned} \tag{3.3}$$

where $a_k = 0.5 * (2 + k)$, and $\gamma_r = \ln(2)/r^2$ are tuning parameters related to rupture distance.

We use a mixture model that sums over all Gaussian kernels $k = 0, 1, 2$ in order to account for all three environmental interaction states. The full sensor model that estimates the occupancy state $\hat{m}_{\mathbf{x}}$ for all points $\mathbf{x}$ within the neighborhood of the robot fingertip $\mathbf{x}_t$ is as follows:

$$p(m_{\mathbf{x}} = 1, \Theta_t) = \sum_{k=0}^{2} \theta_{t,k}(a_k - exp(-\gamma_r \|\mathbf{x} - \mathbf{x}_t\|^2)). \tag{3.4}$$

Eq. 3.4 constitutes a model of tactile measurements within granular materials that accounts for different environmental interaction states through correspondence variables (Θ_t, k, a_k, γ_r). Given a stream of measurements $\mathbf{z}_t$, the sensor model generates a stream of datasets $\mathcal{D}_t = \{\mathbf{x}, \hat{m}_{\mathbf{x}}\}^N$ that can be used to update the occupancy map. The datasets include all points $\mathbf{x}$ within the neighborhood $\mathcal{B}$ of the current location of the robot fingertip $\mathbf{x}_t$.

3.4.3.2 Bayesian Hilbert Maps

Hilbert Maps (HMs) [36] and Bayesian Hilbert Maps (BHMs) [100] were developed to be simpler and faster for continuous occupancy mapping than Gaussian Process Occupancy Maps (GPOMs) [101]. HMs evaluate the occupancy probability based on the computation of fast kernel approximations and eliminate the cubical time complexity of GPOMs.

The key insight in the HM framework is that the input dataset is projected into a high dimensional Hilbert space by kernel approximation κ. This approximation is defined by the inner product of features $\Phi(\cdot)$ as $\kappa(\mathbf{x}, \tilde{\mathbf{x}}) \approx \Phi(\mathbf{x})^T \Phi(\tilde{\mathbf{x}})$. The feature vector is a function of the location $\mathbf{x}$ in the 2-D occupancy map, where $\tilde{\mathbf{x}}$ represents one of D spatially fixed points or "hinges" in the map. The location and spatial distribution of the hinges can be selected according to the number and density of estimates that are desired for a given application. The feature vector is computed as

$$\Phi(\mathbf{x}) = [k(\mathbf{x}, \tilde{\mathbf{x}}_1), k(\mathbf{x}, \tilde{\mathbf{x}}_2), ..., k(\mathbf{x}, \tilde{\mathbf{x}}_D)] \in \mathbb{R}^D. \tag{3.5}$$

Each element of the feature vector is based on Gaussian kernels k, with γ_m as the bandwidth to adjust map smoothness. Although random Fourier features or Nystróm features can be used for occupancy mapping, hinge features based on Gaussian kernels are intuitive and result in superior performance [36]. Thus, we use Gaussian kernels here.

The estimated occupancy probability for an arbitrary query point $p(m_{\mathbf{x}_*} = 1|\mathbf{x}_*, \mathbf{w})$ is computed as a sigmoid function:

$$\sigma(\mathbf{w}^T \Phi(\mathbf{x}_*)) = \frac{1}{(1 + exp(-\mathbf{w}^T \Phi(\mathbf{x}_*))}. \tag{3.6}$$

The vector $\mathbf{w} \in \mathbb{R}^D$ represents the weights that describe the discriminative model $p(m|\mathbf{x}, \mathbf{w})$. For HMs, the weights can be optimized by minimizing a regularized, negative log-likelihood via fast Stochastic Gradient Descent methods [36].

For the BHMs used here, the weights follow a multivariate normal distribution $\mathcal{N}(\mu, \Sigma)$ with a mean μ and covariance matrix Σ. The BHM prior distribution $Q(\mathbf{w})$ is defined as

$$Q(\mathbf{w}) \approx P(\mathbf{w}|\mathbf{x}, m) = \mathcal{N}(\mathbf{w}|\mu, \Sigma), \tag{3.7}$$

where μ and Σ are learned iteratively via Expectation-Maximization [100].

Consider samples $\hat{\mathcal{W}}_q = \{\hat{\mathbf{w}}_q\}$ drawn from the estimates for weights $\hat{\mathbf{w}} \sim \mathcal{N}(\mu, \Sigma)$. The probability of a query point $\mathbf{x}_*$ being occupied by a buried object can be calculated as an empirical mean $\hat{m}_{\mathbf{x}_*}$.

$$\begin{aligned} \hat{m}_{\mathbf{x}_*} = p(m_{\mathbf{x}} = 1) &= \int \sigma(\hat{\mathbf{w}}^T \Phi(\mathbf{x}_*)) p(\hat{\mathbf{w}}) d\hat{\mathbf{w}} \\ &\approx \frac{1}{|\hat{\mathcal{W}}_q|} \sum_{\hat{\mathbf{w}}_q \in \hat{\mathcal{W}}_q} \sigma(\hat{\mathbf{w}}_q^T \Phi(\mathbf{x}_*)) \end{aligned} \tag{3.8}$$

Since the query point $\mathbf{x}_*$ can be any point in the space, we can build a *continuous* occupancy map based on BHM. Additionally, the uncertainty of the occupancy probability $\hat{std}_{\mathbf{x}_*}$ can be approximated in the same way as the empirical standard deviations from BHM.

Once an occupancy map has been generated, the estimated outline of the buried object can be presented as a level curve $\mathcal{C} = \{\mathbf{x}_1, ..., \mathbf{x}_n\} \in \mathbb{R}^2$ of equal occupancy levels $\hat{m}_{\mathbf{x}}$ around the estimated location for the centroid of the buried object $\hat{\mathbf{x}}_{obj}$ at the current timestep t. The estimated centroid $\hat{\mathbf{x}}_{obj}$ is defined as a pair of (x, y) coordinates in the x-y plane of the free surface of the granular materials (Figure 3.1).

3.4.4 Haptic Exploration Strategy

The haptic search for objects buried in granular materials involves two related problems: (i) planning and selection of exploratory movements in order to acquire new sensor measurements, and (ii) estimation of object features, such as centroid and outline, from occupancy maps. The haptic exploration strategy is divided into two phases: before and after the first pair of interactions with a buried object.

Phase 1: Before the first pair of interactions with a buried object, the well-established Coverage Path Planning (CPP) approach [102] is used to generate parallel, straight-line candidate EM paths that cover the entire search region of interest S. As long as the burial

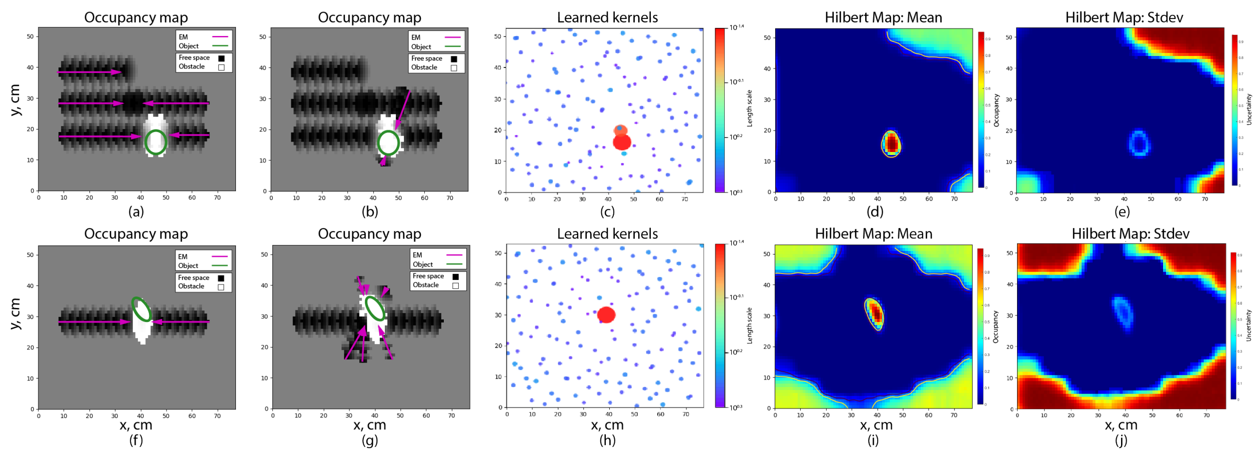


Figure 3.5: Simulation results for the haptic localization of randomly placed objects having circular (top row) and elliptical (bottom row) outlines in the x-y plane: Occupancy maps in log odds are shown for Phase 1 (a,f) and Phase 2 (b,g). Optimal kernel shapes and placements were learned (c,h). Final Bayesian Hilbert Maps are shown for occupancy (d,i) and uncertainty (e,j).

depth of an object overlaps with the depth of a robot fingertip during raking, interaction of the fingertip with the buried object is guaranteed if all CPP candidate paths are traversed. A binary search method is used to generate raking trajectories for the robot fingertip until either granular material jamming ($\theta_{t,1} = 1$) or contact with a buried object ($\theta_{t,2} = 1$) is detected at the fingertip. For our demonstrations, we conservatively stopped the motion of the fingertip as soon as granular material jamming was detected in order to minimize disturbances to the buried object.

After the first interaction with a buried object, the next EM approaches the buried object from the opposite direction along the same straight-line path that led to the first interaction. The fingertip begins the EM from the opposing boundary of the search region in order to establish two interactions with the buried object from opposing directions.

Phase 2: For the remainder of the exploratory movements, we propose a haptic exploration strategy that is aimed at maximizing the localization gain in new tactile measurements, enhancing the efficiency of the haptic search, and minimizing the total number of interactions with the buried object. We use a search strategy that leverages information from the current Bayesian Hilbert Map in order to select the next-best linear exploratory path.

For search efficiency, we exclude areas with high uncertainty by shrinking the search region by 25% and centering the new map around the latest estimate of the buried object's centroid $\hat{\mathbf{x}}_{obj}$. The CPP algorithm is invoked using a randomly generated "sweep angle" that describes the angle of the parallel candidate paths with respect to the x-axis. For search efficiency, we required that the difference between two "sweep angles" was larger than 30°.

We select the next-best EM path $\mathcal{X}^*$ by selecting an $\mathcal{X}_i$ that maximizes a weighted sum of the occupancy probabilities $\hat{m}_{\mathbf{x}}$, the uncertainties related to the estimated occupancy $\hat{std}_{\mathbf{x}}$, and the distance between the candidate path and the estimated location of the object's centroid $\hat{\mathbf{x}}_{obj}$.

$$\underset{\{\mathcal{X}_i\}}{\operatorname{argmax}} \sum_{\mathbf{x}\in\mathcal{X}_i} (w_1 \hat{m}_{\mathbf{x}} + w_2 \hat{std}_{\mathbf{x}} + w_3 \|\mathbf{x}, \hat{\mathbf{x}}_{obj}\|^2) \tag{3.9}$$

The weights w_1, w_2, and w_3 are determined by the size of the current search region. The weight w_1 is set high in order to maximize the occupancy probabilities, while w_2 is set low in order to select an EM path that minimizes the uncertainty in the occupancy map. The third term, scaled by a small w_3, forces the next-best EM path away from the estimated location of the buried object's centroid $\hat{\mathbf{x}}_{obj}$. This helps to prevent entrapment in local minima due to repetitive exploration of the same area.

The haptic search continues autonomously until the uncertainty of the occupancy mapping is acceptable for the application. We stopped the search process when $\hat{std}_{\mathbf{x}}$ within the current search region S fell below a percentage λ of the size of the current search region.

3.5 Experimental Procedure and Evaluation

We demonstrated the framework for a representative haptic exploration task in simulation and with a real robot.

3.5.1 Granular Material Model

To mimic compact soil, the heterogenous granular material environment was comprised of a proprietary mixture of sand, pea gravel, and cement mortar mix. The sand mixture was housed within a 90×90 cm square, 30 cm deep sandbox whose top surface was flush with the floor. The search area was 75×50 cm.

We modeled the failure zone in front of the robot fingertip during vertical raking ($\alpha = 90°$). The shear band angle β of the failure zone depends on the internal angle of friction ϕ_u

and the dilation angle ψ [103] for the granular materials. We measured the internal angle of friction as $\phi_u = 42°$ by performing a triaxial shear test with a sand mixture sample, and estimated the dilation angle as $\psi = \phi_u - 30° = 12°$ [104]. Since numerical simulations [103] indicate that average values of β are closer to "Arthur angles" [105] if $\phi_u > 20°$, the shear band angle β was estimated as 31.5°.

Another important parameter that affects the sand mixture failure zone is the raking depth d. During pilot experiments, we observed that fingertip forces tended to decrease with each subsequent EM, as soil was displaced during raking and soil depth decreased around the buried object. To account for decreases in soil depth throughout each haptic search trial, raking depth d was decreased after each EM for which there is an interaction with a buried object. Raking depth was decreased by a diminishing amount over the course of each trial (first by 1 cm, then by 0.5 cm, 0.25 cm, etc.).

Based on our empirical observations, we formulated the sensor model shape parameters, γ_r and η, as a function of the number of pairwise interactions n with the buried object (Table 3.1) in order to eliminate manual tuning of raking depth throughout each trial.

3.5.2 Haptic Exploration in Simulation

3.5.2.1 Experimental Set-up

Simulations were conducted using the sensor model parameters listed in Table 3.1. We ran two sets of 25 trials, in which each trial consisted of an automated haptic search that ran until the uncertainty-based stopping criterion was met. One set of trials was conducted for each of two 2-D object shapes that were "buried" at randomly generated locations and orientations in the simulated sandbox: circular (diam. = 8 cm) and elliptical (major axis =

Table 3.1: Haptic Exploration Parameters

Sensor Model

Shear band angle β	$31.5°^{\dagger}$
Rupture distance refer to reduced raking depth d	7.1, 5.5, 4.7, 4.3 cm
Rupture distance parameter γ_r	$1/(72 \times 0.82^n)$ ‡
Side crescent angle η	$(50 + 7n)°$ ‡

† calculate β using Arthur angles [105]: $\beta = 45° - (\phi_u + \psi)/4$

‡ n is the # of pairwise interactions with the buried object.

Bayesian Hilbert Map

Smoothing bandwidth γ_m for free movement within sand	0.1
Smoothing bandwidth γ_m for interaction with buried object	0.04
Spatial resolution of hinged features	5 cm

Exploration Strategy

Weights w_1, w_2, w_3 to select the next-best path	1, 0.5, -1
Stopping criterion for uncertainty λ	10%

9 cm, minor axis = 5 cm).

Once the failure zone simulated with the kernel-based sensor model intersected with the boundary of the buried object (green outline in Figures 3.5a, b, f, and g), the environmental interaction state was updated to "granular material jamming", the EM was stopped, and the occupancy map was updated.

3.5.2.2 Experimental Results

Two representative haptic search trials are shown in Figure 3.5, where the top and bottom rows correspond to circular and elliptical objects, respectively. The occupancy maps in the first and second columns show the first and second phases of the haptic search, respectively.

The third column of Figure 3.5 shows the learned kernels after random initialization by applying automorphing kernels and aggregation methods in Hilbert maps [89]. This method was found to be more efficient than stationary learning, where a single length scale is used for all Gaussian kernels. The kernels learned during simulation are plotted as colorized markers (Figures 3.5c and h), where the length scale is indicated by both color and size. For kernel placement, we used uniformly spread hinges with a 5 cm spacing. For areas having high and low occupancy values, we assigned the Gaussian kernel smoothing bandwidth as $\gamma_m = 0.04$ and $\gamma_m = 0.1$, respectively.

Bayesian Hilbert maps for occupancy (mean) and uncertainty (standard deviation) are shown in the fourth and fifth columns of Figure 3.5. The estimated outline for each buried object is shown in yellow for an occupancy value of 0.6. Despite the relatively large 75×50 cm search area, the mean localization error, as root mean squared error (RMSE), was ≤ 2 cm for both objects (Table 3.2).

The directed Hausdorff distance [106] was used to measure the dissimilarity between the estimated object outline and the true simulated object outline. Due to difficulty in estimating the orientation of the elliptical object, the Hausdorff distance for the elliptical object was larger than that for the symmetric circular object.

3.5.3 Haptic Exploration with a Real Robot

3.5.3.1 Experimental Set-up

The robotic system was comprised of a 16-DOF, dual-arm mobile manipulator (HDMS, RE2 Robotics), two 3-finger adaptive robot grippers (Robotiq, Inc.), and a tri-modal laser, 3-D stereo, and video sensor (MultiSense SL, Carnegie Robotics, LLC). Each robot gripper was equipped with wrist-mounted inertial measurement units (BNO055, Bosch Sensortec GmbH) and force/torque sensors (FT 300, Robotiq, Inc.).

For each gripper, two fingers were flexed toward the palm and a single finger was left extended for raking. A single robot fingertip was raked in a straight line at a constant speed of 5 cm/s, and with a depth of 4.4 cm and vertical orientation ($\alpha = 90°$) relative to the free surface of the sand mixture. The buried object was a rigid plastic cylinder having a 9.1 cm diameter and 9.1 cm length. The cylinder was buried 2.5 cm below the free surface of the sand mixture.

Table 3.2: Haptic Exploration Simulation Results

	# of Phase 1 EMs	# of Phase 2 EMs	Centroid RMSE	Hausdorff distance
Circle	2.5 ± 1.2	4.7 ± 1.6	1.5 ± 0.5 cm	1.4 ± 0.5 cm
Ellipse	3.0 ± 1.6	4.0 ± 1.8	2.0 ± 0.3 cm	4.9 ± 0.6 cm

3.5.3.2 Force Observer Model

Since the movement of the end-effector had non-negligible inertial effects on the wrist force-torque measurements, we designed a force observer model based on Newton-Euler equations in order to account for inertial effects and more accurately estimate interaction forces between the robot fingertip and its environment.

The force observer model treated force measurements from the force/torque sensor $\mathbf{f}_{sens}$ as the summation of inertial forces, forces due to gravity, environmental interaction forces $\mathbf{f}_{env}$, and a force/torque sensor error term ϵ_{sens}. The force due to gravity and the inertial force due to acceleration with respect to gravity were combined in the first term of the following equation:

$$\hat{\mathbf{f}}_{sens} = m_e R^c_{IMU} \hat{\mathbf{a}}_{IMU} + \mathbf{f}_{env} + \epsilon_{sens} \tag{3.10}$$

where $\hat{\mathbf{f}}_{sens}$ and $\hat{\mathbf{a}}_{IMU}$ denote low-pass filtered data from the force/torque and IMU sensors, respectively. Using the end-effector's center of mass c as a reference point, R^c_{IMU} is a rotation matrix that reflects alignment error between the IMU and end-effector reference frames. Regression was used to estimate constant values for the mass of the end-effector m_e, the rotation matrix R^c_{IMU}, and the sensor error $epsilon_{sens}$. The actual environmental interaction force $\mathbf{f}_{env}$ that we seek for use as $\mathbf{f}_t$ in Eq. 3.1 were determined using Eq. 3.10.

3.5.3.3 Perception of Environmental Interaction State

Per the data-driven criteria in Figure 3.4, the environmental interaction state $\theta_{t,*}$ was determined based on the filtered force $\mathbf{f}_t$ and its rate of change. To minimize disturbances to the buried object, individual EMs were halted when $\mathbf{f}_t > 80$ N.

3.5.3.4 Experimental Results

The parameters for mapping and haptic exploration with the real robot are summarized in Table 3.1. The haptic search stopping criterion was met after a total of 7 exploratory movements (full experimental trial can be viewed in the Supplemental Video). EM #1 resulted in the first interaction with the buried object with the extended fingertip on the robot's right gripper. EM #2 was attempted from the opposing direction of EM #1 using the robot's left gripper, but was halted due to arm workspace limits. As a result, EM #3 was implemented as a continuation of EM #2 using the robot's right gripper. Thus, EMs #1 and #3 constituted Phase 1 and EM pairs #4-5 and #6-7 completed Phase 2.

Figure 3.6 shows the fingertip forces along with the perceived environmental interaction state as a function of time. Granular granular material jamming was perceived for EMs #1, 3, 4, 6, and 7. By EM #7, the sand depth near the object has decreased, resulting in fingertip forces that are smaller than those for prior EMs when the fingertip is moving freely through the granular materials.

The final Bayesian Hilbert Map is shown in Figure 3.7. The final fingertip force is shown in green for each of the 7 EMs shown in magenta. The estimated outline of the object is shown with a black dashed line for a level curve $\mathcal{C} = \{\mathbf{x} \in \mathbf{X} \colon m_{\mathbf{x}} = 0.6\}$. The outline is approximately 10 cm long, which is reasonable given that the cylinder's true length was 9.1 cm. The narrower 5 cm width of the outline makes sense given that the cylinder was buried 2.5 cm below the free surface of the sand mixture and the robot fingertip was raked at a depth of 4.4 cm. These results suggest that failure zone models for homogeneous granular materials [1, 11] can be applied successfully to densely packed sand mixtures.

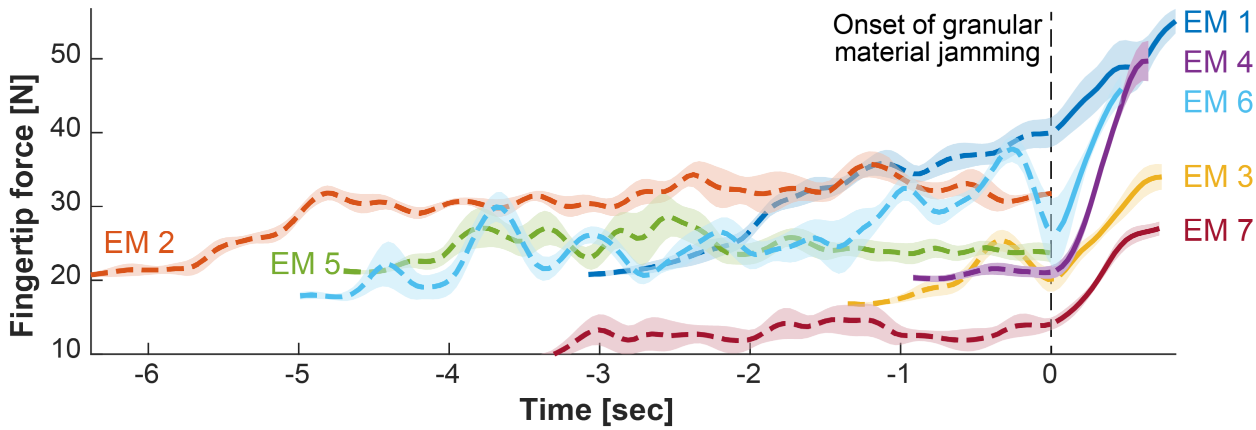


Figure 3.6: Fingertip force for each of 7 EMs: Since fingertip forces fluctuate during interactions with granular materials, dashed and solid lines indicate mean values for free movement through granular materials and during granular material jamming, respectively. Colored bands indicate values that are one standard deviation from the mean. For visual comparison, the onset of granular material jamming is aligned at 0 sec (black dashed line) for EMs #1, 3, 4, 6, and 7. No direct contact with the object was perceived.

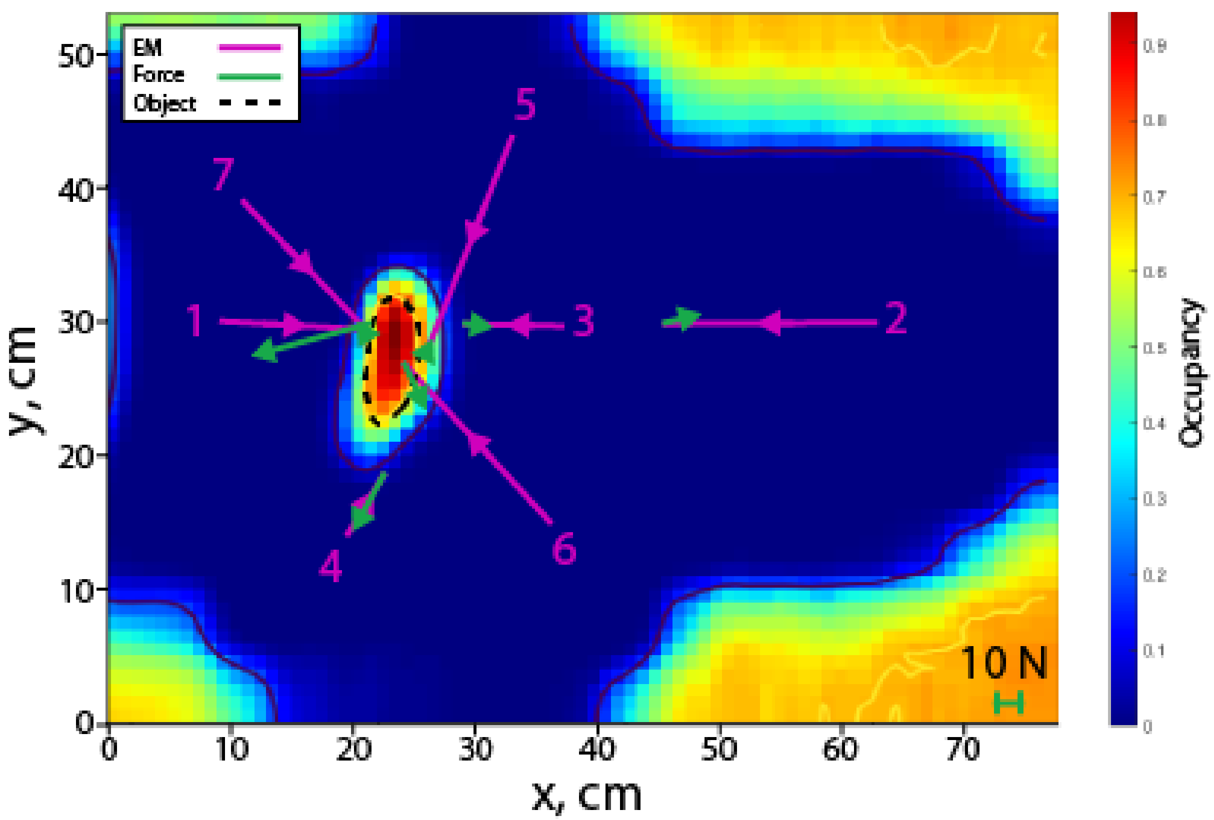


Figure 3.7: Final Bayesian Hilbert Map for a haptic search trial that autonomously ended after 7 EMs (magenta). The last fingertip force for each EM is shown in green. The estimated outline of the buried object for an occupancy level of 0.6 is shown with a black dashed line.

3.6 Conclusion

We have addressed the challenge of haptically locating an object buried within granular materials. First, we integrated localized, sparse tactile spatial information into a global, continuous occupancy map by proposing a kernel-based sensor model that accounts for granular material jamming. Second, we designed a haptic exploration strategy that maximizes localization efficiency from tactile measurements while minimizing interactions with buried objects. We demonstrated our framework for tactile SLAM in granular materials in simulation and with a real robot.

One limitation of this work is that the sensor model is based on "rupture distance" values from soil-tool interactions without jamming. In the future, we would like to incorporate experimental data-driven models of granular material jamming for the heterogeneous soil mixture, the addition of which could further enhance map accuracy. Another limitation is that EMs were planned as straight line paths within a 2-D search area. The use of 3-D curvilinear EMs could enable 3-D maps that provide level curves at arbitrary depths within a search volume. Advanced models for tactile perception in granular materials could enable the development of adaptive robot behaviors [107] for enhanced robot autonomy in extreme, unstructured environments.

CHAPTER 4

Summary and Conclusion

Haptic exploration within granular materials is an essential task for robots designed for locating and interacting with delicate and/or dangerous buried objects. Haptic intelligence is required to carefully maneuver within the granular materials and amongst buried objects. However, because of the complex behavior of granular materials, it is challenging to infer contact states and build models of buried objects and their surrounding granular material environment. The work presented in this dissertation has provided a framework for tactile perception, haptic exploration planning, and map rendering for the localization of an object buried in granular materials.

4.1 Contributions

Supervised learning-based, multimodal tactile perception model: We developed a tactile perception pipeline for estimating fingertip contact state and locating objects buried in granular materials. We established a multimodal architecture based on Recurrent Neural Networks with Long Short-Term Memory units that incorporates multiple tactile sensing modalities (vibration, internal fluid pressure, and electrode impedance as a proxy for finger-pad deformation). The Sparse-Fusion Recurrent Neural Network model architecture learned sparse representations of tactile data and performed robustly on novel test data that included variations in granular material type and particle size, fingertip orientation, fingertip speed, and object location along the linear fingertip trajectory.

Sensor model based on tactile measurements during haptic exploration within granular materials: We developed a sensor model for tactile measurements that accounts for different environmental interaction states between the robot fingertip and granular materials as well as the local flow characteristics of granular materials. The sensor model enabled the fusion of local, sparse tactile measurements into a global map for the planning of future haptic exploration movements.

Tactile-based, continuous occupancy map: We firstly introduced a Gaussian-based belief map representation to visualize tactile percepts. We then advanced the map representation by transitioning to a Bayesian Hilbert Map framework, which enables the generation of a continuous occupancy map. The new method provides a simpler and faster parametric model that eliminates the cubical time complexity of the benchmark Gaussian Process Occupancy Mapping approach.

Tactile SLAM for haptic exploration within granular materials: We developed a novel framework that extended tactile SLAM to haptic exploration within granular materials. The haptic exploration strategy was designed to efficiently locate a buried object and refine its outline while simultaneously minimizing potentially damaging physical interactions with the buried object. We demonstrated our framework in both simulation and with a real robot with granular materials.

4.2 Future Work

4.2.1 Enhance the Generalizability and Robustness of the Sensor Model

In Chapter 3, the kernel-based sensor model incorporates a spatial model of the granular material failure zone ahead of the robot's fingertip. As an essential part of occupancy mapping, the accuracy of the sensor model directly affects the final mapping results. Although erroneous measurements and flaws in the sensor model could be accounted for through the optimization procedure in the Hilbert Map module, we could improve the performance of the foundational kernel-based sensor model in two ways.

First, the current sensor model assumes that the model observed from homogeneous granular materials applies to heterogeneous granular materials. In the future, we could extend the tactile SLAM framework by leveraging dense sensors, such as camera-based tactile sensors (e.g. GelSight[108], DIGIT[109]), which could not only detect clues for edges of buried objects but also provide texture details of the granular materials [110]. Updating parameters in the 3-D failure zone wedge model [1] and criteria for detecting granular materials jamming for different types of granular materials will further improve the generalizability of the sensor model.

Second, the detection of the environmental interaction state of the robotic fingertip plays an important role in the kernel-based sensor model. The ability to estimate the distance between the robotic fingertip and the buried object could be used to improve the accuracy of the sensor model since the measurement range could be defined according to actual distance to the buried object instead of the rupture distance from a granular material failure zone model. Improvements to the sensor model will lead to improved accuracy in the localization of the buried object.

4.2.2 Extend the Continuous Occupancy Map

The ability to model the surrounding space and determine which areas are occupied is of key importance in many robotic applications, ranging from grasping and manipulation to path planning and obstacle avoidance. Occupancy mapping is often hindered by several factors, such as real-time data collection and processing constraints; limited, intermittent, or low quality sensory data; and limited memory capacity, especially for large-scale environments.

The Hilbert Map provides a novel framework that produces an efficient, non-stationary, continuous occupancy mapping function that can be easily queried at arbitrary resolutions. Although the Bayesian Hilbert Map already automatically optimizes the weighting parameters through a probabilistic method, techniques that allow the learning of individual features for different areas of the input space will significantly improve the modeling performance in a complex environment.

In Chapter 3, the Hilbert Map module learned the occupancy map through an optimization process that maximized the likelihood of the tactile sensor data. However, there are several hand-tuned parameters that can significantly influence performance, such as the kernel selection, hinged feature placement, and its length scales. Extensions of the Hilbert Map framework [88, 89] show robustness in the presence of sparse data and significantly better interpolative powers. We could leverage the latest developments in automatic machine learning and probabilistic programming under the Hilbert Map framework to automate the tuning of model parameters.

4.2.3 Employ Adaptive Behaviors for Haptic Exploration within Granular Materials

Endowing robots with adaptive physical interaction skills is a critical requirement of robot autonomy in extreme, unstructured environments. As for haptic exploration within granular materials, we could enhance autonomy for physical interaction in two ways.

First, advanced tactile perception models could equip the robot with a better ability to react and adapt to its environment [111]. In Chapters 2 and 3, exploratory movements of the robot fingertip were planned as straight line paths with the same raking depth, which constrained the learned mapping to a 2-D search area. In the future, we could consider 3-D curvilinear movements of the fingertip. The use of 3-D curvilinear exploratory movements could enable 3-D maps that provide level curves at arbitrary depths within a search volume. As fingertip width and depth affect the model of the granular material failure zone, additional work would be needed to update the sensor model parameters in real-time. We believe a 3-D modeling method would improve the accuracy and efficiency for the robot to identify environment characteristics and find the optimal interaction strategy.

Second, robotic systems based on torque sensing and actuation or variable impedance mechanisms could be used to make robots compliant with respect to their surroundings [112]. In particular, we could leverage self-tuning impedance controllers [113] to regulate the impedance profile based on prior haptic exploration experience when interacting with similar environments.

APPENDIX A

Multimodal Tactile Perception Model

A.1 Empirical Mode Decomposition

The Empirical Mode Decomposition (EMD) [61] approach is based on the assumption that time series data consist of simple intrinsic modes of oscillation. Each of these oscillatory modes is represented by an intrinsic mode function (IMF) that satisfies two basic conditions to be described below: (i) the number of extrema and the number of zero-crossings may differ at most by one, and (ii) the mean value of the local maxima and local minima envelopes is zero.

In contrast to a Fourier analysis, EMD decomposes a time series signal $q(t)$ adaptively by sifting the signal into a series of intrinsic mode functions c_j until only a small monotonic final residue r_n remains. For an arbitrary $q(t)$, the sifting process starts with the identification of all local extrema. Local maxima are then connected with a cubic spline in order to create an upper envelope; local minima are similarly connected to produce a lower envelope. The sifting procedure is repeated on the difference between the signal $q(t)$ and the mean of the upper and lower envelopes until the mean qualifies as an IMF per the two conditions stated previously [61]. Once the iterative sifting process has resulted in a small enough residue r_n, the original signal $q(t)$ can be represented as

$$q(t) = \sum_{j=1}^{n} c_j + r_n, \tag{A.1}$$

The original signal $q(t)$ is thus decomposed in the time domain. The length of each IMF is identical to that of the original signal, which ensures that characteristics of the original signal are preserved, including variations in frequency. This is important for times series signals that are often the result of multiple, natural processes that may or may not be concurrent. Such data features would remain hidden through the use of Fourier domain or wavelet coefficient analyses that filter out specific frequencies.

A.2 Sparse Coding

Sparse coding is widely used for audio signal reconstruction and image processing. The sparsity of nonzero values in the model offers four main advantages: increased storage capacity, representation as explicit signals, faster storage and retrieval of representations, and reduced computational effort [114, 115]. By minimizing an empirical cost, the algorithm learns a dictionary $D = [d_1, \ldots d_k] \in R^{m \times k}$ and a sparse matrix of coefficients $\alpha = [\alpha_1, \ldots \alpha_n] \in R^{k \times n}$ that can be used to approximate a signal $X = [x_1, \ldots x_n] \in R^{m \times n}$ as $\hat{X} = D\alpha$ in a linearized manner (eq. 2.2).

The cost function can be cast as a matrix factorization problem that minimizes errors in the linearized reconstruction of the original signal X with an additional penalty term that induces sparsity.

$$\min_{D \in C, \alpha \in R^{k \times n}} \frac{1}{2} \|X - D\alpha\|_F + \lambda \|\alpha\|_{1,1}, \tag{A.2}$$

The operator $\|\cdot\|_F$ denotes the Frobenius norm. The term $\|\alpha\|_{1,1}$ denotes the l_1-norm of the sparse matrix of coefficients α that is the sum of the magnitude of all coefficients. Per [64], the sparsity-inducing regularization parameter λ was set as $\lambda = 1.2/\sqrt{m}$. The sizes of D and k are tuned according to a trade-off between computing time and objective function values (eq. A.2). To prevent D from having arbitrarily large values, its columns d_1,...,d_k are constrained to have an l_2-norm less than or equal to one. This is achieved by defining C as a convex set of matrices satisfying the constraint: $C \triangleq \{D \in R^{m\times k} s.t. \forall j = 1, \ldots, k, {d_j}^T d_j \leq 1\}$.

A.3 RNN with GRUs

Standard RNNs take temporal sequences of data as inputs and yield high-level representations or labels as outputs. The feedforward neural network structure is enhanced by "recurrent edges" between adjacent timesteps that consider temporal information both forward and backward in time. Such internal loops across timesteps allow temporal information to persist within the network. At time t, nodes with recurrent edges receive input from current data point x_t as well as input from a previous timestep $(t-1)$ from hidden node h_{t-1}. In this manner, the previous input x_{t-1} may influence the softmax output layer y_t through the recurrent connection.

$$h_t = \sigma(W_h \cdot [h_{t-1}, x_t] + b_h), \tag{A.3}$$

$$y_t = softmax(W_y h_t + b_y), \tag{A.4}$$

where σ represents an activation function (sigmoid function) and W_* and b_* are weight and bias parameters, respectively.

The GRU is a simplified LSTM unit and only consists of two gates: a reset gate r and an update gate z. Each gate is composed of a sigmoid neural net layer and pointwise multiplication operation. The reset gate r_t (eq. A.5) decides whether the previous hidden state is ignored and a $tanh$ layer produces a vector of new candidate values $\widetilde{C}_t$ (eq. A.6) that could be added to the current state (eq. A.8). The update gate z_t (eq. A.7) combines the forget and input gate, and enables the GRU to add or remove (i.e. remember or forget) information with respect to the memory cell or cell state C_* (eq. A.8), which is equivalent to the hidden state h since the GRU merges the cell state and the hidden state. In this formulation, the cell state is hard-bounded and helps to store long-term temporal dependencies within the RNN.

$$r_t = \sigma(W_r \cdot [C_{t-1}, x_t] + b_r), \tag{A.5}$$

$$\widetilde{C}_t = tanh(W_c \cdot [r_t * C_{t-1}, x_t] + b_c), \tag{A.6}$$

$$z_t = \sigma(W_z \cdot [C_{t-1}, x_t] + b_z), \tag{A.7}$$

$$h_t = C_t = (1 - z_t) * C_{t-1} + z_t * \widetilde{C}_t, \tag{A.8}$$

A.4 Belief Map Representation

When exploring an unstructured environment in which an object may be buried, there will be instances in which there is uncertainty in the fingertip contact state. Within a given state space, $X \subset R^n$, the probability of whether an object is buried in a particular region can be represented as a spatial distribution $\phi_t(x)$ for all $x_t \in X$ that may change over time. The spatial distribution is initially unknown, but is presumed to be a uniform distribution until fingertip contact state estimates are generated.

During haptic exploration within the space X, the robot was assumed to be able to measure the exact values from the spatial distribution $\phi_t(x)$. In other words, the fingertip contact state and fingertip location were assumed to be true when used to create and update the belief map. The $\phi_t(x)$ distribution was updated through a probabilistic classification method that considers the fingertip contact state estimates $y_t \in [0, 1]$ and their spatial locations $x_t \in R^2$ along a given exploration trajectory. Specifically, we applied Support Vector Classification (SVC) with a radial basis function (RBF) kernel. Spatial information was stored in an RBF kernel K that maps the influence of contact state y_t at its spatial location x_t into an infinite dimensional feature space. The spatial reach, or sphere of influence, of the radial basis function was set by γ in (eq. A.9).

The belief map is also capable of keeping track of the states it has previously sampled. Prior state information is encoded in an SVC score function (eq. A.10) [116], which sums learned mappings *from all previous timesteps*. Finally, the SVC scores were calibrated by Platt Scaling (eq. A.11) [69]. In our case, the spatial distribution $\phi_t(x)$ was transformed into a probability distribution, indicating whether an object is buried at a specific location ($P(y_t = 1|x)$), since the fingertip contact state $y_t = 1$ was assumed to be true during contact with a buried object.

$$K(x, x_t) = exp(-\gamma \|x - x_t\|^2), \tag{A.9}$$

$$\phi_t(x) = \sum\nolimits_t \alpha_t y_t K(x, x_t) + b, \tag{A.10}$$

$$P(y_t = 1|x) \approx P_{A,B}(\phi_t) = \frac{1}{1 + e^{A\phi_t(x)+B}}, \tag{A.11}$$

The variables α_t and b represent the SVC weights and bias, respectively. The Platt scaling parameters A and B are determined by solving a regularized maximum likelihood problem [117].

A.5 Test of SF-RNN Model Performance

Table A.1: Effects of Multimodality on SF-RNN Model Performance: Model accuracy is reported as mean in %. Best performing models are shown in bold. The models were trained and evaluated with cottonwood beads.

Model description	Test accuracy (%)			
SF-RNN trained with 1 modality	Trained and tested with	Accuracy		
	Vibration	95.7		
	Fluid Pressure	85.8		
	Electrode impedance	**97.6**		
SF-RNN trained with 2 modalities		Trained with		
	Tested with	Vibration, fluid pressure	Vibration, electrode impedance	Fluid pressure, electrode impedance
	Vibration	**96.6**	86.6	N/A
	Fluid pressure	91.0	N/A	74.3
	Electrode impedance	N/A	**96.9**	**97.4**
	2 modalities	97.1	98.7	98.5
SF-RNN trained with 3 modalities	Tested with		Accuracy	
	1 modalities	Vibration	90.0	
		Fluid pressure	83.1	
		Electrode impedance	**97.8**	
	2 modalities	Vibration, fluid pressure	95.2	
		Vibration, electrode impedance	**98.6**	
		Fluid pressure, electrode impedance	98.3	
	3 modalities	Vibration, fluid pressure, electrode impedance	98.7	

REFERENCES

[1] W. Swick and J. Perumpral, "A model for predicting soil-tool interaction," *Journal of Terramechanics*, vol. 25, no. 1, pp. 43–56, Jan. 1988. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/0022489888900614

[2] C. K. Wentworth, "A Scale of Grade and Class Terms for Clastic Sediments," *The Journal of Geology*, vol. 30, no. 5, pp. 377–392, Jul. 1922. [Online]. Available: https://www.journals.uchicago.edu/doi/10.1086/622910

[3] H. M. Jaeger, S. R. Nagel, and R. P. Behringer, "Granular solids, liquids, and gases," *Reviews of modern physics*, vol. 68, no. 4, p. 1259, 1996. [Online]. Available: https://journals.aps.org/rmp/abstract/10.1103/RevModPhys.68.1259

[4] J. F. Peters, M. Muthuswamy, J. Wibowo, and A. Tordesillas, "Characterization of force chains in granular material," *Physical Review E*, vol. 72, no. 4, p. 041307, Oct. 2005. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.72.041307

[5] D. Bideau, J. A. Dodds, and C. de physique des Houches, *Physics of Granular Media*. Commack, N.Y: Nova Science Publishers, 1991.

[6] G. Y. Onoda and E. G. Liniger, "Random loose packings of uniform spheres and the dilatancy onset," *Physical Review Letters*, vol. 64, no. 22, p. 2727, 1990. [Online]. Available: https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.64.2727

[7] H. M. JAEGER and S. R. NAGEL, "Physics of the Granular State," *Science*, vol. 255, no. 5051, p. 1523, Mar. 1992. [Online]. Available: http://science.sciencemag.org/content/255/5051/1523.abstract

[8] P. Richard, M. Nicodemi, R. Delannay, P. Ribiere, and D. Bideau, "Slow relaxation and compaction of granular systems," *Nat Mater*, vol. 4, no. 2, pp. 121–128, Feb. 2005. [Online]. Available: http://dx.doi.org/10.1038/nmat1300

[9] C. S. O'Hern, S. A. Langer, A. J. Liu, and S. R. Nagel, "Force Distributions near Jamming and Glass Transitions," *Physical Review Letters*, vol. 86, no. 1, pp. 111–114, Jan. 2001. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.86.111

[10] E. I. Corwin, H. M. Jaeger, and S. R. Nagel, "Structural signature of jamming in granular media," *Nature*, vol. 435, no. 7045, pp. 1075–1078, Jun. 2005. [Online]. Available: http://www.nature.com/doifinder/10.1038/nature03698

[11] M. Kobayakawa, S. Miyai, T. Tsuji, and T. Tanaka, "Interaction between dry granular materials and an inclined plate (comparison between large-scale DEM simulation and three-dimensional wedge model)," *Journal of Terramechanics*, vol. 90, pp. 3–10, Aug.

2020. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0022489819301090

[12] N. Gravish, P. B. Umbanhowar, and D. I. Goldman, "Force and flow at the onset of drag in plowed granular media," *Physical Review E*, vol. 89, no. 4, Apr. 2014. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.89.042202

[13] ——, "Force and Flow Transition in Plowed Granular Media," *Physical Review Letters*, vol. 105, no. 12, Sep. 2010. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.105.128301

[14] A. Schofield and P. Wroth, *Critical State Soil Mechanics*. New York: McGraw-Hill, 1968.

[15] L. Rothenburg and N. Kruyt, "Critical state and evolution of coordination number in simulated granular materials," *International Journal of Solids and Structures*, vol. 41, no. 21, pp. 5763–5774, Oct. 2004. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0020768304003063

[16] J. Aguilar, T. Zhang, F. Qian, M. Kingsbury, B. McInroe, N. Mazouchova, C. Li, R. Maladen, C. Gong, M. Travers, R. L. Hatton, H. Choset, P. B. Umbanhowar, and D. I. Goldman, "A review on locomotion robophysics: the study of movement at the intersection of robotics, soft matter and dynamical systems," *Reports on Progress in Physics*, vol. 79, no. 11, p. 110001, Nov. 2016. [Online]. Available: http://stacks.iop.org/0034-4885/79/i=11/a=110001?key=crossref.ca2966c57b06c47b4fe1725316799370

[17] R. D. Maladen, Y. Ding, P. B. Umbanhowar, A. Kamor, and D. I. Goldman, "Biophysically inspired development of a sand-swimming robot." Georgia Institute of Technology, 2011. [Online]. Available: https://smartech.gatech.edu/handle/1853/44542

[18] S. P. Molner, "The Art of Molecular Dynamics Simulation (Rapaport, D. C.)," *Journal of Chemical Education*, vol. 76, no. 2, p. 171, 1999. [Online]. Available: http://dx.doi.org/10.1021/ed076p171

[19] R. D. Maladen, Y. Ding, C. Li, and D. I. Goldman, "Undulatory Swimming in Sand: Subsurface Locomotion of the Sandfish Lizard," *Science*, vol. 325, no. 5938, pp. 314–318, Jul. 2009. [Online]. Available: http://www.sciencemag.org/cgi/doi/10.1126/science.1172490

[20] F. Qian, T. Zhang, C. Li, P. Masarati, A. M. Hoover, P. Birkmeyer, A. Pullin, R. S. Fearing, and D. I. Goldman, "Walking and running on yielding and fluidizing ground," 2012.

[21] S. Jia and V. J. Santos, "Tactile Perception for Teleoperated Robotic Exploration within Granular Media," *ACM Transactions on Human-Robot Interaction*, vol. 10, no. 4, pp. 1–27, Dec. 2021. [Online]. Available: https://dl.acm.org/doi/10.1145/3459996

[22] B. Beneš, E. Dorjgotov, L. Arns, and G. Bertoline, "Granular Material Interactive Manipulation: Touching Sand with Haptic Feedback," in *WSCG 2006 conference prodeedings*. Plzen, Czech Republic: UNION Agency - Science Press, 2006, p. 7.

[23] V. Purchart, "Interactive manipulation of sand on a TIN terrain model for virtual reality," p. 7, 2012.

[24] J. A. Fishel and G. E. Loeb, "Bayesian Exploration for Intelligent Identification of Textures," *Frontiers in Neurorobotics*, vol. 6, p. 4, 2012. [Online]. Available: http://www.frontiersin.org/Neurorobotics/10.3389/fnbot.2012.00004/abstract

[25] Z. Pezzementi, E. Plaku, C. Reyda, and G. D. Hager, "Tactile-Object Recognition From Appearance Information," *IEEE Transactions on Robotics*, vol. 27, no. 3, pp. 473–487, Jun. 2011. [Online]. Available: http://ieeexplore.ieee.org/document/5752871/

[26] K. Hertkorn, M. A. Roa, and C. Borst, "Planning in-hand object manipulation with multifingered hands considering task constraints," in *2013 IEEE International Conference on Robotics and Automation*. Karlsruhe, Germany: IEEE, May 2013, pp. 617–624. [Online]. Available: http://ieeexplore.ieee.org/document/6630637/

[27] R. B. Hellman, C. Tekin, M. van der Schaar, and V. J. Santos, "Functional Contour-following via Haptic Perception and Reinforcement Learning," *IEEE Transactions on Haptics*, vol. 11, no. 1, pp. 61–72, Mar. 2018. [Online]. Available: http://ieeexplore.ieee.org/document/8039205/

[28] J. C. Gwilliam, Z. Pezzementi, E. Jantho, A. M. Okamura, and S. Hsiao, "Human vs. robotic tactile sensing: Detecting lumps in soft tissue," in *Haptics Symposium, 2010 IEEE*. IEEE, 2010, pp. 21–28. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5444685

[29] J. Fishel, G. Lin, and G. E. Loeb, "SynTouch LLC BioTac Product Manual, v.16," Feb. 2013.

[30] S. Dong, W. Yuan, and E. H. Adelson, "Improved GelSight tactile sensor for measuring geometry and slip," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017, pp. 137–144.

[31] S. Luo, J. Bimbo, R. Dahiya, and H. Liu, "Robotic tactile perception of object properties: A review," *Mechatronics*, vol. 48, pp. 54–67, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957415817301575

[32] B. A. Richardson and K. J. Kuchenbecker, "Improving Haptic Adjective Recognition with Unsupervised Feature Learning," in *2019 International Conference on Robotics and Automation (ICRA)*. Montreal, QC, Canada: IEEE, May 2019, pp. 3804–3810. [Online]. Available: https://ieeexplore.ieee.org/document/8793544/

[33] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A Large-Scale Hierarchical Image Database," in *CVPR09*, 2009.

[34] S. Dasari, F. Ebert, S. Tian, S. Nair, B. Bucher, K. Schmeckpeper, S. Singh, S. Levine, and C. Finn, "Robonet: Large-scale multi-robot learning," in *CoRL 2019: Volume 100 Proceedings of Machine Learning Research*, 2019.

[35] S. Jia and V. Santos, "Multimodal Haptic Perception within Granular Media via Recurrent Neural Networks," in *Proc RSS Workshop on "Tactile Sensing for Manipulation: Hardware, Modeling, and Learning"*, Boston, MA, Jul. 2017.

[36] F. Ramos and L. Ott, "Hilbert maps: Scalable continuous occupancy mapping with stochastic gradient descent," *The International Journal of Robotics Research*, vol. 35, no. 14, pp. 1717–1730, Dec. 2016. [Online]. Available: http://journals.sagepub.com/doi/10.1177/0278364916684382

[37] W. Hong, "Modeling, Estimation, and Control of Robot-Soil Interactions," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, 2001.

[38] C. Tan, Y. Zweiri, K. Althoefer, and L. Seneviratne, "Online Soil Parameter Estimation Scheme Based on Newton-Raphson Method for Autonomous Excavation," *IEEE/ASME Transactions on Mechatronics*, vol. 10, no. 2, pp. 221–229, Apr. 2005. [Online]. Available: http://ieeexplore.ieee.org/document/1420330/

[39] T. S. Majmudar and R. P. Behringer, "Contact force measurements and stress-induced anisotropy in granular materials," *Nature*, vol. 435, no. 1079, pp. 1079–1082, Jun. 2005. [Online]. Available: http://www.nature.com/doifinder/10.1038/nature03805

[40] R. D. Ponce Wong, R. B. Hellman, and V. J. Santos, "Spatial Asymmetry in Tactile Sensor Skin Deformation Aids Perception of Edge Orientation During Haptic Exploration," *IEEE Transactions on Haptics*, vol. 7, no. 2, pp. 191–202, Apr. 2014. [Online]. Available: http://ieeexplore.ieee.org/document/6642031/

[41] G. Cannata, S. Denei, and F. Mastrogiovanni, "Tactile sensing: Steps to artificial somatosensory maps," in *RO-MAN, 2010 IEEE*. IEEE, 2010, pp. 576–581.

[42] J. J. Gibson, "Observations on active touch." *Psychological review*, vol. 69, no. 6, p. 477, 1962.

[43] H. Tappeiner, R. Klatzky, P. Rowe, J. Pedersen, and R. Hollis, "Bimanual haptic teleoperation for discovering and uncovering buried objects," in *2013 IEEE International Conference on Robotics and Automation*. Karlsruhe, Germany: IEEE, May 2013, pp. 2380–2385. [Online]. Available: http://ieeexplore.ieee.org/document/6630900/

[44] N. Mazouchova, P. B. Umbanhowar, and D. I. Goldman, "Flipper-driven terrestrial locomotion of a sea turtle-inspired robot," *Bioinspiration & Biomimetics*, vol. 8, no. 2, p. 026007, Apr. 2013. [Online]. Available: http://stacks.iop.org/1748-3190/8/i=2/a=026007?key=crossref.ccfceefdbb7b84939873c613aa375a5b

[45] H. Marvi, C. Gong, N. Gravish, H. Astley, M. Travers, R. L. Hatton, J. R. Mendelson, H. Choset, D. L. Hu, and D. I. Goldman, "Sidewinding with minimal slip: Snake and robot ascent of sandy slopes," *Science*, vol. 346, no. 6206, pp. 224–229, Oct. 2014. [Online]. Available: http://www.sciencemag.org/cgi/doi/10.1126/science.1255718

[46] J. Aguilar and D. I. Goldman, "Robophysical study of jumping dynamics on granular media," *Nature Physics*, vol. 12, no. 3, pp. 278–283, Mar. 2016. [Online]. Available: http://www.nature.com/articles/nphys3568

[47] X. Xiong, A. D. Ames, and D. I. Goldman, "A stability region criterion for flat-footed bipedal walking on deformable granular terrain," in *Intelligent Robots and Systems (IROS), 2017 IEEE/RSJ International Conference on*. IEEE, 2017, pp. 4552–4559.

[48] M. Cakmak and A. L. Thomaz, "Designing robot learners that ask good questions," in *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction - HRI '12*. Boston, Massachusetts, USA: ACM Press, 2012, p. 17. [Online]. Available: http://dl.acm.org/citation.cfm?doid=2157689.2157693

[49] A. Yamaguchi and C. G. Atkeson, "Stereo vision of liquid and particle flow for robot pouring," in *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*. Cancun, Mexico: IEEE, Nov. 2016, pp. 1173–1180. [Online]. Available: http://ieeexplore.ieee.org/document/7803419/

[50] C. Schenck, J. Tompson, S. Levine, and D. Fox, "Learning Robotic Manipulation of Granular Media," in *Proceedings of the 1st Annual Conference on Robot Learning*, ser. Proceedings of Machine Learning Research, S. Levine, V. Vanhoucke, and K. Goldberg, Eds., vol. 78. PMLR, Nov. 2017, pp. 239–248. [Online]. Available: http://proceedings.mlr.press/v78/schenck17a.html

[51] A. Yamaguchi and C. G. Atkeson, "Differential dynamic programming for graph-structured dynamical systems: Generalization of pouring behavior with different skills," in *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*. Cancun, Mexico: IEEE, Nov. 2016, pp. 1029–1036. [Online]. Available: http://ieeexplore.ieee.org/document/7803398/

[52] A. Yamaguchi, C. G. Atkeson, and T. Ogasawara, "Pouring Skills with Planning and Learning Modeled from Human Demonstrations," *International Journal of Humanoid Robotics*, vol. 12, no. 03, p. 1550030, Sep. 2015. [Online]. Available: http://www.worldscientific.com/doi/abs/10.1142/S0219843615500309

[53] A. Gupta and M. O'Malley, "Design of a haptic arm exoskeleton for training and rehabilitation," *IEEE/ASME Transactions on Mechatronics*, vol. 11, no. 3, pp. 280–289, Jun. 2006. [Online]. Available: http://ieeexplore.ieee.org/document/1642690/

[54] J. M. Walker, A. A. Blank, P. A. Shewokis, and M. K. OMalley, "Tactile Feedback of Object Slip Facilitates Virtual Object Manipulation," *IEEE Transactions on Haptics*, vol. 8, no. 4, pp. 454–466, Oct. 2015. [Online]. Available: http://ieeexplore.ieee.org/document/7080902/

[55] C. Pacchierotti, D. Prattichizzo, and K. J. Kuchenbecker, "Cutaneous Feedback of Fingertip Deformation and Vibration for Palpation in Robotic Surgery," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 2, pp. 278–287, Feb. 2016. [Online]. Available: http://ieeexplore.ieee.org/document/7155529/

[56] K. Bark, J. Wheeler, P. Shull, J. Savall, and M. Cutkosky, "Rotational Skin Stretch Feedback: A Wearable Haptic Display for Motion," *IEEE Transactions on Haptics*, vol. 3, no. 3, pp. 166–176, Jul. 2010. [Online]. Available: http://ieeexplore.ieee.org/document/5453369/

[57] J. D. Brown, A. Paek, M. Syed, M. K. O'Malley, P. A. Shewokis, J. L. Contreras-Vidal, A. J. Davis, and R. B. Gillespie, "Understanding the role of haptic feedback in a teleoperated/prosthetic grasp and lift task," in *2013 World Haptics Conference (WHC)*. Daejeon: IEEE, Apr. 2013, pp. 271–276. [Online]. Available: http://ieeexplore.ieee.org/document/6548420/

[58] Keehoon Kim, J. Colgate, J. Santos-Munne, A. Makhlin, and M. Peshkin, "On the Design of Miniature Haptic Devices for Upper Extremity Prosthetics," *IEEE/ASME Transactions on Mechatronics*, vol. 15, no. 1, pp. 27–39, Feb. 2010. [Online]. Available: http://ieeexplore.ieee.org/document/4815425/

[59] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, and others, "Tensorflow: Large-scale machine learning on heterogeneous distributed systems," *arXiv preprint arXiv:1603.04467*, 2016. [Online]. Available: https://arxiv.org/abs/1603.04467

[60] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in pytorch," 2017.

[61] N. E. Huang, *Hilbert-Huang transform and its applications*. World Scientific, 2014, vol. 16.

[62] A. Pigorini, A. G. Casalii, S. Casarotto, F. Ferrarelli, G. Baselli, M. Mariotti, M. Massimini, and M. Rosanova, "Time-frequency spectral analysis of TMS-evoked EEG oscillations by means of Hilbert–Huang transform," *Journal of Neuroscience Methods*, vol. 198, no. 2, pp. 236–245, Jun. 2011. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0165027011002172

[63] N. E. Huang, C. C. Chern, K. Huang, L. W. Salvino, S. R. Long, and K. L. Fan, "A new spectral representation of earthquake data: Hilbert spectral analysis of station TCU129, Chi-Chi, Taiwan, 21 September 1999," *Bulletin of the Seismological Society of America*, vol. 91, no. 5, pp. 1310–1338, 2001.

[64] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online learning for matrix factorization and sparse coding," *Journal of Machine Learning Research*, vol. 11, no. Jan, pp. 19–60, 2010. [Online]. Available: http://www.jmlr.org/papers/v11/mairal10a.html

[65] D. Erhan, Y. Bengio, A. Courville, P.-A. Manzagol, P. Vincent, and S. Bengio, "Why does unsupervised pre-training help deep learning?" *Journal of Machine Learning Research*, vol. 11, no. Feb, pp. 625–660, 2010. [Online]. Available: http://www.jmlr.org/papers/v11/erhan10a.html

[66] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," *arXiv preprint arXiv:1406.1078*, 2014. [Online]. Available: https://arxiv.org/abs/1406.1078

[67] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *Advances in Neural Information Processing Systems 27*, Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2014, pp. 3104–3112. [Online]. Available: http://papers.nips.cc/paper/5346-sequence-to-sequence-learning-with-neural-networks.pdf

[68] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," in *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, Y. Bengio and Y. LeCun, Eds., 2015. [Online]. Available: http://arxiv.org/abs/1412.6980

[69] J. C. Platt, "Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods," in *ADVANCES IN LARGE MARGIN CLASSIFIERS*. MIT Press, 1999, pp. 61–74.

[70] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, and V. Dubourg, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, no. Oct, pp. 2825–2830, 2011.

[71] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin, "LIBLINEAR: A library for large linear classification," *Journal of machine learning research*, vol. 9, no. Aug, pp. 1871–1874, 2008. [Online]. Available: http://www.jmlr.org/papers/v9/fan08a.html

[72] A. Jain, A. Singh, H. S. Koppula, S. Soh, and A. Saxena, "Recurrent neural networks for driver activity anticipation via sensory-fusion architecture," in *Robotics and Automation (ICRA), 2016 IEEE International Conference on*. IEEE, 2016, pp. 3118–3125. [Online]. Available: http://ieeexplore.ieee.org/abstract/document/7487478/

[73] L. Van Der Maaten, "Accelerating t-SNE using tree-based algorithms." *Journal of machine learning research*, vol. 15, no. 1, pp. 3221–3245, 2014. [Online]. Available: http://www.jmlr.org/papers/volume15/vandermaaten14a/source/vandermaaten14a.pdf

[74] M. Hollins, S. J. Bensmaïa, and E. A. Roy, "Vibrotaction and texture perception," *Behavioural brain research*, vol. 135, no. 1-2, pp. 51–56, 2002.

[75] J. M. Loomis, D. Katz, and L. E. Krueger, "The World of Touch," *The American Journal of Psychology*, vol. 104, no. 1, p. 147, 1991. [Online]. Available: https://www.jstor.org/stable/1422857?origin=crossref

[76] B. S. Peters, P. R. Armijo, C. Krause, S. A. Choudhury, and D. Oleynikov, "Review of emerging surgical robotic technology," *Surgical Endoscopy*, vol. 32, no. 4, pp. 1636–1655, Apr. 2018. [Online]. Available: http://link.springer.com/10.1007/s00464-018-6079-2

[77] G.-Z. Yang, J. Cambias, K. Cleary, E. Daimler, J. Drake, P. E. Dupont, N. Hata, P. Kazanzides, S. Martel, R. V. Patel, V. J. Santos, and R. H. Taylor, "Medical robotics—Regulatory, ethical, and legal considerations for increasing levels of autonomy," *Science Robotics*, vol. 2, no. 4, p. eaam8638, Mar. 2017. [Online]. Available: http://robotics.sciencemag.org/lookup/doi/10.1126/scirobotics.aam8638

[78] O. o. E. M. U.S. Department of Energy, "Research and Technology Roadmap: Robotics and Remote Systems for Nuclear Cleanup," Tech. Rep., Mar. 2018. [Online]. Available: http://rrsd.ans.org/wp-content/uploads/2018/06/DOE-EM-Robotics-Roadmap-March-2018-e-version.pdf

[79] F. M. P. Behbahani, G. Singla–Buxarrais, and A. A. Faisal, "Haptic SLAM: An Ideal Observer Model for Bayesian Inference of Object Shape and Hand Pose from Contact Dynamics," in *Haptics: Perception, Devices, Control, and Applications*, F. Bello, H. Kajimoto, and Y. Visell, Eds. Cham: Springer International Publishing, 2016, vol. 9774, pp. 146–157. [Online]. Available: http://link.springer.com/10.1007/978-3-319-42321-0_14

[80] Li Zhang, Siwei Lyu, and J. Trinkle, "A dynamic Bayesian approach to real-time estimation and filtering in grasp acquisition," in *2013 IEEE International Conference on Robotics and Automation*. Karlsruhe, Germany: IEEE, May 2013, pp. 85–92. [Online]. Available: http://ieeexplore.ieee.org/document/6630560/

[81] R. O'Flaherty and M. Egerstedt, "Optimal exploration in unknown environments," in *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*. IEEE, 2015, pp. 5796–5801.

[82] I. Abraham, A. Prabhakar, M. J. Hartmann, and T. D. Murphey, "Ergodic Exploration Using Binary Sensing for Nonparametric Shape Estimation," *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 827–834, 2017.

[83] A. De Santis, B. Siciliano, A. De Luca, and A. Bicchi, "An atlas of physical human–robot interaction," *Mechanism and Machine Theory*, vol. 43, no. 3, pp. 253–270, Mar. 2008. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0094114X07000547

[84] D. P. Losey, C. G. McDonald, E. Battaglia, and M. K. O'Malley, "A Review of Intent Detection, Arbitration, and Communication Aspects of Shared Control for Physical Human–Robot Interaction," *Applied Mechanics Reviews*, vol. 70, no. 1, p. 010804, Feb. 2018. [Online]. Available: http://appliedmechanicsreviews.asmedigitalcollection.asme.org/article.aspx?doi=10.1115/1.4039145

[85] A. D. Dragan and S. S. Srinivasa, "A policy-blending formalism for shared control," *The International Journal of Robotics Research*, vol. 32, no. 7, pp. 790–805, Jun. 2013. [Online]. Available: http://journals.sagepub.com/doi/10.1177/0278364913490324

[86] A. Kucukyilmaz, T. M. Sezgin, and C. Basdogan, "Intention Recognition for Dynamic Role Exchange in Haptic Collaboration," *IEEE TRANSACTIONS ON HAPTICS*, vol. 6, no. 1, p. 11, 2013.

[87] M. C. Gombolay, R. A. Gutierrez, S. G. Clarke, G. F. Sturla, and J. A. Shah, "Decision-making authority, team efficiency and human worker satisfaction in mixed human–robot teams," *Autonomous Robots*, vol. 39, no. 3, pp. 293–312, Oct. 2015. [Online]. Available: http://link.springer.com/10.1007/s10514-015-9457-9

[88] V. Guizilini and F. Ramos, "Towards real-time 3D continuous occupancy mapping using Hilbert maps," *The International Journal of Robotics Research*, vol. 37, no. 6, pp. 566–584, May 2018. [Online]. Available: http://journals.sagepub.com/doi/10.1177/0278364918771476

[89] R. Senanayake, A. Tompkins, and F. Ramos, "Automorphing Kernels for Nonstationarity in Mapping Unstructured Environments," ser. Proceedings of Machine Learning Research, A. Billard, A. Dragan, J. Peters, and J. Morimoto,

Eds., vol. 87. PMLR, Oct. 2018, pp. 443–455. [Online]. Available: http://proceedings.mlr.press/v87/senanayake18a.html

[90] A. M. Walker, D. P. Miller, and C. Ling, "User-Centered Design of an Attitude Aware Controller for Ground Reconnaissance Robots," *Journal of Human-Robot Interaction*, vol. 4, no. 1, p. 30, Jan. 2015. [Online]. Available: http://dl.acm.org/citation.cfm?id=3109838

[91] M. J. Pearson, C. Fox, J. C. Sullivan, T. J. Prescott, T. Pipe, and B. Mitchinson, "Simultaneous localisation and mapping on a multi-degree of freedom biomimetic whiskered robot," in *2013 IEEE International Conference on Robotics and Automation*. Karlsruhe, Germany: IEEE, May 2013, pp. 586–592. [Online]. Available: http://ieeexplore.ieee.org/document/6630633/

[92] M. Salman and M. J. Pearson, "Advancing whisker based navigation through the implementation of Bio-Inspired whisking strategies," in *2016 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. Qingdao, China: IEEE, Dec. 2016, pp. 767–773. [Online]. Available: http://ieeexplore.ieee.org/document/7866416/

[93] O. Struckmeier, K. Tiwari, M. Salman, M. J. Pearson, and V. Kyrki, "Vita-slam: A bio-inspired visuo-tactile slam for navigation while interacting with aliased environments," in *2019 IEEE International Conference on Cyborg and Bionic Systems (CBS)*, 2019, pp. 97–103.

[94] Z. Lu, Z. Hu, and K. Uchimura, "Slam estimation in dynamic outdoor environments: A review," in *Intelligent Robotics and Applications*, M. Xie, Y. Xiong, C. Xiong, H. Liu, and Z. Hu, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 255–267.

[95] L. Zhang and J. C. Trinkle, "The application of particle filtering to grasping acquisition with visual occlusion and tactile sensing," in *2012 IEEE International Conference on Robotics and Automation*. St Paul, MN, USA: IEEE, May 2012, pp. 3805–3812. [Online]. Available: http://ieeexplore.ieee.org/document/6225125/

[96] C. Strub, F. Worgotter, H. Ritter, and Y. Sandamirskaya, "Correcting pose estimates during tactile exploration of object shape: a neuro-robotic study," in *4th International Conference on Development and Learning and on Epigenetic Robotics*. Genoa, Italy: IEEE, Oct. 2014, pp. 26–33. [Online]. Available: https://ieeexplore.ieee.org/document/6982950

[97] C. Fox, M. Evans, M. Pearson, and T. Prescott, "Tactile slam with a biomimetic whiskered robot," in *2012 IEEE International Conference on Robotics and Automation*, 2012, pp. 4925–4930.

[98] R. Godwin and G. Spoor, "Soil failure with narrow tines," *J. Agricultural Engin. Research*, vol. 22, no. 3, pp. 213–228, Sep. 1977. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/0021863477900440

[99] A. Elfes, "Using occupancy grids for mobile robot perception and navigation," *Computer*, vol. 22, no. 6, pp. 46–57, Jun. 1989. [Online]. Available: http://ieeexplore.ieee.org/document/30720/

[100] R. Senanayake and F. Ramos, "Bayesian Hilbert Maps for Dynamic Continuous Occupancy Mapping," 2017, pp. 458–471.

[101] S. T. O'Callaghan and F. T. Ramos, "Gaussian process occupancy maps," *The International Journal of Robotics Research*, vol. 31, no. 1, pp. 42–62, Jan. 2012. [Online]. Available: http://journals.sagepub.com/doi/10.1177/0278364911421039

[102] E. Galceran and M. Carreras, "A survey on coverage path planning for robotics," *Robotics and Autonomous Systems*, vol. 61, no. 12, pp. 1258–1276, Dec. 2013. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S092188901300167X

[103] B. J. P. Kaus, "Factors that control the angle of shear bands in geodynamic numerical models of brittle deformation," *Tectonophysics*, vol. 484, no. 1, pp. 36–47, Mar. 2010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0040195109004831

[104] W.-K. Hong, "Chapter 3 - the investigation of the structural performance of the hybrid composite precast frames with mechanical joints based on nonlinear finite element analysis," in *Hybrid Composite Precast Systems*, ser. on Civil and Structural Engin., W.-K. Hong, Ed. Woodhead Publishing, 2020, pp. 89–177. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780081027219000030

[105] J. R. F. Arthur, T. Dunstan, Q. A. J. L. Al-Ani, and A. Assadi, "Plastic deformation and failure in granular media," *Géotechnique*, vol. 27, no. 1, pp. 53–74, Mar. 1977. [Online]. Available: https://www.icevirtuallibrary.com/doi/10.1680/geot.1977.27.1.53

[106] A. A. Taha and A. Hanbury, "An efficient algorithm for calculating the exact hausdorff distance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 11, pp. 2153–2163, 2015.

[107] P. Balatti, D. Kanoulas, N. Tsagarakis, and A. Ajoudani, "A method for autonomous robotic manipulation through exploratory interactions with uncertain environments," *Autonomous Robots*, vol. 44, no. 8, pp. 1395–1410, Nov. 2020. [Online]. Available: https://link.springer.com/10.1007/s10514-020-09933-w

[108] W. Yuan, S. Dong, and E. Adelson, "GelSight: High-Resolution Robot Tactile Sensors for Estimating Geometry and Force," *Sensors*, vol. 17, no. 12, p. 2762, Nov. 2017. [Online]. Available: http://www.mdpi.com/1424-8220/17/12/2762

[109] M. Lambeta, P.-W. Chou, S. Tian, B. Yang, B. Maloon, V. R. Most, D. Stroud, R. Santos, A. Byagowi, G. Kammerer, D. Jayaraman, and R. Calandra, "DIGIT:

A Novel Design for a Low-Cost Compact High-Resolution Tactile Sensor with Application to In-Hand Manipulation," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 3838–3845, Jul. 2020, arXiv: 2005.14679. [Online]. Available: http://arxiv.org/abs/2005.14679

[110] L. Zhang, E. Chong, and V. J. Santos, "Locating Buried Objects through Granular Media Jamming," p. 4.

[111] S. Dong, D. K. Jha, D. Romeres, S. Kim, D. Nikovski, and A. Rodriguez, "Tactile-RL for Insertion: Generalization to Objects of Unknown Geometry," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. Xi'an, China: IEEE, May 2021, pp. 6437–6443. [Online]. Available: https://ieeexplore.ieee.org/document/9561646/

[112] N. G. Tsagarakis, D. G. Caldwell, F. Negrello, W. Choi, L. Baccelliere, V. Loc, J. Noorden, L. Muratore, A. Margan, A. Cardellino, L. Natale, E. Mingo Hoffman, H. Dallali, N. Kashiri, J. Malzahn, J. Lee, P. Kryczka, D. Kanoulas, M. Garabini, M. Catalano, M. Ferrati, V. Varricchio, L. Pallottino, C. Pavan, A. Bicchi, A. Settimi, A. Rocchi, and A. Ajoudani, "WALK-MAN: A High-Performance Humanoid Platform for Realistic Environments," *Journal of Field Robotics*, vol. 34, no. 7, pp. 1225–1259, 2017, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/rob.21702. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/rob.21702

[113] P. Balatti, D. Kanoulas, N. Tsagarakis, and A. Ajoudani, "A method for autonomous robotic manipulation through exploratory interactions with uncertain environments," *Autonomous Robots*, vol. 44, no. 8, pp. 1395–1410, Nov. 2020. [Online]. Available: https://link.springer.com/10.1007/s10514-020-09933-w

[114] E. T. Rolls and A. Treves, "The relative advantages of sparse versus distributed encoding for associative neuronal networks in the brain," *Network: Computation in Neural Systems*, vol. 1, no. 4, pp. 407–421, Jan. 1990. [Online]. Available: https://www.tandfonline.com/doi/full/10.1088/0954-898X_1_4_002

[115] H. Lee, A. Battle, R. Raina, and A. Y. Ng, "Efficient sparse coding algorithms," in *Advances in Neural Information Processing Systems 19*, B. Schölkopf, J. C. Platt, and T. Hoffman, Eds. MIT Press, 2007, pp. 801–808. [Online]. Available: http://papers.nips.cc/paper/2979-efficient-sparse-coding-algorithms.pdf

[116] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, Sep. 1995. [Online]. Available: http://link.springer.com/10.1007/BF00994018

[117] H.-T. Lin, C.-J. Lin, and R. C. Weng, "A note on Platt's probabilistic outputs for support vector machines," *Machine Learning*, vol. 68, no. 3, pp. 267–276, Aug. 2007. [Online]. Available: http://link.springer.com/10.1007/s10994-007-5018-6