Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

MIN Faculty
Department of Informatics

State Estimation of an Adaptive 3-Finger Gripper using Recurrent Neural Networks

Yannick Jonetzko

University of Hamburg
Faculty of Mathematics, Informatics and Natural Sciences
Department of Informatics
Technical Aspects of Multimodal Systems

May 28th, 2024

Introduction

Complex $\longleftrightarrow$ Simple

State known $\longleftrightarrow$ State unknown

Expensive $\longleftrightarrow$ Cheap

Motivation

A series of events has led to this work:

- We want to visualize the gripper correctly during teleoperation in VR
- Robotiq Gripper did not work in simulation (Gazebo)
- We mounted tactile sensors to the fingertips

$\rightarrow$ Master thesis on simulation of the gripper

General motivation:

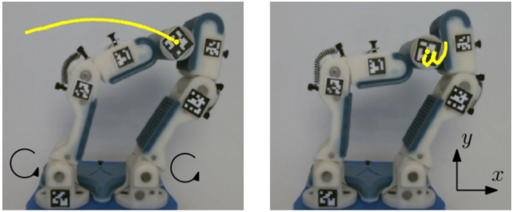
- The real state is interesting for grasp analysis

Related Work

Sintov et al. *"Learning a State Transition Model of an Underactuated Adaptive Hand"*. 2019

Gripper state is described as diffusion map with:

- Object position
- Actuator angles
- Actuator load



The accuracy is measured by the deviation of the object from a trajectory during In-Hand manipulation:

- Average error: 0.53 mm
- Max error: 1.6 mm

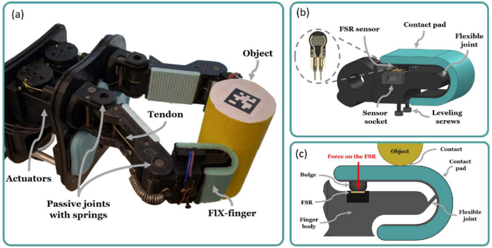
Related Work

Azulay et al. *"Learning Haptic-based Object Pose Estimation for In-hand Manipulation Control with Underactuated Robotic Hands"*. 2022

Hand state and pose estimation of the object with kinesthetic and tactile features:

- Gaussian Processes (GP)
- Fully-Connected Neural Network (FC-NN)
- Long Short-Term Memory (LSTM)

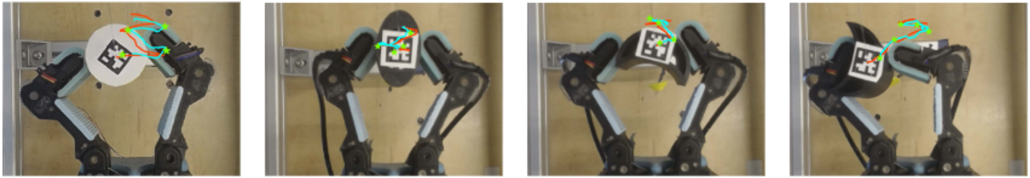
Best results with LSTM



Related Work

Azulay et al. *"Learning Haptic-based Object Pose Estimation for In-hand Manipulation Control with Underactuated Robotic Hands"*. 2022

Model Predictive Control



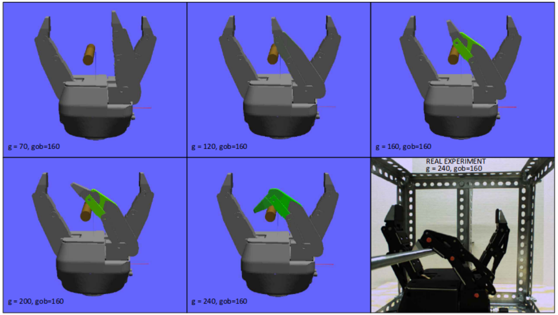
The accuracy is measured by the deviation of the object from a trajectory during In-Hand manipulation:

- Average error: 4.4 mm

Related Work

Franchi et al. *"Technical Report: Use of Hybrid Systems to model the RobotiQ Adaptive Gripper"*. 2014

- Mathematical model of the robotiq gripper
- Ground truth tracking with visual markers
- Accuracy of:
  - Θ1: 0.74 degree
  - Θ2: 2.00 degree
  - Θ3: 1.93 degree

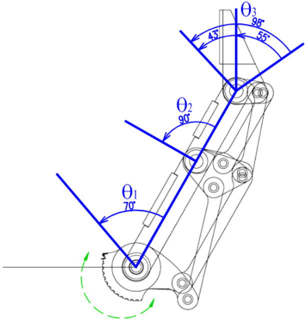


What are we going to do?

- Mount contact sensors to all phalanges
- Build experiment setup with ground truth tracking
- Implement neural network approach
- Conduct experiments
- Compare with Franchi et al. (and master thesis)

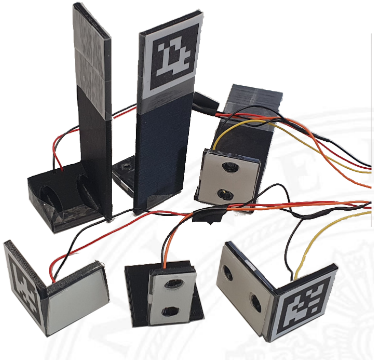
Robotiq 3-Finger Adaptive Gripper

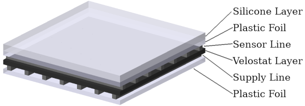
- Three finger
- Adaptive behavior
- Underactuated
  (4 Motors overall)
- 10 DoF
- Exchangeable fingertips and phalanx pads



Tactile Sensors

- Custom built resistive sensors [1]
- One contact per phalanx
- 10 bit readout with 50 Hz
- Normalized between 0 and 1
- Integrated AprilTag holder



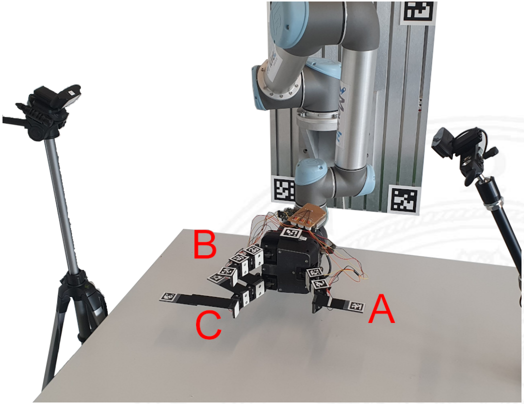


[1] Fiedler et al. *A Low-Cost Modular System of Customizable, Versatile, and Flexible Tactile Sensor Arrays*. 2021

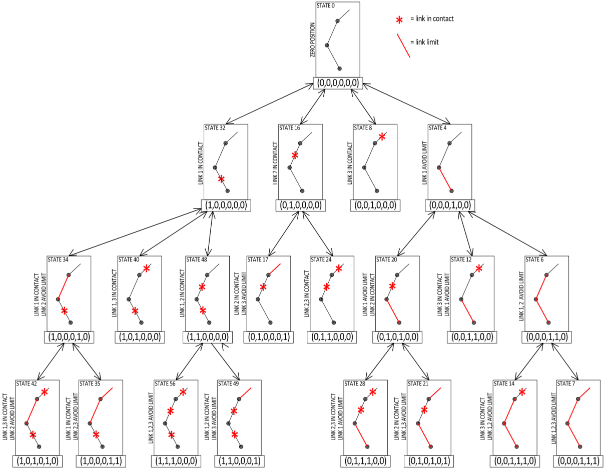
Setup

Ground truth:

- Joint angles calculated with AprilTags
- Tracked by two cameras to reduce occlusion
- Fixed position of end-effector during experiment
- 3 tags per finger
- One base tag



Analytical Approach



Representation of the possible discrete states for one finger[2]

---

[2] Franchi et al. *Technical Report: Use of Hybrid Systems to model the RobotiQ Adaptive Gripper*. 2014

Analytical Approach

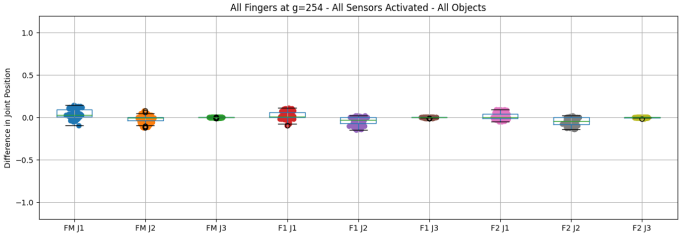
Phase	State tuples	$\Delta\Theta_1$	$\Delta\Theta_2$	$\Delta\Theta_3$	Δg
1	(0,0,0,0,0,0)	$f_1(x,u)$	0	$-f_1(x,u)$	u
1'	(0,0,0,0,0,-1)	$f_1(x,u)$	0	0	u
2	(1,0,0,0,0,0),(0,0,0,1,0,0)	0	$f_2(x,u)$	$-f_2(x,u)$	u
2'	(1,0,0,0,0,-1),(0,0,0,1,0,-1)	0	$f_2(x,u)$	0	u
3	(·,1,0,·,0,0),(·,0,0,·,1,0)	0	0	$f_3(x,u)$	u
4	(·,·,1,·,·,0),(·,·,0,·,·,1)	0	0	0	u

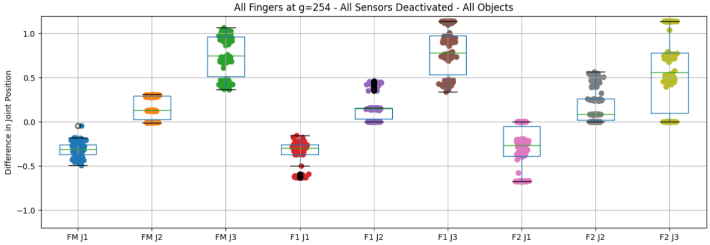
$f_1(x,u) = m_1 u$, with $m_1 = \Theta_{1,max}/140$
$f_2(x,u) = m_2 u$, with $m_2 = \Theta_{2,max}/100$
$f_3(x,u) = m_3(g) u$, with $m_3(g) = \Theta_{3,min} + (\Theta_{3,max} - \Theta_{3,min})/(255 - g)$
With $u \in [-1,1]$ describing the change in g from one time step to the next.

[2] Franchi et al. *Technical Report: Use of Hybrid Systems to model the RobotiQ Adaptive Gripper*. 2014

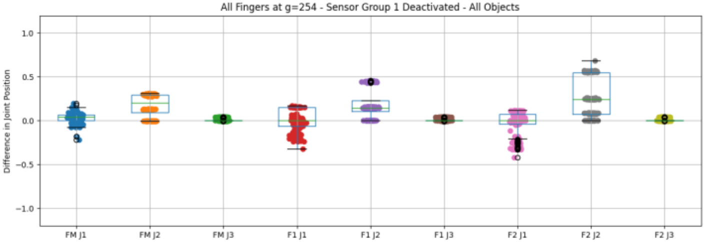
Theresa Alexandra Aurelia Naß, *"Simulation and Joint State Estimation of the Underactuated Robotiq 3-Finger Gripper in Gazebo"*. 2023

- Simulation of gripper in Gazebo
- Fundamental work to this publication
- Estimate the state of the simulated robot
- Test if contact sensors on all 9 phalanges are necessary to accurately estimate the state





All Fingers at g=254 - All Sensors Deactivated - All Objects
Difference in Joint Position
1.0
0.5
0.0
−0.5
−1.0
FM J1
FM J2
FM J3
F1 J1
F1 J2
F1 J3
F2 J1
F2 J2
F2 J3



All Fingers at g=254 - Sensor Group 1 Deactivated - All Objects
Difference in Joint Position
1.0
0.5
0.0
−0.5
−1.0
FM J1
FM J2
FM J3
F1 J1
F1 J2
F1 J3
F2 J1
F2 J2
F2 J3

Shortcomings

- Only works when objects do not move
- Does not distinguish between the fingers

New Approach

Possible approaches:

- IMUs
- Hall effect sensors
- Visual tracking

Solve the problem with neural networks:

- Due to the state dependency, recurrent networks are suitable
- Use Tactile readings and motor positions

Research Questions

Research Question 1

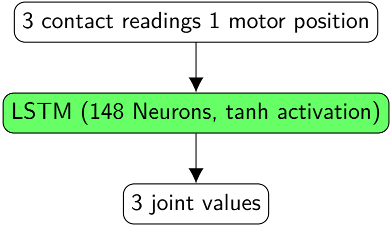
Can a newly designed recurrent neural network approach outperform the existing analytical one for state estimation and compensate object movements?

Research Question 2

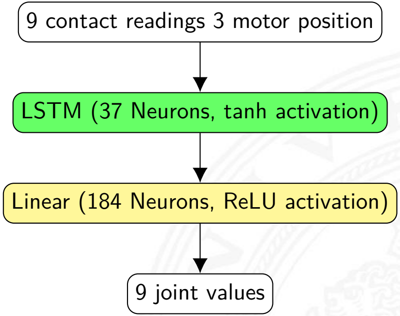
Does an approach that takes all fingers into account at the same time perform better than not differentiating between the fingers?

Network Architecture

Equal Finger Network



Entire Hand Network



The number of layers and neurons, as well as the learning rate and weight decay were determined through hyperparameter optimization.

Experiments

- Record grasp sequences for training
- 20 objects, partly from YCB object set
- Robot is in fixed position grasping from the side
- Each object is grasped three times
  $\rightarrow$ 60 samples for Entire Hand network
  $\rightarrow$ 180 samples for Equal Finger network

Results

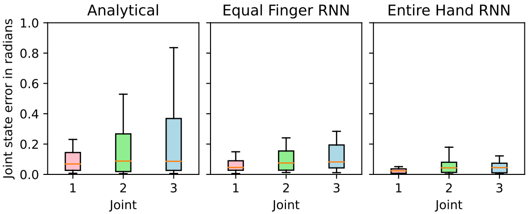
Results: Average Error

Average error in radians

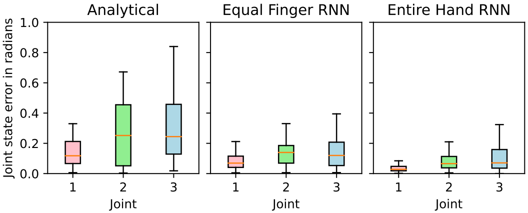
Whole closing motion	Joint 1	Joint 2	Joint 3	Overall
Analytical	0.084	0.145	0.196	0.142
RNN equal finger	0.057	0.093	0.117	0.089
RNN entire hand	**0.026**	**0.048**	**0.047**	**0.040**
End state				
Analytical	0.140	0.275	0.329	0.248
RNN equal finger	0.088	0.149	0.163	0.133
RNN entire hand	**0.045**	**0.085**	**0.107**	**0.079**

Results: Average Error

Whole closing motion:

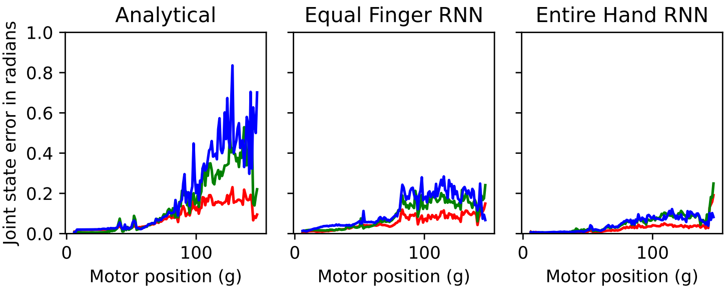


End State:



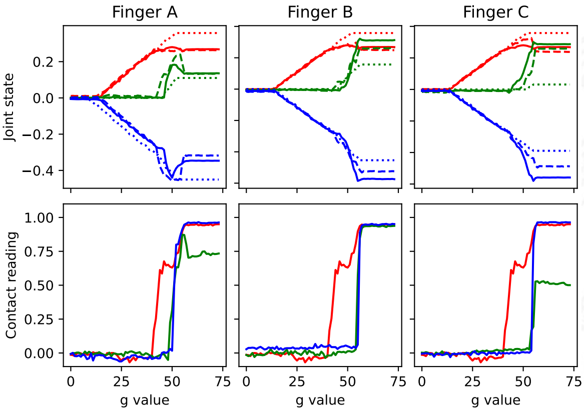
Results: Average Error

Joint difference to ground truth (red: proximal joint, green: middle joint, blue: distal joint)



Results: Complete Grasping Motion

Estimated states and contact readings during a complete grasping motion. (line: Ground truth, dashed: RNN entire hand, dotted: analytical; red: proximal joint, green: middle joint, blue: distal joint)



Discussion

Research Questions:

1. Yes, we can outperform the existing analytical approach
2. Yes, it makes a difference (as expected) to consider the entire hand

Shortcomings:

- Only works during grasps
- The state can not be estimated at any arbitrary time

Future Work

I will probably not continue working on it!

But, possible next steps/future work:

- Use the setup for in-hand manipulation
- Combine with PointNetGPT

Conclusion

- Simulation of the hand
- Implementation of tracking setup
- Implementation of 9 contact sensors on the hand
- Successful state estimation with an accuracy of 2.29 degrees
- We showed that our approach performs better then the state of the art

Thank you for your attention!

Any questions?