A Review Of Recent Research of Tactile Sensing For Robot Grasping

Cognitive Systems Advanced Seminar

B.Sc. Stephan Reith

supervised by
Dr.-Ing. Mohsen Kaboli

Institute for Cognitive Systems
Technical University of Munich

July 16, 2018

Abstract

Reliable grasping and manipulation of objects is an important skill for service robots. The benefit of tactile sensors has been recognized and the development towards anthropomorphic sensors and hands, together with the advances in the broad field of machine learning, create new opportunities to extend common approaches with learning from experience and make generalization to unknown objects easier. This report provides an overview of the recent advances in robot grasping using tactile sensing, in combination with or without additional visual input. Existing methods are summarized and examined for their advantages and disadvantages and a structured overview of the approaches and the hardware provides possible starting points for future work.

Introduction

For service robots, the interaction with objects and the environment is a crucial capability. This comprises a lot of subproblems. It begins with the perception of the target object. A grasp needs to be planned and carried out and the robot needs to be able to estimate wether the grasp is stable or if slip occurs. If the grasp is not stable, the robot decides if it just needs to adjust the grasp slightly or if it needs to regrasp.

Since the ultimate goal is for the robot to use the grasped object, for example a tool, dexterous manipulation and the interaction of the grasped object with the environment are very important. This should be kept in mind and approaches to robot grasping should be assessed, inter alia, in this context as well as their ability to generalize to unknown objects.

Dahiya *et al.* differentiate between "perception for action" [1, p. 2] and "action for perception" [1, p. 2], meaning the robot can either use tactile information to improve grasping and object manipulation, or it can perform exploratory actions to learn about the objects tactile and geometrical properties [2]. This work focuses on the former and aims at giving an overview of the recent research in robot grasping, focusing on the sensor modalities tactile and tactile plus vision.

Tactile and Vision Sensors

Vision is the easiest way for a robot to get an insight into the scene. The robot can recognize and locate objects in his workspace and pointclouds from RGBD-cameras enable the use of geometry-based grasp planners ([3]).

Visual information however is also subject to uncertainty. This can result from poor lightning conditions or the low depth resolution with increasing distance (Kinect) on the one hand. On the other hand can cameras only perceive the facing part of the object ([4]) and objects may be occluded by other objects or even the gripper itself during the grasp. Purely vision based grasping is therefore mainly open-loop control.

Tactile sensors on the inside of the robot hand or on the fingertips enable to close the control loop and provide information about contact forces, friction, surface topology, compliance and pressure.

In combination with concepts like for example the Virtual Frame method ([5]), this allows for an estimation of the objects position, based solely on the position of the fingertips in contact with the object ([6]). Tactile sensors are also key for the detection of slip, the occurance of which indicates an unstable grasp ([7]), and grasps with minimal force.

Many different tactile sensors have been proposed with working principals ranging from capacitive over piezoelectric to optical ([1]). The following sensors have been used especially often in recent works or offer interesting possibilities and are therefore to be presented here.

The GelSight sensor ([8]) and its enhancement the GelSlim sensor ([9]) are optical based sensors. They consist of an elastomer with an overlying membrane, which adapts to the objects surface when pressed against it. A small camera underneath films this membrane, so the sensor returns an image of the objects surface topology. Common image processing methods like convolutional neural networks can be used in combination with this sensor and make the integration with camera images attractive ([10], [11]).

Tactile pads and sensor arrays measure force or pressure. Capacitive pads are sensitive, robust, have a good spatial resolution and are small, so they are usually arranged as arrays ([1]). In combination with the forward kinematics of the hand, they can also determine the three dimensional contact point. Similar to capacitive pads are (piezo-) resistive cells. High sensitivity and low cost ([1]) allow to cover the whole inner hand with resistive cells, however they must be properly calibrated to relate analogue signals to force ([12]).

The SynTouch BioTac is a biomimetic sensor and can measure force, vibration signals and temperature at once. It also makes use of an elastomer, with a fingerprint texture on the outside. The inner layer consists of an incompressable conductive fluid. Electrodes measure the impedance, which correlates with the thickness of the fluid layer and changes according to the applied force. A hydro-acoustic pressure sensor measures the vibration signal of the liquid, with a sensitivity that exceeds human performance ([13]). Embedded thermistors sense the temperature.

Data-Driven Grasping

How tactile data is processed, how the robot hand or gripper is controlled and how grasp stability is determined differs a lot depending on the goal. Approaches range from model-based to data-driven.

Calandra *et al.* [10] train a convolutional neural network (CNN) to estimate the probability of a stable grasp based on the current sensor input and the action taken. Possible next actions can then be sampled and their success evaluated with the network. Levine *et al.* [14] train a similar (purely vision-based) network with real grasps, collected from up to 14 robots in parallel over two months. Hogan *et al.* [11] combine a CNN stability estimator with a point-cloud based grasp planner. After the initial planned grasp, an improved regrasp is estimated using rigid-body transformations on the tactile readings. Guo *et al.* [15] learn to associate rectangular grasping positions with objects.

Such neural networks need a lot of training data and can typically not be trained in simulation, because simulations do not reflect the physical properties and the real sensor data properly, which means they require many robot hours. Their general advantage is that they generalize well for similar enough objects. Nonetheless, tactile sensing as used in these approaches only serves as an additional source of information to learn pick and place tasks end-to-end (motor commands from sensor input) with a two-fingered gripper and top-down grasps. Since no planning is involved, no models or sensor calibration are needed [10], but with increasing complexity of the grasping and manipulation task, pure end-to-end learning becomes more and more impractical.

The before mentioned purely vision-based approach plays a special role in this context, because it does not employ tactile sensing. It does however show that machine learning approaches can be trained from the real-world experience of many distributed robots ([14]), which increases amount and diversity of the available training data.

Grasp Controllers

A different approach is chosen by Li *et al.* [6]. Instead of learning the motor commands from data, they use an impedance controller to grasp the object. Tactile readings are combined into features and a Gaussian Mixture Model determines the stability of the grasp, so the parameters of the controllers can be adapted accordingly. In different work, Li *et al.* [16] learn the parameters for impedance control from human demonstration.

Regoli *et al.* [17] propose a hierarchical controller, which can hold a stable grasp and regulate the grasping force at the same time. Li *et al.* [18] control both the position of an object between two fingers and the force, so that the object position follows a trajectory while restricting the grasping force.

Hybrid Position and Force Controllers usually have better precision than impedance controllers, but the transition between the two modes can cause stability problems [6].

Another important control objective is minimum force control. Kaboli *et al.* [7] use three dimensional contact forces to estimate the friction and prevent grasped objects from slipping without deforming them. This method proved to keep a successful grasp, even when the weight of the object changes. Delgado *et al.* [12] perform a squeezing action to estimate how rigid the object is and determine the grasping force accordingly. A PID-controller in combination with forward kinematics controls the position and force for each finger.

Learning from Experience

Using controllers on a lower level in combination with machine learning methods, which learn how to adapt the respective parameters on a higher level, requires a lot less training experience than end-to-end learning. Dang *et al.* [19] derive a stable grasp from an unstable one using a similarity metric and a grasp database. Bekiroglu *et al.* [20] represent the object characteristics, gripper configuration and the sensor readings in one latent variable, which allows to identify the factors directly, which need to be adjusted to make an unstable grasp stable.

A unified framework from grasp planning to grasp adaption is proposed by Hang *et al.* [21]. A hierarchical fingertip space represents possible fingertip configurations around the object. A grasp planner finds a good hand configuration and reaches it with position control. During the grasp, the controller regulates the grasping force with impedance control and if the grasp is unstable (see [6]), one of the fingers changes its fingertip position.

In this work, the fingertip space is built based on pointclouds of the objects. This limits the generalizeability to new objects, because the pointcloud might not immediately be accessible for new objects, especially if the object is occluded. But the concept of a fingertip space (or similar spaces) has a lot of potential, since control, stability estimation, parameter learning and planning can be combined in one framework. It could be interesting future work to derive such a fingertip space for new objects based on similarity to geometrical and tactile properties of known objects, or learn it from human demonstration.

Reinforcement Learning Methods

A completely different approach is to let the robot learn how to grasp objects itself, namely reinforcement learning. Van Hoof *et al.* [22] use a Markov Decision Process (MDP) to let the robot learn a policy for rolling an object between its fingers. No system analysis or kinematic model is needed, but similar to CNNs, learning requires physical interaction, which is why physical reinforcement learning problems are often simplified. In this case, only an improvement of a hard-coded policy is learned. Chebotar *et al.* [23] learn to adjust the parameters for top-down regrasps based on spatio-temporal feature descriptors, however generalization to new objects is very limited. Stork *et al.* [24] rotate a pen in-hand and use the Predictive State Representation framework to describe the hidden state of their dynamical system from observations. The state consists of possible future tests and Value Iteration is used to learn the optimal policy.

Underactuated Robot Hands

Underactuated and compliant robot hands play a special role and have the advantage that control is a lot simpler, due to the reduced degrees of freedom. Luberto *et al.* [4] equipped one with infrared sensors and close the hand uniformly around the object. The fingers press the object against the palm, which makes grasps immediately more stable than fingertip grasps, but makes it a lot harder to manipulate the objects, since the kinematics of the hand and the position of the object in-hand are unknown. Van Hoof *et al.* [22] use an underactuated hand with one tendon per finger when they learn how to roll an object between fingers using reinforcement learning. This reduces the dimension of the state vector and therefor the total number of states and learning time drastically.

Object Manipulation

As mentioned before, the long-term goal is not just to grasp an object, but actually use it. Stork *et al.* [24] use Predictive State Representation and Value Iteration to learn how to rotate a pen in a two-fingered gripper, by pushing it against a table. The position of the pen is hereby observed from the outside with a camera, the pressure readings are processed into tactile features. Molchanov *et al.* [25] compare different machine learning methods (neural networks, Gaussian Processes and Support Vector Machines) to estimate which point on the surface of an grasped object is in contact with the environment, based on the tactile information of the BioTac sensor. No models are needed, but the error is quite high.

Problems and Future Work

Most approaches employ machine learning methods. Depending on the approach, whether end-to-end learning is used or sensor information is encoded in features, how much training data is actually required varies a lot. But the performance of machine learning algorithms exactly depends on the quality of the datasets. Since real world interaction with the robot is costly, well structured underlying representations of the experience can reduce the amount of training data and increase the variety of unknown objects that can be grasped. Hierarchical frameworks with low-level controllers and high-level planners offer the best compromise between generalizeability on the one hand, and acceptable training effort on the other hand. One good example is presented in section 10.

With the improvement of tactile sensors from pressure sensor pads to multi-modal anthropomorphic sensors it becomes easier to estimate the objects properties, which can be used to categorize experience and apply it to completely unknown objects by exploring them first and then reason grasping configurations and forces.

The approaches which use reinforcement learning are a good example. At the moment, many restrictive simplifications need to be made. The more informative dynamical state representations become, the more simplifications can be relaxed.

Table 1: Summary of Methods

No.	Year	Sensor Modality	Method	Applications
1	2018	Visual and Tactile	Convolutional Neural Network (CNN)	Grasping, Regrasping, Outcome Prediction
2	2014	Tactile	Impedance Controller, Gaussian Mixture Model (GMM), Support Vector Machine (SVM)	Grasp Adaption, Stability Estimation
3	2016	Visual	CNN, Large Scale Experimental Data Collection	Outcome Prediction, Grasping
4	2016	Visual and Tactile	Latent Variable Model	Stability Estimation, Grasp Adaption
5	2018	Visual and Tactile	CNN	Quality Metric, Regrasping
6	2017	Visual and Tactile	CNN	Stability Estimation, Grasping
7	2016	Tactile	PID Control, GMM	Grasp Adjustment, Grip Strength Control
8	2013	Visual and Tactile	Position/Force Controller	In-hand manipulation
9	2014	Tactile	Impedance Control	Grasping, Dexterous Manipulation
10	2016	Visual and Tactile	Hierarchical Fingertip Space, Impedance Control	Grasp Synthesis and Adaption
11	2015	Tactile	Markov Decision Process	Dexterous Manipulation
12	2016	Tactile	Tactile Features, SVM, Policy Learning	Stability Estimation, Regrasping
13	2016	Tactile	Tangential Force Estimation	Slip Detection, Manipulation, Deformation Prevention
14	2014	Visual and Tactile	SVM, K Nearest Neighbors (KNN), Tactile Experience Database	Stability Estimation, Grasp Adjustment
15	2015	Tactile	Kernel Logistic Regression, (Tactile) Image Moments, Prototypical Shapes	Stability Estimation
16	2017	Visual and Tactile	Control Optimization, Gazebo Simulation	Grasping, Pre-Grasp Adaption
17	2015	Visual and Tactile	Predictive State Representation, String Kernel Features, Value Iteration	Dexterous Manipulation
18	2015	Tactile	Locally Weighted Projection Regression, Neural Networks, Force-Derivative, Vibration Signal	Contact Force Estimation, Slip Detection and Classification
19	2016	Tactile	Neural Network, Gaussian Process, Vibration Signal	Object-Environment contact point estimation
20	2017	Visual and Tactile	Tactile Servo Controller, State Machine Task Planner	Minimal Force Grasps, Object Stiffness Estimation, Execution of sequential actions

Contents

1 More Than a Feeling: Learning to Grasp and Regrasp using Vision and Touch

[10] R. Calandra, J. Lin, A. Owens, J. Malik, U. Berkeley, D. Jayaraman, and E. H. Adelson, "More than a feeling: Learning to grasp and regrasp using vision and touch," *no. Nips*, pp. 1–10, 2017

Motivation And Problem Setting

The process of grasping is highly interactive, and can therefore highly benefit from tactile feedback. However, the performance of analytic grasping methods depends on how good the model fits the real world. On the contrary, data-driven methods are so far largely vision-based, which limits them to choosing a grasp in advance only. This paper presents a data-driven, action-conditional model, which uses raw vision and touch input and allows for reactive adjustments during the grasp.

Sensors

This work uses two Microsoft Kinect 2 cameras for visual input from the front and top, as well as two GelSight optical tactile sensors on the two fingertips. These tactile sensors provide raw topological image-like input, as shown in Figure 1. The GelSight sensor consists of a gel with a reflective membrane on the surface and a camera on the bottom, visualizing the deformation of the membrane.

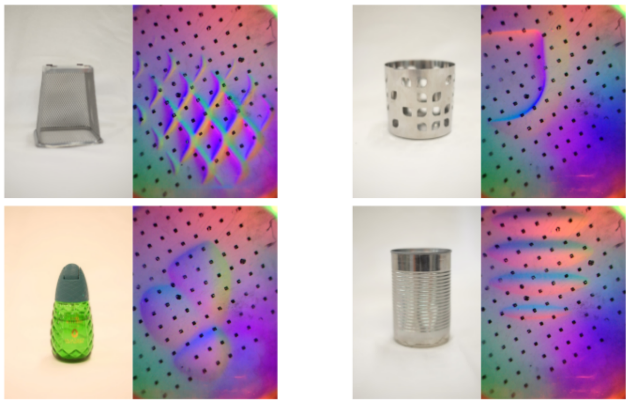

Figure 1: Raw input of the GelSight sensors. [10]

Method

A Markov Decision Process always selects the next action, which has the highest predicted grasp success probability. The possible actions are randomly sampled. This probability is predicted using a neural network architecture. Convolutional sub-networks are used to pre-process the input of each of the two tactile sensors and the camera, respectively. Position and orientation of the gripper are pre-processed in a multi-layer perceptron. These sub-nets are concatenated using another multi-layer perceptron, which output is the predicted probability. This is visualized in Figure 1.

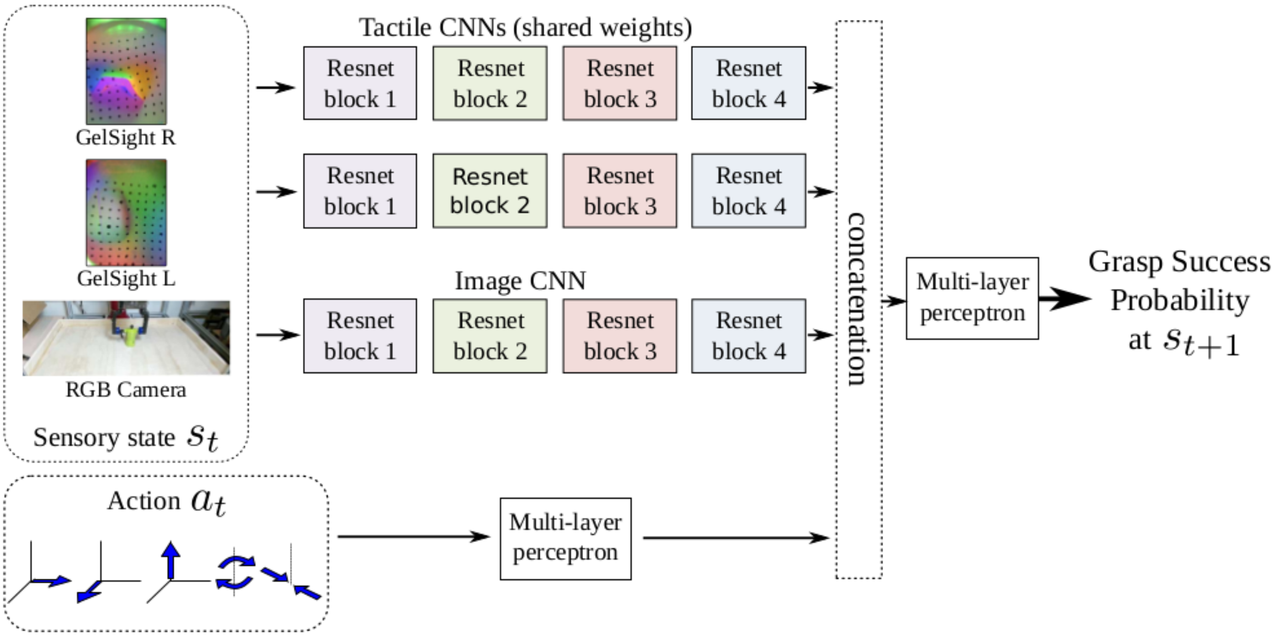

Figure 2: Model network architecture. [10]

Data Collection

The model was trained using the success of automatically collected state-action pairs. The robot executed random grips with a random amount of force, and a deep neural network classifier, based on the tactile input, determined whether the grab was successful or not.

Results

The model indeed learns to use actions to predict the next grip and performs better than visual-only or tactile-only models. Experiments with 11 unseen objects with easy shapes outperform the vision-only and the cylinder method with a 94% average grasp success rate, and 73.6% for objects of difficult shape.
The model seemed to have learned that using more force and a lower contact point is beneficial for stable grasps, whereas for unstable grasps less force is better.
A minimum force optimization had no impact on the performance, while using significantly less grasping force.

Advantages

- reactive grasp adjustments
- no analytical model, engineering of features and no calibration are needed
- force constraints are permitted
- high success rate

Disadvantages

- No correlation between the grasping height and the center of mass seems to be learned
- possible next actions are sampled randomly
- so far no information gathering actions are applied
- untested for cluttered environments

2 Learning of Grasp Adaptation through Experience and Tactile Sensing

[6] M. Li, Y. Bekiroglu, D. Kragic, and A. Billard, "Learning of grasp adaptation through experience and tactile sensing," in *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*, pp. 3339–3346, Ieee, 2014

Motivation And Problem Setting

When executing planned grasps, the grasp might not be stable due to model and control errors. One possibility is to model uncertainties, however more information can be obtained by sensory feedback from the fingertips. This paper proposes a strategy to estimate the grasp stability and to adapt the grasp under uncertainties of the physical object.

Sensors

This paper uses three of the four fingers of the Allegro Hand, with BioTac tactile sensors mounted on the fingertips. Only the pressure data from the sensor is used.

Method

Position and orientation of the grasped object is estimated using the Virtual Frame method and a object level impedance controller is used for grasping. The parameters of the impedance controllers are the grasp stiffness and the distance from the fingertip to the center of the virtual frame (rest length).

In the feature space consisting of the grasp stiffness, rest length and tactile reading for each fingertip, a Gaussian Mixture Model and a Support Vectore Machine are used to estimate the region of stable grasps and classify a grasp as stable or unstable.

If a grasp is unstable and its point in the feature space is close enough to a learned model, the grasp stiffness is adapted according to the model.

If the grasp is not close enough to a model in the feature space, the point is projected to the closest center of the Gaussian components and the respective rest length serves as basis for the adaption of finger one (fingertip impedance controller). This finger is adapted until a similar stable grasp can be found.

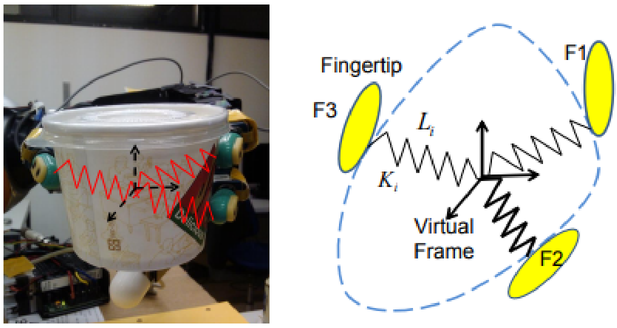

Figure 3: The impedance controller controls position and orientation of the virtual frame instead of the actual object. [6]

Experiment

The robot hand is already preshaped for the grasp, the geometry of the four objects is not known. 250000 stable grasps are recorded with different object weights (they are filled with pepper) and different grasping stiffness. Finally grasps are disturbed by a human or by adding weight and the grasp adaption is compared to no grasp adaption.

Results

When predicting grasp stability, the GMM has a significantly better false positive rate than the SVM, i.e. it mistakes less unstable grasps for stable ones, while the true positive rates are similar.

It is shown, that with the adaption, the grasps can support significantly more weight than without (see table 2).

obj.	cola can	food box	box	cup
w/o	17.2 ± 1.92	12.8 ± 0.84	37.2 ± 2.59	15.0 ± 2.55
with	69.0 ± 6.52	84.0 ± 3.80	121.2 ± 9.20	146.4 ± 5.46

Table 2: The comparison of the supported object weights. [6]

Advantages

- grasps can be adapted dynamically
- no prior geometrical information necessary, uncertainties can be handled
- the grasps can support significantly more weight

Disadvantages

- only finger one is used for exploration
- the grasp can become unstable when finger one loses contact during exploration

3 Learning Hand-Eye Coordination for Robotic Grasping with Large-Scale Data Collection

[14] S. Levine, P. Pastor, A. Krizhevsky, and D. Quillen, "Learning hand-eye coordination for robotic grasping with large-scale data collection," in *International Symposium on Experimental Robotics*, pp. 173–184, Springer, 2016

Motivation And Problem Setting

Robotic manipulation tasks often use heavy advance analysis and planning and have relatively simple control feedback (e.g. trajectory control), whereas humans rely on immediate visual and tactile feedback instead. Inspired by this, this work uses a data-driven approach with end-to-end training to learn the grasp success probability based on visual sensor input and the gripper motion. This is used like a closed loop control system, where the best motor command is continuously computed based on the current sensory input.

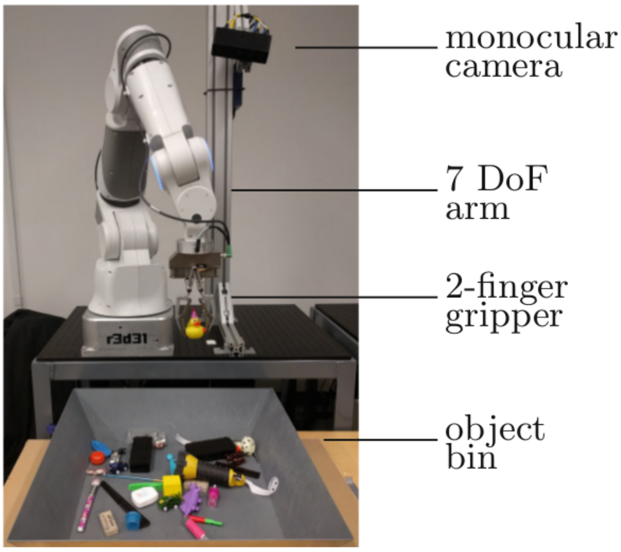

Figure 4: The robotic setup. [14]

Sensors

This method only uses a monocular camera, which observes the workspace of the robot. The camera position relative to the robot is not important and varies. The input images of size 512x512 px are randomly cropped to 472x472 px patches for translation invariance.

Method

A convolutional neural network uses two images as input: the current scene, as well as one reference image before the grasping task without the gripper. The third input is the current motor command. This neural network is used to predict the probability of successful grasping. After the network was trained, possible motor commands are randomly sampled using the cross-entropy method, and their respective probability for success is calculated for the current visual input. The motor command with the highest success probability is executed.

When the success probability for the current visual input in combination without any action is above 90%, the gripper is closed, and when it is less then 50%, the gripper is lifted off the table so a better configuration can be found.

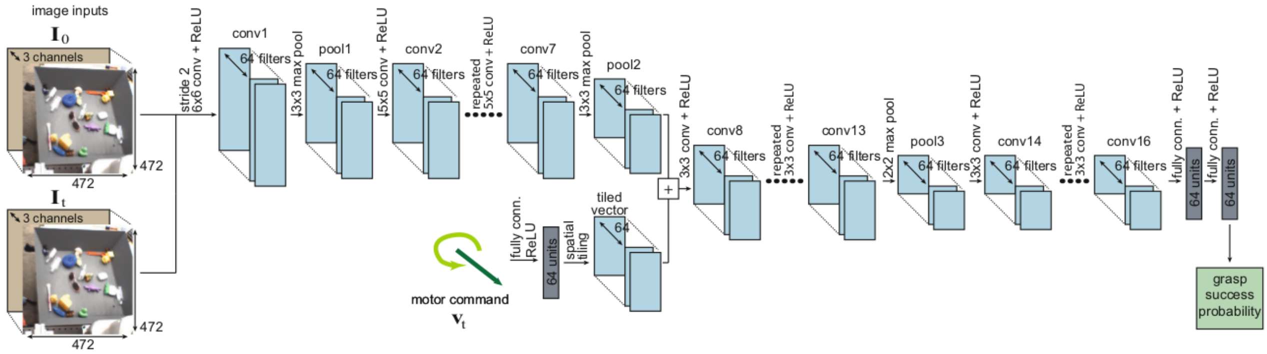

Figure 5: The convolutional neural network architecture. It uses the current and the reference image and the motor command as input and predicts the probability of successful grasping. [14]

Data Collection

Between 6 and 14 real robots performed grasps over two months, picking up objects from a cluttered workspace. The camera position varied, the robots had slight differences in the shape of the gripper and the workspace was regularly shuffled to ensure generalization. Each grasp attempt was divided into timesteps, each of which consists of the observed image, the current gripper movement, and the label wether or not the grasp was successful in the last timestep.

The success of a grasp attempt was automatically evaluated by dropping the grasped object into the bin and comparing images of it before and after the drop.

without replacement	first 10 ($N = 40$)	first 20 ($N = 80$)	first 30 ($N = 120$)
random	67.5%	70.0%	72.5%
hand-designed	32.5%	35.0%	50.8%
open loop	27.5%	38.7%	33.7%
our method	**10.0%**	**17.5%**	**17.5%**

with replacement	failure rate ($N = 100$)
random	69%
hand-designed	35%
open loop	43%
our method	**20%**

Figure 6: Failure rates for each method, with and without replacement, for N grasp attempts. [14]

Results

Objects were picked up from a cluttered bin, with and without replacement, and compared to an open loop variation of their method, using only a pre-grasp image as input to the neural network and moving the robot to the position with the best predicted respective action. It was also compared to a random baseline method and a hand-engineered approach based on depth images.

Figure 6 shows that the proposed method has the lowest failure rates for both scenarios, and that the closed-loop approach outperforms the open-loop approach.

Advantages

- No hand-to-eye camera calibration is required and differences in hardware can be compensated
- No planning is required and mistakes and movements of the objects can be compensated
- A large variety of objects is being learned and 800,000 grasp attempts are used for training
- Could possibly be combined with other sensory modalities, especially the GelSight tactile sensor could possibly integrate well into

the convolutional neural network architecture

- When the next motor commands are sampled, workspace and robot constraits can be taken into account

Disadvantages

- Large amounts of training data from real physical robots is required, which is costly and time consuming

- Only a two finger gripper is used. Extensions for more complicated manipulation with more fingers, in combination with different input modalities could make end-to-end learning impractical, unless data is collected from a large number of deployed robots performing real world tasks.

4 Probabilistic Consolidation of Grasp Experience

[20] Y. Bekiroglu, A. Damianou, R. Detry, J. A. Stork, D. Kragic, and C. H. Ek, "Probabilistic consolidation of grasp experience," in *Robotics and Automation (ICRA), 2016 IEEE International Conference on*, pp. 193–200, IEEE, 2016

Motivation And Problem Setting

All approaches to robotic grasping, analytic or data-driven, as well as sub-problems like grasp adaption and stability estimation, suffer from noisy data, insufficient experience or high dimensionality and often have difficulties with new situations. This paper aims at providing a probabilistic model for selecting and merging relevant information from multiple sensor modalities, and reason about the probability of success.

Sensors

A Kuka articulated arm is equipped with a Schunk SDH hand. Capacitive pads measure the pressure on the three fingers and a RGB camera observes the robots workspace and continuously tracks the pose of all objects.

Method

The tactile characteristics and the gripper pose relative to the object, for successful and unsuccessful grasps respectively, as well as the object identity and gripper orientation are consolidated into one latent variable using the Manifold Relevance Determination model [26]. Each of these factors is weighted using the Automatic Relevance Determination method [27], making feature selection possible.

This learned model is now used to infer a stable from an unstable view by finding the shared parameters (the ones with non-zero weight between the two views). A nearest neighbor search in the training data finds the data point which matches both views and this point is used to project the unstable grasps to a stable one.

Experiments

An object was placed in the robots workspace, the fingers were pre-shaped for an initial grasp, then closed by the robot and the object was lifted to label the grasp stable or unstable. The camera observed to object and gripper positions. Then the robot executed random grasps in the neighborhood of the initial grasp to become robust against small disturbances. A Gaussian Process, which builds a functional relationship between unstable and stable grasps, is used as a baseline method.

Results

The proposed model has a significantly higher success rate than the baseline method for correcting failed test grasps, as shown in figure 8. Unlike the baseline method, the proposed model has learned which factors to alter.

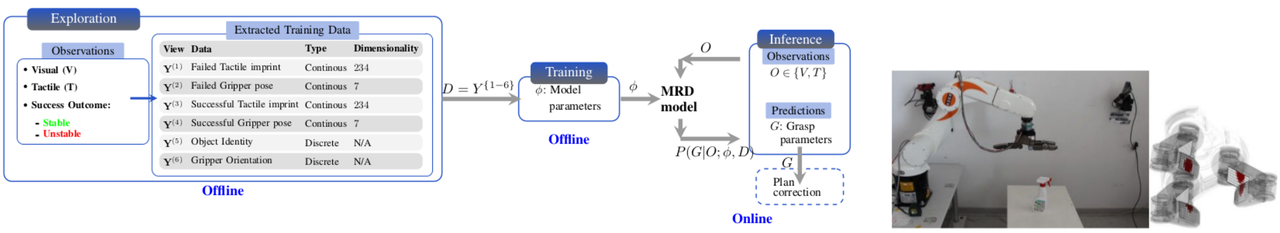

Figure 7: System model overview (left) and experimental robot platform (right). [20]

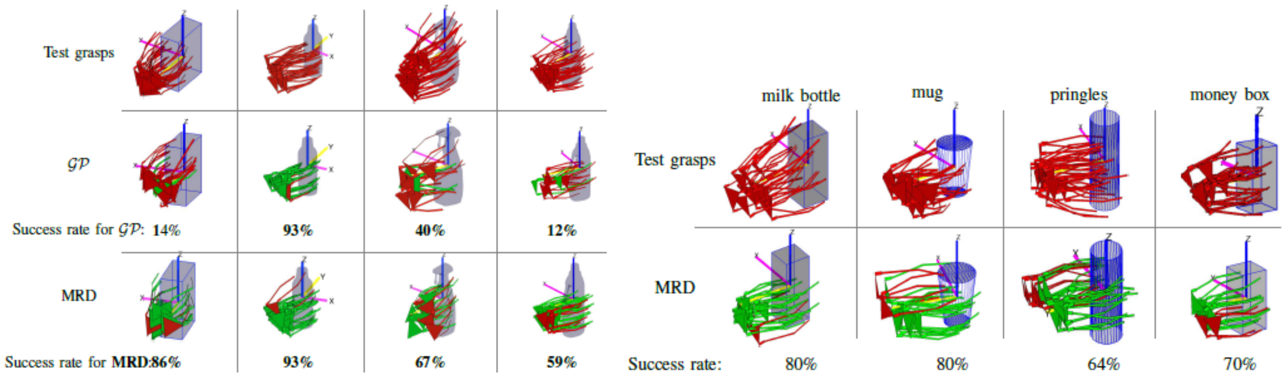

Figure 8: Unstable grasps (top rows) are transferred to stable ones using the respective methods. (Right) Corrections for new objects. [20]

Advantages

- Systematic approach to combine different sensor modalities
- The robot learns from successful as well as from unsuccessful grasps
- The model can generate new stable grasps, which have not been learned
- The model learns which factors to alter
- Can also be used to detect unexpected sensor input and to identify objects

Disadvantages

- Mesh and texture models of all objects are needed to track the pose of all objects, but this does not necessarily limit the general approach
- Only top-down and sideways grasps are considered
- Only four, relatively similar objects have been used for training and testing (see 8), more variation in shape, weight and size would be interesting.

5 Tactile Regrasp: Grasp Adjustments via Simulated Tactile Transformations

[11] F. R. Hogan, M. Bauza, O. Canal, E. Donlon, and A. Rodriguez, "Tactile regrasp: Grasp adjustments via simulated tactile transformations," *arXiv preprint arXiv:1803.01940*, 2018

Motivation And Problem Setting

Tactile sensor information is very beneficial for robot grasping, because the immediate feedback helps to predict the success of an ongoing grasp, which can be used for reactive grasp adjustments. This work uses tactile sensing to evaluate the quality of a grasp and, based on this metric, adjust the grasp such that stability is increased.

Sensors

A two finger gripper is mounted on a robot arm with 6DOF. Each finger is equipped with a GelSlim tactile sensor, which consists of a membrane in front of a camera and produces high-resolution images of the contact surface topology. Two RGBD cameras observe the three bins.

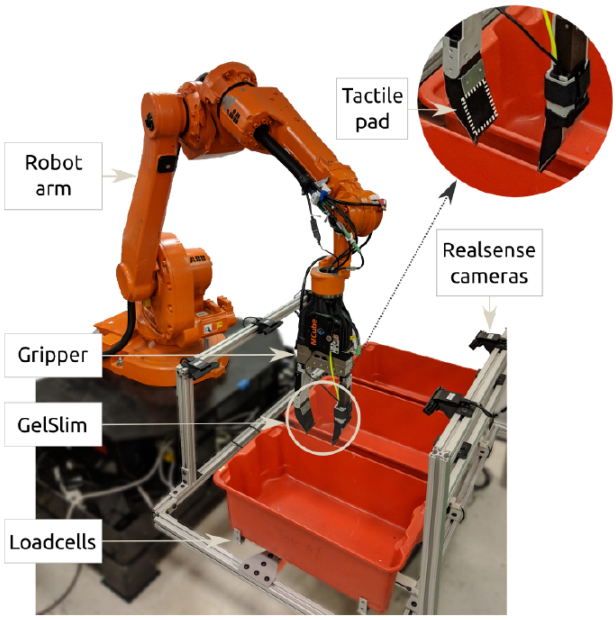

Figure 9: The robot used for the grasping experiments. [11]

Method

A deep convolutional neural network is used to predict the grasp quality. Each of the sensor images is preprocessed using pretrained ResNet50 sub-networks, and the two subnets are connected in an output layer (see figure 10).

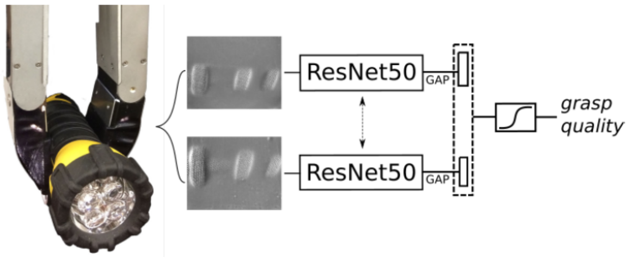

Figure 10: This deep convolutional neural network evaluetes the grasp quality. [11]

After an initial grasp, rigid-body transformation of the sensor image is used to simulate tactile images for all possible regrasps. The regrasp with the best predicted quality is executed (figure 11).

Data Collection

The grasp quality prediction network is trained by executing real self-supervised grasps. A vision-based grasp planner [28] is used for the initial grasp. Infused noise makes the prediction more robust. The robot shakes the grasped object and depending on the time the grasp can still hold the object, the grasp quality is calculated.

Results

The accuracy of the grasp quality prediction is 85% on known objects and 75% on unknown objects, unevenly shaped or heavy being the most difficult objects. The grasp adjustment policy increased the overall accuracy by 14% and outperformed the centroid-centering baseline method.

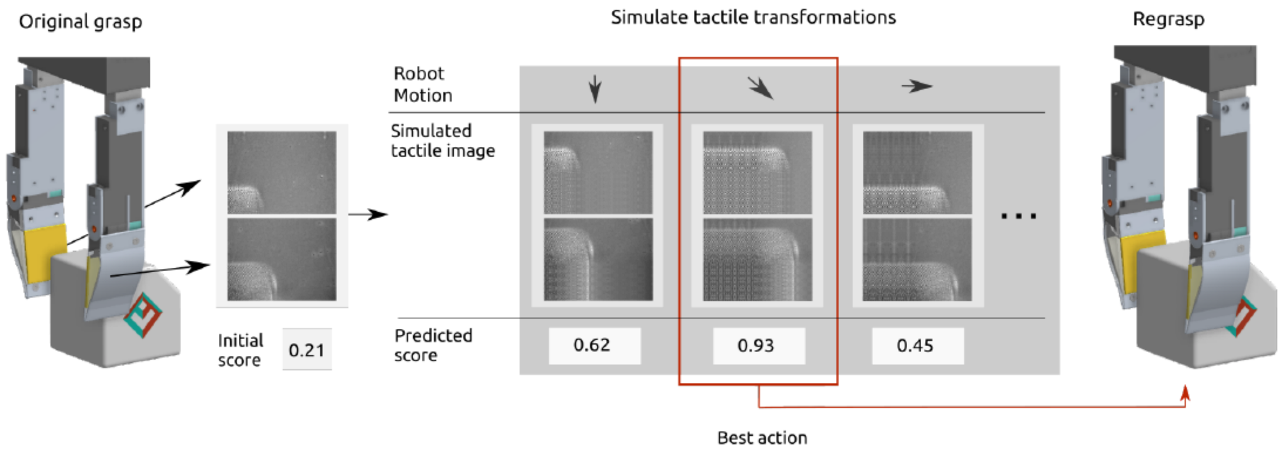

Figure 11: The best regrasp action is chosen using simulated tactile images. [11]

Objects	black pepper (233 g)	heavy play-doh (332 g)	logitech mouse (123 g)	metal bar (255 g)	satin care (276 g)	soy milk (261 g)	tomato can (367 g)	ZzzQuil bottle (237 g)
Images								
No regrasp	61 %	38 %	64 %	75 %	46 %	61 %	36 %	17 %
Tactile-based regrasp	85 %	61 %	78 %	93 %	63 %	83 %	72 %	49 %
Relative Improvement	39 %	61 %	22 %	75 %	37 %	36 %	100 %	188 %

Table 3: Grasping accuracy. [11]

Advantages

- Combination of a data-driven and a model-based approach
- Feedback is only based on the contact area, not the object itself or its geometry
- A wide range of objects with different shapes, texture and hardness was used.

Disadvantages

- The objects position must not change between two grasps, which will not always be satisfied, especially when different grippers are being used
- The gripper needs to leave the object before it can regrasp
- Only local adjustments can be made

6 Robotic grasping using visual and tactile sensing

[15] D. Guo, F. Sun, B. Fang, C. Yang, and N. Xi, "Robotic grasping using visual and tactile sensing," *Information Sciences*, vol. 417, pp. 274–286, 2017

Motivation And Problem Setting

Visual perception can helps robots to discover their environment, while tactile feedback is important for fine motor skills. This paper aims at combining both and proposes a method to detect good grasping spots in an image and uses tactile sensor information to determine the stability of a grasp.

Platform and Sensors

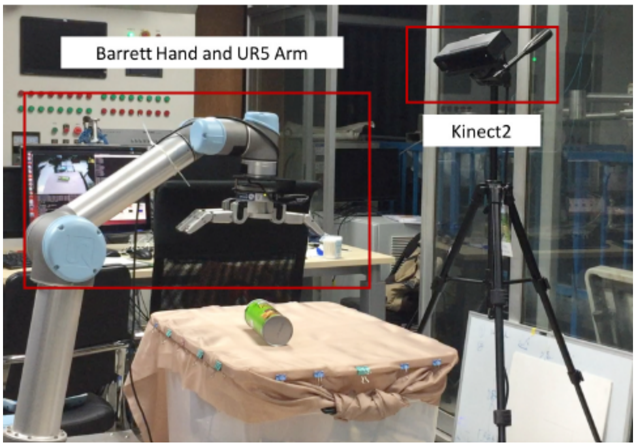

Figure 12: Robot workspace. [15]

A Barrett hand with three fingers is mounted on a UR5 robotic arm. A fixed Kinect2 camera provides depth images of the workspace. It is never explicitly mentioned in the paper, which kind of tactile sensor they use. However, because it is "integrated on the fingertip" [15, p. 11], it can be assumed that the torque sensor from the hand manufacturer is used.

Method

A grasp configuration is represented by a rectangle, as shown in figure 13. A convolutional neural network, consisting of three layers, is trained to determine the parameters of a good grasp rectangle. The first layer of the network extracts features, which determine possible locations. The second layer computes a metric, representing how graspable a reference rectangle is for a feature, and the third layer outputs the parameters of the inferred rectangle, e.g. orientation and size, eventually.

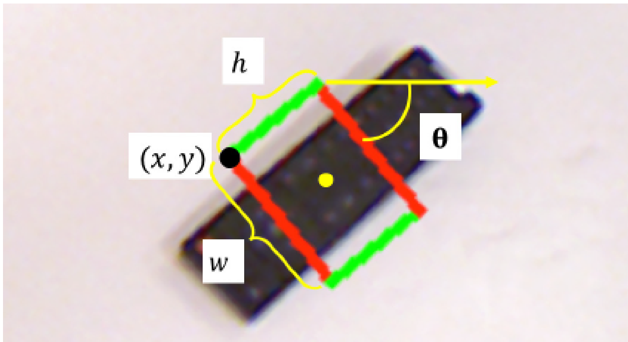

Figure 13: Rectangles are used to represent grasping positions. [15]

Data Collection

The robot collects training data unsupervised, by segmenting the object in the depth image, approaching vertically to the geometrical center of the object, and choosing a random discretized gripper orientation. If the tactile sensor marks the grasps as stable, the object is lifted and shaken to assure stability. The so collected THU dataset comprises 17 different objects.

Results

The public Cornell dataset was used in combination with other approaches and it could be shown that the proposed method outperforms the other methods in terms of grasp detection accuracy on this dataset.

Real grasping experiments have a grasp success rate of 61.2%. in experiments with a different robot and multiple (partially unknown) objects, the robot grasped the object with the strongest rectangle one by one (see figures 14 and 15).

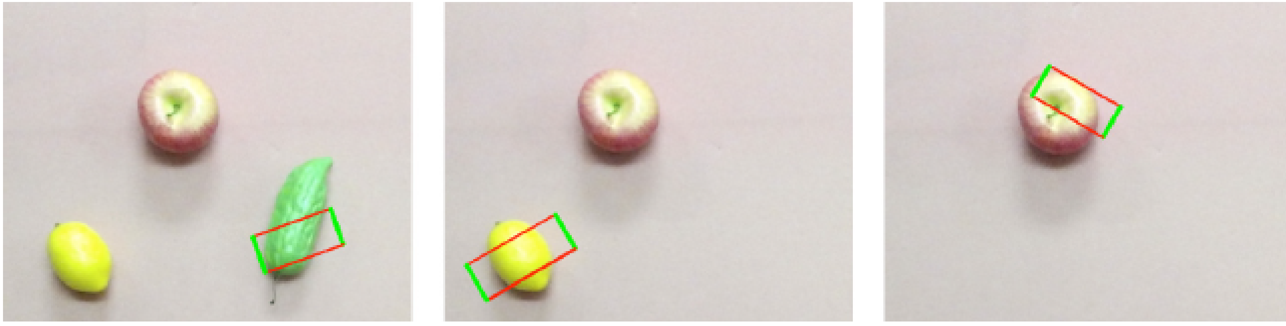

Figure 14: A different robot grasps objects one by one. [15]

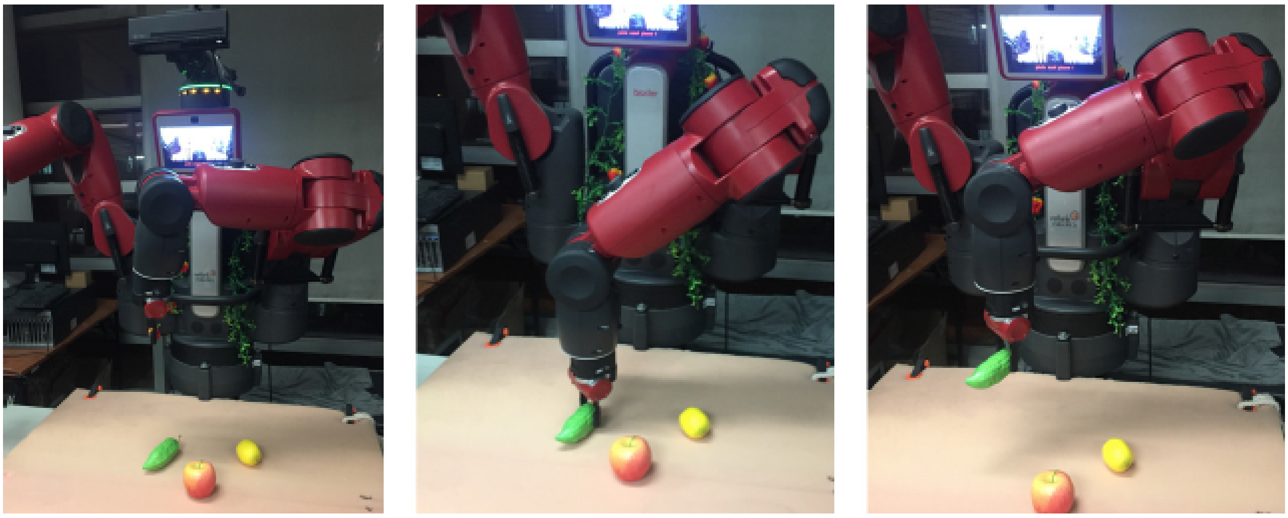

Figure 15: The rectangles corresponding to figure 14. [15]

Advantages

- The proposed method is real time capable
- The method was tested on a different robot and has shown to generalize well

Disadvantages

- Only vertical grasps are considered
- Only grasping positions are predicted in this data-driven approach, the grasping movement is not learned. An accurate model of the robot is therefore still required.

7 Hierarchical grasp controller using tactile feedback

[17] M. Regoli, U. Pattacini, G. Metta, and L. Natale, "Hierarchical grasp controller using tactile feedback," in *Humanoid Robots (Humanoids), 2016 IEEE-RAS 16th International Conference on*, pp. 387–394, IEEE, 2016

Motivation And Problem Setting

Being able to adjust an unstable grasp such that it becomes stable is an important capability for tactile manipulation. Especially when the robot should also be able to handle unknown objects, i.e. no object models are available, grasp preplanning becomes insufficient. This paper proposes a framework which can stabilize a grasp, while still being able to control the grip strength at the same time.

Sensors

Experiments are conducted on the iCub humanoid robot. Three of the five fingers are used, each equipped with twelve capacitive tactile sensors on the fingertips measuring the normal force.

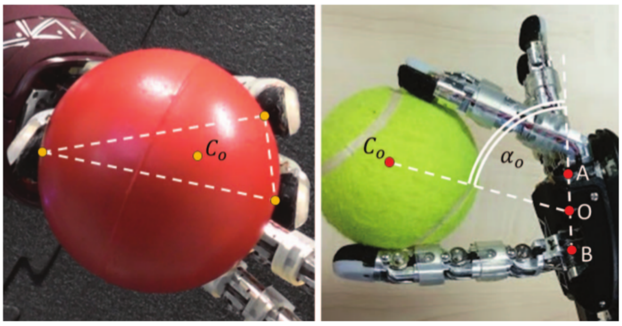

Figure 16: The anthropomorphic hand of the iCub robot. The angle α_0 is used to control the hand configuration. [17]

Method

The low-level PID controller controls the exerted force of each finger. Stable gains for all conditions are determined using classical control theory and by approximating transfer functions for the system. The high level PID controller stabilizes the grasp by controlling the position of the object relative to the palm. The reference position is given by a Gaussian Mixture Model, which learned stable grasps and was trained by human demonstration. Given the parameters of a grasp, the GMM regresses new parameters such that the likelihood of stability is maximized.

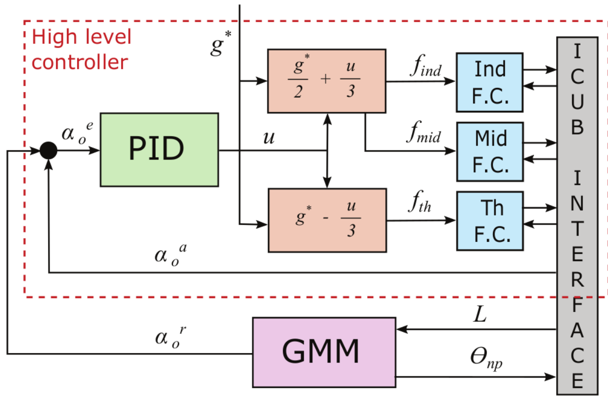

Figure 17: The hierarchical control loop, consisting of low-level PID finger-force controllers, a high level PID controller for the frame-relative object position, and a GMM calculating optimal hand configurations. The desired grip strength $g^\star$ serves as input. [17]

Experiments and Results

A first experiment verified that the high-level controller can track a sinusoidal reference with low error. A second experiment confirmed that the low-level controller follows a reference grip strength accurately. The grasp adjustments regressed by the GMM were tested for stability by shaking the gripper and proofed to be more stable than the baseline method (figure 18). Unknown objects of different shapes, sizes and stiffness were used for the experiments. The GMM was trained by human demonstration and good grasps were chosen by hand.

Success rate	Octopus	Tennis ball	Bottle	Sponge
Our method	5/5	4/5	5/5	5/5
Baseline	4/5	1/5	2/5	4/5

Figure 18: The number of times each object remained stable in the gripper after being shaken. Soft objects are generally easier to grasp. [17]

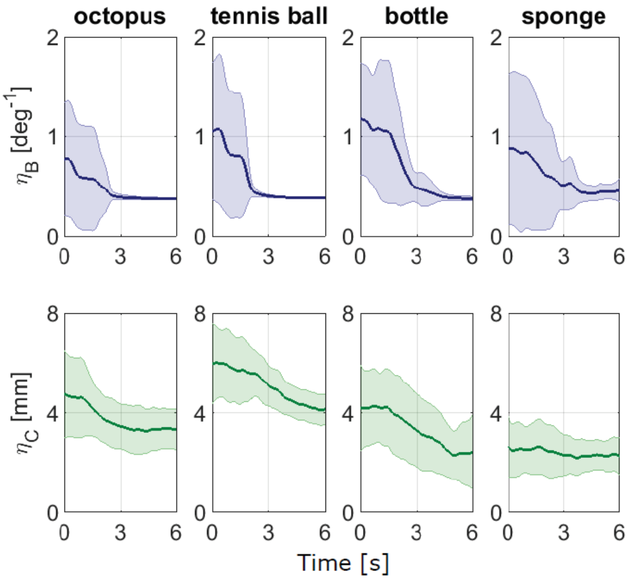

Figure 19: Bounds penalty index (top) and contact penalty index (bottom) during the grasp adjustments. [17]

The last experiment introduced a bound penalty index and a contact penalty index, which weight how far away the joints are from their limits and how close the fingertips are to the object center, respectively. Figure 19 shows that the GMM "move[s] the joints far from their bounds and the points of contact close to the center of the fingertips" [17, p. 393].

Advantages

- Achieves a stable grasp and grasp force control at the same time
- Works well with unknown objects of different shapes, sizes and stiffness
- Due to the underlying controllers, the GMM needs to learn less data

Disadvantages

- Good grasp configurations used to train the GMM were obtained by hand for each training object.

8 Integrating vision, haptics and proprioception into a feedback controller for in-hand manipulation of unknown objects

[18] Q. Li, C. Elbrechter, R. Haschke, and H. Ritter, "Integrating vision, haptics and proprioception into a feedback controller for in-hand manipulation of unknown objects," in *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*, pp. 2466–2471, IEEE, 2013

Motivation And Problem Setting

In-hand object manipulation remains to be a challenging task and often requires detailed knowledge of the objects properties. Other approaches have difficulties when the grasp unexpectedly becomes unstable or are simply based on torque control, instead of hybrid position/force control. This paper combines visual and tactile information to move the object in the hand while maintaining a stable grasp, using a hybrid position and force controller.

Sensors

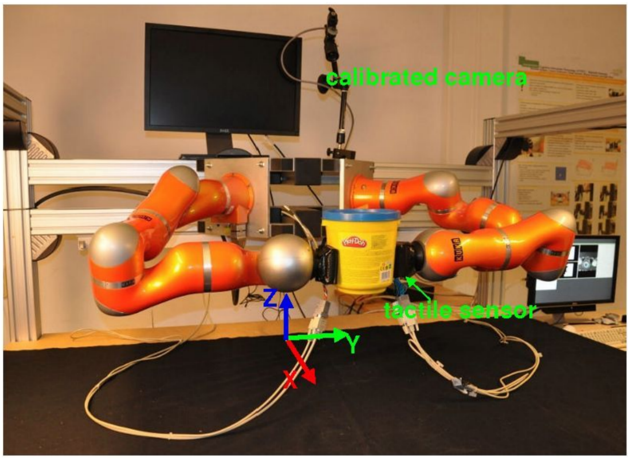

Figure 20: Two robot arm serve as a large gripper. [18]

Instead of a two-fingered gripper, two KUKA LWR arms are used with 16x16 piezo-resistive tactile sensor arrays as end-effectors (figure 20). These measure the contact forces and emulate the fingertips. Connected component analysis is used to identify the contact region within the tactile image and converted to a 3D position using forward kinematics. A calibrated monocular camera in combination with markers is used to estimate the object pose.

Method

The method consists of a position and a force controller. The target object positions are given as a trajectory and a simple PI-Controller follows the goal positions. From this, the positional adjustments on each contact point are calculated.

The force controller maintains a stable grasp. All contact force directions should meet in the objects center and the magnitudes need to add up to zero. From this requirement, the desired contact forces are derived. How the force planner determines the desired force for each object (to prevent slip without damaging the objects) was not mentioned.

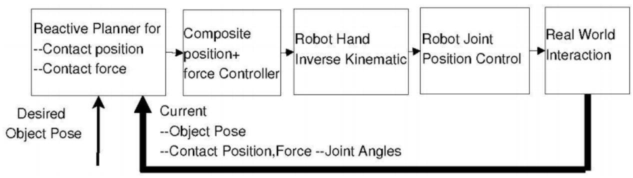

Figure 21: The control loop takes the desired object pose as input. [18]

Experiment

The robot estimates the objects position and approaches it with the fingertip until contact. Slip is detected by the camera. Then the arms translate and rotate the object.

Results

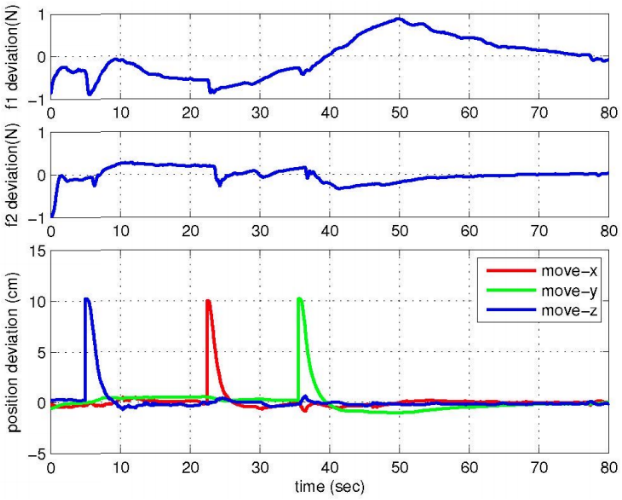

Figure 22: Force and position control errors during the movement. [18]

As shown in figure 22, all control errors converge to zero during the movement. The force control errors for each arm start to converge after the goal position is reached.

Advantages

- The robot does not need to know the friction and geometry of the object
- Provides a simple low-level position-control scheme for known objects, which could be adapted for two fingered grippers and could perform well in combination with high-level planners

Disadvantages

- Markers are used to obtain the object pose. This does not necessarily limit the general approach, but requires camera calibration and additional object recognition and pose estimation would be needed to generalize for unknown objects
- Slip is detected using the camera. A proper force planner is needed additionally.
- All controller gains need to be tuned by hand
- The tactile sensors need force calibration

9 Learning Object-level Impedance Control for Robust Grasping and Dexterous Manipulation

[16] M. Li, H. Yin, K. Tahara, and A. Billard, “Learning object-level impedance control for robust grasping and dexterous manipulation,” in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, pp. 6784–6791, IEEE, 2014

Motivation And Problem Setting

Hybrid position and force control methods, like the one discussed in the previous paper, are very precise. Problems arise when the controller switches during contact and the force overshoots [16]. Also the problem of how to choose the appropriate contact force remained unsolved. The idea behind Impedance control on the other hand is a spring-mass-damper system, which outputs the necessary force for an input motion. This paper learns the object-level impedance from human demonstration and extends the controller with tactile-feedback.

Robot and Sensors

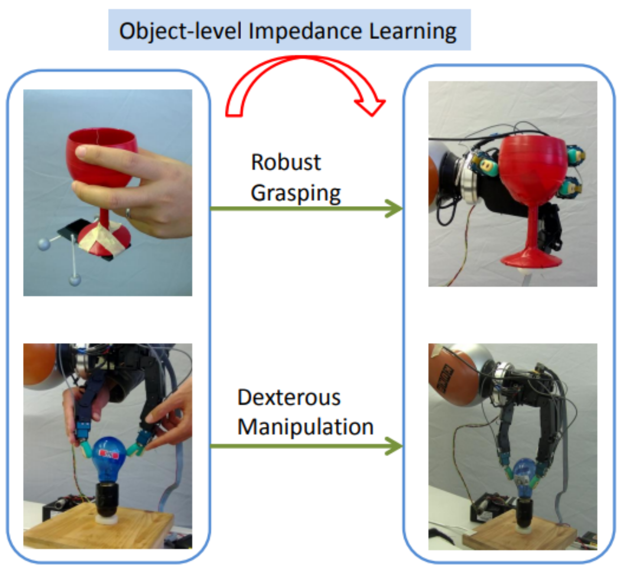

Figure 23: The Allegro hand used for grasping. The objects impedance is learned from human demonstration. [16]

This work uses three of the four fingers of the Allegro hand. Calibrated SynTouch tactile sensors on each fingertip measure contact force and position.

Method

The impedance controller consists of a grasp controller, which regulates the stable grasping force, and an object manipulation controller, which calculates the contact forces needed to reach a desired object position. The virtual frame method is used to estimate the pose of the object in the hand from using measured contact points, when fingers occlude the object for cameras.

The impedance is learned by human demonstration and a Gaussian Mixture Model allows to derive impedances based on the likelihood of the pose of the virtual frame.

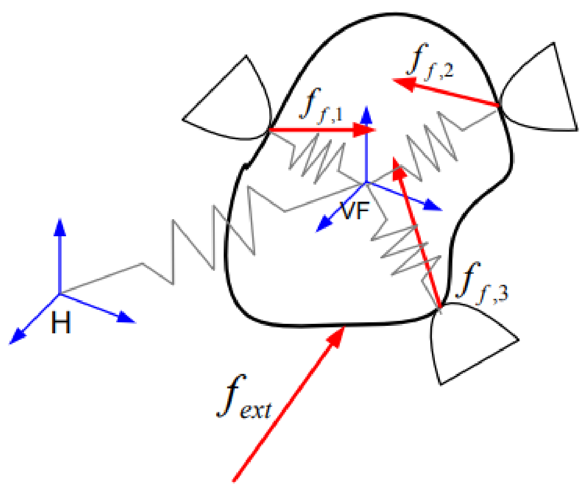

Figure 24: Springs represent the impedances of the object controller (between the virtual and the reference frame) and the grasp controller (between the fingers and the virtual frame). [16]

Experiments and Results

In the first experiment, the robot learns a robust grasp from human demonstration using only the fingers. Then a human pushes the hand to test compliance.

In a second experiment, the human demonstrates how to replace a light bulb. The robot observes the reference trajectory and the impedance. The learned results are shown in figures 26 and 27, respectively.

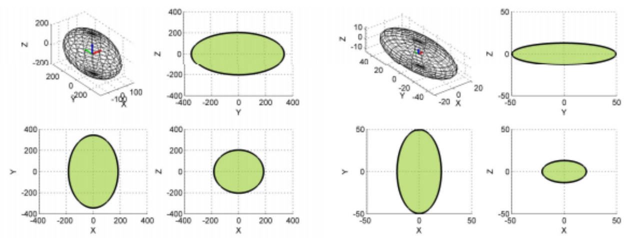

Figure 25: Relative translational (a) and rotational (b) stiffness for a cup. [16]

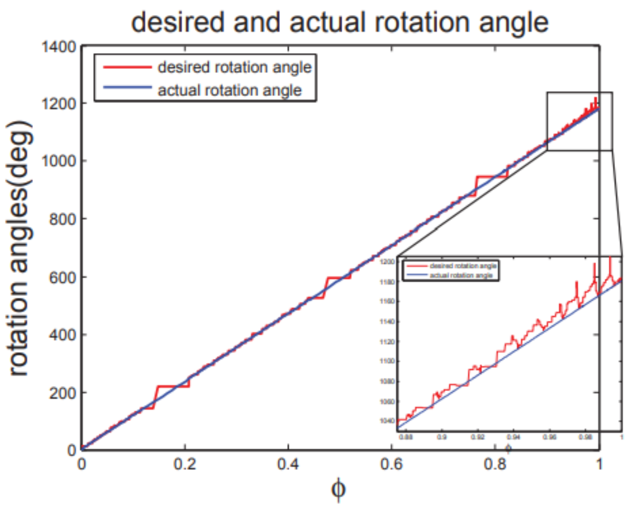

Figure 26: The learned reference trajectory for one trial. [16]

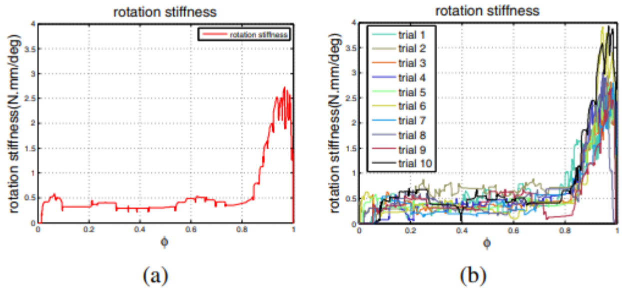

Figure 27: The learned reference object stiffness for one trial (a) and ten trials (b). [16]

Advantages

- This method offers an alternative to the previous paper with more compliant grasps
- Trajectories and object impedance can be learned from human demonstration
- Higher level object specific grasp planners could use this so they need to learn less parameters, compared to end-to-end learning

Disadvantages

- No grasp stability analysis
- initial grasps are predefined, it remains open which grasp configurations can actually carry out the desired impedance

10 Hierarchical Fingertip Space: A Unified Framework for Grasp Planning and In-Hand Grasp Adaptation

[21] K. Hang, M. Li, J. A. Stork, Y. Bekiroglu, F. T. Pokorny, A. Billard, and D. Kragic, "Hierarchical fingertip space: A unified framework for grasp planning and in-hand grasp adaptation," *IEEE Transactions on robotics*, vol. 32, no. 4, pp. 960–972, 2016

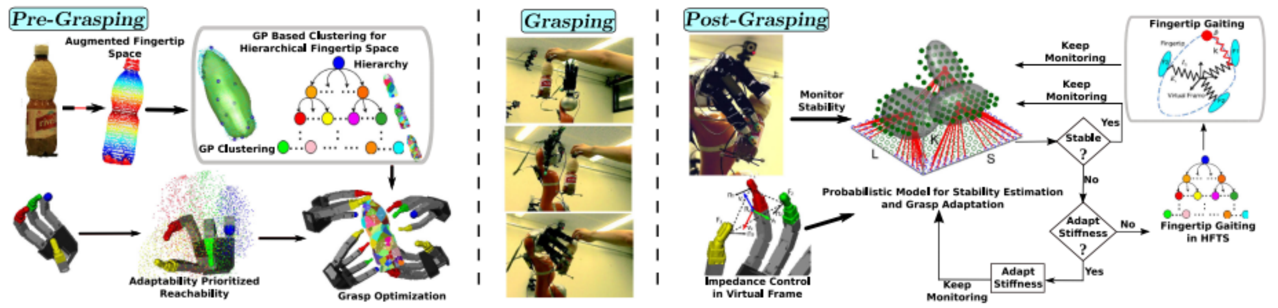

Figure 28: Overview of the complete framework. [21]

Motivation And Problem Setting

The previous paper showed the general usefulness of an impedance controller for object grasping and manipulation, but lacked a higher level grasp planner. This paper combines grasp synthesis, grasp adaption using impedance control and regrasping in one framework and has the goal of keeping the grasp stable even if the object changes weight or is perturbed.

Sensors

An Allegro hand is mounted on a KUKA LWR arm. SynTouch sensors on three fingertips measure contact force and locations. The OptiTrack system tracks the motion of the grasped object.

TODO

Method

The concept of the hierarchical fingertip space (HFTS) is introduced to encode contact locations and hand configurations with a Gaussian Process evaluating their similarity for the hierarchy.

The framework consists of three phases. First a valid grasp is synthesized from the HFTS using surrogate-based optimization on the criterias stability (contact-based force closure analysis), reachability (best reachability measure of all randomly sampled hand configurations) and adaptability (calculated by decomposing the hand Jacobian). In the second phase, position control is used to execute the grasp.

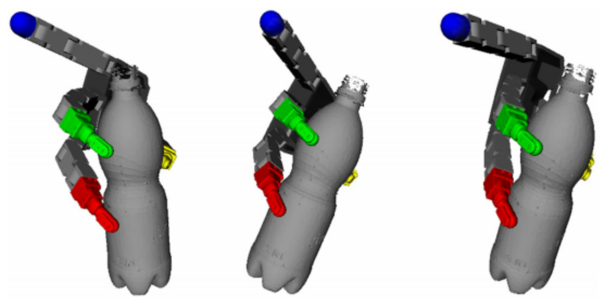

Figure 29: All hand configurations have the same contact points, but the leftmost has the highest adaptability. [21]

In the third phase, tactile sensing is incorporated and the grasp is adapted (i.e. impedance is increased) using the impedance control and virtual frame concept discussed in [6]. If that is not sufficient due to upper force limits, breadth-first search on the HFTS and the current tactile input are used to choose which fingers repositioning best stabilizes the grasp.

Experiments and Results

For the first experiment all six objects are grasped and then filled with pepper to test the maximum supported weight. [6] serves as a baseline method. The results in figure 31 shows that the proposed method can support the highest weight.

The second experiment measures the maximum acceleration the grasp can resist by shaking the object with linear acceleration. Figure 32 shows that this method can support more weight and higher horizontal and vertical accelerations than the baseline methods.

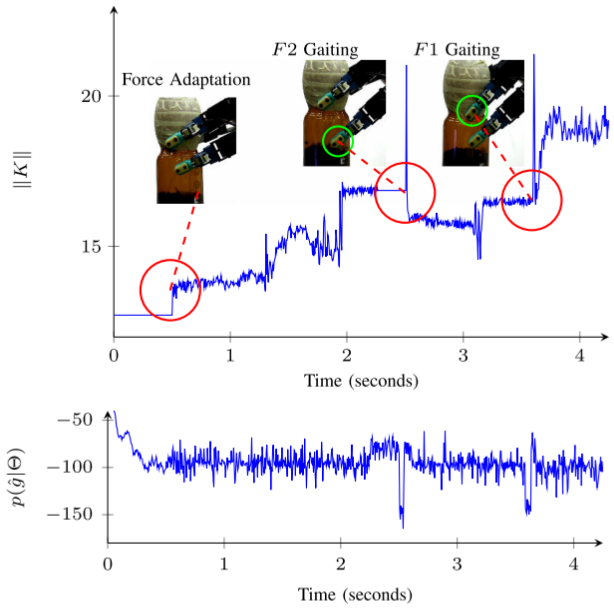

Figure 30: Fingertip gaiting while the bottles weight increases. Grasp stiffness (top) and stability likelihood (bottom) over time. [21]

Object	Weight	Without	With [9]	Improved
bottle1	34	55.1 ± 7.11	153.1 ± 12.31	165.3 ± 13.27
bottle2	39	62.8 ± 6.63	102.3 ± 13.38	121.3 ± 9.91
jug	112	125.3 ± 14.90	147.4 ± 9.62	162.1 ± 13.12
rivella	24	36.0 ± 6.96	76.5 ± 9.4	92.7 ± 7.45
milk	34	63.5 ± 8.20	151.8 ± 7.24	157.4 ± 8.35
spray	63	75.7 ± 7.21	102.2 ± 6.02	121.6 ± 7.15

Figure 31: Supported object weights in gram, with mean and standard deviation, without grasp adaption, with the grasp adaption proposed in [6] and for this method. [21]

Advantages

- Combines grasp synthesis, stability estimation, adaption and finger gaiting
- Impedance control during finger gaiting allows the finger to keep contact while sliding into the new position
- Non-planar object shape improves the grasp
- Grasps can be synthesized offline for each object using point clouds

Disadvantages

- So far the initial grasp stiffness is hard coded and point clouds are needed for training, which limits the generalization. However, an active learning strategy is planned for future work.

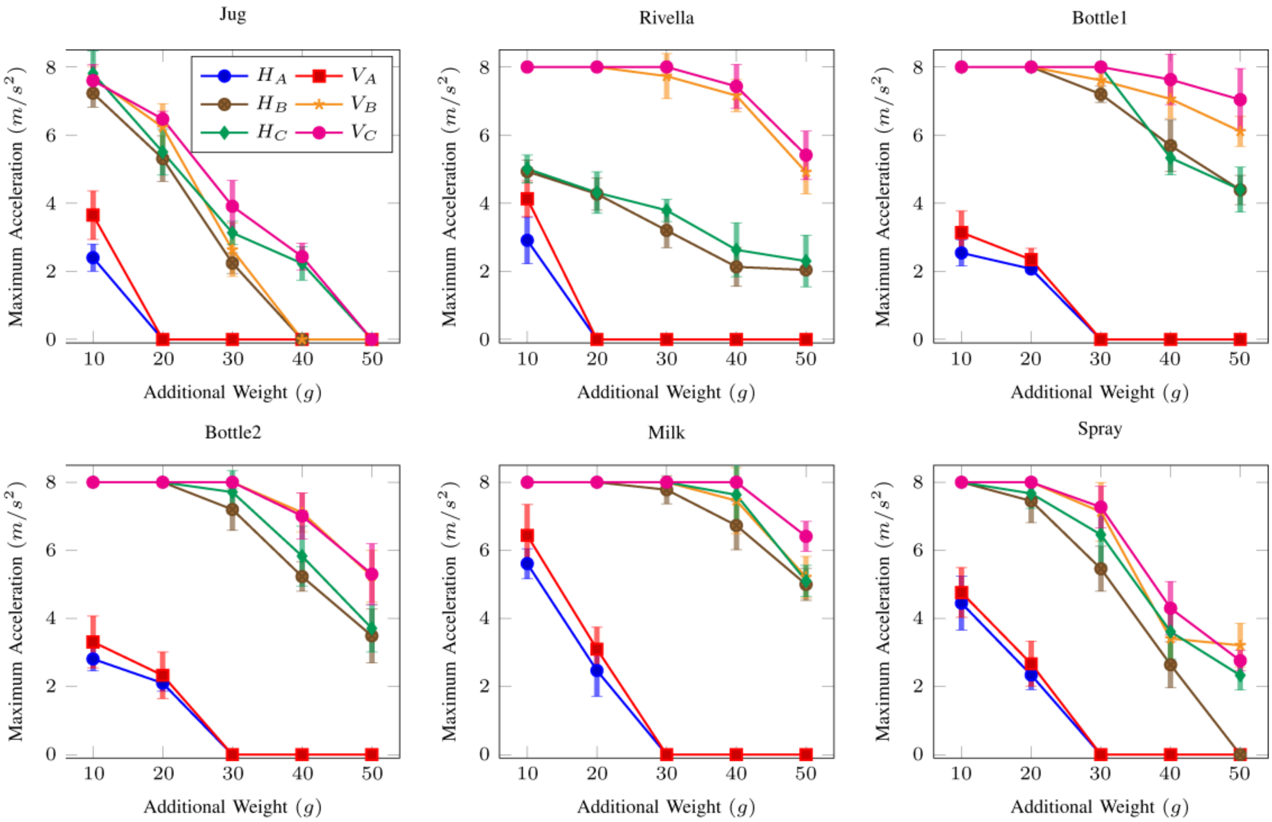

Figure 32: Maximum horizontal (H), and vertical (V) acceleration and weight each method can support. A: without grasp adaption, B: method proposed in [6], C: this method. [21]

11 Learning Robot In-Hand Manipulation with Tactile Features

[22] H. Van Hoof, T. Hermans, G. Neumann, and J. Peters, "Learning robot in-hand manipulation with tactile features," in *Humanoid Robots (Humanoids), 2015 IEEE-RAS 15th International Conference on*, pp. 121–127, IEEE, 2015

Motivation And Problem Setting

Most in-hand manipulation algorithms are either analytical, or need to know the kinematics of the hand or the contact points and thus are model based. For underactuated compliant hands however, kinematic models are not feasible. This work uses reinforcement learning to roll an object between the fingers of a compliant hand.

Sensors and Platform

Figure 33: ReFlex robot hand. [22]

The ReFlex underactuated compliant robot hand is equipped with nine MEMS[1] pressure sensors on each of the two out of three used fingers.

[1] Micro Electro Mechanical System

Method

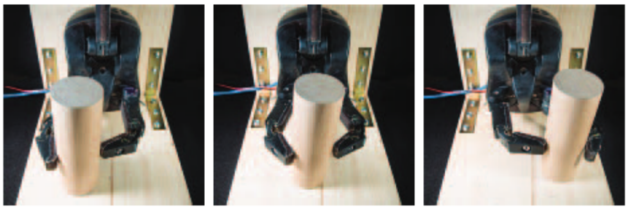

Figure 34: Two fingers roll the object from left to right to a goal position. [22]

The rolling task is modeled as a Markov Decision Process with a continuous state-action space. The state is six-dimensional and consists of the hand configuration and the encoded tactile measurements and the actions are the applied velocities for the two fingers. Reward is given based on the distances to the goal and the desired finger pressure and high action magnitudes are punished. Non-Parametric Relative Entropy Policy Search solves the MDP (figure 35).

Algorithm 1 The NPREPS algorithm

Require: Initial explorative policy $\tilde{\pi}_0$
  for $i = 1, \ldots, \text{max\_iteration}$ **do**
    generate roll-outs according to $\tilde{\pi}_{i-1}$
    minimize kernel-based dual:
      $\eta^*, \boldsymbol{\alpha}^* \leftarrow \arg\min g(\eta, \boldsymbol{\alpha})$ Eq. 6
    calculate kernel-based Bellman errors:
      $\boldsymbol{\beta}_j \leftarrow (\mathbf{K}_{\text{sa}} + \lambda \mathbf{I})^{-1} \mathbf{k}_{\text{sa}}(\mathbf{s}_j, \mathbf{a}_j)$ Eq. 7
      $\delta_j \leftarrow \mathcal{R}_j + \boldsymbol{\alpha}^{*T} \left(\tilde{\mathbf{K}}_{\text{s}} \boldsymbol{\beta}_j - \mathbf{k}_{\text{s}}(\mathbf{s}_j) \right)$ Eq. 7
    calculate the weighting factors:
      $w_j \leftarrow \exp(\delta_j / \eta^*)$ Eq. 8
    determine cost-sensitive GP:
      $\tilde{\pi}_i(\mathbf{a}|\mathbf{s}) = \mathcal{N}(\mu(\mathbf{s}; \mathbf{w}), \sigma^2(\mathbf{s}; \mathbf{w}))$ Eq. 9
  end for

Figure 35: Non-Parametric Relative Entropy Policy Search. [22]

Experiments and Results

An initial policy is executed on two training objects. After every ten iterations, the policy is optimized according to the algorithm in figure 35. Each of the adapted policies is tested on new objects for generalizability. The initial policy (hard-coded, but without introduced noise) is used as a baseline method.

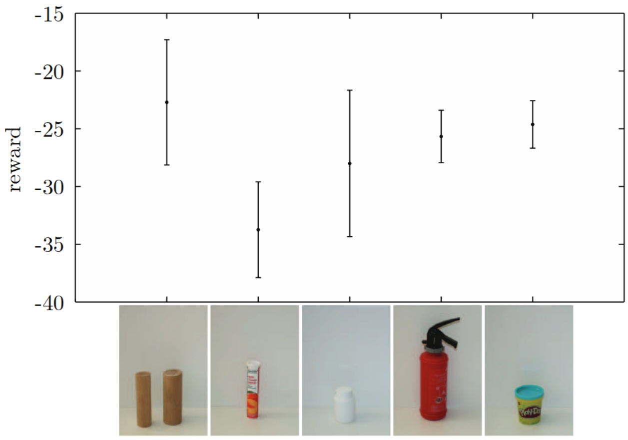

Figure 36: The wooden cylinders on the left are used for training and the other objects for testing the ability to generalize. Except for the second object (which is very small), the mean and standard deviation are comparable and the learned policy generalizes. [22]

Advantages

- Works for underactuated hands, no kinematic model needed

Disadvantages

- Reinforcement learning requires physical interaction for learning and even small state-action spaces need many episodes for training. The problem was simplified in this case, because only a refinemend of a hard-coded policy was learned
- The method only generalizes within small bounds, the objects need to be similar in shape and size

12 Self-Supervised Regrasping using Spatio-Temporal Tactile Features and Reinforcement Learning

[23] Y. Chebotar, K. Hausman, Z. Su, G. S. Sukhatme, and S. Schaal, "Self-supervised regrasping using spatio-temporal tactile features and reinforcement learning," in *Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on*, pp. 1960–1966, IEEE, 2016

Motivation And Problem Setting

Reinforcement methods learn by interacting with the environment and their biggest advantage is that they are able to find a trade-off between short term and long term reward based on the cummulative experience of many autonomous robots. However, simulations usually do not reflect the real world well enough, which means physical interactions are required and efficient state-action spaces are crucial. This paper uses spatio-temporal tactile input to predict the stability of a grasp and uses this prediction for regrasps.

Sensors

A three fingered Barrett hand is equipped with BioTac sensors on the fingertips. They consist of electrodes with an overlying liquid, so the distributed impedances depict the deformation of the surface. Figure 37 shows how the electrodes are arranged in a tactile image.

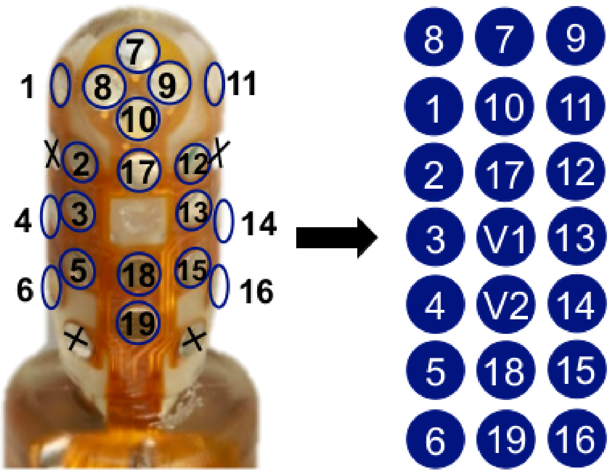

Figure 37: The electrodes of the BioTac sensor are arranged as tactile image. [23]

Method

The Spatio-Temporal Hierarchical Matching Pursuit algorithm builds a dictionary of features for a sequence of tactile readings during the grasp. These feature descriptors are clustered with a Support Vector Machine for their stability. The relative entropy policy search infers the regrasp configuration by learning weights for a linear combination of the feature descriptors. Principal Component Analysis reduces the dimensionality. Stable grasps are rewarded 0.5 if the object slips during shaking and 1.0 if the object does not slip.

Experiments and Results

Figure 38: Experimental setup. The bowl helps put the object back in the same position when dropped. [23]

The robot executes top down grasps on Cylindrical objects (figure 38) using the force gripping algorithm in, . The gripper shakes the object to assure stability. This process is repeated autonomously for a 1000 grasps.

The table in figure 39 shows the accuracy of the stability prediction for different sensor modalities. The electrodes are a good choice considering that the combination of all modalities increases the accuracy by less than one percent for the same number of grasps.

Features	# of grasps	Avg. Accuracy, %
Electrodes	500	90.73
Finger angles	500	80.55
Strain gages	500	84.91
Hand orientation	500	74.00
Force-torque	500	69.63
BioTac finger forces	500	88.55
All features	500	91.09
All features	1000	93.00

Figure 39: The feature descriptor applied to different sensor modalities. [23]

After a random grasp is executed and predicted unsuccessful, the current policy is sampled for regrasps. The policy converges after eight policy updates (figure 40). The learned policy is then tested on a new object (figure 38). Multiple regrasps increase the success rate above 90%.

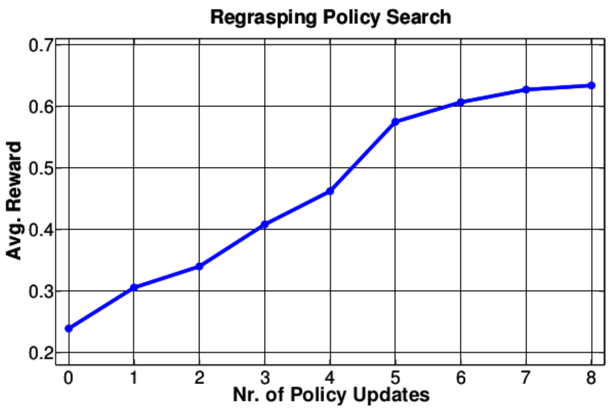

Figure 40: Policy convergence. The policy is updated every 100 regrasps. [23]

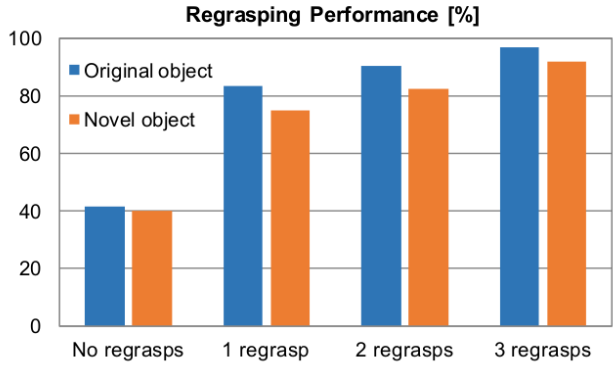

Figure 41: Evaluation of the success rate of the learned policy on the training and the test object. [23]

Advantages

- The reinforcement algorithm learns parameters instead of a controller, which decreases the size of the state-action space and therefore learning time
- Reinforcement learning helps to estimate the information gain from different sensor modalities

Disadvantages

- Cylindrical objects in combination with a bowl helps the object to return to the original position when it slips out of the grasp. However this does not generalize well, a vision system could extract the new object pose instead.
- The robot cannot generalize to new or unseen objects

13 Tactile-based Manipulation of Deformable Objects with Dynamic Center of Mass

[7] M. Kaboli, K. Yao, and G. Cheng, "Tactile-based manipulation of deformable objects with dynamic center of mass," in *Humanoid Robots (Humanoids), 2016 IEEE-RAS 16th International Conference on*, pp. 752–757, IEEE, 2016

Motivation And Problem Setting

Real world object manipulation requires the robot to not just stabely grasp unknown, deformable objects, but also to keep the grasp stable during manipulation, when the center of mass or even the weight changes. This paper proposes a framework which uses slip detection for minimal force grasps and a strategy to handle heavier objects, when increasing force alone does not suffice to keep the grasp stable.

Sensors and Setup

OptoForce OMD-20-SE-40N tactile sensors attached to a Robotiq gripper on a UR10 arm provide contact force measurements for three directions. The three fingers are position controlled.

Method

The weight force of the object is estimated as the sum of the downward forces during stable state and helps define an upper force limit. When slip is detected (tangential force jumps), the friction is estimated as ratio of the tangential and normal force and the fingers close one step. If there is no slip, the fingers open until slip is detected. This way, the exerted force is kept within the lower and upper bounds. If the maximum allowed applied force need to be exerted for heavy objects, the object is rotated to reduce the weight force by the factor $cos(\theta)$.

Experiments and Results

Experiments using minimal force grasps and the proposed framework were carried out and compared to the built-in grasping method without tactile sensing. The built-in method could only detect a minimal force of 15 N and grasped all objects without slip but with severe deformation.

During the minimum grasping force experiment, the object weight was increased by pouring in rice or the object was rotated. No large deformation occured, but slips could not be prevented (figure 42).

Figure 42: Minimum grasping force experiment. Minimal deformation, but slip occured. [7]

The proposed framework was able to prevent slip with minimal deformation during object rotation and weight increase (figure 43). When the weight force increased over the upper bound, the controller rotated the object to keep it stable (figure 44). Figure 45 clearly shows that deformation is kept minimal by this framework.

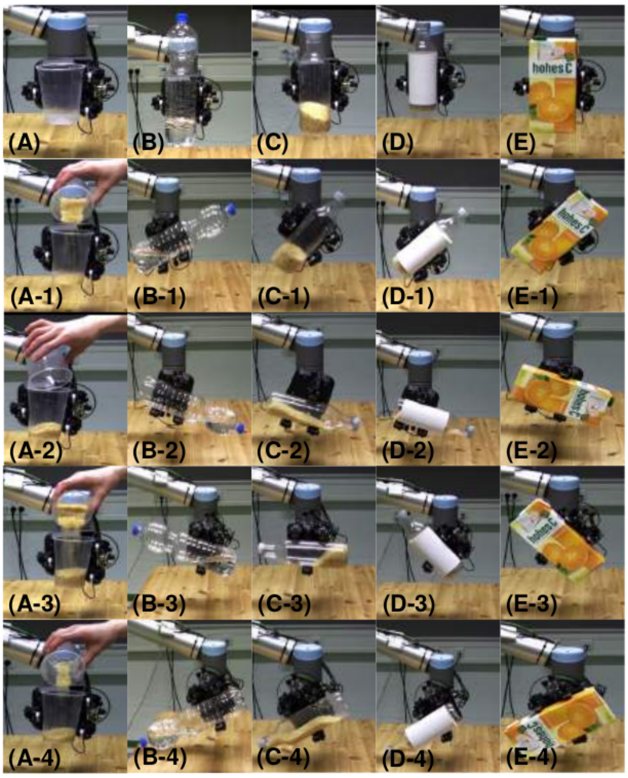

Figure 43: Experiment with the proposed framework. Minimal deformation and no slip during weight increase and object rotation. [7]

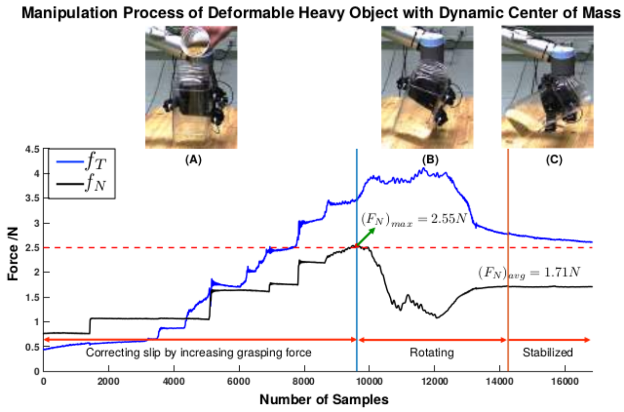

Figure 44: Normal and tangential forces during manipulation. [7]

Advantages

- Gripper closes the fingers until contact so it can adapt to different shapes and sizes of objects
- Weight and surface friction do not need to be known beforehand
- Slip detection is not disturbed by vibration of the robot
- real-time capable

Disadvantages

- The upper bound for the normal force is set manually. A squeezing action like the one presented in section 20 could be a possible extension.

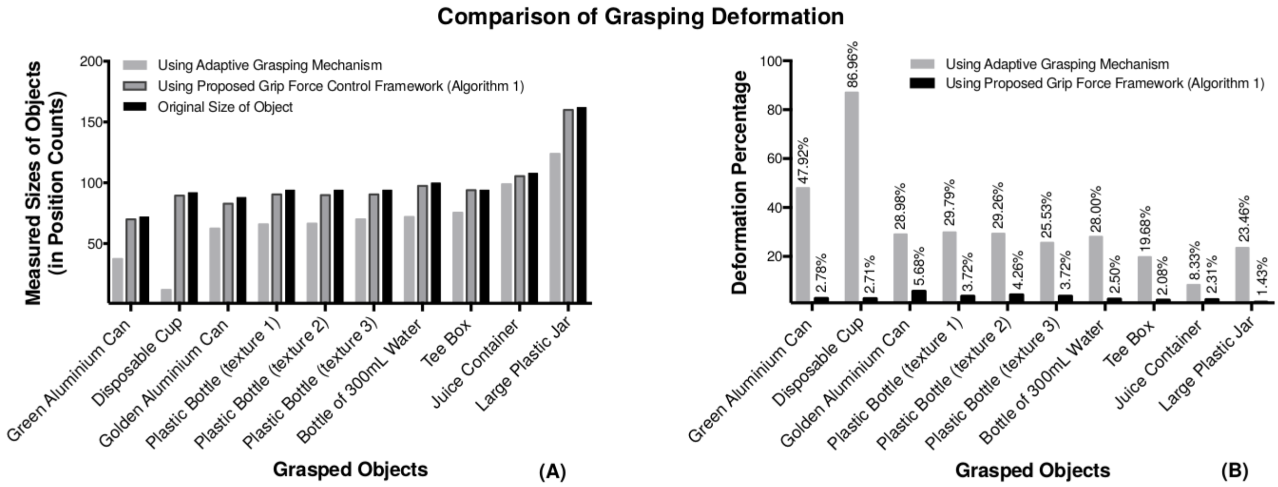

Figure 45: Deformation for the built-in and the proposed methods, measured in position counts of the finger. [7]

14 Stable grasping under pose uncertainty using tactile feedback

[19] H. Dang and P. K. Allen, “Stable grasping under pose uncertainty using tactile feedback,” *Autonomous Robots*, vol. 36, no. 4, pp. 309–330, 2014

Motivation And Problem Setting

Simulations can be helpful to learn and synthesize grasp configurations, but when applied on a real robot the position uncertainty often causes grasp failures. This paper proposes a pipeline, which estimates stability after an initial grasp and adjusts the hand in order to make the unstable grasps stable.

Sensors and Setup

Four tactile pads on the fingertips and the palm were mounted on a Barrett hand on a Staubli robot arm. The GraspIt! simulator was used for simulation experiments. A Kinect was used to get the 6D pose of real objects.

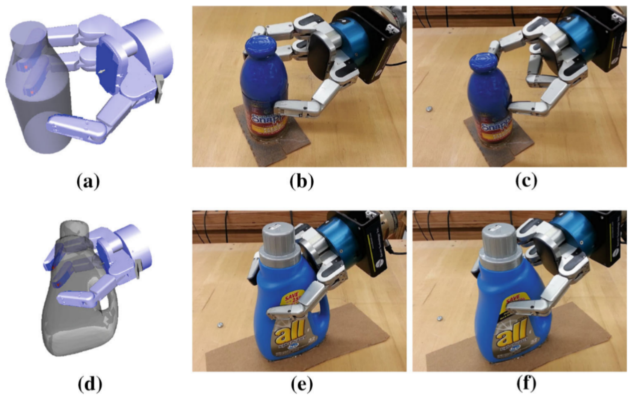

Figure 46: The Barrett hand. Grasps planned in simulation (a,d) either succeed (b,e) or fail (c,f) with real world pose uncertainty. [19]

Method

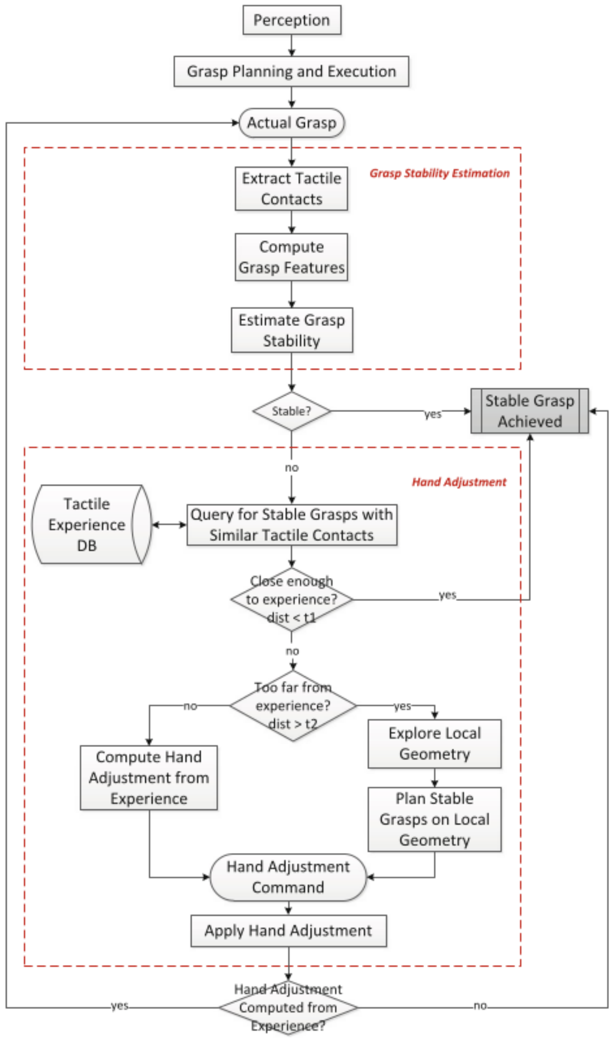

Figure 47: Proposed algorithm for achieving stable grasps. [19]

Figure 47 gives an overview of the algorithm. A Bag-of-words model represents a set of contacts (clustered using kmeans) in a dictionary, which is used to train a SVM stability classifier. A tactile experience database stores stable grasps and the corresponding tactile contacts. Similarity is based on the contact location, derived from forward kinematics. Hand adjustments are computed by sampling similar hand configurations around k nearest (stable) neighbors. If the distance to stable grasps is too big, a mesh of the local geometry will be reconstructed by fitting the contact point cloud to a quadratic function. Then the Eigengrasp planner generates a new stable grasp on the local geometry.

Data Collection

Grasps are obtained from the Columbia Grasp Database and GraspIt! simulated the tactile feedback. A metric based on epsilon quality and volume quality labeled the grasps stable or unstable for training of the SVM.

Experiments and Results

Test accuracy of the stability estimator in simulation was 81%, and 84.6% on the real robot for the objects in figure 48. Results of experiments with the complete algorithm and the real robot (grasp, lift, shake) can be seen in figure 49.

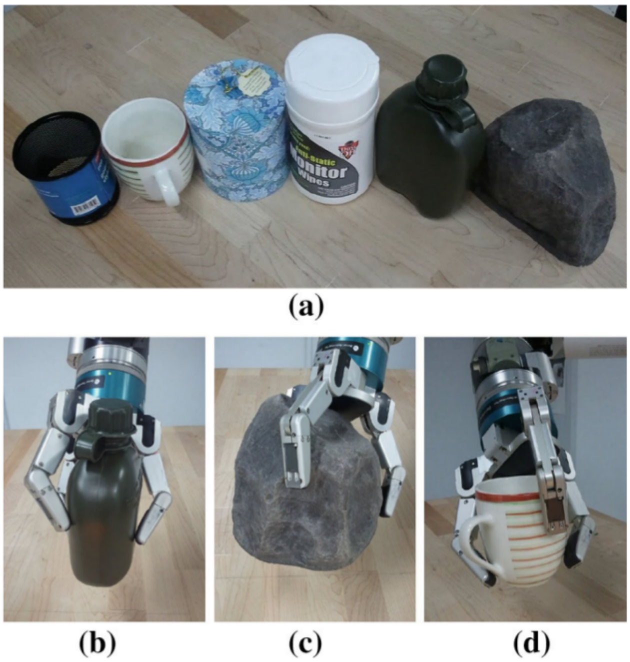

Figure 48: The objects used in the experiments (a) and stable grasps (b, c, d). [19]

Object	# Of grasps	Avg. # adj.	Lift-up[a]	Score
Snapple	10	5.5	10	1.0
Box	10	3.3	10	0.95
Detergent	10	2.1	8	0.75
Cup	10	2.4	10	0.95
Rock	10	3	10	0.95

[a] The robot hand successfully lifts up the object. The object could have moved during the grasping and lift-up procedure

Figure 49: Results of experiments on the real robot. [19]

Advantages

- The algorithm could be used with every hand contained in the Columbia Grasp Database without large adjustments.

- No assumptions about the shape are necessary.

Disadvantages

- The hand needs to move away and reapproach the object for the grasp adjustment.

- The use of kmeans requires the number of clusters (and therefore the dimensionality of the feature vector) to be set manually and hand tuned. The same applies to the chosen stability thresholds.

- The generated grasps do not take the use of the objects under consideration and "do not make sense" from the viewpoint of a human (figure 48).

- Failed grasps can move the object.

15 Learning the Tactile Signatures of Prototypical Object Parts for Robust Part-based Grasping of Novel Objects

[29] E. Hyttinen, D. Kragic, and R. Detry, "Learning the tactile signatures of prototypical object parts for robust part-based grasping of novel objects," in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*, pp. 4927–4932, IEEE, 2015

Motivation And Problem Setting

Many grasp planners work on 3D-models of objects. Pure vision can only perceive the facing part of the object. This work extends a part-based grasp planner with a grasp stability classifier for prototypical shapes.

Platform and Sensors

Figure 50: Tactile sensor pads on the inside of the three-fingered Robotiq gripper (fingers and palm) measure distributed force. [29]

Method

The tactile measurements are represented as tactile images. Moment features representing total pressure and contact location are calculated for each tactile image. Together with the joint angles they form the 17-dimensional features for Kernel Logistic Regression. One model is trained for each grasping prototype.

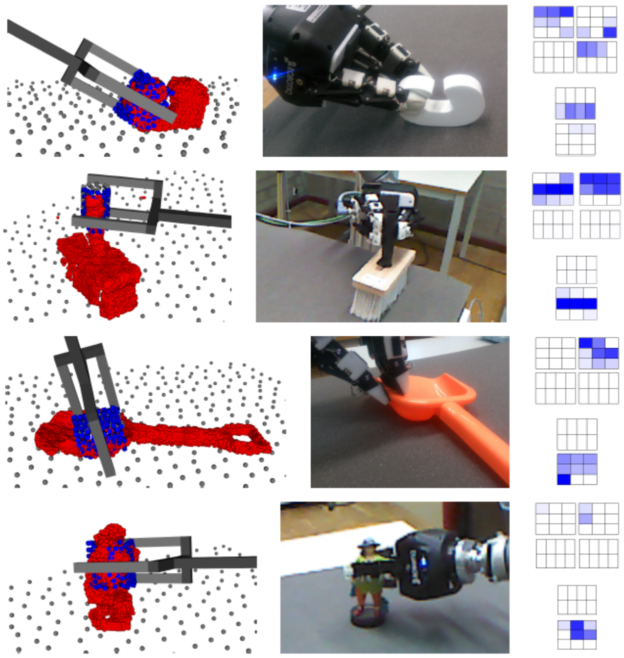

Figure 51: Two successful (top) and two unsuccessful (bottom) grasps, generated from the grasp planner with the respective tactile readings. [29]

Experiments and Results

Objects of different shapes and sizes were grasped using the part-based planner. The robot lifted the grasped objects and humans pushed the object to evaluate stability. 32 stable and unstable grasps where collected for three prototypes. A baseline classifier was trained without the prototype shape information. Classification accuracy was 89% for the optimal number of features.

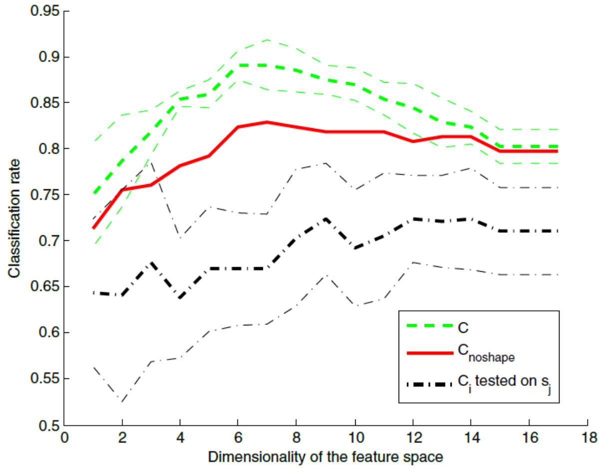

Figure 52: Using 6-8 out of the 17 features yields the best classification results. Green is the average of the classifiers, red the baseline and black the classifiers applied to the "wrong" prototypes. Thin lines indicate standard deviation respectively. [29]

The fact that classifiers, which are trained for one prototype and are asked to evaluate stability for a different prototype (black line in figure 52), have a lower accuracy "is a strong indication that shape information is relevant to stability estimation" ([29, p. 5]).

Advantages

- In contrast to convolutional neural networks for example, this approach is simpler and requires less training data

Disadvantages

- Requires an additional part-based grasp planner
- A new classifier needs to be trained for new shape prototypes

16 Enhancing Adaptive Grasping Through a Simple Sensor-Based Reflex Mechanism

[4] E. Luberto, Y. Wu, G. Santaera, M. Gabiccini, and A. Bicchi, "Enhancing adaptive grasping through a simple sensor-based reflex mechanism," *IEEE Robotics and Automation Letters*, vol. 2, no. 3, pp. 1664–1671, 2017

Motivation And Problem Setting

In contrast to fully actuated hands, under-actuated compliant hands offer the advantage of simpler control and grasps are generally more robust, because the complete inner side of the fingers touches the object instead of just the fingertips. This work incorporates tactile sensing for such a hand to increase robustness against visual uncertainties of the object position.

Sensors

Figure 53: Pisa/IIT Softhand equipped with infrared sensors. [4]

A Kuka LWR arm is used with a Pisa/IIT SoftHand. This hand has 19 degrees of freedom, but only one is actuated. The thumb, index- and ringfingers are equipped with Avago HDSL9100 infrared sensors. An Asus Xtion Prolive RGB-D camera provides a point cloud of the object.

Method

The point-cloud of the RGB-D camera is used to approach the object. For the baseline method the objects center is approximated and for the proposed method, the best pre-grasp configuration is calculated from matching the observed point-cloud to an object data-base and evaluating possible pre-grasp shapes in Gazebo. If all sensors now see the object, the grasp is refined, such that all fingers close around the object uniformly. This is an optimization problem over the measured distances and the control input.

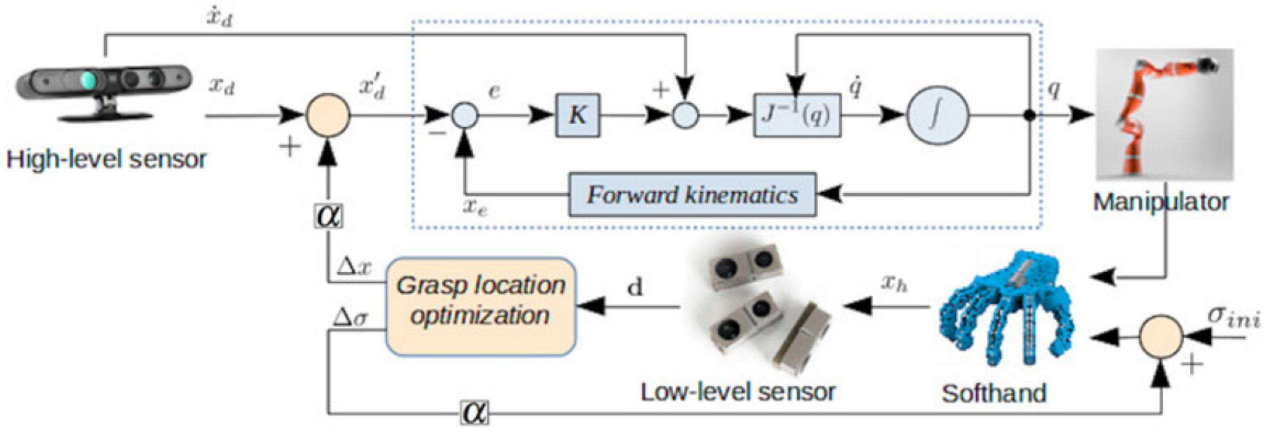

Figure 54: Control algorithm. [4]

Experiments and Results

Five objects were grasped either from their positions observed from the camera, or after the object position was perturbed after the observation. The pre-grasp configuration was calculated using the object center estimation (M1), using the database (M2) or from the IR-sensor based refinement, respectively.

Results show that using proximity sensors can compensate uncertainties in the objects position and lead to successful grasps.

	Nominal Pose $(\mathfrak{N})$			Perturbed Pose $(\mathfrak{P})$	
Object and reference pose	Blind (M1)	Blind (M2)	IR-guided (M1)	Blind (M2)	IR-guided (M2)
Cylinder	0/3	2/3	3/3	0/3	2/3
Baby cup, pose (a)	0/3	3/3	2/3	0/3	2/3
Baby cup, pose (b)	1/3	3/3	3/3	0/3	2/3
Paper box, pose (a)	0/3	3/3	2/3	0/3	2/3
Paper box, pose (b)	1/3	3/3	3/3	0/3	2/3
Polytope	1/3	3/3	3/3	0/3	3/3
Duck, pose (a)	1/3	3/3	2/3	0/3	2/3
Duck, pose (b)	0/3	3/3	2/3	0/3	2/3
Total	4/24	23/24	20/24	0/24	17/24

Figure 55: Succesful grasps for all objects and methods. [4]

Advantages

- Infrared sensors are low-cost, small and robust.
- Uniform closing of all fingers reduces unwanted changing of the objects position.
- Uncertainties in object position can be compensated.
- Underactuated hands allow for simpler grasping controllers.

Disadvantages

- Infrared sensors suffer from cross-talk, here only one sensor is activated at a time.
- Sensor calibration is needed.
- Different reflection characteristics of objects could worsen performance.
- Each object in the grasp database needs to be added using experiments. Does not generalize well.
- Underactuated hands can grasp but do not allow in-hand object manipulation.

17 Learning Predictive State Representation for In-Hand Manipulation

[24] J. A. Stork, C. H. Ek, Y. Bekiroglu, and D. Kragic, “Learning predictive state representation for in-hand manipulation,” in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*, pp. 3207–3214, IEEE, 2015

Motivation And Problem Setting

In-hand manipulation of objects requires a good representation of the dynamical system consisting of the hand, the object and the environment, which can be used for planning. Partially Observable Markov Decision Processes and Hidden Markov Models model the probability distribution over unobserved (latent variable) states. Predictive State Representations (PSR) in contrast are action-conditional and depend on history and less on accurate prior models. This work incorporates tactile information into a PSR model to rotate a pen in-hand.

Sensors and Platform

Experiments are carried out on the Willow Garage PR2 robotic platform, which comprises a two-fingered gripper with pressure sensor arrays on the fingertips and cameras.

Figure 56: The gripper rotates the pen in-hand by pushing it against the table. [24]

Method

The action space consists of six different actions, which rotate the pen clockwise and counterclockwise from different starting angles, by pushing the pen against the table (figure 56). The observation space consists of pressure readings (image moments) and the gripper-marker configuration, obtained from the camera and blob detection. The Transformed Predictive State Representations learning algorithm is used, which transform the parameters of a linear PSR to a lower dimensional space. String Kernel features solve the problem that different sequences of pen angles can lead to similar observations. Negative reward is given for actions leading to states with high uncertainty and positive reward is given for reaching goal configurations. Value Iteration learns the policy.

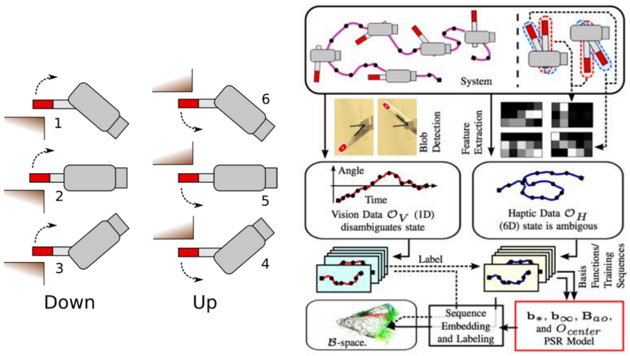

Figure 57: The six different actions (a) and a system overview (b). [24]

Experiments and Results

Experiments show that the PSR sufficiently tracks the system dynamics (figure 58, first two rows). In a second experiment, different goal states were given and the learned policy was able to reach them (figure 58, bottom row).

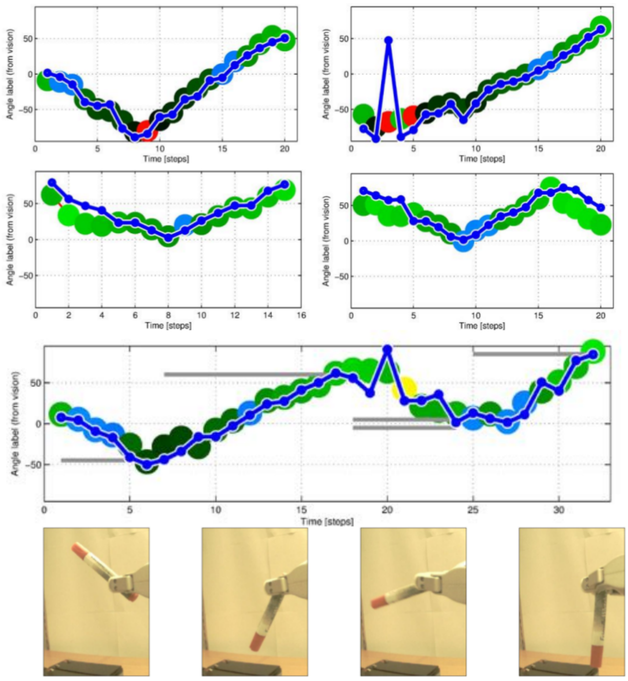

Figure 58: First two rows: The pen angle ground truth (green) and the PSR state estimations (blue). Bottom rows: Goal configurations and system states for the second experiment. [24]

Advantages

- Models continuous system dynamics and resolves disambiguity of tactile readings.
- Goal configurations can be given and reached.
- Distance-based heuristics are possible.

Disadvantages

- Training data needs to cover all manipulation goals, which limits generalizability.
- Sensors need calibration.

18 Force Estimation and Slip Detection/Classification for Grip Control using a Biomimetic Tactile Sensor

[30] Z. Su, K. Hausman, Y. Chebotar, A. Molchanov, G. E. Loeb, G. S. Sukhatme, and S. Schaal, “Force estimation and slip detection/classification for grip control using a biomimetic tactile sensor,” in *Humanoid Robots (Humanoids), 2015 IEEE-RAS 15th International Conference on*, pp. 297–303, IEEE, 2015

Motivation And Problem Setting

Service robots need to be able to stabely grasp unknown objects, and slip detection is an important part of it. When slip is detected, the grasping force can be adapted accordingly. This work uses biomimetic sensors and a machine learning approach to estimate grasping force and detect slip as well as a grasp adaption controller.

Sensors

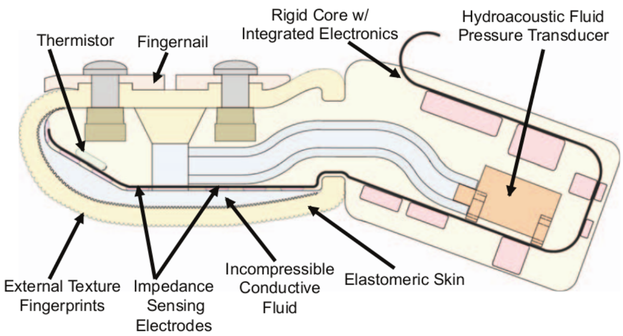

Figure 59: The BioTac tactile sensor can measure force, pressure and temperature. Its structure is inspired by the human sense of touch. [30]

This work uses the BioTac tactile sensor and the three-fingered Barrett hand.

Method

Force is estimated in all three directions by relating it to the impedance changes of the sensor. Different approaches are compared to learn the necessary scaling factors and weights: Weighted sum of normals with linear regression, locally weighted partial least squares regression, a neural network with one hidden layer and one with three hidden layers.

Two methods for slip detection are investigated: The first detects if the derivative of the tangential force becomes negative, and the second method uses the pressure sensor to measure vibrations on the contact surface. Both report slip if a threshold is exceeded. Another neural network is trained to distinguish linear and rotational slip using 100 consecutive electrode measurements as input features.

The controller is position-controlled while the fingers close around the object, and force-controlled to lift it up. Minimal necessary grasping force is calculated as soon as slip occurs by estimating the friction coefficient from the measured tangential and normal force.

Experiments and Results

In order to compare the different methods, the tactile sensor was pressed against a force plate and the true force and the electrode signals were recorded. The comparison of the four force estimation approaches is shown in figure 60.

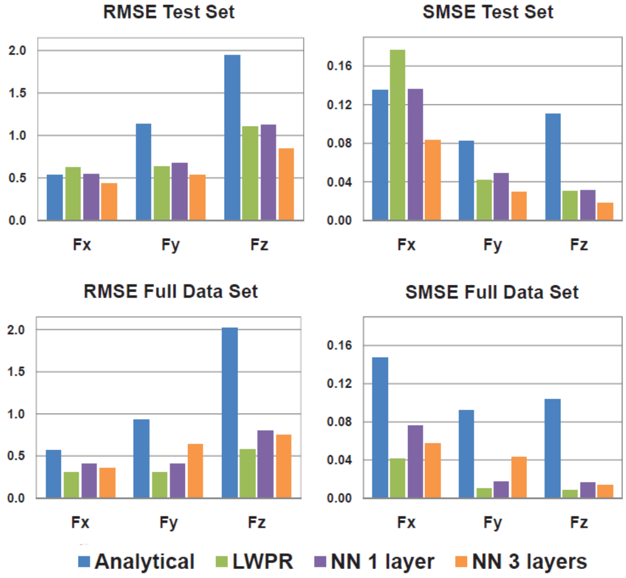

Figure 60: Root Mean Squared Error and Standardized Mean Squared Error. The analytical approach performs worst. Low training errors and high test errors of methods two and three suggest overfitting. The last method performs best. [30]

In a second experiment, slip was induced by slightly opening the fingers around the two grasped objects (a plastic jar and a wooden block), and the two methods were compared (61). An IMU was attached to the grasped objects for ground truth.

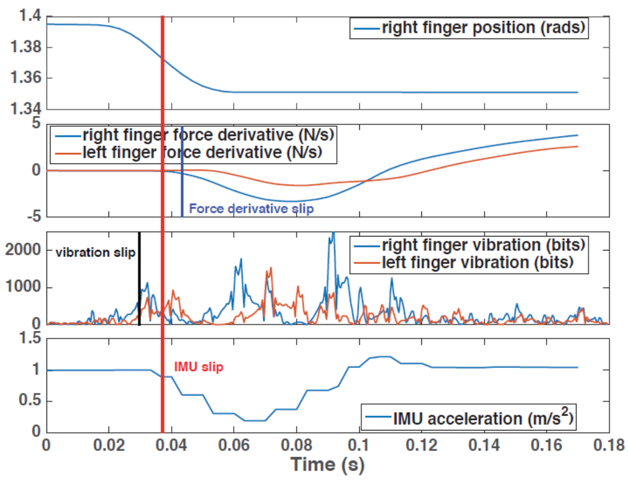

Figure 61: Results of 20 slips per object. The vibration based approach detected slip even before the IMU. [30]

The slip classification network was trained by grasping four objects on the edge or close to the center, to provoke either rotational or linear slip. At the time slip happens (IMU), the classifier has an accuracy of 80%, which increases over time 62.

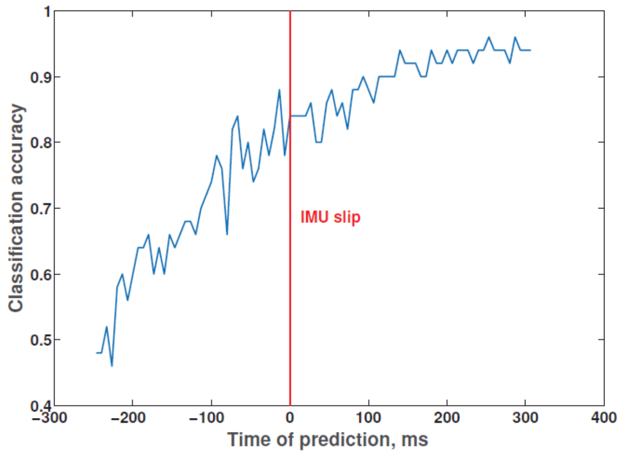

Figure 62: Accuracy of the slip classifier. [30]

The controller was able to detect and prevent slip for a plastic jar, which weight was increased by 300g during the grasp.

Advantages

- Slip detection and classification is simple and fast.
- The methods could also be used to detect collision of the object and the environment.

Disadvantages

- Multiple thresholds are set manually, such as the vibration threshold for slip detection (set to twice the baseline vibration). This could potentially limit generalizeability.

19 Contact Localization on Grasped Objects using Tactile Sensing

[25] A. Molchanov, O. Kroemer, Z. Su, and G. S. Sukhatme, “Contact localization on grasped objects using tactile sensing,” in *Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on*, pp. 216–222, IEEE, 2016

Motivation And Problem Setting

Grasping is a very important ability for service robots. However, it is just the first step. If a robot picks up a tool, he also needs to be able to use it, which means he interacts with the environment via the grasped object. This paper presents a data-driven approach to localize the contact pose.

Sensors

This work uses the BioTac biomimetic tactile sensor, as shown in figure 59, mounted on a Barrett hand with three fingers. This sensor can measure the modalities force, vibration and temperature and has elastomeric skin and fingerprints.

Method

Contact is detected if the high-pass filtered pressure vibration signal passes a threshold. DBSCAN merges jitter into one event. The high (PAC) and low (PDC) frequency pressure vibrations and the electrode signals, for each finger and over several timestamps respectively, form the feature vector. A Neural Network and a Gaussian Process are compared for regression from these features to the position and orientation of the contact point.

As second approach, the contact estimation is seen as a classification problem over the discretized pose space. The same neural network is used, with an output layer modified for classification. One classifier for every pose parameter is trained and this approach is compared to a Support Vector Machines (SVM) with Spatio-Temporal Hierarchical Matching Pursuit Features (see section 12).

Experiments and Results

The ground truth of the contact events for training was obtained by tapping the grasped objects with the rod in figure 63. Around 15100 samples have been collected for 18 objects.

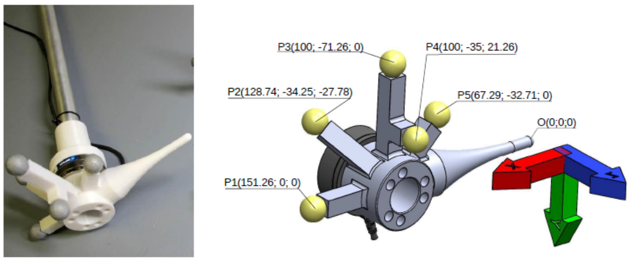

Figure 63: The Vicon marker-based tracking system is used to obtain the ground truth of the contacts. [25]

Results are shown in figures 64 and 65. The large errors for the regression are ”probably caused by ambiguities in the mapping between features and the estimated contact parameters” ([25, p. 5]). It can be seen that the pressure vibration signals generally worsen the performance. The neural network classifier with electrode input only outperforms the other approaches.

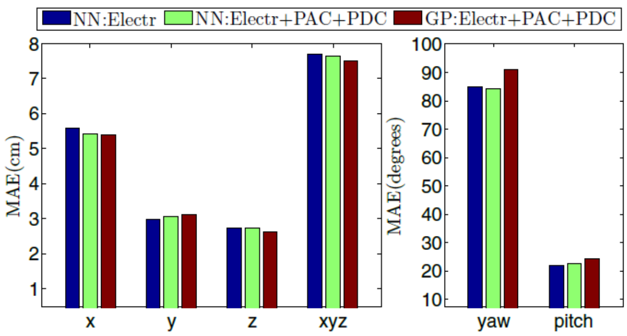

Figure 64: Mean absolute error for both regression methods and different sensor modalities. [25]

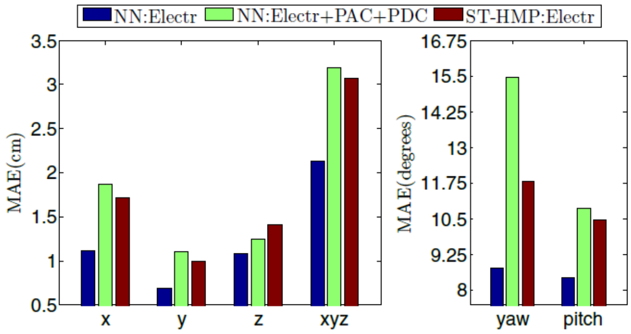

Figure 65: Mean absolute error for both classification methods and different sensor modalities. The Discretizition is 1cm and 5°. [25]

Advantages

- The neural network classification with electrodes only is robust to the size of the time window.
- Wrongly classified discrete pose bins are mainly confused with their direct neighbors and not totally off.

Disadvantages

- A mean absolute error of 2cm is still quite high.
- A minimum of 5 timestamps as input feature still result in relatively low error. With the sampling frequency of 100Hz this still causes a delay of 50ms minimum just for the sensor reading, which makes real-time contact localization more difficult.
- Generalizeability for new objects has not been investigated and could potentially be difficult, since no relationship to the objects properties is taken into account. Every new object needs manual training with many contact points so far.
- The object surface was modified to prevent slip. Slip and other sources of vibration sources could potentially worsen the performance and were not taken into account. Then again they did not contribute to the performance anyway and where not taken into account in the end.
- Only single short taps were considered as contacts.

20 Adaptive tactile control for in-hand manipulation tasks of deformable objects

---

[12] A. Delgado, C. A. Jara, and F. Torres, "Adaptive tactile control for in-hand manipulation tasks of deformable objects," *The International Journal of Advanced Manufacturing Technology*, vol. 91, no. 9-12, pp. 4127–4140, 2017

Motivation And Problem Setting

Many approaches to robot grasping either need models, which limits the diversity of objects that can be grasped, or are learned from experience, which requires lots of sample grasps and limits the manipulation to previous experience. This paper proposes a framework to grasp and manipulate unknown objects with various stiffnesses using an anthropomorphic robot hand.

Sensors and Setup

The anthropomorphic Shadow Hand with 20 degrees of freedom is mounted on a Mitsubishi PA10 robot arm. The Kinect sensor perceives the objects in the workspace. The Tekscan Grip tactile sensor is a resistive sensor and measures the contact pressure for each of the cells on the fingers and palm.

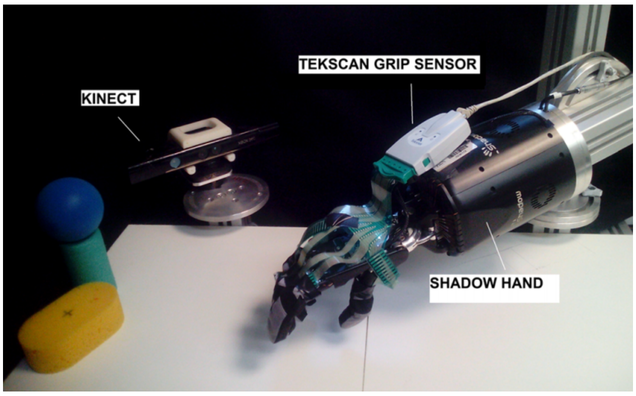

Figure 66: The robot setup. [12]

Method

The Kinect camera perceives the position of the object for an initial grasp. Then the fingers increase the gripping force (torque commands) and measure the displacement (forward kinematics) and the force for each finger, to estimate how deformable the object is. The fingers return to the initial configuration and the now measured force is compared to the intially measured force to distinguish elastic from plastic objects.

Tactile Servo Control controls the position and magnitude of the applied force for each finger, as given by the task planner. Slip is prevented by relating the positional change of the tactile imprints to the applied normal force.

The task planner is a state machine, which executes a sequence of basic actions. A linear dependency between the explored stiffness and the displacement is approximated to estimate the desired contact force.

Experiments and Results

Experiments have been conducted with a plastic bottle, a carton package, a sponge, and a ball, to test the different parts of the framework. Figure 67 shows how the force curve for the squeezing action on the elastic ball returns to the start value, whereas the force curve of the carton package becomes zero after squeezing (figure 68), which indicates a plastic object.

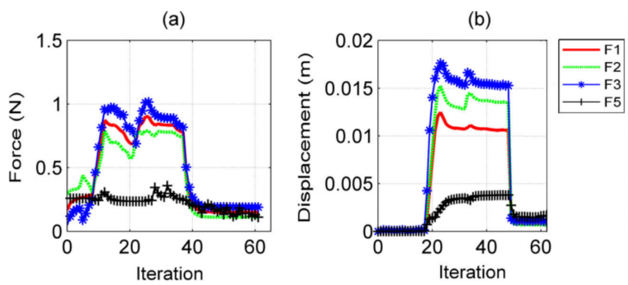

Figure 67: Force and displacement of each finger when the ball is squeezed. [12]

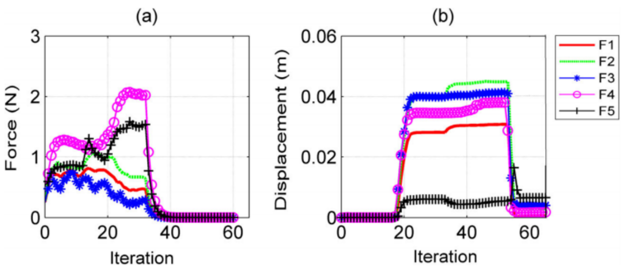

Figure 68: Force and displacement of each finger when the carton package is squeezed. [12]

When a deformable cube is being grasped, the controller reaches an equilibrium after 3.5 s (figure 69). Figure 70 shows how the task planner is used to execute a sequence of actions.

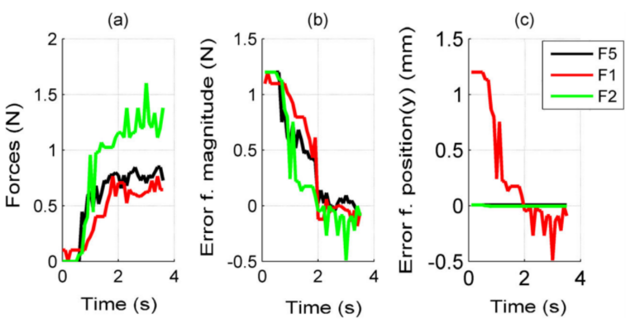

Figure 69: The controller stabelizes a top-down grasp with three fingers after 3.5 s. [12]

Advantages

- An exploratory squeezing action determines the stiffness.
- The state machine task planner allows simple to program sequential actions.
- The rigidity can be estimated separately for every finger.
- The approach could be extended to explore the shape of the object as well.

Disadvantages

- Cumbersome calibration of the resistive tactile sensors is needed.
- The maximum contact force is applied to the object to measure deformability. However, the maximum force could already destroy plastic objects. Instead, the hand could for example squeeze the object multiple times with increasing force to minimize possible damage.
- The controller stabilizes the system relatively slow.

Action	Force computation	Target contact force (N)
Move the object	f_m_{tgt} = MIN_F	0.25
Drag the object with max. deformation = 0.02 m	$f\_m_{tgt} = (d_{max} * f_{av}) / d_{av}$	0.51
Lift the object with max deformation = 0.01 m	$f\_m_{tgt} = (d_{max} * f_{av}) / d_{av}$	0.38
Squeeze the object with target deformation = 0.04 m	$f\_m_{tgt} = (d_{max} * f_{av}) / d_{av}$	1.02

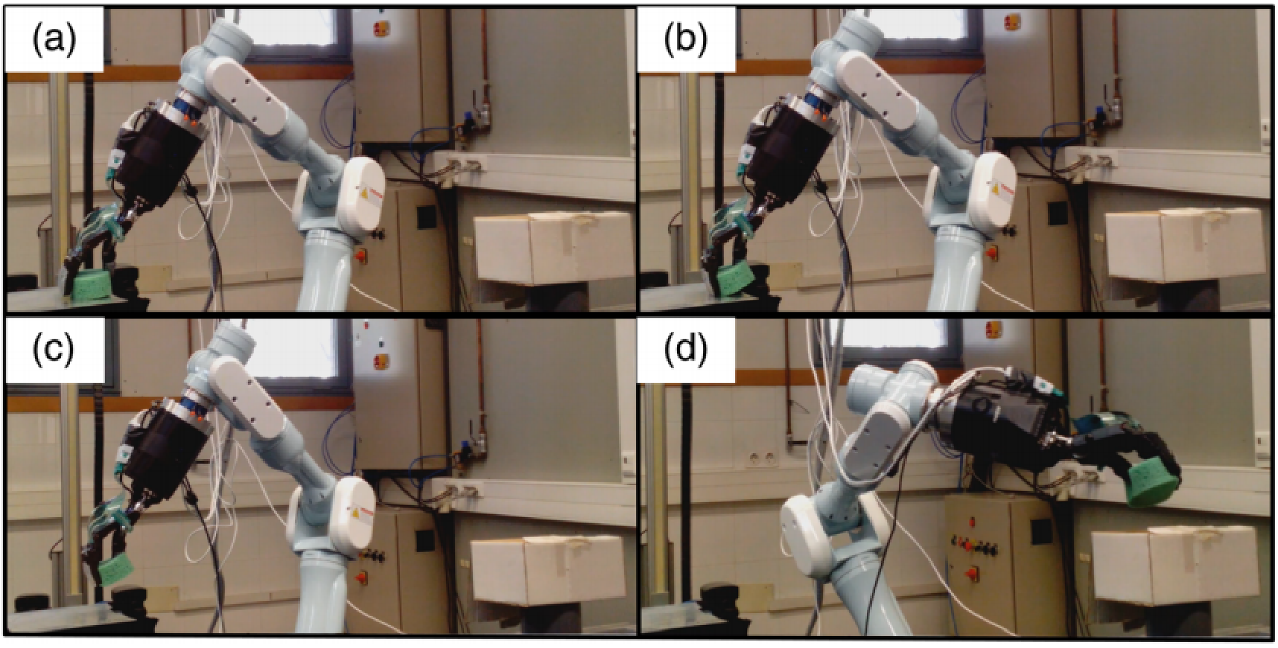

Figure 70: The controller executes the actions with the target force given from the task planner. [12]

References

[1] R. S. Dahiya, G. Metta, M. Valle, and G. Sandini, "Tactile sensing—from humans to humanoids," *IEEE transactions on robotics*, vol. 26, no. 1, pp. 1–20, 2010.

[2] M. Kaboli, D. Feng, K. Yao, P. Lanillos, and G. Cheng, "A tactile-based framework for active object learning and discrimination using multimodal robotic skin," *IEEE Robotics and Automation Letters*, vol. 2, no. 4, pp. 2143–2150, 2017.

[3] J. Bohg, A. Morales, T. Asfour, and D. Kragic, "Data-driven grasp synthesis—a survey," *IEEE Transactions on Robotics*, vol. 30, no. 2, pp. 289–309, 2014.

[4] E. Luberto, Y. Wu, G. Santaera, M. Gabiccini, and A. Bicchi, "Enhancing adaptive grasping through a simple sensor-based reflex mechanism," *IEEE Robotics and Automation Letters*, vol. 2, no. 3, pp. 1664–1671, 2017.

[5] K. Tahara, S. Arimoto, and M. Yoshida, "Dynamic object manipulation using a virtual frame by a triple soft-fingered robotic hand," in *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, pp. 4322–4327, IEEE, 2010.

[6] M. Li, Y. Bekiroglu, D. Kragic, and A. Billard, "Learning of grasp adaptation through experience and tactile sensing," in *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*, pp. 3339–3346, Ieee, 2014.

[7] M. Kaboli, K. Yao, and G. Cheng, "Tactile-based manipulation of deformable objects with dynamic center of mass," in *Humanoid Robots (Humanoids), 2016 IEEE-RAS 16th International Conference on*, pp. 752–757, IEEE, 2016.

[8] W. Yuan, S. Dong, and E. H. Adelson, "Gelsight: High-resolution robot tactile sensors for estimating geometry and force," *Sensors*, vol. 17, no. 12, p. 2762, 2017.

[9] E. Donlon, S. Dong, M. Liu, J. Li, E. Adelson, and A. Rodriguez, "Gelslim: A high-resolution, compact, robust, and calibrated tactile-sensing finger," *arXiv preprint arXiv:1803.00628*, 2018.

[10] R. Calandra, J. Lin, A. Owens, J. Malik, U. Berkeley, D. Jayaraman, and E. H. Adelson, "More than a feeling: Learning to grasp and regrasp using vision and touch," *no. Nips*, pp. 1–10, 2017.

[11] F. R. Hogan, M. Bauza, O. Canal, E. Donlon, and A. Rodriguez, "Tactile regrasp: Grasp adjustments via simulated tactile transformations," *arXiv preprint arXiv:1803.01940*, 2018.

[12] A. Delgado, C. A. Jara, and F. Torres, "Adaptive tactile control for in-hand manipulation tasks of deformable objects," *The International Journal of Advanced Manufacturing Technology*, vol. 91, no. 9-12, pp. 4127–4140, 2017.

[13] J. A. Fishel and G. E. Loeb, "Sensing tactile microvibrations with the biotac—comparison with human sensitivity," in *Biomedical Robotics and Biomechatronics (BioRob), 2012 4th IEEE RAS & EMBS International Conference on*, pp. 1122–1127, IEEE, 2012.

[14] S. Levine, P. Pastor, A. Krizhevsky, and D. Quillen, "Learning hand-eye coordination for robotic grasping with large-scale data collection," in *International Symposium on Experimental Robotics*, pp. 173–184, Springer, 2016.

[15] D. Guo, F. Sun, B. Fang, C. Yang, and N. Xi, "Robotic grasping using visual and tactile sensing," *Information Sciences*, vol. 417, pp. 274–286, 2017.

[16] M. Li, H. Yin, K. Tahara, and A. Billard, "Learning object-level impedance control for robust grasping and dexterous manipulation," in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, pp. 6784–6791, IEEE, 2014.

[17] M. Regoli, U. Pattacini, G. Metta, and L. Natale, "Hierarchical grasp controller using tactile feedback," in *Humanoid Robots (Humanoids), 2016 IEEE-RAS 16th International Conference on*, pp. 387–394, IEEE, 2016.

[18] Q. Li, C. Elbrechter, R. Haschke, and H. Ritter, "Integrating vision, haptics and proprioception into a feedback controller for in-hand manipulation of unknown objects," in *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*, pp. 2466–2471, IEEE, 2013.

[19] H. Dang and P. K. Allen, "Stable grasping under pose uncertainty using tactile feedback," *Autonomous Robots*, vol. 36, no. 4, pp. 309–330, 2014.

[20] Y. Bekiroglu, A. Damianou, R. Detry, J. A. Stork, D. Kragic, and C. H. Ek, "Probabilistic consolidation of grasp experience," in *Robotics and Automation (ICRA), 2016 IEEE International Conference on*, pp. 193–200, IEEE, 2016.

[21] K. Hang, M. Li, J. A. Stork, Y. Bekiroglu, F. T. Pokorny, A. Billard, and D. Kragic, "Hierarchical fingertip space: A unified framework for grasp planning and in-hand grasp adaptation," *IEEE Transactions on robotics*, vol. 32, no. 4, pp. 960–972, 2016.

[22] H. Van Hoof, T. Hermans, G. Neumann, and J. Peters, "Learning robot in-hand manipulation with tactile features," in *Humanoid Robots (Humanoids), 2015 IEEE-RAS 15th International Conference on*, pp. 121–127, IEEE, 2015.

[23] Y. Chebotar, K. Hausman, Z. Su, G. S. Sukhatme, and S. Schaal, "Self-supervised regrasping using spatio-temporal tactile features and reinforcement learning," in *Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on*, pp. 1960–1966, IEEE, 2016.

[24] J. A. Stork, C. H. Ek, Y. Bekiroglu, and D. Kragic, "Learning predictive state representation for in-hand manipulation," in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*, pp. 3207–3214, IEEE, 2015.

[25] A. Molchanov, O. Kroemer, Z. Su, and G. S. Sukhatme, "Contact localization on grasped objects using tactile sensing," in *Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on*, pp. 216–222, IEEE, 2016.

[26] A. Damianou, C. Ek, M. Titsias, and N. Lawrence, "Manifold relevance determination," *arXiv preprint arXiv:1206.4610*, 2012.

[27] R. M. Neal, *Bayesian learning for neural networks*, vol. 118. Springer Science & Business Media, 2012.

[28] A. Zeng, S. Song, K.-T. Yu, E. Donlon, F. R. Hogan, M. Bauza, D. Ma, O. Taylor, M. Liu, E. Romo, *et al.*, "Robotic pick-and-place of novel objects in clutter with multi-affordance grasping and cross-domain image matching," *arXiv preprint arXiv:1710.01330*, 2017.

[29] E. Hyttinen, D. Kragic, and R. Detry, "Learning the tactile signatures of prototypical object parts for robust part-based grasping of novel objects," in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*, pp. 4927–4932, IEEE, 2015.

[30] Z. Su, K. Hausman, Y. Chebotar, A. Molchanov, G. E. Loeb, G. S. Sukhatme, and S. Schaal, "Force estimation and slip detection/classification for grip control using a biomimetic tactile sensor," in *Humanoid Robots (Humanoids), 2015 IEEE-RAS 15th International Conference on*, pp. 297–303, IEEE, 2015.